WORKSHOP PROCEEDINGS OF THE 8TH INTERNATIONAL CONFERENCE ON INTELLIGENT ENVIRONMENTS

Ambient Intelligence and Smart Environments

The Ambient Intelligence and Smart Environments (AISE) book series presents the latest research results in the theory and practice, analysis and design, implementation, application and experience of *Ambient Intelligence* (AmI) and *Smart Environments* (SmE).

Volume 13

Recently published in this series

Vol. 12. T. Bosse (Ed.), Agents and Ambient Intelligence – Achievements and Challenges in the Intersection of Agent Technology and Ambient Intelligence
Vol. 11. J.C. Augusto et al. (Eds.), Handbook of Ambient Assisted Living – Technology for Healthcare, Rehabilitation and Well-being
Vol. 10. J.C. Augusto et al. (Eds.), Workshop Proceedings of the 7th International Conference on Intelligent Environments
Vol. 9. B. Gottfried and H. Aghajan (Eds.), Behaviour Monitoring and Interpretation – BMI – Well-Being
Vol. 8. R. López-Cózar et al. (Eds.), Workshop Proceedings of the 6th International Conference on Intelligent Environments
Vol. 7. E. Mordini and P. de Hert (Eds.), Ageing and Invisibility
Vol. 6. G. van den Broek et al. (Eds.), AALIANCE Ambient Assisted Living Roadmap
Vol. 5. P. Čech et al. (Eds.), Ambient Intelligence Perspectives II – Selected Papers from the Second International Ambient Intelligence Forum 2009
Vol. 4. M. Schneider et al. (Eds.), Workshops Proceedings of the 5th International Conference on Intelligent Environments
Vol. 3. B. Gottfried and H. Aghajan (Eds.), Behaviour Monitoring and Interpretation – BMI – Smart Environments
Vol. 2. V. Callaghan et al. (Eds.), Intelligent Environments 2009 – Proceedings of the 5th International Conference on Intelligent Environments: Barcelona 2009

ISSN 1875-4163 (print)
ISSN 1875-4171 (online)

Workshop Proceedings of the 8th International Conference on Intelligent Environments

Edited by

Juan A. Botía, *Universidad de Murcia, Spain*
Hedda Schmidtke, *Carnegie Mellon University in Rwanda and USA*
Tatsuo Nakashima, *WASEDA University, Japan*
Mohammed R. Al-Mulla, *Kuwait University, Kuwait*
Juan Carlos Augusto, *University of Ulster, United Kingdom*
Asier Aztiria, *Mondragon Unibertsitatea, Spain*
Matthew Ball, *University of Essex, UK*
Victor Callaghan, *University of Essex, United Kingdom*
Diane J. Cook, *Washington State University, United States of America*
James Dooley, *University of Essex, UK*
John O'Donoghue, *University College Cork, Ireland*
Simon Egerton, *Monash University, Australia*
Pablo A. Haya, *Universidad Autónoma de Madrid, Spain*
Miguel J. Hornos, *Universidad de Granada, Spain*
Eduardo Morales, *National Institute for Astrophysics, Optics and Electronics, México*
Juan Carlos Orozco, *ACELAB Industrial Automation, México*
Otniel Portillo-Rodríguez, *Universidad Autónoma del Estado de México, México*
Alejandro Rodriguez-González, *Universidad Carlos III, Spain*
Oscar Sandoval, *Instituto Tecnológico de Orizaba, México*
Paolo Tripicchio, *Scuola Superiore Sant'Anna, Italy*
Minjuan Wang, *San Diego State University, USA*
Victor Zamudio, *Instituto Tecnológico de León, México*

IOS Press

Amsterdam • Berlin • Tokyo • Washington, DC

ISBN 978-1-61499-079-6 (print)
ISBN 978-1-61499-080-2 (online)
Library of Congress Control Number: 2012941077

Publisher
IOS Press BV
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 687 0019
e-mail: order@iospress.nl

Distributor in the USA and Canada
IOS Press, Inc.
4502 Rachael Manor Drive
Fairfax, VA 22032
USA
fax: +1 703 323 3668
e-mail: iosbooks@iospress.com

PRINTED IN THE NETHERLANDS

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

Preface

For the fourth time, the Intelligent Environments Conference hosted a range of international and multidisciplinary workshops. The conference itself was held for the eighth time in 2012 at the historic World Heritage UNESCO city of Guanajuato (Mexico), following a series of highly successful conferences that were organized in Colchester, UK (2005), Athens, Greece (2006), Ulm, Germany (2007), Seattle, USA (2008), Barcelona, Spain (2009), Kuala Lumpur, Malaysia (2010), and Nottingham Trent University, United Kingdom (2011). The workshops provide a forum for scientists, researchers and engineers from both industry and academia to engage in discussion on newly emerging, or rapidly evolving topics that have yet to reach the level of maturity associated with conferences or journals. This year we are pleased to include the following workshops:

WISHWell'12 The 4th International Workshop on Intelligent Environments Supporting Healthcare and Well-being discussed novel research regarding all aspects of pervasive healthcare solutions. The workshop was organized by Juan Carlos Augusto (University of Ulster, United Kingdom) and John ODonoghue (University College Cork, Republic of Ireland).

AITAmI'12 The 7th workshop on Artificial Intelligence Techniques for Ambient Intelligence aimed at gathering researchers in a variety of AI subfields together with representatives of commercial interests to explore the technology and applications for ambient intelligence. This year's edition presented a special focus on context and situation understanding. The workshop was organized by Asier Aztiria (University of Mondragon, Spain), Juan Carlos Augusto (University of Ulster, United Kingdom), and Diane Cook (Washington State University, USA).

WOLSIE'12 The Workshop on Large Scale Intelligent Environments was organized by James P. Dooley (Essex University), Matthew H. Ball (Essex University), and Mohammad Al-Mulla (Kuwait University). The aim of the workshop was to go towards intelligent environments beyond the size of normal experimental labs.

iDR'12 The First International Workshop on Intelligent Domestic Robots provided a forum to explore the latest research in domestic robots, with particular emphasis on robots that are functionally designed to interact and communicate with people. The workshop was organized by Simon Egerton (Monash University, Sunway Campus, Malaysia) and Eduardo Morales (National Institute for Astrophysics, Optics and Electronics, Mexico).

WOFIEE'12 The 1st Workshop on Future Intelligent Educational Environments was organized by Victor Callaghan (Essex University), Minjuan Wang (San Diego State University), and Juan Carlos Augusto (University of Ulster, United Kingdom). The focus on the workshop is on the increasingly crucial role intelligent environment technology is playing in the delivery of education.

IMIASH'12 The 1st International Workshop on Intelligent Multimodal Interfaces Applied in Skills Transfer, Healthcare and Rehabilitation investigated the strong integration between humans and technology as an opportunity to investigate human behavior through this interaction and to analyze potentialities and advantages when the multimodal interfaces are used in the field of skills transfer, healthcare and rehabilitation. The organizers were Oscar Sandoval (Instituto Tecnolgico de Orizaba, Mexico), Alejandro Rodriguez-Gonzlez (University Carlos III of Madrid, Spain), Paolo Tripicchio (Perceptual Robotics Labortory (PERCRO) Scuola Superiore Sant Anna, Italy), and Otniel Portillo-Rodrguez (University of Mexico State, Mexico).

WoRIE'12 The 2nd Int. Workshop on Reliability of Intelligent Environments was organized by Miguel J. Hornos (University of Granada, Spain), Juan Carlos Augusto (University of Ulster, United Kingdom), and Pablo A. Haya Coll (Escuela Politcnica Superior IIC-UAM, Spain). The event brought together developers and researchers to focus on all aspects of the development process that can contribute to make Intelligent Environment systems safer and to provide methodologies that can increase the confidence in these developments.

AIE'12 The 1st International Workshop on Improving Industrial Automation Using the Intelligent Environments Paradigm was organized by Juan Carlos Orozco (ACE-LAB Industrial Automation, México) and Victor Zamudio (Instituto Tecnológico de León, México). It addressed the oportunities brought by the intelligent environments paradigm to industrial automation. The format of this workshop was different than the rest. It had no paper based presentations in order to foster discussion among participants coming from the industry rather than from an academic context.

As is evident from this list, the IE'12 workshops addressed a diverse set of topics. The formats are equally diverse ranging from regular lectures to practical sessions.

Of course, we would like to take the opportunity to thank everybody who made these proceedings possible. We are especially grateful to our organizing committees, without whom these workshops would not have been possible. We are also grateful to the local organizers who worked tirelessly behind the scenes to make these events a success. Of course we reserve our special thanks for the researchers, who do the work, achieve the advances, set the research agenda and then come to Intelligent Environments workshops to present and discuss their insights. Through these proceedings we hope to be able to share not only these insights but also the unique, inspiring spirit of the IE workshops.

Juan A. Botía, Hedda Schmidtke, Tatsuo Nakashima
General chairs of the IE'12 Workshops

Contents

1st International Workshop on Intelligent Domestic Robots (iDR'12)

1st Workshop on Future Intelligent Educational Environments (WOFIEE'12)

1st International Workshop on Intelligent Multimodal Interfaces Applied in Skills Transfer, Healthcare and Rehabilitation (IMIASH'12)

2nd Int. Workshop on Reliability of Intelligent Environments (WoRIE'12)

This page intentionally left blank

4th International Workshop on Intelligent Environments Supporting Healthcare and Well-Being (WISHWell'12)

Juan Carlos Augusto (University of Ulster, UK)

John O'Donoghue (University College Cork, Republic of Ireland)

This page intentionally left blank

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

Introduction to the Proceedings of WISHWell'12

This section contains the papers presented at the 4th International Workshop on Intelligent Environments Supporting Healthcare and Well-being (WISHWell'12).

The focus of this workshop is to bring together researchers from an array of disciplines to contribute towards the advancement of Healthcare environments (within the hospital and the home) which are extremely complex and challenging to manage from an IT and IS perspective. Such health environments are required to cope with an assortment of patient conditions under various circumstances with a number of resource constraints. Pervasive healthcare technologies seek to respond to a variety of these pressures by integrating them within existing healthcare services. It is essential that intelligent pervasive healthcare solutions are developed and correctly integrated to assist health care professionals in delivering high levels of patient care. It is equally important that these pervasive solutions are used to empower patients and relatives for self-care and management of their health to provide seamless access for health care services.

The program for this edition is enhanced by the invited speaker Jesus Favela (Computer Science Department, CICESE, Mexico) who will make a presentation on *Activity, Behavior and Context: The ABC of Pervasive Healthcare Research.* The focus for this year will be the discussion on how our community can become more integrated and effective.

We deeply appreciate the Intelligent Environment 2012 conference organizers for their help on hosting this event. As a final note, we wish to express our sincere thanks to the Program Committee for their thorough reviews and support along this year and the previous editions.

Juan C. Augusto and John O'Donoghue
Co-chairs WISHWell'12

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-4

Activity, Behavior and Context: The ABC of Pervasive Healthcare Research

Jesus FAVELA
Computer Science Department, CICESE, Mexico

Abstract. Pervasive healthcare aims at applying mobile and ubiquitous computing technologies to address some of the major issues facing healthcare, such as the aging of the population, an increase in the prevalence of chronic diseases and escalating costs. One of the main characteristics of pervasive healthcare systems is their reliance on contextual information such as using the location of a nurse in a hospital ward to provide her with information relevant to the task at hand, or reminding a patient to take his medication when he returns home and his pillbox has not been opened during the day. Advances in activity recognition have made possible a second generation of pervasive healthcare solutions using activity-aware computing to identify the activities of daily living being performed by an elder and assist her with the task, or to monitor the intensity of the physical activity performed by a patient undergoing treatment. Moving forward this paper describes a new wave of pervasive healthcare applications, which are becoming possible due to progress in the estimation of user behavior. These behavior-aware systems can be used to measure the effects of medication in patients suffering from mental disorders, provide feedback in rehabilitation therapies or to perceive abnormal behaviors that provide early signs of an ailments such as Alzheimer's disease.

Keywords. Behavior-aware computing, pervasive healthcare, activity recognition

Introduction

With the aging of the population, a significant increase in the prevalence of obesity and chronic diseases, and escalating costs, the provision of healthcare services faces major challenges in the coming years. These challenges demand patients to play a more active role in the management of their disease and rehabilitation process, and from the general population to be more conscious of their health habits and wellbeing. Recent advances in ICT, and particularly in mobile and ubiquitous computing, promise to address some of these issues by assisting in the monitoring of patients for disease management, providing opportune information to caregivers and healthcare professionals for timely treatment, and deploying the core infrastructure for people to have information that allows them to make healthier decisions and play a more active role in their own healthcare and wellbeing.

Pervasive healthcare refers to the application of pervasive or ubiquitous computer technologies for the diagnosis, treatment and prevention of disease, illness and other physical and mental impairments. This area has become one of the most active fields of research and application within ubiquitous computing. Advances in this field include smart home technologies that assist elders in performing activities of daily living [1]; mobile wellbeing applications that use the sensors in mobile phones to measure energy expenditure and persuade users to improve their eating habits [2, 3]; and pervasive hospital environments that provide physicians with the latest laboratory results when approaching the bed of a patient [4].

In this paper I describe the evolution of Pervasive Healthcare considering three main periods characterized by the underlying information inferred and used by these systems: Context, Activity and Behavior, arguably the main building blocks of pervasive healthcare systems. The first period is characterized by the use of contextual information directly measured from sensors or inferred from low-level data. An example of this first-generation of pervasive healthcare solutions is the use of location information in a hospital information system to provide hospital staff with information relevant to the task at hand. The second generation of pervasive healthcare research has been made possible by advances in activity recognition. Examples of activity-aware applications include assisting elders with activities of daily living or detecting activity periods to monitor the amount exercise performed by a user. These applications rely on the estimation of the user's activity, such as whether the person is walking versus running.

We are at the dawn of a new wave of pervasive healthcare solutions that can take advantage of behavior information inferred from users and communities. An example of this type of application is recognizing wandering behavior in a person suffering from dementia to notify a caregiver. I describe some potential healthcare applications of behavior-aware computing and some of the main technical challenges that need to be tackled in pervasive healthcare research in order to realize these solutions. Both activity and behavior can be considered context, and thus, applications that adapt to changes in these parameters are, strictly speaking, context-aware. Yet, there has been a clear evolution over the last few years from ubiquitous applications that use lower-level context, such as location, towards those that rely the recognition of activities, such as walking, and gradually towards applications that depend on the recognition of human behaviors, such as wandering.

1. Context-aware computing in healthcare

Context-aware computing refers to an application's ability to adapt to changing circumstances and respond based on the context of use [5]. A context-aware system has the capacity to perceive and capture the world that surrounds the user, and to adapt its behavior to provide information and services that are useful and relevant to that place and time [6]. In this sense, context is any information that can be used to characterize the situation of a given entity, which can be a person, place or object that is considered relevant for the interaction between a user and the application [6]. Context can be characterized at different levels of abstraction. Low-level context is deduced directly from sensors; were a 'sensor' refers not only to sensing hardware but also to every data source that provides usable contextual information, such as temperature, luminosity, sound, movement, air pressure, environmental humidity, user profile, preferences and schedule. Whereas high-level context is abstract and inferred from low-level information, for example, activity, location or user mood [7].

Context is difficult to use for several reasons [8]. First, capturing low-level contextual information requires the use of sensors and computing devices. Context must be abstracted to make sense for the application. For instance, the ID of a mobile user must be abstracted into the user's name or role. Finally, context is dynamic, i.e., a location-based system must update its display as the user moves, which require tracking the user's location by gathering information from multiple sensors, and using techniques that estimate the user's location or guess the route that the user will follow.

The design and development of context-aware applications faces two major challenges. First, it is important to identify which information is worth capturing and how contextual information could be automatically detected to support the user's tasks. Second, it is essential to provide adequate methods for the presentation of information and services in support of the user's ongoing activities.

The incorporation of robust location estimation technology in mobile phones has made location-based systems ubiquitous. Location information can be used to provide services and information to mobile users. For instance, in a hospital setting where highly mobile hospital staff executes activities depending on their location, location-based systems can provide opportune information, such as, access to patient's records when the user is near to the patient's bed or inform the user of the location of an electrocardiograph.

Context-aware computing has been applied to a variety of healthcare environments, such as supporting communication among hospital workers [9, 10], aiding patients to adhere to their medication regimes [11], or alerting about risks of drug-drug interaction [12].

To illustrate the use of context-aware computing in healthcare consider the smart hospital environment (iHospital) described in [13, 14]. The iHospital is a vision of a highly interactive workplace, where hospital staff can access relevant medical information trough a variety of heterogeneous devices, and can collaborate with colleagues taking into account contextual information. To realize this vision several pervasive healthcare applications were developed that altogether use contextual information to: provide awareness of the location of hospital staff and artifacts [10]; support collaboration and opportunistic encounters through context-aware communication [15] (see Figure 1); adapt and personalize information to the user's or artifact's context (e.g., role, location, laboratory studies status) [15]; and support the peripheral monitoring of patients [17].

Key to the development of the iHospital is accurate and inexpensive indoor location estimation. This was achieved using a backpropagation neural network that maps the strength of radiofrequency signals received by mobile devices to the location of the user carrying the device [18]. The approach uses information about the neighbors surrounding the location of the user to simulate previous time instant guesses and alleviate the hopping trajectories of users. The results obtained using this approach average 1.3m distance error during continuous motion, which favorably compares with the 2m average error obtained when not using information of neighbor locations.

Figure 1. A medical intern receives a context-aware message and collaborates with a nurse on a public display.

2. Activity-aware computing for healthcare

Activity-aware technologies aim at allowing "smart environments" to respond proactively to the needs and intentions of their users by automatically inferring the activity being performed by the user [19]. Activity-aware applications thus need intelligent capabilities to be adaptive and reactive to users' tasks and learn from user's actions in order to provide opportune and reliable services based on their goals.

Human activities are complex and dynamic. Activities at different levels of granularity can be performed concurrently and they might be interwoven. Hence, the development of activity-aware applications requires appropriate representations of computational activities that are relevant to the services provided by the system. An additional challenge for the development of activity-aware computing is the development of robust approaches for activity recognition.

COACH (Cognitive Orthosis for Assisting with aCtivites in the Home) is an example of an activity-aware application aimed at assisting in the completion of activities of daily living [1]. The system aims at assisting people who suffer from dementia perform tasks, such as handwashing. The system uses a video camera to identify the position of the hands and task objects such as the soap or towel, to infer the state of the activity and provide appropriate prompting. A planning system determines the current step of the activity and decide the appropriate response of the system. Knowledge about the user's abilities is used to adapt the prompting.

Activity recognition is currently one of the more active areas of research in ubiquitous computing. Numerous projects have explored algorithms, methods and devices for the automatic recognition of users' activities. These approaches to inferring people's activities range from estimating low-abstraction activities (such as chewing or sitting) to estimating more complex activities such as whether the person is leaving for work later than usual. For instance, detecting the manipulation of certain objects using RFID tags has been used to estimate activities of daily living (ADL), such as preparing a drink or taking a medication [20]. The SEER (Selective Perception Architecture for Activity Recognition) system used a hidden markov model (HMM) to estimate activities at a higher level of abstraction (such as whether a user is attending a meeting) [21]. The inputs to SEER include audio, video, and computer interactions captured through sensors distributed in an office environment.

One of the trends pushing activity recognition research is the significant increase in available data, due to the incorporation of sensors such as GPS, accelerometers and microphones, in mobile phones. The ubiquity of these devices also makes them ideal platforms for the deployment of pervasive healthcare solutions. Relevant examples to healthcare include the use of accelerometer data to reliably determine if a person is standing, walking or running, and to estimate energy expenditure [22]. The automatic estimation of these activities is relevant for wellness applications that measure periods of moderate or intense activity. In addition, the audio captured by microphones in mobile devices can also be used to detect the semantic location of a person (e.g. at a restaurant, on a bus) and infer activities such as "eating", "transportation", "conversation", etc. This field of research known as computational auditory scene recognition can also be used to measure the intensity and frequency of social interaction [23, 24]. Data from multiple sources can be used to achieve higher recognition precision.

Decision trees and Hidden Markov models have been used to combine evidence from heterogeneous sources. Figure 2. shows an example of a two-level HMM used to recognize hospital staff performed activities [25]. The model was trained with data gathered from close to 200 hours of observation of hospital activities and interactions. The data includes information about the location of the hospital worker, the artifacts s/he uses to perform a task, and the roles of the people with whom s/he interacts (e.g., nurse, patient) [26]. The model achieved an accuracy of 92.6% distinguishing among 6 different activities such as patient care, clinical case assessment, and information management. Interestingly, the HMM attained better performance than human experts who were asked to recognize the activity with the same data.

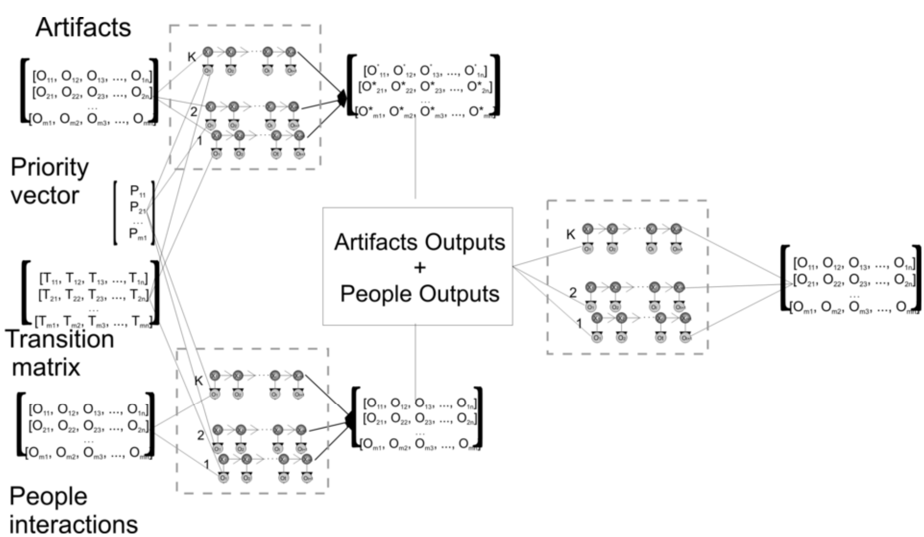


Figure 2. Architecture of HMM used to estimate the activities performed by hospital staff [25].

Different activity-aware applications tolerate different recognition accuracies for providing adequate activity-aware responses. An approximate estimation of energy expenditure using accelerometer data might be more than appropriate if the alternative is self-report on the amount of daily exercise performed by a user. Yet, some activities might be more difficult to classify than others. For instance, an accelerometer worn at the waist greatly underestimates activity counts during cycling. With regards to the recognition of the activities performed by hospital workers an application might be more tolerant to errors while inferring the current availably of a nurse in order to decide whether to interrupt her or no, than to errors registering that a nurse gave a medication to a patient. The accuracy achieved by the HMM described above might be adequate for the first application, but not for the second one. Thus, the selection of the activity recognition strategy should be closely associated to the design of the activity-aware application.

3. Towards behavior-aware computing

Behavior is the response of a system or organism to various stimuli in its environment. In sociology, behavior is considered as not being directed at other people, thus being the most basic of human action. Analyzing behavior is relevant to healthcare since different pathologies are a consequence of behavior or are shaped by it. For instance, unhealthy eating habits and/or lack of exercise lead to obesity, which in turn might lead to hypertension and diabetes. On the other hand, mental disorders such as Alzheimer's disease, Parkinson's disease or autism are characterized by behaviors such as wandering, hand tremor, or failure to socialize. Indeed, the recognition of these behaviors often leads to their early detection and diagnosis.

Over the last few years there have been important advances in automatic behavior recognition. ALEX Pentland has explored the use of "honest signals" that can to reliably predict human behavior [27]. These signals, which are often perceived unconsciously by a person, can be reliably estimated using sensors such as accelerometers and microphones. He has proposed four honest signals for which he has found significant correlation with a variety of behaviors. These signals are: *mimicry*, the reflective copying of one person by another; *activity*, the movement that indicates interest and excitement; *influence*, the extent to which a person matches the other's speech; and *consistency*, the fluidity of speech and motion, which is perceived by others as expertise [27]. Additional research has explored the feasibility of recognizing other behaviors relevant to healthcare such as stress [28], mood [29], or personality traits [30]. Conscientiousness, for instance, has been found to be linked to longevity and to have wide-ranging effects on health-relevant activities [31]. Thus, a behavior-aware medication reminder might be designed to use different persuasion mechanisms for conscientious patients than for a user who tends to be careless.

In the rest of this section I describe some of the challenges facing the development of behavior-aware computing and potential applications in healthcare and wellbeing.

3.1. Challenges in behavior-aware computing

The research challenges of behavior-aware computing are similar to the ones faced in the development of context and activity-aware applications. These are: indentifying meaningful behavior-aware applications; representing behavior information; determining what behavior information is relevant to the application and how it should adapt to changes in behavior; and recognizing behavior. Behavior-aware computing research can be motivated by application requirements (top-down), such as detecting problematic or abnormal behavior in persons with dementia to notify their caregiver, or by technological advances (bottom-up) that make feasible inferring certain types of behavior.

3.2. Applications of behavior-aware computing in healthcare

There are numerous healthcare applications for systems that can adapt to changes detected in the behavior of the user.

3.2.1. Detecting influenza epidemics

A study that collected data from 70 students in a University campus for a period of 4 months provides evidence that data obtained from mobile phones could be used to detect changes in behavior associated to symptoms of influenza [32]. In particular, subjects that suffered symptoms of influenza experienced lower activity and entropy levels and a decrease in communication frequency at different times of the day. Detecting these changes in behavior could be used to detect when a person has a flue.

Understanding the dynamics of social networks can also be used to predict the spread of diseases. The Reality Mining project has shown how social relationships can be automatically uncovered with information gathered from the sensors in mobile phones [33]. Research being conducted to predict mobility patterns and known locations can be used by public health authorities to implement appropriate responses to epidemic breaks.

3.2.2. Risky or abnormal behaviors: Wandering

It has been estimated that around 60 percent of those suffering from dementia wander at some point, and half of those not found within a day suffer serious injuries or death. Wandering behavior is unique to a given person in a particular context or situation [34]. The factors that affect wandering include those related to the background of the person such as health status, personality, cognitive decline, and socio-demographic characteristics, as well as factors related to the context of the situation including personal and physical needs and social aspects of the environment [34].

A straightforward approach to detect wandering behavior would be to recognize random walking from location data. This would work for some instances of wandering. However, some of the most problematic episodes of wandering are due to direct walking, when the wanderer aims to the place were they used to work or live, without feeling lost nor disoriented.

LaCasa (Location and Context-Aware Safety Assistant) is an early prototype of a system that incorporates background and contextual information to detect wandering behavior and provide appropriate assistance to the person with dementia and/or her caregiver [35]. The system is based on a partially observable Markov decision process, or POMDP, a stochastic model of a temporal dynamic process that can handle partial obsevability. The model includes a set of preferences encoded as utilities for the persons involved, and computes a policy of assistance, such as prompting the user or calling the caregiver, that optimizes over these preferences given the state and the dynamics of the world in the long term. The data used by the POMDP is obtained thru opportunistic sensing using a mobile phone (e.g. known locations and common routes), as well as background information gathered from interviews with the person with dementia and her caregiver.

3.2.3. Early diagnosis of dementia or frailty

Frailty is as a multidimensional syndrome, which involves loss of physical and cognitive reserves and ultimately an increase of vulnerability [36]. Frailty is commonly assessed with a collection of instruments and tests. Several of these tests involve self-reporting to questions such as: "during the last week how many days you walked continuously for at least 10 minutes?" or "how frequently you speak with your friends?". More accurate responses to these questions could be obtained from the analysis of data obtained from the sensors available in mobile phones. For instance, voice activity detection could be used to assess the strength of the social network of an older adult or the nature of a conversation [24]. Further analysis of the audio signal and the amount of time each participant talks could be used to infer depression or anxiety.

Fatigue is another factor associated to frailty. Unobtrusively monitoring gait speed over a period of time, for instance in a route frequently followed by the user, could provide early evidence of fatigue. In [37] the use of ambient videogames is proposed to measure muscle fatigue, which has been associated with sensation of fatigue, which is considered a predictor of disability and even mortality.

4. Conclusions

Inferring contextual information and recognizing the activity being performed by a person have been two of the cornerstones of ubiquitous computing research. Numerous advances in these areas have been motivated by applications in healthcare and wellbeing, giving rise to the field of Pervasive Healthcare. A new wave of pervasive healthcare research is being made possible by the ubiquity of mobile phones that incorporate sensing technology and advances in behavior recognition. Thus, making behavior, along with context and activity, the third building block of pervasive computing research.

References

[1] J. Boger, J. Hoey P. Poupart, C. Boutilier, G. Fernie, A. Mihailidis. A planning system based on Markov decision processes to guide people with dementia through activities of daily living. *IEEE Trans Inf Technol Biomed*. 10(2):323-33. 2006.

[2] K.H. Connelly, A.M. Faber, Y. Rogers, K.A. Siek, T. Toscos. Mobile applications that empower people to monitor their personal health. *e & i Elektrotechnik und Informationstechnik*, 123(4):124–128, 2006.

[3] K.A. Siek, *et al*. When Do We Eat? An Evaluation of Food Items Input into an Electronic Food Monitoring Application. *In Pervasive Healthcare 2006*, pp. 1–10, 2006.

[4] M.D. Rodríguez, J. Favela, E.A. Martínez, M..A. Muñoz. Location-aware access to hospital information and services. *IEEE Transactions on Information Technology in Biomedicine* 8(4): 448-455. 2004.

[5] A.K. Dey. Understanding and Using Context. *Personal and Ubiquitous Computing* Vol. 5, No.1, 4-7. 2001
[6] G.D. Abowd, M., Ebling, G., Hunt, H., Lei & M-H. Gellersen, H.-W. Context-aware computing. *IEEE Pervasive Computing*, 1(3), 22-23. 2002.
[7] P. Prekop, M. Burnett. Activities, Context and Ubiquitous Computing. *Computer Communications*, Vol 26, No. 11, 1168-1176. 2003.
[8] A.K. Dey. Providing Architectural Support for Building Context-Aware Applications. PhD. Thesis. Georgia Institute of Technology, 2000.
[9] J.E. Bardram. Applications of context-aware computing in hospital work: examples and design principles, In *Proc. of the ACM symposium on Applied computing*. 2004, pp. 1574 - 1579.
[10] M. Munoz, *et al*. Context-aware mobile communication in hospitals. *IEEE Computer*. Vol. 36, No. 8, pp. 60–67. 2003
[11] D. Asai, *et al.* Context-aware reminder system to support medication compliance. *Proc. IEEE SMC*, Anchorage, Alaska 2011: 3213-3218
[12] J.D. Duke and D. Bolchini. A successful model and visual design for creating context-aware drug-drug interaction alerts. *AMIA Annu Symp Proc*. 2011;2011:339-48.
[13] J. Bardram, H. Baldus, J. Favela. Pervasive Computing in Hospitals, Chapter No. 3, In *Pervasive Computing in Healthcare*. Bardram, J., Mihailidis, A., Wan, D. (eds). CRC Press. Pp. 49-77. 2006.
[14] J. Favela, A.I. Martínez, M.D. Rodríguez, V.M. González. Ambient Computing Research for Healthcare: Challenges, Opportunities and Experiences. *Computación y Sistemas*.Vol. 12, No. 1, pp. 109-127. 2008.
[15] A.D. Mejía, J. Favela, A.L. Moran. Understanding and Supporting Lightweight Communication in Hospital Work. *IEEE Trans. on Information Technology in BioMedicine*, Vol. 14, No. 1, pp 140-146. 2010.
[16] J. Favela, M.D. Rodríguez, A. Preciado, V.M. Gonzalez. Integrating Context-aware public displays into a Mobile Hospital Information System, *IEEE Trans. on Information Technology in BioMedicine*. 2004. Vo., 8, No. 3. pp. 279-286.
[17] M. Tentori, J. Favela, D. Segura. Chapter 8: Monitoring hospital patients using ambient displays. In *Mobile Health Solutions for Biomedical Applications*. Medical Information Science Reference. 2009.
[18] L.A. Castro and J. Favela. Reducing the Uncertainty on Location Estimation of Mobile Users to Support Hospital Work. *IEEE Transactions on systems, man and cybernetics--Part C*, Vol. 38, No. 6: 861-866. 2008.
[19] M. Tentori, J. Favela. Activity-aware computing for healthcare. *IEEE Pervasive Computing*, 2008. 7(2): p. 51-57.
[20] M. Philipose, K.P. Fishkin, M. Perkowi. Inferring Activities from Interactions with Objects, *IEEE Pervasive Computing*, vol. 3, No. 4, 2004, pp. 50–57.
[21] N. Oliver, A. Garg, E. Horvitz. Layered Representations for Learning and Inferring Office Activity from Multiple Sensory Channels, *Computer Vision and Image Understanding*, vol. 96, no. 2, 2004, pp. 163–180.
[22] M. Mathie, A. Coster, N. Lovell, B. Celler. Detection of daily physical activities using a triaxial accelerometer. *Medical and Biological Engineering and Computing*. Vol. 41, No. 3, 296-301. 2003.
[23] V. Peltonen, *et al.*, Computational Auditory Scene Recognition, In *IEEE International Conference on Acoustics, Speech, and Signal Processing*. - Orlando, Florida, USA ,2002. - May 13 - 17.
[24] D. Zhang, et al., Modeling Individual and Group Actions in Meeting with Layered HMMs, *IEEE Trans. Multimedia*, vol. 8, no. 3, 2006, pp. 509–520.
[25] D. Sanchez, J. Favela, and M. Tentori, Activity recognition for the Smart Hospital. *IEEE Intelligent Systems*, 2008. 23(2): p. 50-57.
[26] J. Favela *et al*., Activity Recognition for Context-Aware Hospital Applications: Issues and Opportunities for the Deployment of Pervasive Networks, *Mobile Networks and Applications*, vol. 12, no. 2-3 2007, pp. 155–171.
[27] A. Pentland. Honest signals: how social networks shape human behavior. In *Proc. ACM Multimedia 2011: 583-584*. 2011
[28] J. Hernandez, R.R. Morris, R.W. Picard. "Call Center Stress Recognition with Person-Specific Models," In Proceedings of the Affective Computing and Intelligent Interaction, Memphis, October 9-12, 2011.
[29] A. Matic, A. Papliatseyeu, A. Gabrielli, V. Osmani, O. Mayora-Ibarra. Happy or moody? why so? In: Proc. Ubi-Health (2010)
[30] G. Chittaranjan, J. Blom, D. Gatica-Perez. Who's Who with Big-Five: Analyzing and Classifying Personality Traits with Smartphones. In Prox. of ISWC 2011: 29-36
[31] H.S. Friedman, *et al*. Childhood conscientiousness and longevity: Health behaviors and cause of death. Journal of Personality and Social Psychology, Vol 68(4): 696-703. 1995.

[32] A. Madan, M. Cebrián, D. Lazer, A. Pentland. Social sensing for epidemiological behavior change. In *Proc*. *UbiComp 2010*: 291-300. 2010

[33] N. Eagle, A. Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing* 10(4): 255-268. 2006.

[34] D. Algase, C. Beel-Bates, E., Beattie. Wandering in Long-Term Care. *Annals of Long-Term Care: Clinical Care and Ageing* 2003. Vol 11, No. 1, pp 33-39.

[35] X. Yang, J. Hoey, E. Quintana, and J. Favela. LaCasa Location and Context-Aware Safety Assistant. *Pervasive Healthcare 2012*. San Diego, CA., May 2012.

[36] K. Rockwood, *et al*., A global clinical measure of fitness and frailty in elderly people, *CMAJ*, vol. 173, no. 5, pp. 489-495, 2005.

[37] I. Zavala-Ibarra and J. Favela. Ambient videogames for assessing muscle disease related to aging. *Pervasive Healthcare 2012*. San Diego, CA., May 2012.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-14

System Multi-Agents for the study of basic emotions and their effect on personality

Cinthya SOLANO ARAGÓN [a, 1], Arnulfo ALANIS GARZA [b] Miguel Ángel LÓPEZ [b], María del Rosario BALTAZAR [c], Víctor Manuel ZAMUDIO [c]

[a] *PhD student in Computer Science Tijuana Institute of Technology Tijuana, Baja California, México*

[b] *Department of Systems and Computer Tijuana Institute of Technology Tijuana, Baja California, México*

[c] *Research Division, Technology Institute of Leon, Tijuana BC, México*

Abstract— We report the development of an architecture for a Multi-Agent Social Sound which is based on the human personality in the search for empathy among persons focused on mobile robotics for evaluation of working groups in different controlled environments, focusing specifically on how they affect emotions to personality in a situation before making the stimulus act (emotion), taking on the work in the psychology of personality which Raymond Cattell[13] developed a tool called the 16 PF for the study of personality in a short time, emotions using the methodology of human primary emotions of Paul Ekman[12] who lets us know that there are six primary emotions Joy, Surprise, Anger, Fear, disgust and sadness, the paradigm of Intelligent Agents.

Keywords. Personality; Emotions; Multi-Agent System; Intelligent Agents; Robotic; Empathy

Introduction

Since the beginning of biologically inspired robots, researchers have been fascinated with the possibility of robots that can behave socially, ie they interact with each other and why not with humans.

The researchers began to apply principles like stigmergía was first described by Pierre-Paul Grasse as a way to explain how social insect societies collectively produce complex behavior patterns and physical structures, while each individual appears to work on their own. Deneubourg and his colleagues pioneered the first experiments in simulation and physical stigmergía in the early 90's. [7]

K. Dautenhahn and A. Billard, in its publication "Bringing up robots or-the psychology of socially intelligent robots: from theory to implementation" proposed the following definition: Social robots are agents that are part of a diverse group: a society of robots or humans. They can recognize each other and participate in social interactions, which have histories (perceive and interpret the world in terms of their own experience), and explicitly communicate and learn from each other. An agent is anything that perceives its environment through sensors and acting upon that environment through effectors [10]. Castelfranchi [4] defines social intelligence as part of AI systems and models designed for reasoning and acting in the unpredictable world with limited and determined in real time and with limited resources. Thus, socially intelligent agents are agents that not only from the standpoint of observers behave socially, but also are able to recognize and identify other agents (or humans) and establish and maintain relationships with them [10].

1. State of Art

In this section definitions are taken from the literature, necessary for the understanding of work and related works that provide something similar to our proposal.

In the section 1.1 and 1.2 are two works which have worked to apply them primary and secondary emotions in virtual human MAX and EMMA, the first project started with MAX can show emotions playing a card game with the computer or even a human being and in the second project create the second virtual human to interact with MAX and EMMA to instil emotions empathy between them.

1.1. Affective computing with primary and secondary emotions in a virtual human

Shows the WASABI ([W]ASABI [A]ffect [S]imulation for [A]gents with [B]elievable [I]nteractivity) Affect Simulation Architecture, in which a virtual human's cognitive reasoning capabilities are combined with simulated embodiment to achieve the simulation of primary and secondary emotions. In modeling primary emotions we follow the idea of "Core Affect" in combination with a continuous progression of bodily feeling in three-dimensional emotion space (PAD space), that is subsequently categorized into discrete emotions. In humans, primary emotions are understood as onto-genetically earlier emotions, which directly influence facial expressions. Secondary emotions, in contrast, afford the ability to reason about current events in the light of experiences and expectations. By technically representing aspects of each secondary emotion's connotative meaning in PAD space, we not only assure their mood-congruent elicitation, but also combine them with facial expressions, that are concurrently driven by primary emotions. [3]

1.2. Empathy/Based Emotional Aligment for a Virtual Human: A Three/Step Approach

Allowing virtual humans to align to others' perceived emotions is believed to enhance their cooperative and communicative social skills. In our work, emotional alignment is realized by endowing a virtual human with the ability to empathize. Recent research shows that humans empathize with each other to different degrees depending on several factors including, among others, their mood, their personality, and their social relationships. Although providing virtual humans with features like affect, personality, and the ability to build social relationships, little attention has been devoted to the role of such features as factors modulating their empathic behavior. Supported by psychological models of empathy, we propose an approach to model empathy for the virtual human EMMA—an Empathic MultiModal Agent—consisting of three processing steps: First, the Empathy Mechanism by which an empathic emotion is produced. Second, the Empathy Modulation by which the empathic emotion is modulated. Third, the Expression of Empathy by which EMMA's multiple modalities are triggered through the modulated empathic emotion. The proposed model of empathy is illustrated in a conversational agent scenario involving the virtual humans MAX and EMMA. [6]

These projects give us a perspective of what exists in projects related to ours but we will use not only emotions but also personality profiles for the relationship between two entities (empathy).

1.3. Cattel's personality theory

Cattell argues that personality will predict what a person in a given situation. The concept of Cattell's personality traits are talking about (for Cattell trait is the tendency to react relatively constant) as a frame of mind that comes from the consistent observation of a behavior, and are identifiable by factor analysis. [13]

Cattell believed that there is an underlying personality structure of the language that describes the features, represented as follows in formula 1.

$$R = f\ (S, P) \qquad (1)$$

The trait is the basic structural element of the theory of Cattell, defined as a trend, relatively permanent and comprehensive, to react in a certain way. Implies a certain regularity of behavior over time and situations.

1.4. Emotions

Emotions are reactions to the information (knowledge) that we receive in our relations with the environment. The intensity of the reaction based on subjective assessments we make about how information received will affect our wellbeing.

What is involved in these assessments are subjective prior knowledge, beliefs, personal goals, perceptions, environmental challenge, and so on. An emotion depends on what is important to us. If the emotion is intense can cause intellectual or emotional disorders (phobia, stress, depression). [14]

Ortony and Turner [2] collected a wide range of research on the identification of basic emotions, Anger, Disgust, Fear, Joy, Sadness, Surprise.

1.5. Empathy

In psychology the way a person can communicate their emotions and they can be understood by someone else is related to the concept of empathy.

They have various definitions taking the next two most representative for the concept which represents empathy for this investigation which are: "The action and the ability to understand, be aware, be sensitive or alternative way to experience the feelings, thoughts and experiences of another, without these feelings, thoughts and experiences have been reported in an objective or explicit". [8]

"Mental function that allows us to not be self-centered and see things from another point of view. Through them we can feel part of the experience of others and develop shared experiences". [9]

With these definitions we can realize that in many occasions or situations in an ordinary day of empathy is that we generate or change our mood and makes us have different reactions to each of the stimuli around us.

2. Development

After learning the foundations held for this project we develop our architecture, experiments and results obtained so far.

One goal of our project is to study how emotions are an important factor in making decisions as depending on the personality profile of each person may be more susceptible to such emotions.

To perform these experiments describe the architecture with which we worked.

As shown in Fig. 1 Multi-Agent System Personality Profile consists of four Intelligent Agents each with their respective tasks as described below: [5]

ACP (Agent Coordinator Personality) is the coordinator of the Multi-Agent System Personality Profile.

AAwareness (Agent Awareness) has the responsibility to establish a dialogue with the Multi-Agent System for Social Emotional emotion perceived demand and the degree of emotion recognition, the 16 factors that make up a personality profile to send this information to a fuzzy system where depending on the emotion and the 16 degree active personality factors may vary.

AFactors (Agent Factors) is responsible for obtaining the 16 factors of personality profile and send them to the Agent Consciousness.

AnvoPP (new Agent Personality Profile) this agent receives the new values of the fuzzy system, and is responsible for sending them to Clustering for the conduct of the working groups.

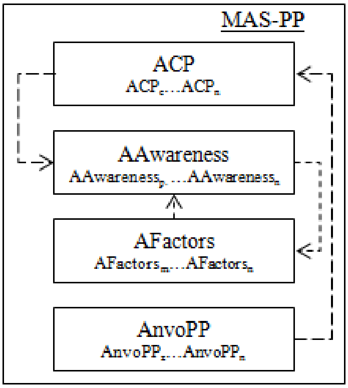


Figure 1 Multi-Agent System Profile Personality.

In Fig. 2 we can see that the MAS-PP is part of an agent called Cognitive Agent (AC) where it interacts with the Multi-Agent System Social Emotional (MAS-SoE) that consists of three agents, an agent to determine the profile of emotion, an Agent for determining the group to which the emotion and the last agent that acts as coordinator of the Multi-Agent system Social-Emotional also uses a neural network for recognizing emotions. The data are reported from the MAS-SoE to MAS -PP is the recognized emotion (joy, surprise, sadness Anger, Disgust, Fear), the profile of emotion (Shortly Efficient, Fair, Efficient) and to which group belongs emotions (Positive, Negative). [11]

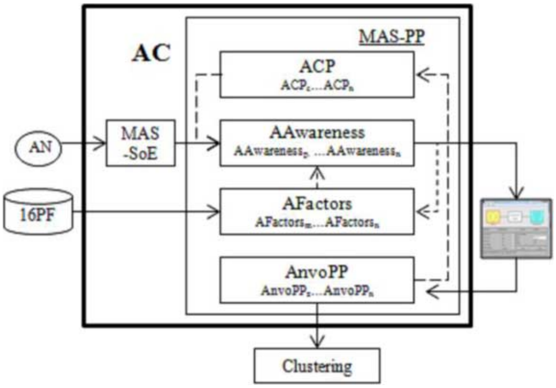


Figure 2 Cognitive Agent.

Multi-Agent system was analyzed and designed using UML, and respecting Gaia FIPA standards, was developed on the JADE platform and interacts with Matlab for using neural networks in the recognition, in this case, emotion recognition, we use Fuzzy Logic to generate new values of the factors of personality profiles that are affected depending on the emotion.

We use fuzzy systems for the global dimensions of personality to which it belongs according to the values of each of the factors in the personality profile [1], the degree of empathy between the emotion that has the robot at the time and which recognizes another robot in the environment have two fuzzy systems as mentioned, and then presents the structure of each.

In Figure 3 we show the structure of the fuzzy system of global dimensions, as input to the fuzzy system we have 16 entries are the factors that make up a personality profile and to output the five global dimensions, what is sought is to know in which these dimensions Total houses the personality profiles for groups to work together.

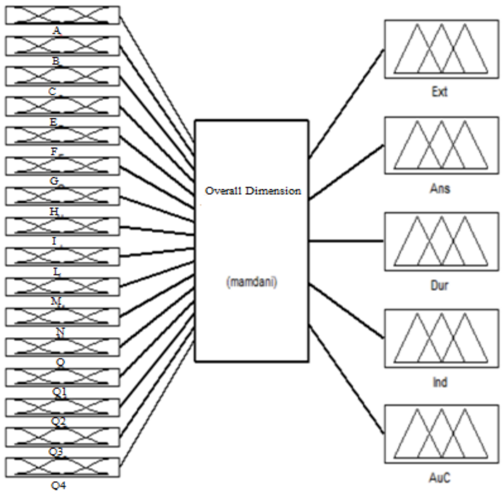


Figure 3 Fuzzy system of overall dimensions.

The five global dimensions are part of the tool 16 PF, [1] known by the names of Extraversion, Anxiety, Toughness, Independence and Self Control, each one takes into account the polar low and high.

In Figure 4 we show the structure of the fuzzy system of empathy, there are 3 inputs to the fuzzy system are perceived emotion of the robot in the environment, the degree of the emotion and excitement of the robot tries to find the working group and as output empathy, what is sought is to ascertain the level of empathy that exists between emotion recognition and emotion of the robot to start the search.

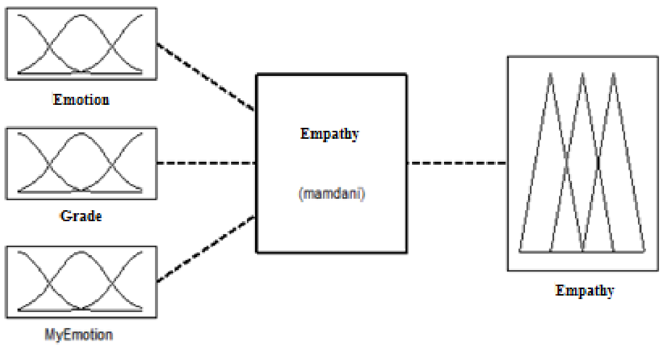


Figure 4 Fuzzy system of Empathy.

Multi-Agent System Personality Profile (SMA-PP) sends the 16 factors of personality profile to the fuzzy system of global dimensions and get the global dimensions that were activated by factors profile for that specific personality profile, the SMA -PP sends the thrill and excitement level of the previously recognized by the SMA-SoE [11], the emotion is activated robot that is starting the search for group work, these three data are sent to the fuzzy system of empathy and returns a level of empathy between emotions to SMA-PP, for later use by a clustering technique.

3. Results

Our experimental database were obtained by 16 PF tool because it is a tool for understanding personality profiles in the short term, use a population of 45 individuals ranging in age from 22-35 years, average economic level, graduate students.

Experiments were conducted with personality profiles obtained by applying an excitement in the air, was varied the number of individuals participating in the experiment below shows the results, it is noteworthy that our base case is an individual with an emotion. The data presented here are based on measures of the factors of personality profiles used in the tool 16 PF where they take from 1-10 to represent the value of each factor and these values are granulated into three parts, the first call "pole low" with a range of 1-3.5, "average" with a range of 3.5-7.5 and "pole high" with a range of 7.5-10, where the factor scores falling within the tall poles are those ay yield information from the literature.[1]

In the Table 1 is a sample of personality profiles, where we see the variations of the profiles to have different emotions, took the two emotions most strongly as ranked by Ekman. [12]

Table 1 We shown the personality profile of 20 subjects with two different emotions.

Factors	Emotions	Subjects																			
		I-1	I-2	I-3	I-4	I-5	I-6	I-7	I-8	I-9	I-10	I-11	I-12	I-13	I-14	I-15	I-16	I-17	I-18	I-19	I-20
A	Anger	5	4.3	8	7.1	6	5	7	6.7	5	4.3	8	7.1	6	5	7	6.7	5	4.3	8	7.1
	Disgust	5	5	8	8	6	6	7	7	5	5	8	8	6	6	7	7	5	5	8	8
B	Anger	10	10	6	6	3	3	6	6	10	10	6	6	3	3	6	6	10	10	6	6
	Disgust	10	10	6	6.8	3	4	6	6.7	10	10	6	6.8	3	4	6	6.7	10	10	6	6.8
C	Anger	4	4	3	3	6	6	8	8	4	4	3	8	6	6	8	8	4	4	3	3
	Disgust	4	4	3	3	6	6	8	8	4	4	3	2	6	6	8	8	4	4	3	3
E	Anger	1	2	4	5	7	7.5	7	7.7	1	2	4	10	7	7.5	7	7.7	1	2	4	5
	Disgust	1	1	4	4	7	7	7	7	1	1	4	6	7	7	7	7	1	1	4	4
F	Anger	1	3	6	6	4	4	7	7	1	1	6	5	4	4	7	7	1	1	6	6
	Disgust	1	8	4	4	4	4	7	7	1	1	6	6	4	4	7	7	4	7	7	6
G	Anger	8	8	3	3	7	7	10	4	4	8	3	3	7	7	8	8	8	8	3	3
	Disgust	8	8	3	3	7	7	8	8	8	8	3	3	7	7	6.5	5	7.5	8	3	3
H	Anger	1	1	7	7	6	6	7	8	4	1	7	7	6	6	8	8	1	1	7	7
	Disgust	1	1	7	7	6	6	8	8	1	1	7	7	6	6	8	8	6	3	6	7
I	Anger	10	10	7	4	3	4	4	4	10	10	10	5.5	4	4	4	4	10	10	10	10
	Disgust	10	10	10	10	4	4	4	4	10	10	10	10	4	4	4	4	10	10	10	10
L	Anger	6	6.5	7	7.3	4	4.8	3	3.9	6	6.5	7	7.3	4	4.8	3	3.9	6	6.5	7	7.3
	Disgust	6	6.8	7	7.5	4	5	3	4	6	6.8	7	7.5	4	5	3	4	6	6.8	7	7.5
M	Anger	5	5	9	9	5	5	4	4	5	5	9	4	5	5	4	4	5	5	9	9
	Disgust	5	4.5	9	8	5	4.8	4	3.6	5	4.5	9	1	5	4.8	4	3.6	5	4.5	9	8
N	Anger	2	2	8	8	4	4	4	4	2	2	8	2	4	4	4	4	2	2	8	8
	Disgust	2	2	8	8	4	4	4	4	2	2	8	8	4	8	4	3	2	2	9	5
O	Anger	5	5	3	7	8	4	3	3	5	5	8	2	4	4	3	3	5	5	8	8
	Disgust	5	5	8	8	4	4	3	3	5	5	8	10	4	4	3	3	5	5	8	8
Q1	Anger	6	6	7	7	8	8	4	4	6	6	7	6	8	8	4	4	8.5	6.8	7.3	7
	Disgust	6	6	7	7	8	8	4	4	6	6	7	5	8	8	4	4	6	6	7	7
Q2	Anger	8	8	8	8	6	6	6	6	8	8	8	2	6	6	6	6	3	6	8	8
	Disgust	8	8	8	8	6	6.5	4.8	5	8	8	8	5	6	7	6	6	8	8	8	8
Q3	Anger	4	4	5	5	6	6	6	6	4	4	5	6	6	6	5	6	6	4	5	5
	Disgust	4	4	5	5	6	6	6	6	4	4	5	8	6	6	6	6	4	4	5	5
Q4	Anger	10	10	7	7.6	5	5.7	4	5	10	10	7	4	5	5.7	4	5	10	10	7	7.6
	Disgust	10	9	7	6.5	5	4.8	4	4	10	9	7	9	5	4.8	4	4	10	9	7	6.5

A different experiments are presented below, where there are the profiles of different individuals with different emotions to those shown in Table 1.

3.1. Joy

This emotion applied to a sample of persons, in Figure 5 we appreciate the first person. On the left we have the values 1 to 10 which is the value in decatipes.

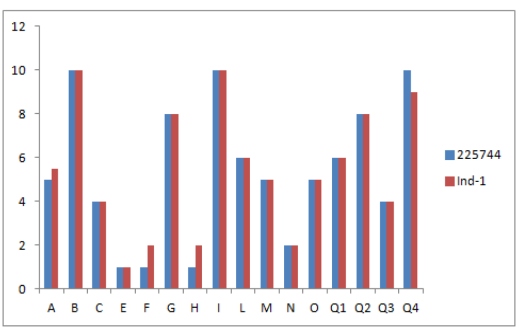


Figure 5 Graph Person number 1 with Joy emotion.

At the bottom we have shown how each of the 16 factors that shape the personality profile. In Figure 6 we have the personality profile obtained in blue denoting the initial data without being affected by any emotion at the time and red when perceiving emotion of Joy.

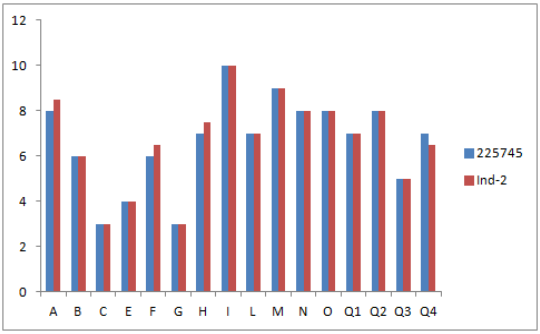


Figure 6 Graph Person number 2 with Joy emotion.

Viewing the results we see that the factors A (Warmth), F (Liveliness), H (Social Boldness), Q4 (Tension) there was an increasing shift factors A, F, H and decreasing Q4 which according to the poles [2] each of the factors of the individual behavior would be more positive.

3.2. Anger

This emotion applied to a sample of persons, in Figure 7 we appreciate the first person. On the left we have the values 1 to 10 which is the value in decatipes.

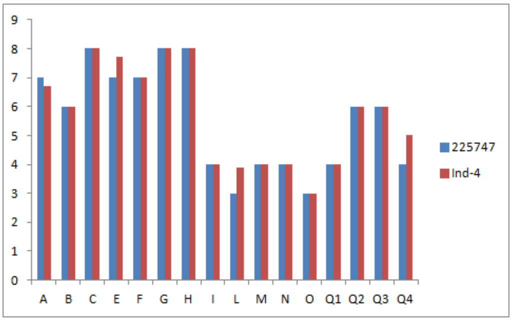


Figure 7 Graph Person number 4 with Anger emotion.

At the bottom we have shown how each of the 16 factors that shape the personality profile. In Figure 8 we have the personality profile obtained in blue denoting the initial data without being affected by any emotion at that time and in red at the moment of perceiving the emotion of anger.

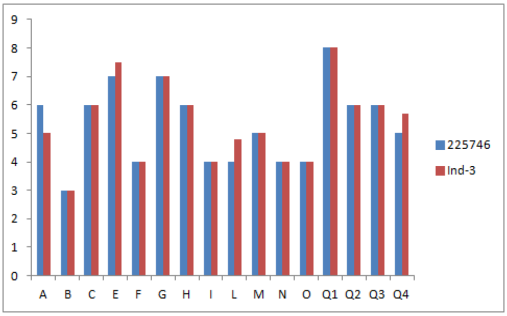


Figure 8 Graph Person number 3 with Anger emotion

Viewing the results we see that the factors A (Warmth), E (Dominance), L (Vigilance), Q4 (Tension) there was an increasing shift factors Q4, E, L and decreasing the factor A which according to the poles [2] of each of the factors of the individual behavior would be more positive.

3.3. Surprise

This emotion applied to a sample of persons, in Figure 9 we appreciate the first person. On the left we have the values 1 to 10 which is the value in decatipes.

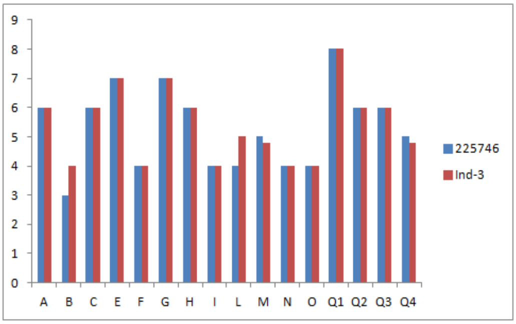


Figure 9 Graph Person number 3 with Surprise emotion.

At the bottom we have shown how each of the 16 factors that shape the personality profile. In Figure 10 we have the personality profile obtained in blue denoting the initial data without being affected by any emotion at that time and in red at the moment of perceiving the emotion of surprise.

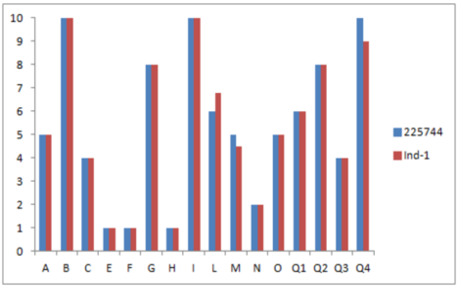


Figure 10 Graph Person number 1 with Surprise emotion.

Viewing the results we see that factors B (Reasoning), M (Abstractedness), L (Vigilance), Q4 (Tension) there was an increasing shift factors B, L and decreasing in the factors M and Q4 which according to the poles [2] of each of the factors of the individual behavior would be more positive.

Conclusions

Now in Figure 11 we contrast the personality profile taken from the database which we call normal profiles are affected by the emotions of Joy, Anger and Surprise.

At the bottom we have shown how each of the 16 factors that shape the personality profile, we have taken the personality profile of the database denoted in blue, red when perceiving emotion joy, green when perceiving emotion Ira and purple when perceiving emotion of surprise.

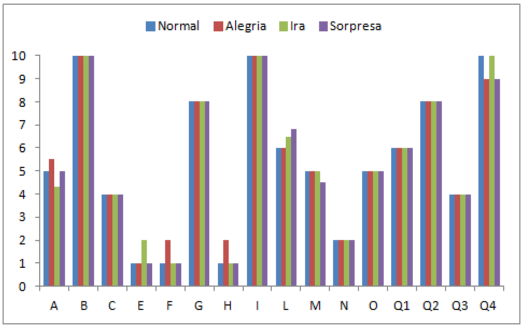


Figure 11 Graph Person number 1 without emotion Joy,Anger and Surprise emotions.

With these results we can realize that emotions have a really great influence on the personality, that even the excitement depending upon the circumstances we are perceiving or modify any of the factors that make up a personality profile, and the changes depend on the group in which the relevant emotion either positive or negative, to improve performance or make decisions.

It is noteworthy that these results are shown only the interaction of emotion with a personality profile of a profile group because the results and interaction could occur in different ways.

References

[1] A. Benito Mariscal. Estudio de la quinta edición del 16 PF. Tesis Universidad Politécnica de Catalunya (2009), 39, 51-89.
[2] A. Ortony, and T. J Turner, What's basic about basic emotions? Psychological Review, (1990), 315-331.
[3] C. Becker-Asano, and I. Wachsmuth, Autonomous Agents and Multi-Agent Systems, Affective computing with primary and secondary emotions in a virtual human, Volume 20, Number 1, 32-49, 2010.
[4] C. Castelfranchi, Modelling social action for AI agents. Artificial Intelligence, Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (1997), 1567-1576.
[5] C. Solano, A. Alanis, Sistema Multi-Agente Social Racional Basado en Personalidad Humana, XI Congreso Nacional de Ingenieria Electrica y Electronica del Mayab(2011), 698-707.
[6] H. Boukricha, I. Wachsmuth, and KI - Künstliche Intelligenz, Empathy-based Emotional Alignment for a Virtual Human: A Three-Step Approach, Volume25, Number 3, 195-204, 2011.
[7] J.L. Deneubourg, S. Goss, N. Franks, A. Sendova-Franks, C. Detrain, and L. Ch-retien, First Conference on Simulation of adaptative Behavior: From Animals to Animats, The dynamics of collective sorting: Robot-like ants and ant/like robots, MIT Press, 1991.
[8] J.L. González de Rivera, ASMR Revista Internacional, Ecpatía y empatía, Psiquis, Volume 4, Number 2, 2004.
[9] J.L. González de Rivera, Crisis Emocionales, Espasa-Calpe, Madrid, (2005).
[10] K. Dautenhahn, The art of designing socially intelligent agents: science, fiction and the human in the loop. Applied Artificial Intelligence: An International Journal (1998), 573-617.
[11] L. Amador, and A. Alanis, Diseño de un Sistema Multi–Agente Social-Emocional para la creación de grupos con emociones primarias similares.(SMA-SoE), 33o Congreso Internacional de Ingeniería En Electronica Electro 2011, Instituto Tecnológico de Chihuahua 12 Al 14 de Octubre 2011, Chihuahua, Chih,. Mexico, (2011).
[12] P. Ekman, An argument for basic emotions, Cognition and Emotion, 6(3-4), 169-200, 1992.
[13] R. B. Cattell, and S. E. Krug, The number of factors in the 16FP: A review of the evidence with special emphasis on methodological problems. Educational and Psychological Measurement, Volume 46, Number 3, 509-522, 1986.
[14] R. Bisquerra, Educación emocional y bienestar, Praxis, Barcelona, 2000.
[15] S. J. Russell, and P. Norvig, Artificial Intelligence: A Modern Approach (2nd Edition) Prentice Hall, 2002.

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-26

Ambient Intelligence Used For Fall Detection

Vanessa SALDIVAR[a,1], Víctor M. ZAMUDIO[a], Rosario BALTAZAR[a], Arnulfo ALANIS[b], Miguel LOPEZ[b]
[a]*Division of Postgraduate Studies and Research, Leon Institute of Technology, Leon, Guanajuato, Mexico*
[b]*Tijuana Institute of Technology, Tijuana Baja California, Mexico*

Abstract. In this article we present a comparative study between a Bayesian classifier and the method SPC (Statistical Process Control). The Bayesian classifier is included in the virtual machine WEKA, and SPC is manly used to monitor processes in different domains. These two methods are applied to a set of data collected from a SunSPOT in order to detect falls. A comparative analysis of these two approaches was performed in order to have the best strategy to be apply in real scenarios.

Keywords. Fall detection, Bayesian classifier, statistical process control.

Introduction

Ambient intelligence has multiple areas such as control, energy, healthcare. In this area human-computer interaction, ubiquitous computing and artificial intelligence techniques are used in the design of advanced technologies to support people. An important trend in healthcare is recognizing human behaviour using simple sensors to infer the physical and cognitive status of people and therefore assist them in an appropriate way to reduce risks associated with aging or disease [1].

In Mexico there is a big effort to develop and implement technologies focused on health using the paradigm of intelligent environments [2], however these strategies are on an early stage, compared to other countries such as Spain, Italy, Russia [3] [4] [5] [6]. Mexico is a country where public policy plays a very important role, and different strategies should be applied in order to encourage a healthy way of life (free medical insurance, programs encouraging people to do exercise, etc.) [3].

In this paper we propose an analytical model based on artificial intelligence techniques (data classifier) and wireless sensors to detect if the person carrying a sensor suffers a fall. This research is in early stage, and is part of a larger project aimed to provide technological solutions to improve the quality of life of the elderly and people suffering from illness such as Alzheimer and other dementias.

[1] Corresponding Author: Vanessa Saldivar, Avenida Tecnologico S/N Fracc. Industrial Julian de Obregon, Leon, Gto., Mexico; E-mail: enav.813@gmail.com

A fall is defined as the occurrence of an event that inadvertently causes a person to go to the floor or below the level that originally was. The 80% of falls in elderly occur inside of their home, and 20% occur outside. This is very important, as most falls in elderly are a symptom of a disease [7].

The data used in this work correspond to body movements collected from a triaxial accelerometer embedded on a SunSPOT [8]. A processing feature extraction technique using linear prediction (LPC) was applied to the signals, in order to classify the (processed and unprocessed) data. This classification was made using the Bayesian classifier integrated in the virtual machine WEKA and the method called Statistical Process Control, which is mainly used to track product quality and/or services.

In Spain [4] an application for mobile devices was developed, in order to assist primary caregivers of elderly, and help them to determine if the person carrying the device is in motion or at rest [4]. Other efforts have been focus on using the accelerometers to recognize and classify activities, such as walking, changing position, smooth movement, standing, sitting and lying [9].

1. Data Collection

SunSPOT is a small device with wireless communication and battery powered developed by Sun Microsystems Laboratories [8]. The SunSPOT is composed of three elements: battery, radio processor card and sensor board. The sensor board has a 3-axis accelerometer (x, y, z), used in our experiments. The units of the accelerations are G (units of gravity) [10]. The data obtained from the SunSPOT are stored in a text file, along with the device identification number, a sequence number, time (in milliseconds) when the data was obtained, and the strength of the signal (see Figure 1).

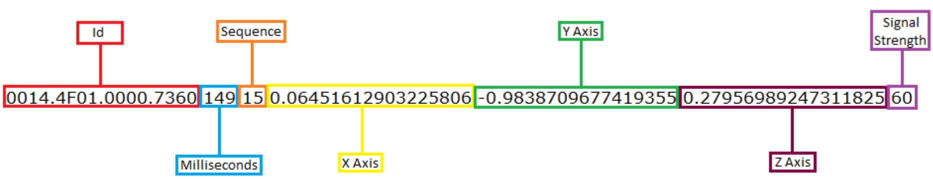


Figure 1. Data received from SunSPOT.

The sampling frequency of the received data is 98 data per second. The data can be sent to a computer in order to be monitored in real time (see Figure 2).

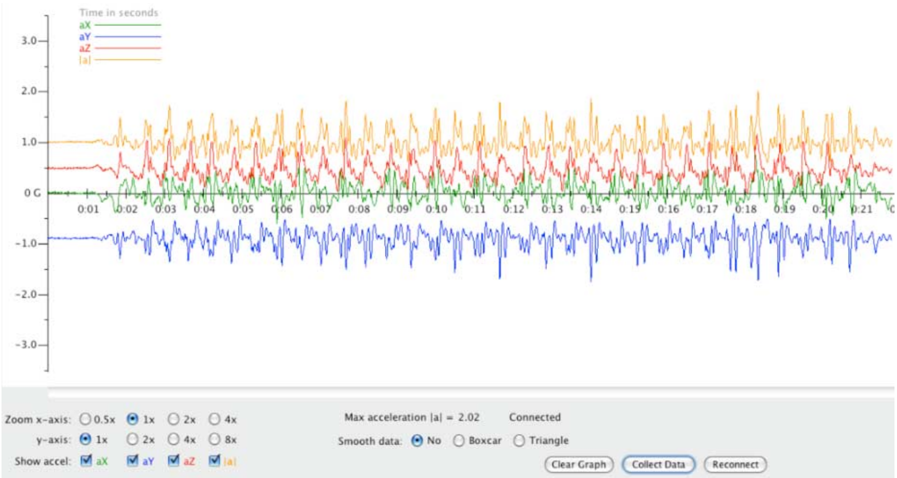


Figure 2. SunSPOT graphics acceleration.

We run experiments with 11 people who are in the age range of 22 to 30 years old, all of them with different heights, mass and gender.

The participants were asked to perform different actions, such as sitting, bending over and tripping, together with different types of falls (forward, back, side, tripping). With this a bi-class problem (falling and not falling) was set.

The tests were performed with two databases, the first consisting of the raw data, i.e. as they were received from the SunSPOT and for the second database a pre-processing step was included, in order to obtain the coefficients of linear prediction (LPC). This process can be seen on Figure 3.

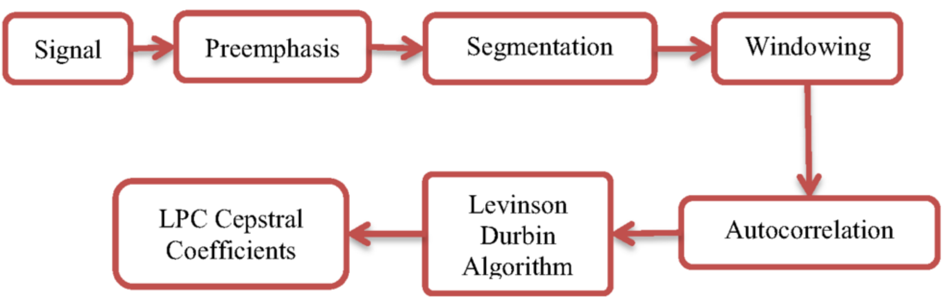


Figure 3. Signal processing.

With the linear prediction coefficients obtained we can pass from the Figure 4 to the Figure 5, where the data corresponding to the biggest acceleration are highlighted.

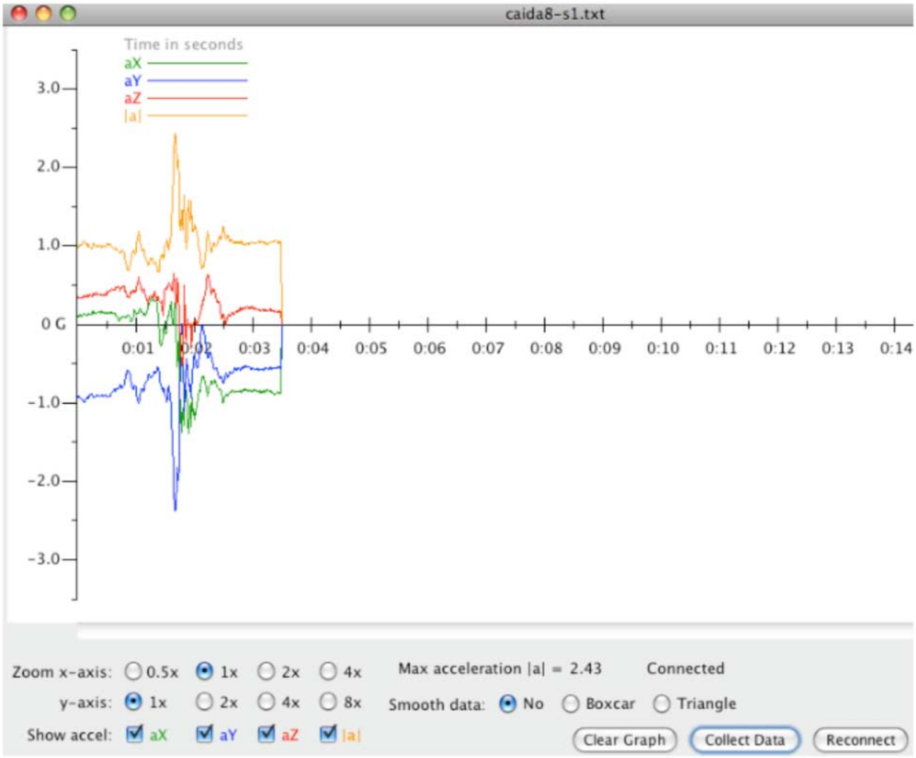


Figure 4. Graph with data received by the SunSPOT.

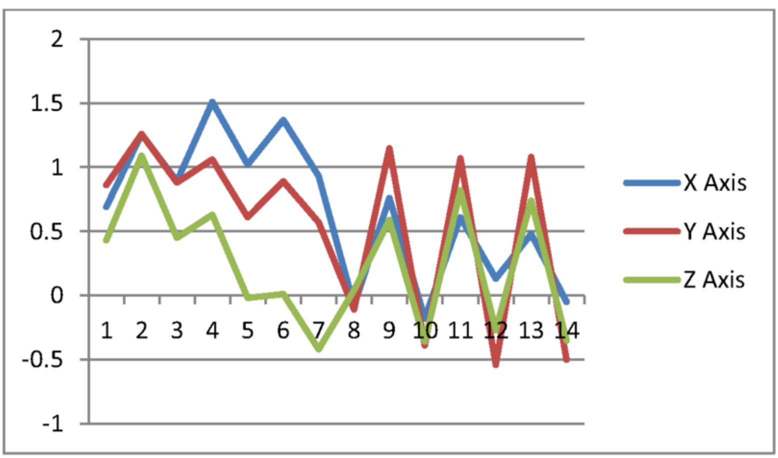


Figure 5. Graph with LPC from figure 4.

2. Bayesian classifier

Weka is a collection of algorithms developed by the University of Waikato (New Zealand) implemented in Java. Weka contains tools to perform transformations and classification of data, regression, clustering, association and visualization. For this project we are interested on the classification algorithm, in particular the Bayesian classifier which is based on statistical theories of learning [11].

The Bayesian classifier is based on Bayes' formula (see Eq. (1)) [12].

$$P\left({}^{w_i}/{}_{x}\right)=\frac{p\left({}^{x}/{}_{w_i}\right)P(w_i)}{p(x)};\ \ i=1,2,\ldots,N \tag{1}$$

Where:

- $P\left({}^{w_i}/{}_{x}\right)$ Is the probability that the class is w_i given the feature vector x.
- $p\left({}^{x}/{}_{w_i}\right)$ Is the probability for x given the class w_i.
- $P(w_i)$ Is the prior probability of class w_i.
- $p(x)$ Is the probability that the vector of features x is presented.
- x Are the characteristic vectors of each element of the N classes $w_1, w_2, \ldots, w_N$.

The Bayes decision rule dictates that for a given x, we can minimize the likelihood of error if you decide that x belongs to $P\left({}^{w_i}/{}_{x}\right) > P\left(w_j/x\right)$, for $i \neq j$ [12].

This classifier it's going to be used to classify two types of data, the first will be the raw data, i.e. as they are received from the SunSPOT and the second type will be the same data already processed with the coefficients of linear prediction (LPC) mentioned on the Figure 3.

With our experiments we want to evaluate the performance of the Bayesian Classifier, and the importance of the preprocessing phase (using the LPCs).

3. Statistical process control

The use of statistical process control involves the following assumptions [13]:

- Once the process is in operation under established conditions, it is assumed that the variability of results in the measurement of a product quality characteristic is due only to random causes, inherent to each particular process.
- The random that acts on the process generates a hypothetical universe of observations (measurements) that has a normal distribution.
- When any assignable cause resulting in additional deviations in the results of the process, we say the process is out of control.

The role of statistical process control is permanently checking if the results emerging from the measurements agree with the first two hypotheses [13].

The implementation of a program of statistical process control involves two stages [13]:

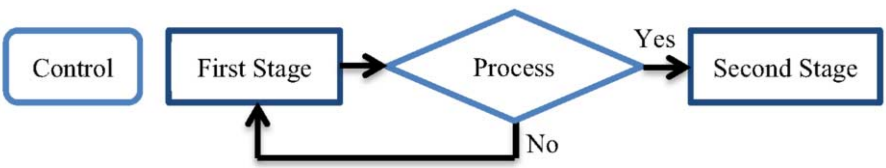


Figure 6. Statistical control for a process.

In the first stage (Figure 6) are collected a certain number of measurements, which calculates the average and standard deviation [13].

$$\bar{x} = \frac{\sum x_i}{N} \; y \; \sigma = \sqrt{\frac{\sum(\bar{x}-x_i)^2}{N}} \quad (2)$$

Then the control limits are calculated as follows:

Lower limit $= \bar{x} - 3\sigma$; Upper limit $= \bar{x} + 3\sigma$.

These limits arise from the assumption that the distribution of observations is normal. In general, we used levels of 2 or 3 sigma around the mean. In a normal distribution, the range of 3 sigma around the mean corresponds to a probability of 0.998.

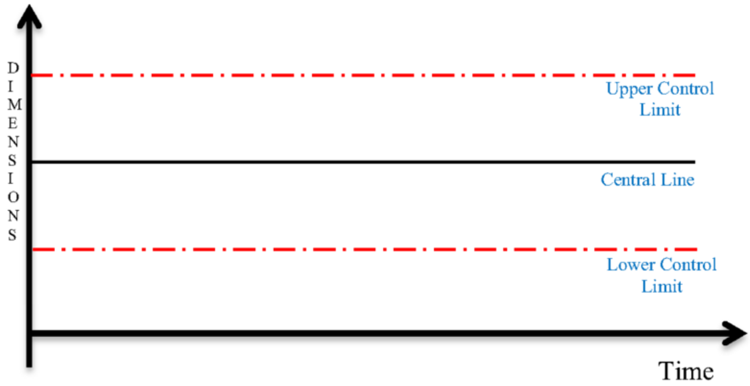


Figure 7. Control graph.

On the first stage a graph is drawn with two straight lines at the height of the control limits obtained (Figure 7).

In the second stage (Figure 6), the observations that emerge from the process are shown in the graph, and controlled by verifying that they are within the limits, and that there are not non-random patterns [13].

The 99.8% of observations should be within the limits of 3 sigma around the mean. These control limits are chosen so that if the process is under control, almost all the points on the graph will be contained between these limits. While the points are within the limits no action is required, because it is assumed that the process is under control. However, a point outside the control limits is interpreted as evidence that the process is out of control [13].

The statistical process control is commonly used to determine if a process has a certain level of quality in our problem, it is going to help us to know if a person suffer a fall or not.

This process use data obtained previously to create a control area. Once the control area is established the statistical process control is ready to work.

In our problem SPC will be used to monitor the acceleration: if the acceleration is out of the control area, we have found a fall .

4. Testing and Experimental results with Bayesian classifier

For testing we used two databases: the first corresponds to the data received from the SunSPOT and has a total of 242,611 records divided into two classes (fall, not fall). The class fall has 116,611 records and the class not fall has 126,000 records.

For training we used 70% of the total of records (169,828 records) and the reminder was used for classification (72,783 records).

The second database corresponds to the pre-processed data, (using the LPC technique) with 9,224 records. This database has two classes: fall and not fall and each class have 4,612 records.

For training with the second database we used 70% of the total records (6,456 records) and the remainder was used for classification (2,768 records).

With the Bayesian classifier were obtained the confusion matrices mentioned on Table 1 for both databases.

Table 1. Confusion matrices obtained with the Bayesian classifier

Processed Data			**Not Processed Data**		
	Fall	Not fall		Fall	Not Fall
Fall	1092	2128	Fall	43222	38299
Not Fall	877	2359	Not Fall	18669	69636

The classification rate obtained for the pre-processed data was 53.45% and for the other (not processed) data 66.46%.

For the second test the same amount of data were used. However a different selection of patterns for training was used. The selection of these patterns was performed at random form (see Table 2).

Table 2. Confusion matrices obtained with the Bayesian classifier

Processed Data			**Not Processed Data**		
	Fall	Not fall		Fall	Not Fall
Fall	645	2590	Fall	42915	38675
Not Fall	431	2790	Not Fall	18868	69368

The classification rate obtained for the processed data was 53.21% and for the not processed data was of 66.12%.

5. Testing and Experimental Results with Statistical Process Control

The performance of the SPC can be better observed in graphical form in graphic form, the following tests were performed with the data of each individual account. Here are only the results obtained from the data of one person.

The training was conducted with information of the class fall for one person, which is a total of 10,913 data, obtained the limits showed on Table 3.

Table 3. Limits obtained from the data of falls of a person.

	X Axis	**Y Axis**	**Z Axis**
Lower Limit	-2.149731081861608	-1.950906797651994	-0.9298955015846748
Upper Limit	2.0761550272596683	1.0605875670608702	1.7515120208982125

The limits from the Table 3 are going to determine the control area which is going to help us to know if a person suffers a fall or his/her moves doesn't represent a risk.

From Figures 8 to 10, the behavior of the class not fall on each axis (X, Y, Z) is shown together with the limits showed in Table 3.

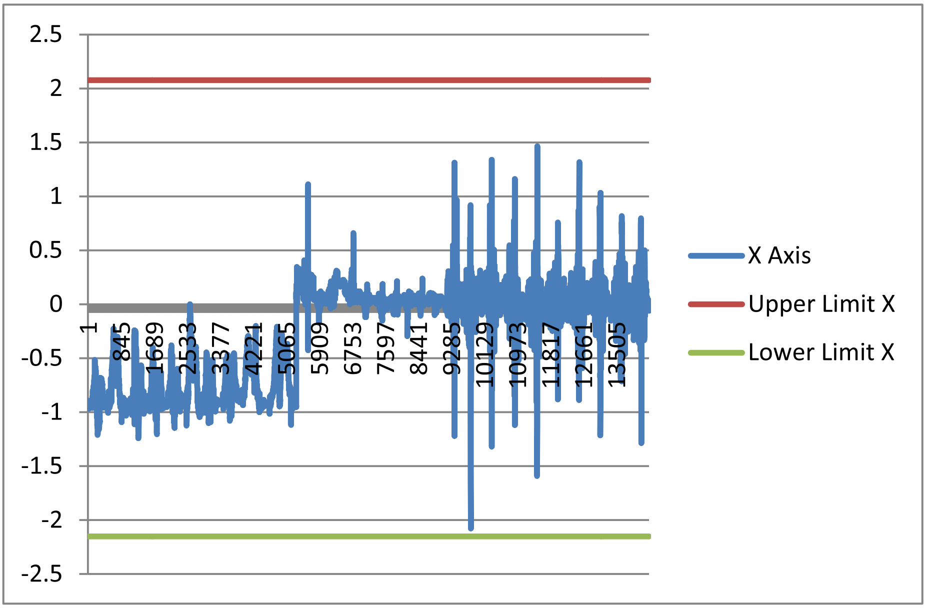


Figure 8. Accelerations in the X axis with their limits for class not fall.

On the Figure 8 we can observe that all the graphed data are between the limits which means that there are no presented any fall even if in the other axis some of the data exceed the limits.

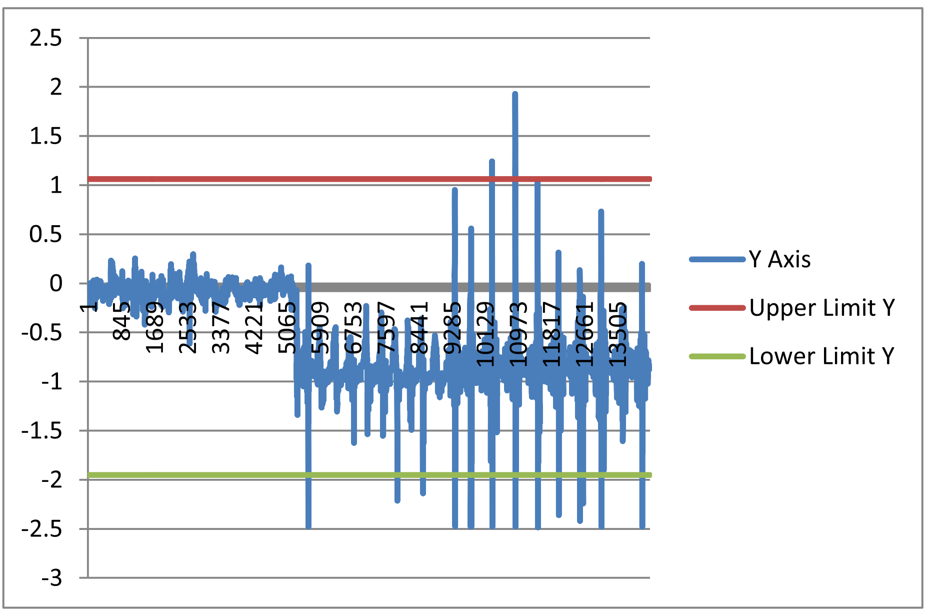


Figure 9. Accelerations in the Y axis with their limits for class not fall.

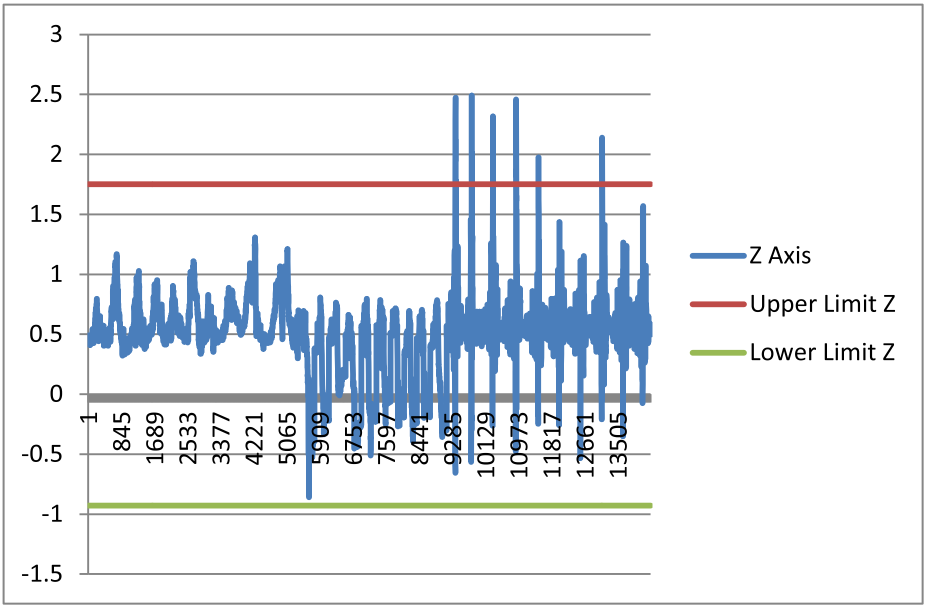


Figure 10. Accelerations in the Z axis with their limits for class not fall.

With the statistical process control to determine if a fall is presented the acceleration of the three axes (x, y, z) need to be located out of the control area.

So keeping that idea, if we observe the graphics above, we can see that there's no presence of a fall because in the Figure 8 all the data of acceleration are between the limits obtained.

From Figures 11 to 13, the behavior of the class fall on each axis (X, Y, Z) is shown together with the limits showed in the Table 3.

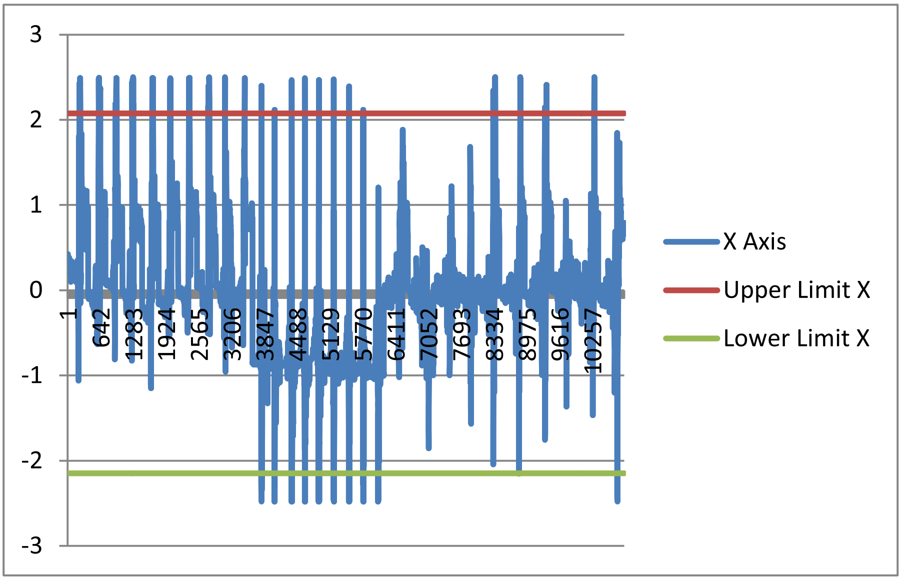


Figure 11. Accelerations in the X axis with their limits for class fall.

We can observe that several data exceed the limits, so this can means that several falls were presented, to determine if a fall effectively it's presented we still need to see if in the axis Y and Z on the same time the data also exceed the limits.

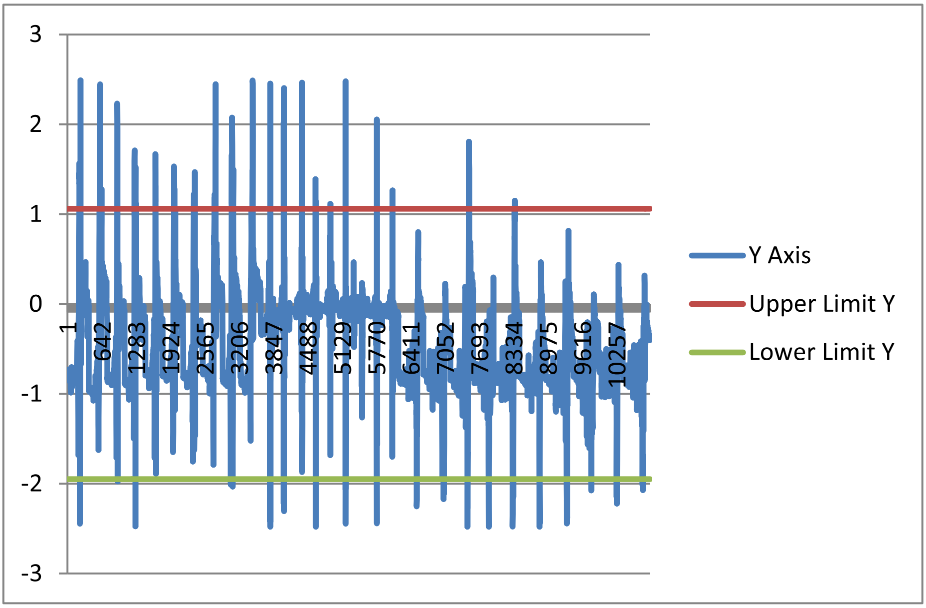


Figure 12. Accelerations in the Y axis with their limits for class fall.

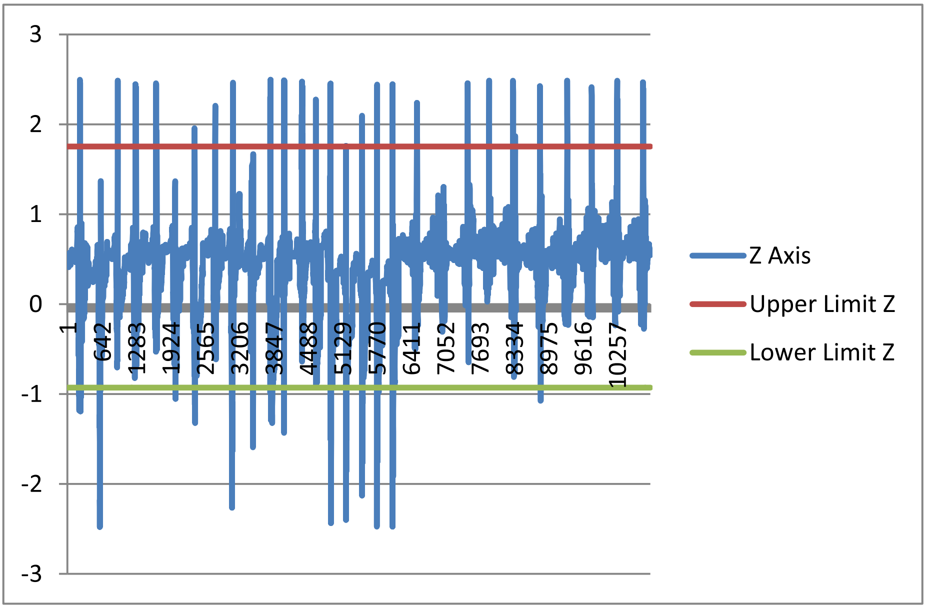


Figure 13. Accelerations in the Z axis with their limits for class fall.

As it's indicated above to determine if a fall occurs, the acceleration in three axes must excel in the control area found; otherwise it indicates no presence of a fall. So in the Figures 11 to 13 it's easy to observe that a several falls are presented, because in several moments the data of the acceleration of the three axes exceeds the limits obtained.

6. Conclusions

From the experiments and analysis performed, it was found that the Bayesian classifier using raw data (directly received from the SunSPOT), showed a classification rate of 66%. Clearly these results must be improved if the system is going to be applied on real scenarios. One strategy to improve the performance of the Bayesian classifier could be increasing the number of experiments performed.

Although the method called statistical process control is applied to quality control, the method showed optimal results on the testing that was performed to observe if the data corresponding to the accelerations which do not represent a fall are inside of the area of control obtained. The advantage of this method is that it can be customized, calibrating the normal activities in order to generate a control area.

From Figures 8 to 11, it can be seen that all the accelerations are located between the limits of the control area. In these figures, the data correspond to the class not-fall. According to the previous, SPC shows a good performance detecting falls, as the original accelerations corresponding to the class not fall are inside the control area.

With the SPC method was found a classification rate of 67%, even when this rate is higher than the rate found with the Bayesian classifier, still can be improve with the application of fuzzy logic on the limits, in this way the limits that determine de control area will not cut the signals of acceleration.

References

[1] M. Aleksandar, V.O. and O.M. Ibarra, Ambient Aware Assistance (A3) for supporting Alzheimer Patients and Caregivers, *5th International Symposium on Ubiquitous Computing and Ambient Intelligence UCAmI*, Riviera Maya, México, del 5 al 9 de Diciembre , 2011

[2] G. R. Luis Miguel, Red temática, *Envejecimiento social y desarrollo social*. Available at: http://www.redenvejecimiento.com/

[3] D. Laura, S. Sandra and V. Atanacio, Analysis of changes in the association of income and the utilization of curative health services in Mexico between 2000 and 2006, *BMC Public Health*, Vol. 11, 2011.

[4] R. Enrique, S. Javier, C.F.O., G. Jaime, A. Mariano, H. Rocío, V. Yolanda, C. Diana and G. P. Azucena, New Platform for Detecting Diary Physical Activity Patterns and Measuring Emotional Feedback using Mobile Technologies in Clinical Pain Therapies, *5th International Symposium on Ubiquitous Computing and Ambient Intelligence UCAmI*, Riviera Maya, México, del 5 al 9 de Diciembre , 2011

[5] C. Lee, D. Nordstedt and S. Helal, Enabling smart spaces with OSGi, *Pervasive Computing, IEEE*, IEEE, 2003, Vol. 2(3), pp. 89-94.

[6] A. Juan Carlos, C.D. Nugent, S. Martin and C. Olphert, Towards Personalization of Services and an Integrated Service Model for Smart Homes Applied to Elderly Care, *Proceedings of International Conference on Smart Homes and health Telematic*, pp. 151-158. Sherbrooke, Canada, July 4-6, 2005.

[7] H.G. Espinola, Caídas en el adulto mayor, *Boletín de la escuela de medicina*, 2000, Vol. 29, pp. 1,2.

[8] SunSPOT, What is a Sun SPOT Device?, Disponible en: http://www.sunspotworld.com/docs/general-faq.html

[9] S. Zhang, P. McCullagh, C. Nugent and H. Zheng, Activity Monitoring Using a Smart Phone's Accelerometer with Hierarchical Classification. *Intelligent Environments (IE)*, 2010 Sixth International Conference on 2010, pp. 158-163

[10] SunLabs, SunSPOT Theory of Operation, 2009.

[11] D.G. Morate, Manual de WEKA. Disponible en: http://www.cs.waikato.ac.nz/ml/weka/.

[12] S. Theodoridis and K. Koutroumbas, *Pattern recognition*, Academic Press, 2006.

[13] J.A. Hervás, Manual de control estadístico para la gestión de procesos.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-37

PAC-Means: *clustering algorithm based on c-Means technique and associative memories*

Luis C. PADIERNA[1], Martha B. GONZÁLEZ and Leoncio A. ROMERO
Technological Institute of León, México

Abstract. In this study a partitional clustering technique is proposed. Our proposal is a variant of the c-Means algorithm that replaces its traditional minimum-distance classifier by a classifier based on associative memories. The variant was compared against the original version by applying both techniques to three datasets belonging to the UCI Machine Learning Repository: Iris, Wine and Pima Indian Diabetes. As a comparison criterion, an intracluster-spread index was used. Results obtained in experimental tests show that, when applied to certain databases, the PAC-Means technique overcomes to the c-Means algorithm.

Keywords. Associative Memory, Pattern Classifier, Diabetes

Introduction

Non-supervised pattern classification aims to assign some entities into one of several clusters, in such a way that entities in one group are similar among them, but different to entities clustered in other groups [1].

The c-Means algorithm, developed by McQueen in 1976, is widely known among non-supervised classification techniques. This algorithm begins selecting cluster centers randomly or by *a priori* information. Afterwards, each entity is assigned to the group with the nearest center. Finally, centers are updated with the mean of entities in the same group and the process is repeated until some stop criterion is fulfilled [2, 3].

In addition, the c-Means algorithm utilizes a minimum-distance classifier for carrying out the assignment of the entities into different groups. However, the minimum-distance classifier could be replaced by any other supervised classifier.

The supervised classification is targeted to find a function that determines the group to which an entity being analyzed belongs to. This function is commonly determined by means of probabilistic, structural, syntactic, neural, or fuzzy sets techniques [4].

The Pattern Associative Classifier (PAC) was developed in 2003 showing significant results on supervised classification [5-8]. However, applications in which PAC algorithm has been used for clustering have not been detected yet.

With the objective of establishing an alternative technique for pattern clustering, in this study the non-supervised pattern associative classifier **PAC-Means** is described and tested.

[1] Corresponding Author.

The next section presents the approach to which the proposed algorithm belongs to. In section 3, PAC-Means is formally described. The datasets used and the experimental methodology followed for the assessment of both algorithms: c-Means and PAC-Means, are described in section 4. Finally, the results of experimental tests and several relevant aspects are discussed in section 5.

1. Associative Classification Approach

The pivotal concept in associative approach is **associative memory**, an association store of input patterns *x* and output patterns *y*, where *x* and *y* are *n*-dimensional feature vectors such that $x \in \Re^n$ and $y \in \Re^m$.

This store allows the retrieval of a *y* by presenting its associated input pattern *x*, which might be noisy or incomplete, and in such a case it is represented by $\tilde{x} \in \Re^n$ .The associative memory design problem is defined as follows [9]:

Given a finite set of *k* associations $A = \{(x^\xi, y^\xi) \mid \xi \in \{1,2,\ldots,k\}\}$, determine a mapping *G* such that $G(x^\xi) = y^\xi$, noise tolerant, i.e., $G(\tilde{x}^\xi) = y^\xi$.

Seminal Works about associative memory are based on the learning matrix published by Steinbuch in 1961 [10]. This memory can work as a classifier of binary patterns if the output patterns are chosen correctly. However, binary patterns restriction makes it impractical for current classification applications. In 1972, Teuvo Kohonen presented a correlation matrix which eliminates this restriction by allowing real numbers as components of its input vectors; however, these vectors must be orthonormal [11]. PAC is a hybrid of the above two techniques that eliminates both orthonormality and binary patterns constraints, and at the same time improves the two techniques capacities trough an ingenious translation method [5]. The operation of this classifier is now shown.

2. Pattern Associative Classifier

Let *X* be an input pattern set with cardinality *n*, where each pattern has real values in its *p* components and is grouped in one of *m* classes, the PAC algorithm steps are:

1. Each input pattern x^ξ belonging to the *j* class is associated with vector y^ξ, which consists of zeros except for its *j*-th coordinate where the value is one. This is, output pattern takes the form

$$y^\xi = \begin{pmatrix} y_1^\xi = 0 \\ y_2^\xi = 0 \\ \vdots \\ y_j^\xi = 1 \\ y_{j+1}^\xi = 0 \\ \vdots \\ y_m^\xi = 0 \end{pmatrix} \tag{1}$$

2. Compute the mean vector $\bar{x}$ from *X*

$$\bar{x}_i = \frac{1}{n}\sum_{\xi=1}^{n} x_i^{\xi}, \quad \forall i \in \{1,2,\cdots,p\} \tag{2}$$

3. Take the mean vector coordinate as a center for a new set of coordinate axes and move all the input patterns to the new system using (3):

$$x^{\xi'} = x^{\xi} - \bar{x}, \quad \forall x^{\xi} \in X \tag{3}$$

4. Apply the learning phase. For all $a \in A$ in the form ($x^{\xi'}, y^{\xi}$), calculate the outer product $y^{\xi} \times (x^{\xi'})^t$ that expresses a matrix of dimensions $m \times p$:

$$M^{\xi} = y^{\xi} \times \left(x^{\xi'}\right)^t = \begin{pmatrix} y_1^{\xi} \\ y_2^{\xi} \\ \vdots \\ y_m^{\xi} \end{pmatrix} \times \left(x_1^{\xi'}, x_2^{\xi'}, \cdots, x_p^{\xi'}\right) \tag{4}$$

5. Add the *n* matrixes for obtaining the classification memory *M*. At this point the learning phase finishes. The resulting memory stands for the mapping *G* above described.

$$M = \sum_{\xi=1}^{n} M^{\xi} \tag{5}$$

6. Start the retrieval phase. The first step is to move all input pattern to be classified z^{ξ} (where $z^{\xi} = x^{\xi}$ or $z^{\xi} = \tilde{x}^{\xi}$) to the new axes:

$$z^{\xi'} = z^{\xi} - \bar{x}, \quad \forall z^{\xi} \tag{6}$$

7. To find the class to which the presented input pattern $z^{\xi'}$ belongs, the y^{ξ} vector coordinates must be obtained. Where y^{ξ} is associated with the original input pattern x^{ξ} , and $y_j^{\xi} \in \{0, 1\}$. The *j*-th coordinate y_j^{ξ} of the class vector y^{ξ} is then gotten by the expression

$$y_j^{\xi} = \begin{cases} 1 \; si \; \sum_{i=1}^{p} m_{ji} \cdot z_i^{\xi} = \bigvee_{h=1}^{m} \left[\sum_{i=1}^{p} m_{hi} \cdot z_i^{\xi} \right] \\ 0 \; in \; other \; case \end{cases} \tag{7}$$

Where $\vee$ is the operator “max” and m_{ji} the *j,i*-th component of the *M* matrix. At this point the classification process finishes. The classified pattern z^{ξ} , belongs to the *j*-th class indicated by $y_j^{\xi} = 1$.

2.1. C-Means Classifier

One definition of this classifier is the following according to [12, 13]: Let t_{max} be the maximum number of allowed iterations, it the current iteration of the algorithm, Z_j^{it} the center of the j-th class in iteration it, m the number of desired classes and S_j the size of the j-th group.

1. Assign $it = 0$ and establish m different initial centers randomly selected from X.

$$Z_j^0 = x^{\xi}, \quad 1 \le j \le m \tag{8}$$

2. Classify $\{x^{\xi}\}_{\xi=1}^{N}$ based on the group centers $\{Z_j^0\}_{j=1}^{m}$ and the minimum distance classifier. For $1 \le j \le m$ denote by $\{x^{(\xi,j)}\}_{\xi=1}^{S_j}$ the patterns grouped around Z_j^0.

3. Calculate new centers for the partition generated in the previous step and the intra-cluster spread index (I).

$$Z_j^{it+1} = (\sum_{\xi=1}^{S_j} x^{\xi}) / S_j \tag{9}$$

$$I_{it} = \sum_{j=1}^{m} \sum_{x^{\xi} \in \{C_j\}} \left\| x^{\xi} - Z_j^{it+1} \right\|^2, \quad 1 \le j \le m \tag{10}$$

4. Assign $it = it + 1$. If Z_j^{it} is equal to Z_j^0, send to output Z_j, S_j, $\{x^{(\xi,j)}\}_{\xi=1}^{S_j}$ and finish. Otherwise, if $it > t_{max}$ send the message "Number of iterations exceeded"; else, assign $Z_j^0 = Z_j^{it}$ and go to step 2.

3. PAC-Means

The algorithm described in this section is a hybrid of PAC algorithm and c-Means technique named PAC-Means, its definition is the following.

Let X be an input pattern set with cardinality n, where each pattern has real values in its p components; t_{max} the maximum number of allowed iterations; it the current iteration of the algorithm; Z_j^{it} the center of the j-th class in iteration it; m the number of desired classes and S_j the size of the j-th group:

1. Assign $it = 0$ an establish m different initial centers randomly selected from X, using the equation (8).

2. Associate a vector y^{ξ} to each class center, employing the equation (1):

$$A = \{(Z_j^0, y^{\xi})\}, \; 1 \le j \le m \tag{11}$$

3. Calculate the mean vector of the class centers, denoted by $\bar{c}$:

$$\bar{c} = \frac{1}{m}\sum_{j=1}^{m} Z_j^0 \tag{12}$$

4. Move all centers to a new axes set, which has $\bar{c}$ as origin:

$$Z_j^{0'} = Z_j^0 - \bar{c}, 1 \leq j \leq m \tag{13}$$

5. Start the learning phase. For all $a \in A$ in the form ($Z^{0'}, y^{\xi}$), calculate the outer product $y^{\xi} \times (Z_j^{0'})^t$ that expresses a matrix of dimensions $m \times p$:

$$M^{\xi} = y^{\xi} \times \left(Z^{0'}\right)^t = \begin{pmatrix} y_1^{\xi} \\ y_2^{\xi} \\ \vdots \\ y_m^{\xi} \end{pmatrix} \times \left(Z_1^{0'}, Z_2^{0'}, \cdots, Z_p^{0'}\right) \tag{14}$$

6. Add the *m* partial memories (M^{ξ}) to obtain an associative memory classifier *M*. At this point the learning phase finishes.

$$M = \sum_{\xi=1}^{m} M^{\xi} \tag{15}$$

7. Start the retrieval phase. First step requires to move all patterns to be classified ($x^{\xi} \in X$) to the new axes system.

$$x^{\xi'} = x^{\xi} - \bar{c}, \quad \forall x^{\xi} \tag{16}$$

8. Assign the pattern $x^{\xi'}$ to the group represented by the center $Z_j^{0'}$, which is identified by its associated pattern y^{ξ}. For this association, the *j*-th center is the *j*-th component of the vector y^{ξ} obtained by the equation (7).

9. Calculate new centers for the partition generated in the previous step and the intra-cluster spread index (*I*) using the equations (9) y (10).

10. Assign $it = it + 1$. If Z_j^{it} is equal to $Z_j^{0'}$, send to the output Z_j, S_j, $\{x^{(\xi,j)}\}_{\xi=1}^{S_j}$ and finish. Otherwise, if $it > t_{max}$ send the message "Number of iterations exceeded"; else, assign $Z_j^{0'} = Z_j^{it}$ and go to step 3.

The PAC-Means algorithm finishes in step 10. The key of this technique and main difference with c-Means is the associative classification by means of the PAC algorithm.

4. Materials and Methods

This section presents three datasets and the methodology utilized for comparing the algorithms c-Means and PAC-Means. The Iris, Wine and Pima datasets belong to the machine learning repository of the University of California, Irvine, [14].

4.1. Iris Dataset

This dataset is one of the most used in pattern recognition. It contains three classes and 50 rows of each class making reference to a certain kind of plant. Only one class is linearly separable from the others. Table 1 shows a statistical summary of this dataset.

4.2. Wine Dataset

This dataset was the result of a chemical analysis carried out on wines in a region of Italy, selected from three different crops. Three variants of Wine were analyzed from which 13 different samples were collected. A statistical summary of this dataset is shown in Table 2.

4.3. Pima Indian Diabetes Dataset

This dataset contains 768 women registers of at least 20 years old, 268 of which are diabetic and the remaining 500 non diabetic. Each register has eight features, which are reported in Table 3 together with a brief statistical summary previous published in [8].

Table 1. Statistics of the Iris dataset

Feature	Min_Max	Mean	Stn.Dev.
Length of the sepal (cm)	4.3 - 7.9	5.8	0.82
Wide of the sepal (cm)	2 - 4.4	3.05	0.43
Length of the petal (cm)	1 - 6.9	3.76	1.76
Wide of the petal (cm)	0.1 - 2.5	1.2	0.76

Table 2. Statistics of Wine dataset

Feature	Min –Max	Mean	Stn. Dev.
Alcohol	11.03-14.83	13	0.81
Malic acid	0.74-5.8	2.34	1.12
Ash	1.36-3.23	2.37	0.27
Alcalinity of ash	10.6-30	19.49	3.34
Magnesium	70-162	99.74	14.28
Total phenols	0.98-3.88	2.29	0.63
Flavanoids	0.34-5.08	2.03	0.99
Nonflavanoid phenols	0.13-0.66	0.36	0.124
Proanthocyanins	0.41-3.58	1.59	0.57
Color intensity	1.28-13	5.06	2.32
Hue	0.48-1.71	0.96	0.228
OD280/OD315	1.27-4	2.61	0.71
Proline	278-1680	746.89	314.9

Table 3. Statistics of the Pima dataset

Feature	Min_Max	Mean	Stn.Dev
Number of times pregnant	0-17	3.8	3.4
Plasma glucose	0-199	120.9	32.0
Diastolic blood pressure (mm Hg)	0-122	69.1	19.4
Triceps skin fold thickness (mm)	0-99	20.5	16.0
Two hour serum insulin (mmU/ml)	0-846	79.8	115.2
Body mass index	0-67.1	32.0	7.9
Diabetes pedigree function	0.07-2.4	0.5	0.3
Age (years)	21-81	33.2	11.8

4.4. Methodology

To monitor the performance of the PAC-Means algorithm, experimental tests were carried out according to the following configuration:

1. Two classifiers were employed, c-Means and PAC-Means. Each classifier was applied to the three datasets: Iris, Pima y Wine.
2. Both classifiers were adjusted with the same initial parameters; the number of desired groups (m) was equal to the number of real groups in each dataset: three for Iris and Wine datasets and two for Pima. Two stop criteria were defined, the maximum number of allowed iterations (it_{max}) set to 20 and an error tolerance $\theta = 0.005$ with respect to the change of the index (I) between two consecutive iterations.
3. All the registers in datasets and all the features for each register were employed in every test.
4. In this study an execution consists in applying one classifier to a dataset. For each possible combination of classifier and dataset, 1000 executions were done. The results of these executions were descendingly sorted by an intracluster-spread index (I), defined by the eq. (10). The lower this index, the better the clustering process is.

5. Results

A summary of the results obtained from experimental tests is presented in Table 4. This table provides information about the maximum, minimum, mean and standard deviation of our comparison criterion, the intra-cluster spread index.

In all the tests, PAC-Means reaches greater maximum indexes than c-Means, but also got lower minimum indexes, which means that our proposal explores different areas of the solution space and, because of this, it could form better clusters.

In average, c-Means shows better performance. However, it should be considered that the mean of PAC-Means is widely affected by some outlier values.

In addition to Table 4, Figures 1 to 3 compare de performance index reached by the c-Means and PAC-Means algorithms on Iris, Wine and Pima datasets, respectively.

Figure 1 presents the behavior of c-Means and PAC-Means when both are applied to the Iris dataset. This figure shows how c-Means indexes remain constant on two sections of the whole test meanwhile PAC-means vary on a wide range of values, reaching in some cases, better indexes than c-Means.

Table 4. Statistics of the c-Means and PAC-Means Performance

Classifier / Dataset	$I_{máx}$	I_{min}	$I_{average}$	$I_{stnd.dev}$
PAC-Means / Iris	696.78	**76.49**	146.35	81.45
c-Means / Iris	90.29	77.92	90.29	25.47
PAC-Means / Pima	**3.33 E7**	**4,781,423**	**5,213,639**	**1,543,363**
c-Means / Pima	5,142,388	5,142,353	5,142,373	4.92
PAC-Means / Wine	3.25 E7	2,375,842	5,262,198	3,730,073
c-Means / Wine	2,654,653	2,370,681	2,428,273	108,797

Figure 2 shows-c-Means and PAC-Means applied to the Wine dataset. In this figure is possible to observe the superiority of c-Means in most of the executions. Nevertheless, PAC-Means reaches similar indexes in about 20 percent of the test, and as it happened on the Iris dataset, PAC-Means got some minimum indexes lower than c-Means.

Finally, Figure 3 provides strong evidence in favor of the hypothesis that PAC-Means algorithm can overcome c-Means when both are applied to the same dataset under similar conditions. The evidence is easy to evaluate because in about 98 percent of the test PAC-Means reaches lower indexes than c-Means.

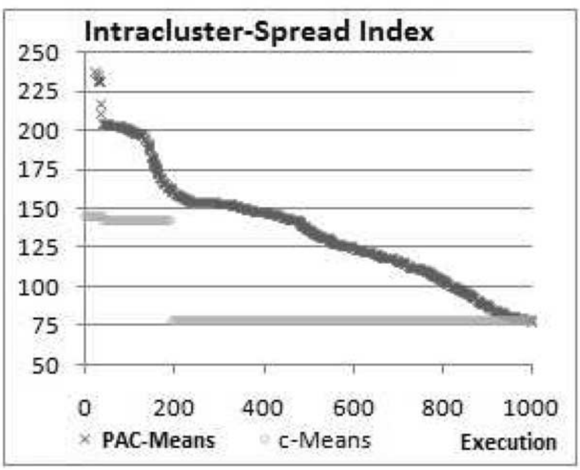


Figure 1. Performance on Iris dataset

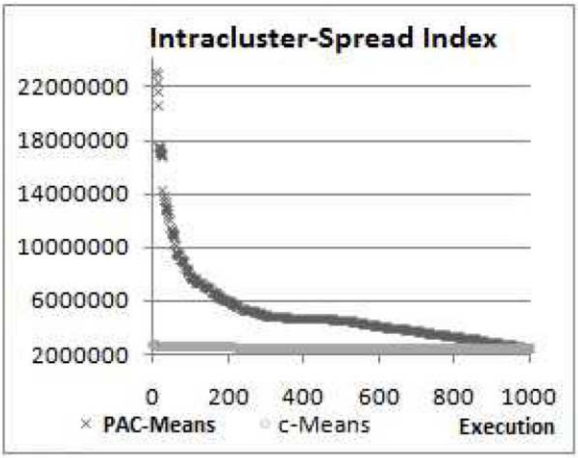


Figure 2. Performance on Wine dataset

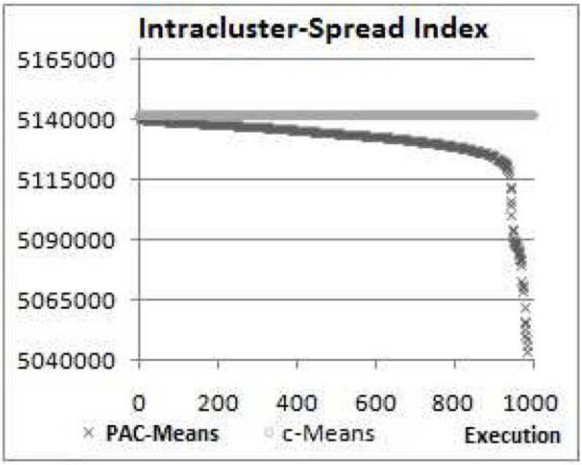


Figure 3. Performance on Pima dataset

5.1. Conclusions

In this study a non-supervised pattern classifier was proposed as a hybrid of c-Means algorithm and associative memories. The proposal was compared on similar conditions against one of the most widely used algorithms for clustering.

Table 4 shows than, in average, c-Means has better performance than PAC-Means algorithm. However, PAC-Means reached lower indexes than c-Means algorithm in two out of three tests; which is evidence that is possible to get better clusters than those grouped by c-Means.

Results in experimental tests show that PAC-Means overcomes c-Means when both algorithms are applied to Pima dataset and evaluated with an intracluster-spread index. These results are encouraging and allow us, in consequence, to propose the PAC-Means algorithm as a competitive alternative for pattern clustering.

Moreover, the associative memory employed in PAC-Means can be used later as a supervised classifier. The results and performance showed by PAC-Means allow us to propose it as an alternative for pattern clustering. Further studies will be presented in future articles.

Acknowledgements

The authors want to thank the financial support and motivation provided by the Technological Institute of León.

References

[1] H. Park, J. Lee and C. Jun, A K-means-like Algorithm for K-medoids Clustering and its Performance, *Proceedings of ICCIE* (2006), 102-117.
[2] J. MacQueen, *Some methods for classification and analysis of multivariate observations in Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 1967.
[3] S. Das, A. Abraham and A. Konar, *Metaheuristic Clustering*, Springer-Verlag, Berlin, 2009.
[4] L. Ogiela and M. Ogiela, *Cognitive Techniques in Visual Data Interpretation*, Springer, Berlin, 2009.
[5] R. Santiago, *Clasificador Híbrido de Patrones basado en la Lernmatrix de Steinbuch y el Linear Associator de Anderson Kohonen (Master thesis in Spanish)*, Instituto Politécnico Nacional-Centro de Investigación en Computación, México DF, 2003.
[6] I. López, *Clasificador automático de alto desempeño (Master thesis in Spanish)*, Instituto Politécnico Nacional-Centro de Investigación en Computación, México DF, 2007.
[7] S. Montero, J.L. Díaz-de-León and C. Yánez, *Clasificador asociativo de patrones: avances teóricos en el XII Congreso Internacional de Computación*, Instituto Politécnico Nacional-Centro de Investigación en Computación, México, DF, 2003.
[8] L. Padierna, R. Santiago, A. Zamarrón and V. Zamudio, *Pattern Associative Classifier Applied to Diabetes Mellitus Diagnosis in the 3er. Congreso Internacional en Ciencias Computacionales*, Universidad Autónoma de Baja California, Ensenada BC, 2010.
[9] P. Sussner and M. Valle, *Fuzzy Associative Memories and Their Relationship to Mathematical Morphology in Handbook of Granular Computing*, Wiley and Sons Publication, Chichester, 2008.
[10] K. Steinbuch, Die Lernmatrix, *Kybernetik* (1961), 36-45.
[11] T. Kohonen, Correlation Matrix Memories, *IEEE Transactions on Computers* (1972), 353-359.
[12] M. Friedman and A. Kandel, *Introduction to Pattern Recognition*, World Scientific, London, 1999.
[13] J. Marqués-de-Sa, *Pattern Recognition: Concepts, Methods and Applications*, Springer, Germany, 2001.
[14] A. Frank and A. Asuncion, *UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]*, School of Information and Computer Science at the University of California, Irvine, 2010.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-47

Diagnostic Aid System based on triaxial high sensitivity MEMS for detection and characterization of specific neurodegenerative diseases

BUSTAMANTE-BELLO, Rogelio [a], RAMIREZ-RODRIGUEZ, Arleni E. [b], YÉPEZ-PÉREZ, Luis [c]
Centro de Microsistemas y Biodiseño
Tecnológico de Monterrey, Campus Ciudad de México. Mexico City
[a] *Researcher, Department of Electronic and Mechanics Engineering.*
rbustama@itesm.mx
[b] *Research Center Assistant. arleni_ra@hotmail.com*
[c] *PhD Student, Computer Science. lyepez@gmail.com*

Abstract— The system presented in this paper aims to serve as a low cost tool for detection and characterization of different neurodegenerative diseases by measuring the skeletal muscle movement in the patients´ fingers. Conditions like Parkinson's, Multiple Sclerosis, psychiatric diseases, alcoholism and drug abuse have specific movement frequencies, which could be graphed. The system looks to make possible the identification of the disease mapped according to the characteristic graph, reducing the number of tests needed for this matter. Research focused, it aims to measure the disease's intensity and progress as well as the effect of medicine treatment and therapy.

Keywords: Neurodegenerative, MEMS Application, Micro-accelerometer, biomedical systems

Introduction

Over the course of the years, human being's life expectancy has increased considerably, and with this the importance of treating certain diseases that lower our life's quality. Out of all the possibilities, a particular group of diseases is called movement disorders, which is closely related to the degeneration of the nervous system and therefore, to the inability to control motor skills. These pathologic conditions are commonly found in people that are around sixty-five years old; however, recent studies have shown that the symptoms may be diagnosed since the potential patient is forty years old [1].

Diseases known as movement disorders are typically characterized by the following symptoms: muscle stiffness, tremor, slowness of movements, loss of

This work was supported by Tecnológico de Monterrey, Campus Ciudad de México and developed by the "Centro de Investigación en Microsistemas y Biodiseño" (CIMB).

voluntary movement (skeletal or muscular) and postural instability. Within this group of medical conditions we can consider Parkinson´s disease, Multiple Sclerosis, Syndenham Chorea and Huntington's disease, among various others. Additionally, involuntary movement of the limbs caused by chronic alcoholism and drug addiction may also be found amongst a given patient [2].

Different approaches to quantify and to qualify such tremors have been done [3, 4, 5]. With the aid of electromyography, computer-tracking tasks, infrared sensors, accelerometry technology using microelecromechanical system (MEMS), and electromagnetic position sensors engineers have tried to aid medical research.

Currently there are two main empirical tests used for data collection. One is to ask the patient to tap the palm of his hand with the index and the middle fingers of the opposite hand as quickly as possible during a given time, and afterwards the frequency is calculated. In the other available test, the patients' force is measured by asking him or her to apply pressure on a squishing object with his hand.

The system presented in this paper aims to serve as a tool for the diagnosis of specific neurodegenerative diseases, remarking Parkinson's disease, Huntington's disease, chronic alcoholism and multiple system atrophy, as well as an aid for obtaining data for further research focused on the development of new technologies to treat these sicknesses. To achieve these objectives, our device focuses on measuring the physical variable of skeletal muscle movement in the patients´ fingers. Using the data obtained by accelerometry it becomes possible to distinguish the intrinsic characteristics of a tremor found in Parkinson's disease form other neurodegenerative diseases or conditions (natural movements, alcoholism and drug abuse) that cause movement disorders. It is our intention to develop an important tool for clinical studies to analyze the effectiveness of the drugs used in treatments and therapies. Additionally, the system's user-friendly Graphic User Interphase (GUI) is intended to make the data collection easier, allowing doctors and medical staff to track the evolution of the patient's symptoms and generate new databases to provide useful information to the scientific community, as well as helping to do an accurate evaluation of new drugs and treatments. A PCT was filled and published with the design of the system in order to protect the intellectual property [6].

1. State of the art

There are several engineering applications in which tilt measurements are required. From earthquakes to motor systems, tilts and other harmonic movements must be measured by certain devices to avoid malfunctions. Following this basis, on the medical environment, tilt is used for making blood pressure monitors more accurate. Certainly, this particular case can be used to quantify tremors. Gradually there have been novel applications that use this technology, commonly found in a commercial use on mobile devices [7, 8]. Even if the logic and hardware used for collecting data inside these devices is based in accelerometry, the results are not as precise as medical researches need to be, mostly because the device is placed on top of the patient's hand reducing the stability and veracity of the measurement. The advantage of using this type of devices could only be found in the fact that is the wireless Internet connection that enables data transfer.

Other approaches like 3-D image processing [3] and electromagnetic sensors have been able to accomplish the main goal of quantifying tremors, but the lack of precision is still present. Wherever is the lack of a three-axis measurement device or the usage of a low-frequency enabled accelerometer, errors and misinterpretation of data are commonly found. Therefore, a device able to operate with a higher precision is needed, as tremors and involuntary movement shown in lower frequencies are registered in any human being.

The use of accelerometery to quantify tremors has been proven efficient by Dr. Warren Grundfest [9] and indicated as a viable choice. Their device can be placed easily over the entire hand. However, since the weight of said device has to be taken into consideration (as it can has some effect in the patient's pathology), not only the collected data to be registered has to be corrected, but simultaneously the natural device's weight may produce a tremor to the patient, since the patient has to produce a force with his hand to counter the device's weight.

After all the considerations made previously [10], triaxial MEMS accelerometers are chosen for this project (MMA7361). It consists of two surface micromachined capacitive sensing cells (g-cell) and a signal conditioning ASIC contained in a single integrated circuit package [11]. Its lightweight and durability gives the MMA7361 additional statements to consider it a smart choice for the project.

The ASIC uses switched capacitor techniques [11] to measure the g-cell capacitors and extract the acceleration data from the difference between the two capacitors. The ASIC also signals conditions and filters (switched capacitor) the signal, providing a high level output voltage that is radiometric (linear scale of the offset voltage and sensitivity with applied voltage).

Other key feature is the ability to choose from different sensitivities depending on the requirements for optimum performance. The g-Select pins allow this change depending on the application chosen, where g is the gravitational constant 9.81 m/s^2

Another aspect in which the MMA7361 becomes useful for our needs is in the sensitivity of the device. As described on the Freescale's application note [12], an accelerometer consists delivers an output voltage depending on the tilting registered from -90º to 90º with reference of the original position. This input-output behavior is typically graphed in a sinusoidal wave, with the slope of the graph being the sensitivity (or resolution) of the sensor. Thus, the steeper the graph is the better sensitivity it has. This device turns out to have the highest sensitivity considering its low cost.

The sensitivity chosen for this project is 800mV/g at 1.5g. The output voltage can be calculated in relation with the degree of the tilt obtained according to the following equations

$$V_{out} = V_{offset} + \left(\frac{\Delta V}{\Delta g} \times 1.0g \times sin\theta\right)$$

Where:

V_{out} = Accelerometer output in volts

V_{offset} = Accelerometer 0g offset

$\Delta V / \Delta g$ = Sensitivity

1g = Earth's gravity

θ = tilt

Finally, a crucial aspect mentioned before regarding the accelerometer used versus the previous projects [4, 10] is the frequency response in which he MMA7361 is working. This allows us to identify tremors and muscle movement in the range of 1 to 16 Hz, making it possible for us to identify different types of muscular movement in the X, Y, and Z-axis corresponding to different neurodegenerative diseases.

2. Development

The system [6] described in this paper consists of at least four accelerometers attached to a glove and a one more which could be attached to the wrist or another place where the specialist may wish, a 16 bit USB – 6211 data acquisition card (DAC) from National instruments® and a laptop computer with a LabView based software. Four fingers are monitored to get more information and being able to distinguish between different diseases with some similar patterns.

Figure 1 shows the previous and the most recent version of the glove. The most recent version uses 4 accelerometers plus one attached to the wrist, which can be attached to a different place is the specialist required to do that.

Since the DAC module has the capacity to receive 16 analogic inputs, we use the entrances A17-A32 of the device, which represent the five accelerometers with the three axes, giving as a result fifteen analogic signals.

Figure 1: Glove with accelerometers attached to fingertips (a) original version and (b) last version

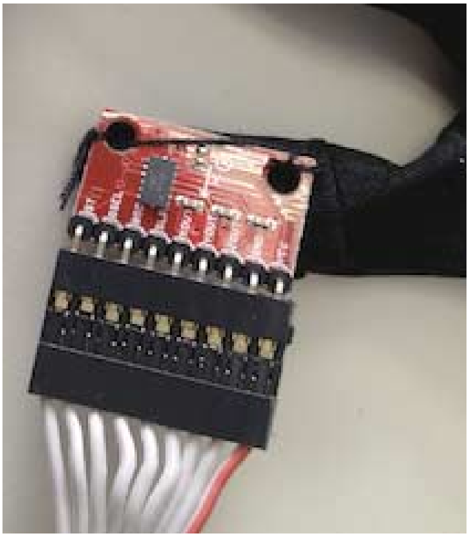


Figure 2: MMA7361 Accelerometer attached to glove's fingertip

The acquisition frequency is set at 1 kHz in which 200 repetitions can be read, resulting in 200,000 samples per second, as proven with the Nyquist Theorem. Afterwards, the signals leave the DAC as a combined signal, thus we have to decompose the resulting signal axis-by-axis (x, y, z) and finger-by-finger. Fifteen groups of data are then created, three per finger. To process the information, the only additional consideration to be done is the fact that the y-axis (the vertical one) is affected naturally by the gravity, so during the calculation we must subtract 1G of acceleration to the total calculus.

To avoid the noise in the system (given the fact that our device works on the millivolt order), a pass band filter is included on the design. As described on Figure 3, such filter should avoid frequencies below 0.001 Hz (DC current noise) and allow frequencies below 16 Hz, in such way that we can avoid environmental noise of 50 and 60 Hz.

The human being hand's articular dependence results in the fact that fingers cannot move independently at different frequencies. Taking advantage of this fact the accelerometers are placed on the fingertips of the glove, allowing us to measure independent movement on each finger. The advantage of this procedure is that different diseases have characteristic tremors. This fact gives us the faculty to associate all the accelerations of every finger on the X, Y and Z-axis respectively. This association results in a bigger quantity of sample signals per axis of movement. In this case we can average all the frequencies per axis and have a resultant graphical result like the one on Figure 3.

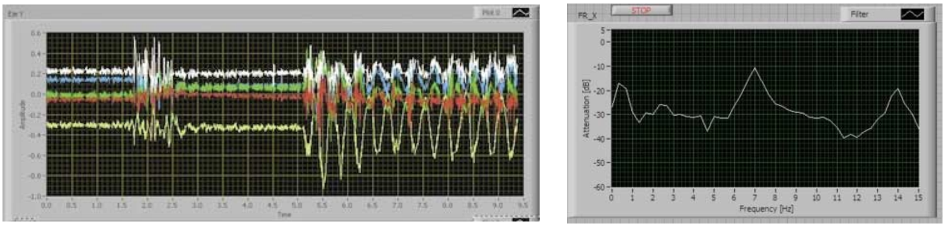


Figure 3: (a) Time and (b) frequency spectrum of the average of the x-axis signal

Figure 4: Visual part of the software revealing a match on a given disease with the measured frequencies

Finally, since the frequency measurement is done by pondering all axes, it is possible to elaborate an arithmetical average of each axis, obtaining a final frequency which can be compared with the descriptive frequencies of each pathology out of the described neurodegenerative diseases. This comparison is elaborated by AND logic in which the software itself recognizes the frequencies and compares it to the possible disease diagnosed. The software has been programmed to detect 4 possible pathologies: Parkinson's disease, Huntington's disease, chronic alcoholism and multiple system atrophy. Actually the system does a simple frequency comparison to choose the best possible disease. A better classification system is actually a work in progress to achieve a most reliable result, using modern classification techniques.

Table 1 shows a small exempt of the possible neurodegenerative diseases that may be diagnosed with the automatic feature within the software. The duration of the measurement is, by default, 10 seconds. The software has the enabled option to save each axis' measurements on a spreadsheet to be analyzed on any given time by a doctor. A database is created to help monitor the disease.

Table 1. Some neurodegenerative diseases that may be diagnosed

		Parkinson's disease	Chronic alcoholism	Huntington's disease	MSA-P	MSA-C
Frequency range (Hz)	1					
	2		X	X		
	3		X	X	X	
	4	X	X	X	X	
	5	X	X		X	X
	6	X			X	X
	7					X
	8					X
	9					X
	10					
Tremor	On rest	X	X	X	X	
	Postural / Action		X			X
	Kinetic		X			
Limbs involved	Dominant superior limb	X	X	X	X	X
	Non dominant superior limb	X		X	X	X
	Dominant inferior limb	X		X		
	Non dominant inferior limb	X		X		
	Other	X		X	X	
Synchrony	Synchronous		X			
	Asynchronous	X		X		
Symmetric	Freq		X			
	Range		X			
Asymmetric	Freq	X		X	X	
	Range	X		X	X	

3. Test and results

Initial tests were made with the original glove at the "Instituto Nacional de Ciencias Básicas y Nutrición Salvador Zubirán" with the help of the Neurologists and implementation of a clinical protocol created for this purpose. Tests were performed on patients who have already been diagnosed with a neurodegenerative disease.

There have been two main tests conducted at this institution, one for patients with Parkinson's disease and one for patients with Multiple Sclerosis.

There were two tests on a Parkinson's patient; one is shown in Figure 5. Here we get an average value detected, inside the values predicted in Table 1.

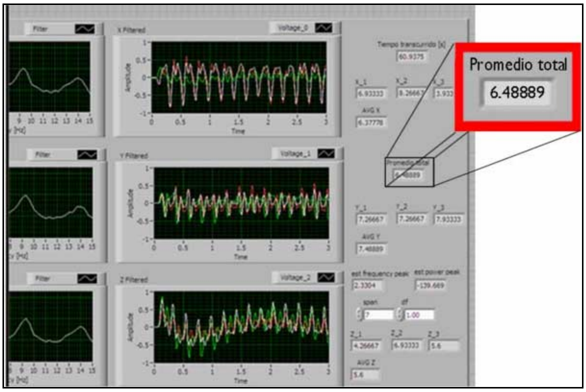


Figure 5: The average frequency detected is 6.48889 Hz

This clearly indicates that the results are well within the known frequency for Parkinson's disease tremors, based on Table 1. The second test was conducted with different conditions as the patient had already taken his medication, which diminished the intensity of the tremors. Results illustrate the effects of said medicine in the frequency and signal graphed. From these results doctors can then learn of the effects of a specific type of medicine in their patients, as it was clear in this case.

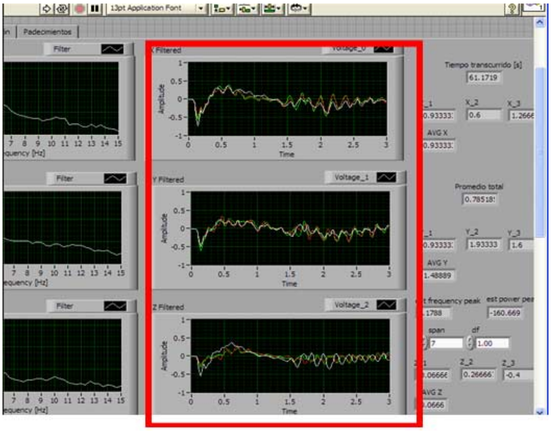


Figure 6: Different signal pattern for Multiple Sclerosis

The next case is one where the patient shows symptoms of Multiple Sclerosis. This disease is characterized by degeneration of the central nervous system and has involuntary muscle movement among its symptoms. Its frequency range is 4 to 5 Hz.

The patient had a particular form of Multiple Sclerosis, which presented tremors as she tried to reach for an object.

The results shown in Figure 6 show a different signal pattern than the Parkinson's patient, even though both diseases' frequency is between 4 to 6 Hz range.

Finally, the system may work with both hands (two gloves), but having in mind this equipment will be a mobile system used in a clinical environment, and in several cases, the tests required to be performed in just one hand, the system may disable by software the second glove. This last operation allows the extended use of the system if should be operated by batteries.

4. Conclusion and future work

The system operation is user friendly, as anyone can operate it. The procedure is done fast with 20 seconds to measure and evaluate the data obtained. It is planned to be compatible with any computer with Windows and Linux as OS, preferably with a laptop as to ensure portability. With the usage of MEMS accelerometers, tremors induced by the system are eliminated as the weight of said MEMS can be discarded, and will not cause them. The procedure planned for the measurement is not invasive; the patient just need to wears a glove.

Doctors in medicine, researchers and students can benefit from using this system; it has the advantage to be low cost compared to the other systems in existence [4]. It can be used to detect neurodegenerative diseases, distinguish them from one another helping the doctor with the diagnosis, monitoring the disease in a patient, and researchers can see the effects of new drugs and therapies with the data obtained.

The classification algorithm is being improved as an ongoing work. Labview code is being used to be able to pinpoint the disease

References

[1] Lepoutre A, Devos D, Blanchard-Dauphin A, et al. "A specific clinical pattern of camptocormia in Parkinson's disease." Journal Neurolorogy Neurosuryg. Psychiatry. 77 (11): 1229–34.

[2] Drake RL, Vogl W, Mitchell AWM. (2005). "GRAY Anatomia para estudiantes." *1ª ed.* México: Elsevier.

[3] Yu Su, Charles R. Allen, David Geng, David Burn, Una Brechany, G. Duncan Bell, and Roger Rowland. "3-D Motion System ("Data Gloves"): Application for Parkinson's Disease". (2003)

[4] M. Welsh, P. Bonato et at. "Monitoring Motor Fluctuations in Patients with Parkinson's Disease Using Wearable Sensors". IEEE Transactions on Information Technology in Biomedicine, Vol. 13, No.6 (2009)

[5] K.Grandez, G.Solas, P.Bustamante, B.Sedano, "Sensor Device for Testing Activities in Parkinson and ALS Patients." Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on-NO PERMISSIONS (2010).

[6] International filled patent PCT/MX2011/000077 (WO 2012/002791 A2); http://www.wipo.int/patentscope/search/es/WO2012002791, date: january 5, 2012.

[7] Robert LeMoyne, Timothy Mastroianni, Michael Cozza, Cristian Coroian, and Warren Grundfest. "Implementation of an iPhone as a Wireless Accelerometer for Quantifying Gait Characteristics". Engineering in Medicine and Biology Society (EMBC), (2010).

[8] Robert LeMoyne, Timothy Mastroianni, Michael Cozza, Cristian Coroian, and Warren Grundfest, "Implementation of an iPhone for Characterizing Parkinson's Disease Tremor Through a Wireless Accelerometer Application". Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE

[9] Robert LeMoyne, Timothy Mastroianni, Michael Cozza, Cristian Coroian, and Warren Grundfest. “Quantification of Parkinson’s disease characteristics using wireless accelerometers”. Complex Medical Engineering, 2009. CME. ICME International Conference on (2010)

[10] Alessio Murgia, Vincent Kerkhofs, Hans Savelberg, and Kenneth Meijer. “A Portable Device for the Clinical Assessment of Upper Limb Motion and Muscle Synergies”. Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE.

[11] Freescale Semiconductor MMA7269QT Datasheet (2006).

[12] Freescale Semiconductor AN3107 Application Note (http://www.freescale.com/files/sensors/doc/app_note/AN3107.pdf)

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-57

Assesment of Respiratory Diseases through Acoustic GMM Modeling

Pedro MAYORGA[a,1], Cristopher DRUZGALSKI[b], Oscar Hugo GONZÁLEZ[a], Hernán Silverio LOPEZ, and Marco Antonio CRIOLLO[a]
[a]*Instituto Tecnológico de Mexicali, Mexicali, B.C, México*
[b]*Electrical Engineering,CSULB,CA, USA*

Abstract. The purpose of this paper is to present a method that utilizes lung sounds (LS) for quantitative assessment of patient health, based on the fact that LS are relates to respiratory disorders. Traditional asthma evaluation methods may involve auscultation and spirometry. However, improved diagnosis opportunities can be offered by utilizing sensitive electronic stethoscopes (now widely available), and the application of quantitative signal analysis methods. In this context, we carried out experiments using normal LS from both the RALE repository and recordings from students. In this paper, we propose an acoustic evaluation methodology based on the Gaussian Mixed Models (GMM) that should assist in broader analysis, identification, and differentiation of LS. Additionally, frequency domain analysis of peculiar sounds reflected during wheezes and crackles, and their differentiation from normal respiratory sounds should assist in improved diagnosis.

Keywords. Asthma, auscultation, crackles, Gaussian Mixed Models (GMM), stethoscope, wheezing.

Introduction

Environmental conditions and related air pollution are considered as critical factors contributing to the increase in respiratory diseases, including asthma, which continues to be a growing problem [1]. Asthma experts claim that the adventitious sounds are not a sufficient element to diagnose asthma; however, lung sounds (LS) continue to be fundamental for diagnoses of respiratory diseases [2, 3]. Medical practice often relates wheezes to asthma and/or obstructive chronic diseases [2-5]. For this reason, LS recognition is critical to describe different adventitious sounds. Studies have demonstrated that in most tests, the child's parents have not been able to identify wheezing episodes [6-8]. Furthermore, it is essential to have a well-documented clinical history, preferably including adventitious sound assessment [6, 8-10]. Auscultation performed with traditional stethoscopes has drawbacks, because the sound is highly influenced by the stethoscope's characteristics, which is often designed to eliminate or weaken low frequency sounds which the human audition cannot perceive; furthermore, auscultation has a high dependence on the physician's skills to distinguish between different sound patterns [4-7].

[1] Corresponding Author.

Digital tools provide a more precise way to identify children with early development of abnormal pulmonary sounds and to strengthen objectivity in medical diagnoses [4-7]. Previous experiments related to the assessment of human hearing show an age-associated deficiency at detecting any noise during an auscultation [8]. Digital analysis methods, such as pattern recognition and digital signal processing have shown more accurate results. Presently, different scientific communities like Computerized Respiratory Sound Analysis (CORSA), the International Lung Sounds Association (ILSA), promote the use of digital methods for respiratory disease detection [4, 5, 8-10].

1. Database and Sound Characteristics

In general, there are two distinct types of acoustic indicators associated with respiration: *tracheal sounds* (TS) and *vesicular sounds* (VS). TS can be heard over the trachea area; they have relatively high volume and a bandwidth of 0-2 kHz, but their frequency spectrum components decline rapidly for values over 850-900 Hz. VS can be found over the chest, with no immediate proximity to the central airways; their frequency ranges from 0 to 600 Hz [4,5]. Some of the normal LS are typically below 400 Hz [4, 5, 9-12]. Heart sounds are sometimes filtered off from LS by using a high-pass filter; however, they are overlapped with LS at certain frequencies (Figure 1).

In order to obtain wide frequency range coverage in LS recording, the signals must be obtained through an electronic stethoscope with a frequency range from 20 Hz to 5 kHz. The sample rates commonly used are as low as 4 kHz and as high as 22.05 kHz. The analysis length is typically between 20 ms and 50 ms, which mean that if the sample changes to 10 kHz, a signal block should have 256, 512 or 1024 samples if a Fourier transform is intended [4,8].

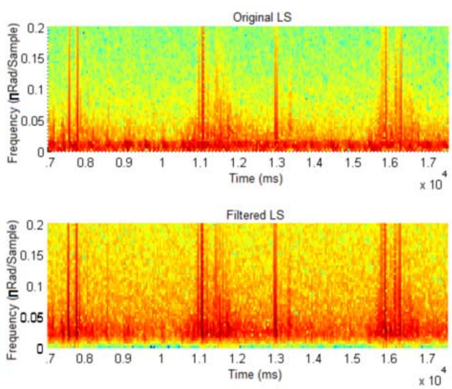


Figure 1. Example of a Crackles Lung Sound (*upper*) and its filtered version (*lower*).

1.1. Adventitial Sounds

For our goals, only the following adventitious sounds are described:

- Crackles: Discontinuous LS, with duration of less than 20 ms and frequency range is typically from 100 to 2000 Hz [4, 5, 9-12].

- Wheeze: Their dominant frequency is within the 100-2000 Hz range and their time length ranges from 80 to 250 ms [4, 5, 9-12].

1.2. RALE Lung Sounds Database

For evaluation purposes, the RALE database, developed by the University of Manitoba, Winnipeg, Canada, was utilized. It is a repository of recordings obtained from patients who exhibited normal breath, crackles, wheezes and other lung sounds. RALE contains over 50 labeled recordings for modeling, and 24 non-labeled recordings for system tests. These signals were high-pass filtered at 7.5 Hz to suppress any DC offset by using a first-order Butterworth filter. Additionally, a second eighth-order low-pass Butterworth filter at 2.5 kHz was applied to avoid aliasing. All RALE signals were sampled at 10 kHz.

2. Acoustic GMM Models for Respiratory Sounds

In recognition systems, the consideration of a large number of signals for each case is essential. They are typically divided in two partitions: the first one is used for acoustic model computation, while the second is used for evaluation. For a small database, a cross-validation method is used and recommended [13].

2.1. Training Step and Acoustic Model Generation

In this approach, LS are parameterized, performing a pre-emphasis with a FIR filter (Finite Impulse Response), applied to a 30 ms Hamming window every 10 ms. Next, the FFT algorithm (Fast Fourier Transform) is applied to the signal frame by frame, then a module is calculated over these results, and finally this module is multiplied by the Mel or Bark scale. This data representation is known as Mel Frequency Cepstral Coefficient characteristic vector (MFCC) [14].

A GMM (Gaussian Mixed Model) is characterized by its means, covariances and weights; each case is represented by a GMM model (λ). A useful and user-friendly medical system must include a pathology models *codebook* to allow discrimination in LS recordings in order to provide expected assessment of the patient's health, and characterization of possible pathologies [11-17]. The acoustic models are obtained for each case or pathology in the training step. The MFCC's are extracted from the recordings set with a given length for a possible pathology or signal that reflects normal conditions. Once all vectors are gathered, the Maximum Expectation algorithm (EM) is used to estimate the corresponding acoustic models. An extended explanation of this process can be found in [13, 14]. Vector Quantification is a very efficient alternative for initialization purposes [14]. A Gaussian mixture density is a weighted sum of M component densities, as depicted by the following equation:

$$p(\vec{x}|\lambda) = \sum_{i=1}^{M} m_i b_i(\vec{x}) \quad (1)$$

Where $\vec{x}$ is a D-dimensional random vector (MFCC font-end), i=1,..,M are the component densities, and m_i, i=1,...,M, are the mixture weights. Each component density is a D-variate Gaussian function [11, 13, 14] as follows:

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma_i|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(\vec{x} - \vec{\mu_i}) \Sigma_i^{-1}(\vec{x} - \vec{\mu_i}) \tag{2}$$

Here $\vec{\mu}$ is the mean vector. The mixture weights satisfy that $\sum m_i = 1$. The whole GMM model is parameterized by the mean vectors, covariance matrices and mixture weights for Gaussian densities. The model that collects the parameters is represented by the next expression [14, 15]:

$$\lambda_i = \{ \mathrm{m}_i, \vec{\mu_i}, \Sigma_i\}_{i=1,\cdots,M} \tag{3}$$

For LS recognition, each case is represented by a GMM and is designated as a model λ for these LS. It is reasonable to assume that MFCC respiratory signals can be characterized by a parameterized model. The spectral shape and the histograms obtained from the data clusters can be represented by several weighted Gaussian densities. Each Gaussian has its mean $\vec{\mu_i}$ and the variations of the average spectral shape can be represented by the covariance matrix Σ_i. This quantification type is important because the recorded signals not have exactly the same shape, even if they come from the same patient or auscultation point. A signal of the same pathology must be recorded with multiple patients to obtain robust models. In our experiments, the classes correspond to the data obtained from acoustic events associated with normal and adventitious LS. The signal's complexity could influence the number of densities in the GMM models.

2.2. Evaluation Step

For model optimization purposes, the GMM model calculation and the evaluation were calculated with a variable number of densities (from 1 to 20), and the best tradeoff between the best results and the lowest density numbers in the mixture was selected. Using the Bayes decision rule and discarding $p(\vec{x})$because it remains constant during the maximization process [13, 16], the fundamental formula which presents the best hypothesis in the automatic signal recognition is as follows:

$$C = arg_{1\leq i\leq I} maxp\ p(X|\lambda_I)p(\lambda_i) \tag{4}$$

Due to the large number of recordings, a large number of characteristic vectors are obtained. By assuming statistical independency among the MFCC vectors, the rule to select the best hypothesis leads to:

$$\prod_{t=1}^{T} p(\vec{x_t}|\lambda_i) > \prod_{t=1}^{T} p(\vec{x_t}|\lambda_r);\ r = 1, \ldots, I \tag{5}$$

In the voice-processing scientific context and statistical pattern recognition, $\prod_{t=1}^{T} p(\vec{x}_t|\lambda_i)$ is known as the *similarity function* [14, 16]. But as each recording involves many acoustic vectors, it is necessary to simplify calculations by avoiding computational overflows. Given that the *log* function is monotonous, it is valid to apply it in equation (5). The product operation is transformed into an addition, and the new expression is known as the *rule of maximum similarity decision*:

$$L(\lambda_I) = \sum_{t=1}^{T} \log p(\vec{x_t}|\lambda_i) \tag{6}$$

This last equation is used in the comparative process, which is in fact the input signal being associated with the acoustic model in our codebook with the most likely one. The cross-validation evaluation (CV) was applied in our experiments due to the unfortunately limited number of recordings for each class or case [4, 5, 9, 10, 15-17]. In particular, we had from 4 to 7 different recordings for each of *adventitious* signals as shown in section 4 and 5.

The partitions were built by selecting three recordings for model computation, while the last one was used for evaluation. These partitions were then alternated, which means that for each new evaluation the three recordings' configuration changed to allow the fourth one to be used for testing, and so on until four evaluations for each case were completed; in other words, four complete compositions were made with four signals per case.

3. Results

The evaluations of signals were made in two ways. The first one (which is defined as a *reference* or REF) was carried out by calculating the acoustic models with a four-signals-per-case set and evaluating with the same signals to check if the system is at least able to recognize the signals that were used to calculate the acoustic model. The second approach was cross-validation (CV), used to verify the system with different signals than those applied in the learning step (*i.e.* training set), which makes this approach more objective, as shown in Figure 2. In the CV approach, a single signal is ignored from its set during model computation, and it is used for evaluation instead.

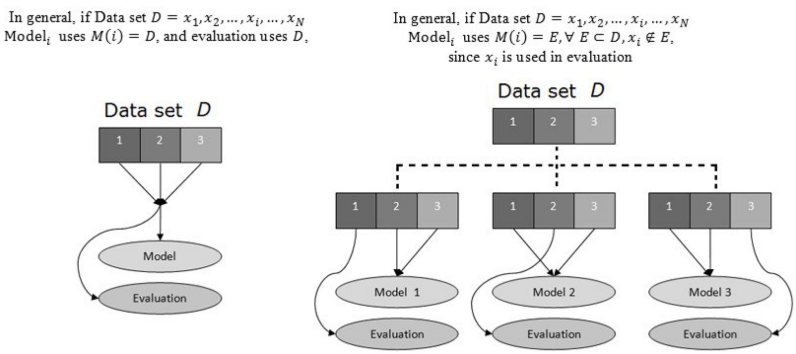


Figure 2. Evaluation algorithms: REF (*left*) and CV (*right*).

One of the first experiments with LS recordings was done to calculate different-sized GMM models, extracting observations from the RALE database as MFCC vectors, and then calculating a *codebook* using normal, crackles, wheeze and asthmatic signals, as shown in Table 1. Another experiment was performed to get the minimum necessary GMM model size (Table 1 and Table 2). Basically, the vector and model sizes are linearly increased during the experiment, while considering the inclusion or exclusion of the derivatives in the MFCC vector as well. MFCC coefficients are related to the resulting energy from a set of applied filters, linearly distributed up to the 1000 Hz order, and logarithmically from there on. The first and second derivatives displayed significant noise sensitivity, which lead to wrong classifications with crackles, whereas this is not the case with the other signals. Nevertheless, derivatives were useful to check the signal evolution and dynamics.

The best behavior results were obtained by applying 11 Gaussians (GMM11), but changes were not substantial compared to GMM9 and GMM10; however, it is evident that the results improved when MFCC vectors were greater than four coefficients for modeling, excluding some irregular "behaviors" (for GMM9, and stabilizing in GMM10). This is reasonable because bigger vectors include more significant energy bands, as shown in Table 2 (MFCC=13).

Table 1. Minimal model parameters for recognition (GMM densities and MFCC coefficients) and recognition percentage. Key: GD = *Gaussian densities*; MFCC = *MFCC coefficients*; REF/CV = *Reference/Cross-Validation evaluation*; Recog. % = *Recognition percentage.*

	Minimal acceptable parameters				**Recog. %**					
	# GD		# MFCC (w/o Δ's)		9 GD		10 GD		11 GD	
	REF	CV	REF	CV	REF	CV	REF	CV	REF	CV
Normal	3	6	3	7	92.3	48	92.3	44.2	92.3	46.1
Crackles	3	6	3	7	100	90.3	100	98	100	98
Asthma	3	6	3	7	100	50	100	51.9	100	50
Wheeze	3	6	3	7	84.6	25	84.6	25	84.6	26.9

4. Discussion

In the voice recognition context, a large quantity of data is necessary in order to compute robust models [4, 5, 14-17]; likewise, a significant quantity of LS recordings is necessary for them to represent their class (robust models). Unfortunately, the RALE database lacks an adequate number of recordings to fully represent a patient's respiratory disease or airways acoustic details. However, the respiratory system's organic function serves as a useful source of acoustic information for classification methodology and validation. Additionally, RALE has a small number of recordings of adventitious signals, and they are not correlated to particular age groups, individual characteristics, or a specific body location. Despite these limitations, the experiments show over 50% accuracy in CV, which confirms the potential of this methodology.

Through an exhaustive analysis in the REF evaluations, models with the higher quantity of available recordings were computed for each case. Tables 3 and 4 show only GMM model configurations that returned uncertainties (wrong hypothesis) during the evaluation step. Tables 3 and 4 show the quantity of coefficients included in the MFCC, quantity of densities included in GMM, and the evaluation signals (N, C, A, W or S) as well; considering an input signal to recognize (IN) and the system's output hypothesis (OUT). The uncertainties were: one with four-density models (W-C), three with five-density models (N-W, N-A, W-A), and two with seven-density models (N-W, N-W). Once again, the models with less MFCC coefficients and less GMM densities produced more recognition errors (4 and 5 MFCC coefficients, 4 and 5 GMM densities). Moreover, the signals with the worst grade of recognition were normal and Wheeze, but overall performance improved when their model coefficients and densities were increased. One explanation is that models with more MFCC coefficients and densities represent the acoustic characteristics and the fine structure of the respiratory tracts more efficiently. In most evaluations, the models with the worst recognition were those computed for normal breathing and wheeze; the reason for this might be that normal breathing has a less defined pattern than those of the adventitious signals. Since the particular anatomic characteristics among children, adults and seniors influence the frequency characteristics of the signals, a large number of diverse recordings from different patients, in diverse situations, in the same body location, and preferably for different age groups, would be necessary.

Table 2. Recognition efficiency (percentage) with MFCC vectors, with and without derivatives (deltas), for average values obtained from 1-20 densities GMM models.

	13 Coef.		13 Coef. (With Δ's)		13 Coef. (With ΔΔ's)	
	REF	CV	REF	CV	REF	CV
Normal	100%	46.2%	95%	48.7%	92.3%	51.2%
Crackles	100%	90%	100%	87.5%	100%	86.2%
Asthma	100%	61.2%	100%	53.7%	100%	50%
Wheeze	100%	30%	100%	31.2%	84.6%	32.5%

Table 3. Uncertainty summary (*i.e.* wrong identifications) in REF evaluation. Example: C-W indicates that a Crackles signal was wrongly identified as Wheeze.

Configuration	Number of Recordings			
	4	5	6	7
4 MFCC ; 4 GMM	W-C	N-C N-A W-A	0	0
5 MFCC ; 5 GMM	0	0	0	0

Table 4. Signal configuration and Quantity.

Configuration	Recording and Signal used			
	4	5	6	7
4 MFCC ; 4 GMM	N,C,A.W,S	N,A,W	N,W	N,W
5 MFCC ; 5 GMM	N,C,A.W,S	N,A,W	N,W	N,W

The best results, regardless of the experiment, were obtained with crackles; this signal was the easiest one to recognize. This suggests they have a characteristic frequency band with a good energy level contribution, which makes them easy to differentiate among asthma-related signals, wheezes and normal signals. In other words, crackles' characteristic peaks contribute statistically enough to tell them apart from the others. Consequently, this contributes greatly to the variance and the mean, resulting in a more robust GMM model.

Asthma-related signals are very complex; actually, the concept itself evolves and diverges. According to numerous opinions, diagnosing asthma just from adventitious sounds is not accurate enough. Since our asthma recordings are few and heterogeneous, they do not have conclusive distinctive statistics, nor bands with particular energy contribution either. Nevertheless, we obtained over 50% accuracy in every case (Table 1, Table 2), which encourages us to use these techniques for asthma assessment.

The melodic nature of wheezes was not properly captured in the statistics and energy bands, as shown in Tables 1 and 3. The derivatives could improve the results by highlighting a particular band or a smooth evolution of a signal. This kind of signal deserves a deeper analysis in future work.

The uncertainties in Table 3 and 4 are mainly due to a similarity between normal and wheeze signals statistical parameters. In an effort to differentiate these sounds, a normal model was calculated by using more LS recordings from healthy subjects in order to confirm that the lack of normal signals was the reason. This idea was confirmed by computing a model, which includes the additional normal signals to build the normal LS model, evidently obtaining fewer uncertainties (Table 5).

Table 5. Recognition results using an improved normal LS model calculated with RALE plus LS signals. *Uncertainties* indicate the number of wrong identification per signal type. Experiment configuration is: 10 Gaussian densities, 13 MFCC vectors.

LS type	No. of available recordings	No. of correct recognitions	Uncertainties			
			N	C	A	S
Normal	36	34	-	1	0	1
Crackles	4	4	0	-	0	0
Asthma	5	5	0	0	-	0
Stridor	7	7	0	0	0	-

5. Conclusion

The presented GMM models could improve the diagnostic classification of auscultation exams in cases of children under five years-old, or in situations where the auscultation is performed by a person with limited skill or experience. These promising results demonstrate that this alternative allows us to avoid ambiguity in adventitious signal recognition. During our experiments, the model size was varied to find the best tradeoff between accuracy and computational cost. Normally, the best results were obtained with 9, 10 and 11 Gaussian mixtures. Thus, we found these technologies have a high potential in airways disease diagnosis and overall health problems recognition. In our experiments, 52.5% accuracy was achieved by applying cross-validation evaluation; while our reference recognition accuracy achieved 98.75%.

Even though, the results were improved by applying a pre-processing stage and using the RALE database plus recorded LS signals. An extended database is mandatory, and this fact is evidenced in the expanded normal LS model experiment (Table 5). The database must contain an extended number of recordings to accomplish a refined analysis of the most important bands, to help determine an optimal set of MFCC coefficients. These considerations lead us to more robust and representative acoustic models for a given normal or adventitious LS. Additional LS recordings such as those associated with coughing will expand the experiments' capabilities and the scope of applications as well.

References

[1] B. Franklin, R. Brook, W. Cascio, Y. Hong, G. Howard, M. Lipsett, R. Luepker, M. Mittleman, J. Samet, S. Smith, I. Tager, Air Pollution and Cardiovascular Disease: A Statement for Healthcare Professionals From the Expert Panel on Population and Prevention Science of the American Heart Association; Circulation. 2004; 109: 2655-2671

[2] A. Abbas, A. Fahim, An Automated Computerized Auscultation and Diagnostic System for Pulmonary Diseases; *Computer Methods and Programs in Biomedicine*; Vol. 97, Issue 2, Feb. 2010, 141-150

[3] S. Charleston-Villalobos, G. Martinez-Hernandez, R. Gonzalez-Camarena, G. Chi-Lem, J. Carrillo, T. Aljama-Corrales, Assessment of multichannel lung sounds parameterization for two-class classification in interstitial lung disease patients; *Computers in Biology and Medicine*; Vol. 41(7), 2011, 473-482; DOI:10.1016/j.compbiomed.2011.04.009

[4] P. Mayorga, C. Druzgalski, O.H. González, A. Zazueta, M.A. Criollo, Expanded Quantitative Models for Assessment of Respiratory Diseases and Monitoring , PAHCE 2011, Rio de Janeiro. March 2011; ISBN 978-1-4244-6291-9; DOI: 10.1109/PAHCE.2011.5871938

[5] P. Mayorga, C. Druzgalski, J. Vidales, Quantitative Models for Assessment of Respiratory Diseases. PAHCE 2010, Lima, Perú, March 2010. IEEE ISBN: 978-1-4244-6291-9.

[6] A. Bush, Diagnostico de Asma en Niños Menores de Cinco, *Primary Care Respiratory Journal* (2007), 16(1):7-15., 1471-4418, General Practice Airways Group. DOI :10.3132/pcrj.2007.00001.

[7] M. Rossi, A.R.A. Sovijärvi, P. Piirilä, L. Vannuccini, F. Dalmasso, J. Vanderschoot, Environmental and subject conditions and breathing manoeuvres for respiratory sound recordings, Eur Respir Rev 2000; 10: 77, 611–615, ISSN 0905 – 9180.
[8] A.R.A. Sovijärvi, J. Vanderschoot, J.E. Earis, Standardization of computerized respiratory sound analysis, Eur Respir Rev 2000; 10: 77, 585, ISSN 0905 – 9180.
[9] C. Druzgalski, Potentials and Barriers of Extensive Auscultatory Databases 27th International Conference on Lung Sounds, September 12-14, 2000, Sweden.
[10] C. Druzgalski, Distributed Analysis of Signal Integrity Impediments in Respiratory Acoustic Signatures WC2003 - The World Congress on Medical Physics and Biomedical Engineering. August 21-30, 2003. Sydney, Australia.
[11] L.J. Hadjileontidais, Biosignal and compression Standards, M-Health Emerging Mobile Health systems, Topics in Biomedical Engineering', Springer 2006 International Book Series, pp. 277-292, ISBN 0387-26558-9.
[12] G. Charbonneau, E. Ademovic, B. Cheetham, L.P. Malmberg, J. Vanderschoot, A.R.A. Sovijärvi, Basic techniques for respiratory sound analysis, *Eur Respir Rev 2000*; 10:77, 625–635, ISSN 0905- 9180.
[13] W.L. Martinez, A.R. Martinez, Computational Statistics Handbook with MatLab, 2nd Edition, Chapman & Hall/CRC, 2008, ISBN 1-58488-566-1.
[14] D. Pearce, An Overview of ETSI Standards Activities for Distributed Speech Recognition Front-ends, AVIOS 2000: The Speech Applications Conference, San Jose, CA, USA, May 22-24, 2000.
[15] I. Sen, M. Saraclar, Y.P. Kahya, Acoustic mapping of the lung based on source localization of adventitious respiratory sound components; EMBC 2010 Buenos Aires, Pages: 3670 – 3673; ISBN: 978-1-4244-4123-5; DOI: 10.1109/IEMBS.2010.5627651
[16] B. Lu, H. Wu, S. Shih, F. Chong, M. Hsueh, Y. Chen, Combination of frequency and amplitude-modulated model for the synthesis of normal and wheezing sounds; 2011; Australasian Physical & Engineering Science in Medicine; Vol. 34 (4) 449-457, DOI: 10.1007/s13246-011-0105-1
[17] Z. Moussavi, Respiratory sound analysis [Special Issue] - IEEE Engineering in Medicine and Biology Magazine, 2007; Vol. 26 (1); DOI: 10.1109/MEMB.2007.289116

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-66

Evaluating the User Experience of a Cognitive Stimulation Tool through Elders' Interactions

Alberto L. MORÁN [1] and Victoria MEZA-KUBO
Universidad Autónoma de Baja California

Abstract. As part of our efforts to improve the effective design of suitable CS applications for the elderly, during the last three years we have been performing an ensemble of studies, which have enabled us to understand the nature of CS activities and to envision how we might support them by means of Ambient Intelligence (AmI) technologies. Furthermore, we have proposed UCSA, a framework for the design of usable CS applications for worried-well elders. In order to validate the framework, we have developed and evaluated the usability of several prototypes by means of TAM-based questionnaires. In this work, we present the results of a further evaluation of one such application, which focuses on user experience (UX). We identified which kinds of interactions occurred among the different participants, along with their form and functions, as a means to gather evidence about the application's user experience. Finally, we discuss the implications of these findings regarding the framework's UX perspective, along with possible directions of future work.

Keywords. Cognitive stimulation, user experience, elders, interaction form and function

Introduction

Research in Ubiquitous Computing (Ubicom) in the Healthcare domain, pursues the idea of "independent living" for the elderly through the seamless integration of ambient intelligence (AmI) at home [1]. AmI refers to technological environments that are sensible and that respond to the presence of its inhabitants, in order to provide them with support for their activities of daily living.

Most of these AmI environments, consider the existence of frail or dependent elders that suffer from an age-related cognitive disease, and propose technology to cope with their limitations. However, fewer works are interested in taking a preventive approach to try to reduce the elder's probability of developing a cognitive disease or to try to delay the apparition of its symptoms [2].

Although currently there is no cure for some cognitive diseases (e.g., Alzheimer's disease) [3], non-pharmacological treatments such as cognitive stimulation (CS), when performed frequently and adequately, reduce the risk of suffering from the disease or

[1] Corresponding Author: Universidad Autónoma de Baja California, km 103 carr. Tijuana-Ensenada, Ensenada, B.C. México; E-mail: alberto.moran@uabc.edu.mx

improve the cognitive state of the elder [4, 5]. CS interventions are a type of mental exercises for the cognitive functions that help maintaining the elder's cognitive state.

Unfortunately, the lack of preventive programs, and of the use of early decline detection mechanisms, causes that elders seek support until the advanced stages of the disease, when the positive effects of the intervention are few or inexistent.

Thus, the introduction of technology is promising in terms of making CS interventions accessible as an everyday activity at the elder's home, and as a monitoring tool for the specialist. However, the design of these environments faces several challenges to achieve its use by the elderly.

As a result from our previous work, we have proposed UCSA, a design framework that defines the concept of usable CS applications, and which provides a set of design guidelines to facilitate the conception and implementation of this kind of applications [6]. A usable CS application is one that is perceived by its users as useful, easy to use and enjoyable.

In order to validate the UCSA framework, we developed and evaluated Abueparty, a collocated board-based game that was designed for "at home" CS following the design guidelines proposed in the UCSA framework [7]. In a first stage, Abueparty was evaluated using subjective techniques, through the use of questionnaires and interviews to participants, to gather the elders' perception regarding the application's usefulness, ease of use and user experience.

In this paper, we present the results of a second evaluation that had the aim of gathering evidence on more specific user experience (UX) aspects of the Abueparty application. To do so, we performed a further analysis on the use of the application, identifying the interactions that occurred during an Abueparty session of actual use, and determining whether they support different aspects of UX, including whether the application allows its users to enjoy the activity, achieve satisfaction and be engaged, and this without producing anxiety or frustration to them.

The paper is organized as follows: Section 1 presents the related work on the UX concept and its evaluation, while Section 2 presents the overall methodology used to perform the study. Section 3 describes the framework, a scenario of use and the proposed system. Section 4 describes the evaluation process, the UX goals and the metrics to validate them, and the observed interactions. Section 5 discusses whether and how the observed interactions support the different UX aspects during actual use of the application. Finally, Section 6 presents our concluding remarks as well as our directions of future work.

1. User Experience and its evaluation

UX is defined as "the perceptions and responses of an individual resulting from the use and / or anticipated use of a product, service or system" [8]. Basically, UX refers to the experience that a person has when interacting with a product under particular conditions.

In recent years, obtaining the UX has begun to be an important research topic in the HCI community [9]. However, understanding and evaluating UX has not yet been established due to the variety of parameters that can be used to determine whether the use of a product can be considered as a successful UX. These parameters depend for example, on the kind of person involved (e.g. children, young adults, older adults, people with disabilities, etc.), the type of product (e.g. production, entertainment, education) as well as on other cultural and social factors [10].

Therefore, depending on the specific context of the product and population in which UX is going to be evaluated, different methods have been used to capture these experiences, e.g. interviews, observation, questionnaires, diaries, storytelling and prototyping, among others [11] or combinations of them. A brief description of some of these follows.

Self-report techniques have been used to assess the UX in mobile applications [12]. Such techniques are employed in settings where expressing emotions verbally can be complicated. Some of the techniques used include Self Assessment Manikin and pictograms, which include cards with faces expressing different emotions.

Observation techniques have been used to analyze nonverbal expressions of the subjects, as these can inform about their emotions; including those about experiences participants may be unaware during the evaluation [10]. For example, Arhippainen and Tähti evaluated the UX of the use of mobile applications in preset scenarios within the work environment [10]. In addition to observation, the techniques used included interviews with participants. In this case, each of the sessions was videotaped for later analysis. The interviews sought information about the participant's past experience, current emotions, and about application's aspects such as mobility, adaptability and context of use. The observation sought information about emotions and experiences through facial expressions and behavior in general.

Psychological techniques have also been used to measure UX. For example, Mandrik et al. uses psychological techniques to infer emotions during the use of entertainment technologies, with devices to measure the respiration rate, electrocardiogram, galvanic skin response and electromyography, in addition to the use of questionnaires to obtain information about the perception of participants regarding boredom, challenge, frustration and enjoyment [13].

Regarding our work, in a first evaluation we assessed UX using questionnaires and interviews, which captured the perception of older adults when using the application [6]. However, a known limitation of these techniques, based on information provided by users, is that users tend to respond based on what they think the researcher wants to hear, or tend to be insincere because they feel evaluated, or because they have forgotten details of their user experience [10].

Therefore, in this paper we aim at complementing our previous evaluation by using indirect observation of the videos recorded during the actual CS sessions with Abueparty. In particular, we propose to determine the user experience based on the features of interactions among participants, by considering aspects such as socialization, motivation, support, feedback, and competition, as described next.

2. Methodology

In order to understand how AmI technologies could support CS activities, during the last three years we have been performing an ensemble of studies following a user-centered design methodology. Based on this understanding, we proposed UCSA a design framework that includes a set of design guidelines for the development of Usable CS Applications.

Furthermore, to validate the framework we designed and developed Abueparty, a usable CS application that consists of a board-based game and several mini-games that promote different cognitive functions, and which allows for collocated interaction. This evaluation was conducted through the use of questionnaires and interviews to participants, to gather the elders' perception regarding the application's usefulness, ease of use and user experience.

In this work, with the aim of performing a further validation of the framework, we present an evaluation of Abueparty from a UX perspective. We used indirect observational analysis techniques to gather data, and coded gestures, corporal language and verbal expressions from interactions as a rich data source [14]. This allowed us to identify the form and function of interactions among the participants engaged in the CS activity (elders), and those guiding it (caregivers or facilitators).

A summary of the UCSA framework and the proposed Abueparty application is provided next.

3. The UCSA Framework and the Abueparty application

3.1 UCSA, a design framework for Usable Cognitive Stimulation Applications

The framework introduces the concept of usable cognitive stimulation applications. A usable CS application is one that is designed to be: cognitively useful to the elderly, easy to use by the elderly, and pleasurable and satisfying to use to the elderly.

This way, the framework proposes an ensemble of design guidelines from three perspectives: usefulness, ease of use and user experience. A brief description of these guidelines follows.

DG1. Design activities that foster the cognitive stimulation of elders. To promote a cognitive function, traditional activities and tasks recommended by therapists can be implemented with technological support following a "reuse" approach, where technology is used to replicate, and perhaps augment, these activities and tasks. Also, new activities that take advantage of the use of technology can be designed.

DG2. Assess and record the elder's performance on the activity to allow the early detection of cognitive decline. The use of technology in the CS activities of elders allows for the automatic recording and identification of their performance. This helps maintaining a current and historical record that could help determining whether there is a decline in their cognitive status.

DG3. Provide awareness information about the elders' performance. The immediate assessment of the performance on the activities allows for the provision of real-time feedback to the elderly and to make them aware of their results. It also allows notifying caregivers or specialists in cases where the elderly require immediate assistance.

DG4. Present contents in an appropriate manner that eases perception and understanding. Due to the decline in the sensory (e.g. vision, hearing) and cognitive (e.g. working memory) capabilities of the elders, it is necessary that the contents be presented in an appropriate manner to be easily perceived by them.

DG5. Facilitate application learnability. Given the elders' lack of familiarity with technology, it is important that applications make intuitive how to use them to facilitate their learning. The elders' prior knowledge about how to interact with known activities and materials can be re-used.

DG6. Afford direct manipulation of materials. Mechanisms for appropriate interaction according to the capabilities of the elderly and the features of the activity should be sought. The manipulation of input devices is usually one of the main technology interaction barriers for the elderly.

DG7. Facilitate the inclusion of and interaction with co-located and remote participants. Appropriate collaboration mechanisms for activities that require the integration of other participants (local or remote) should be considered.

DG8. Select the type of CS activities and materials according to the elder's preferences profile. CS can be carried out through different activities and materials that can be supported by technology. Choosing the type of activity according to the elder's preferences increases the likelihood that they would have fun and a pleasant experience while performing the activities.

DG9. Integrate motivators "as part of" the activity (internal motivators) according to the elder's preferences profile. Each activity is defined by a series of elements (e.g. graphical, physical, tasks and means of interaction) that if appropriate to the elder's preferences, could favor the elder's engagement in the activity.

DG10. Integrate motivators "around" the activity (external motivators) for support and socialization. There are situations external to the activity that may favor an elder's satisfaction or amusement, for instance, the presence of a friend or relative, performing the activity in a specific place or at a certain time, presenting information about his/her performance or achievements that generate satisfaction and pride, among others.

DG11. Avoid anxiety and frustration during the CS activity. The cognitive abilities of each elder are different; therefore, each activity should have different levels of complexity to allow elders to perform the activities without much problem. This seeks to prevent elders from suffering anxiety or frustration. In addition, in the presence of confusion or mistakes immediate and adequate feedback should be provided.

Further details about the framework are provided elsewhere [6].

3.2 Scenario of use

The following scenario illustrates the use of Abueparty as part of an AmI system that promotes entertaining activities as "at home" CS activities. Abueparty, a usable CS application, is designed following the guidelines proposed in the UCSA framework.

This scenario considers a group of elders playing together using a game board in a collocated fashion.

"It is Saturday night, and the Smith's are receiving the visit of the López, a tradition after Mr. and Mrs. Smith retired 2 years ago. All of them are worried-well elders. After their usual chit-chat and cup of coffee, the intelligent system detects that they are talking about playing a game, so it considers appropriate to invite them to play a board game, and offers them a few choices. They decided to play a session with the Abueparty system, Mrs. Smith and Mrs. López favorite. They move to the living room where the intelligent system recognizes and welcomes both couples and initializes the touch and tangible interfaces for the Abueparty game. The couples play a session of two complete games, which are monitored by the intelligent system. During the game, the two couples interact among them to provide feedback and motivation, and they make jokes about who is the funniest and who is going to win the game this time. Information regarding their performance is logged and then sent to a specialist who analyzes this data to monitor their cognitive wellness. Next Monday, the digital portrait at the Smith's living room receives a message and displays a beautiful flower in bloom; which indicates that the Smith's are doing cognitively well".

3.3 Abueparty, a usable cognitive stimulation application

Based on this scenario, and on the UCSA framework, we developed a multiplayer game that implements a game board similar to the "Snakes and Ladders" Mexican game (see Fig. 1). The main game integrates several mini games that implement cognitive challenges that users have to complete in order to advance through the board to the finish line [7].

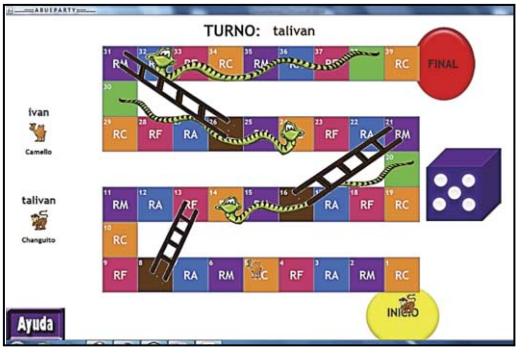


Figure 1. Abueparty main game board.

There are three types of cognitive challenges: i) Musical Challenges (shown as RM in the board), involve music or a song. Currently, there are Sing-it and Guess-it challenges; ii) Artistic Challenges (RA), involve acting or sketching. Currently, there are Act-it and Draw-it challenges; iii) Coordination Challenges (RC), involve coordination skills. Currently, there are Build-it and Match-it challenges.

Abueparty was conceived following the ensemble of design guidelines proposed in the UCSA framework. A description of the features included, according to the design guidelines, is presented in [6].

Regarding implementation, Abueparty implements touching, pressing and sliding interfaces to allows its use by elders. Touch interaction is implemented using a standard touch monitor (DELL SX2210T). Pressing and sliding interfaces are implemented by means of a custom tangible control (Fig. 2). The control implements four pressing buttons and two sliding controls using two touch sensors, two force sensors, and two slider sensors.

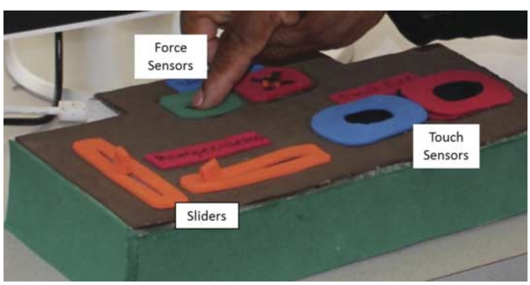


Figure 2. Abueparty custom tangible control.

3.4 Abueparty's Usability evaluation

The aim of this first evaluation was to obtain evidence regarding the validity of the UCSA framework, particularly on the adequacy of the design guidelines used to inform the development of Abueparty as a Usable CS application.

To do so, we conducted an experiment of actual use of Abueparty with a group of elders (our target end users), and obtained their opinion regarding the perceived usefulness, ease of use, experience of use and intention of use of the proposed system. The elder's perception was obtained using an on-exit TAM-like questionnaire. Questions were answered using a 5-point Likert scale, ranging from (1) "Completely Disagree" to (5) "Completely Agree", and additional open questions were included in order to obtain details regarding the adequacy of mini games, controls and functionalities.

Evaluation results indicate that all participants perceived the system as useful (4.56/5), as easy to use (4.39/5), and enjoyable (4.67/5), and all participants have a positive intention of use regarding the system (4.53/5). Furthermore, 94% of the subjects stated that they would use it two or more days per week, and 74.46% of these would use it for one or more hours per day. More information on these results is provided elsewhere [6]. This way, these results of successfully applying the framework for the design of a usable CS application, provide promissory preliminary evidence towards its adequacy for the design of this kind of applications.

4. Abueparty's UX evaluation

In order to further validate the UX from Abueparty's use, we applied indirect observation to gather information about the interactions that occurred among participants during a session of actual use.

4.1 Process

We observed the activity of two groups of four elders while they played the game. A facilitator assisted them during the sessions. Each session lasted 25 minutes, which resulted in a total of 50 minutes of recorded video. Later we analyzed the interactions that occurred in these two sessions to identify the form and the functions they served.

The session started with a 5-minute introduction to the game from the facilitator, where the goal and rules of the main game and mini games were explained to the four participants.

After this introduction, elders were allowed to play the game with minor interventions from the facilitator in order to let them interact with the application and among them, actually concentrate on the task, and freely express their thoughts and perceptions through their behaviors.

4.2 UX Metrics for the evaluation of Abueparty

In order to evaluate Abueparty, we considered the following goals: engagement, enjoyment, satisfaction and anxiety or frustration. These goals were validated through an analysis of the interactions among participants during the Abueparty activity. A brief description of these metrics follows:

- *Engagement*. To determine whether the activity has engaged the elderly, we considered the participant's involvement in the activity, taking into account the type and frequency of their interactions. To further engage the elderly in the activity, the framework proposes to integrate external motivators, e.g., activities that promote socialization.

- *Satisfaction*. To favor a UX that generates satisfaction in elders, the framework recommends integrating motivational elements as part of the activity, according to their preferences. Integrating challenges and a competition level, may be satisfying. In the Abueparty activity, we can measure satisfaction by the participants' expressions of joy, e.g. celebrating their own successes, or those of other participants.

- *Enjoyment*. To validate whether participants enjoyed the activity, we considered their behaviors during the interactions, and analyzed the verbal and nonverbal information available, including laughter, gazes, funny gestures and social interactions outside the activity, but originated from the tasks performed. We considered all these as evidence of enjoyment.

- *Low anxiety or frustration*. Since every elder has different skills and knowledge, it is important that tasks are suitable to them according to their capabilities, in order to reduce anxiety or avoid frustration, especially in activities that involve time or competition.

4.3 Interactions observed during the CS session.

Based on the observational study, we identified a set of interactions that occurred among the Abueparty game participants, as well as some of the functions of these interactions (see Table 1). A brief description of these interactions, as well as their identified functions and forms, follows.

- *Elder – Elder interactions.* This type of interaction occurs between elders who are participating in the activity. Any elder may start the interaction. Identified functions include: Feedback and help, Encouragement, Social communication (social topics outside the CS activity), Competition, and Turn coordination.

- *Elder – Facilitator interactions*. This type of interaction occurs between an elder and the facilitator, with the elder starting it. The interaction starts with a call from the elder to the facilitator, who responds and attends the request. Identified functions include: Request for help or Confirm an action before executing it.

- *Facilitator – Elder interaction*s. This type of interaction also involves the elder and the facilitator, but the latter initiates the interaction. Identified functions include: Give instructions to the elder, or Provide feedback when required.

Table 1. Interaction types, identified actors and functions of a Group Collocated session.

Actors involved	Interaction functions
Elder - Elder	Feedback and help between elders Encouragement between elders Turn coordination between elders Social communication between elders Competition between elders
Elder – Facilitator	Request for help Confirm an action
Facilitator – Elder	Give instructions to elders Feedback and help to elders

4.4 Examples of the form of the observed interactions

1) Feedback and help between Elders. During a game, when B accomplishes her mini game, the system presents a happy face to reward her achievement and congratulate her. In order for the [modal] window to disappear and allow the players to continue with the game, B has to touch the "Accept" button. Other player (R) notices it and tells her:

B: *And now?* … [The modal window is waiting for an input]
R: *Well done B! Now you have to touch the "Accept"* [button].
Once B notices this, she says:
B: *Oh, that's true! ... Done!*

2) Encouragement between elders. During a game, L is presented with a "Guess it" musical challenge. L answers it correctly, and another elder (A), who is sitting beside her, congratulates her:

A: *Yes! That's it!* [While giving her a hug and patting her back].
Next, D and E congratulate her and give her a round of applause.
D and E: *Well done, L!* … [While applauding].

3) Competition between elders. During a game, B keeps challenging H to do better than her on the various mini games. While in a "Sing-it" mini game, H does not know a song that includes the word "Clock", and M tries to help her.

M: *Do you remember?* … [Humming] … [While looking at H]
B: *Hey M! ... Stop it! ...* [Smiling] …
M: [Humming]
B: *Don't tell her! ...* [Less smiling] …
B: *H, you have to choose your own song!* [In a serious tone]
Finally, H sings a song, and conquers the mini challenge.
H: *Come on! That was easy!* [While smiling]
B: *That was not fair! ... They helped you!* [Still in a serious tone]

4.5 Frequency of Interactions in Group Collocated CS sessions

After identifying the different types of interactions, we quantified them based on the actors involved, and based on the functions they served. During the two sessions (50 minutes in total), there were 127 interactions among the participants.

As shown in Fig. 3, from these 127 interactions, 78 (61.42%) were Elder-Elder interactions, 22 (17.32%) were Elder-Facilitator interactions, and 27 (21.26%) were Facilitator-Elder interactions. As expected, elders were the most active during the session, initiating 100 (78.74%) of them.

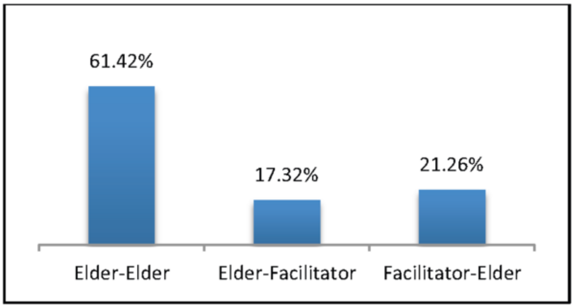


Figure 3. Frequency of interactions among actors.

Furthermore, from the 78 Elder-Elder interactions (see Fig. 4), 28 (35.90%) were for Feedback provision, 25 (32.05%) were for Encouragement between elders, 11 (14.10%) were for Turn coordination, 8 (10.26%) were for Social communication, and 6 (7.69%) were for Competition among elders.

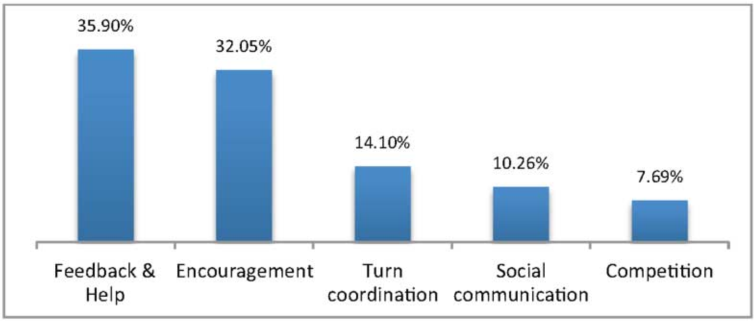


Figure 4. Frequency of functions in Elder-Elder interactions.

In the case of the 22 Elder-Facilitator interactions (see Fig. 5), 13 (59.09%) were to Confirm an action, while the remaining 9 (40.91%) were to Request help to the facilitator. In the case of the 27 Facilitator-Elder interactions, 16 (59.26%) were to Provide feedback, while the remaining 11 (40.74%) were to Give them instructions.

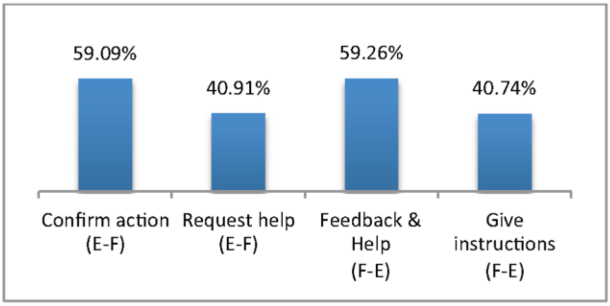


Figure 5. Frequency of functions in Elder-Facilitator (E-F) and Facilitator-Elder (F-E) interactions.

5 Discussion

The type and frequency of interactions observed during the session allowed us to assess the UX of participants during application's use, i.e. determine whether they experienced engagement, satisfaction, and enjoyment while playing the game, while lowering anxiety and avoiding frustration.

As shown in the previous section, elders initiated most of the observed interactions (78.74%). This is not surprising, as there were 4 elders and only one facilitator in each session. This provides evidence that elders were engaged in the activity, as the application required them to be active and involved in the activities.

Furthermore, on the one hand, regarding Elder-Elder interaction, the fact that 61.42% of all interactions were performed among elders only provides evidence that the proposed application reduces the elders' anxiety and frustration, and allowed them to play the game by themselves most of the time. This requires elders to understand the activities proposed in the application, and its interaction modalities, as to play the game without the intervention of the facilitator. To achieve this, elders had to provide Feedback and to help others as to allow them to complete their tasks – as evidenced by the 35.90% of these interactions.

In addition, the activities of the proposed application fostered encouragement (32.05%) and competition (7.69%) interactions among participants, generating elders' satisfaction. These required elders to engage in the game, which in turn required them to focus on the activities and on the user experience they provide, rather than on the means used to provide them. This in turn, provides evidence on the adequacy of the interaction modalities and user interfaces of the application to suit the needs of most participants.

Additionally, as a result of this participants' engagement in the activities and social interactions, 14.10% of all interactions were devoted to indicate players when it was their turn to play (Turn coordination).

Also, the proposed application allowed elders to have fun during the game, and allowed for social communication (10.26%) interactions. Participants were at easy while playing the game; they even started to tell jokes and anecdotes about them and the other participants – laughing was a common behavior during the game. As for the previous points, this required participants to be engaged in the game, and required the proposed application to be used in a transparent manner.

6 Conclusions

In this work we have considered a prevention approach through CS as entertainment activities. Using as a starting point a framework for the design of usable CS applications, and a game that promotes the CS of various cognitive functions, we assess the UX of this application through an analysis of the interactions that occurred during a collocated CS session. These results provide evidence that the tool favors interaction among the elderly, and fosters feedback, motivation and social interaction among them. These kind of interactions and their frequency provide evidence about the UX achieved by participants during a CS session, and allows users to have fun, be satisfied about its use and engaged in the activity, while lowering anxiety and avoiding frustration. These results also highlight the role of the caregiver as feedback and instructions provider to support the proper execution of CS activities. Finally, these findings provide further support to the promissory results of the usability evaluation of the Abueparty application [6].

As future work, we will analyze additional videos of CS activity in this modality (125 minutes more), and in other modalities that we have explored (e.g., CS sessions with support from a remote family member), to identify, quantify and contrast interactions in those CS scenarios. This will allow us to continue evaluating the UX aspects of the framework, refining it, and exploring its use in the development of additional usable CS applications.

Acknowledgments

We thank the participation of Sociedad Gerontológica de Baja California. We also thank Ivan Carrillo, Melissa Ibarra, Casandra Chévez and Luis Miguel Jiménez, who developed Abueparty and assisted in this UX evaluation. This work was funded by Universidad Autónoma de Baja California under grants 0216 XV Convocatoria Interna de Proyectos de Investigación and 0212 Proyectos de Servicio Social 2011.

References

[1] E. Aarts (2004). "Ambient intelligence: a multimedia perspective." IEEE Trans. Multimedia 11(1): 12-19.
[2] N. Scarmeas and Y. Stern (2004). "Cognitive reserve: Implications for diagnosis and prevention of Alzheimer's disease." Current Neurology and Neuroscience Reports 4: 374-380.
[3] L. M. Stevens et al.(2001). "Alzheimer Disease." JAMA 286(17): 2194.
[4] R. S. Wilson, C. F. Mendes de Leon, et al. (2002). "Participation in Cognitively Stimulating Activities and Risk of Incident Alzheimer Disease." JAMA: The Journal of the American Medical Association 287(6): 742-748.
[5] K. Ball, D. B. Berch, et al. (2002). "Effects of Cognitive Training Interventions With Older Adults." JAMA: The Journal of the American Medical Association 288(18): 2271-2281.
[6] V. Meza-Kubo and A. L. Morán (2012). "UCSA: A Design Framework for Usable Cognitive Systems for the Worried-well". To appear in Personal and Ubiquitous Computing. DOI 10.1007/s00779-012-0554-x
[7] V. Meza-Kubo and A. L. Morán (2011) AbueParty: An Everyday Entertainment System for the Cognitive Wellness of the Worried-Well. In: Proc. UCAmI 2011.
[8] ISO DIS 9241-210:2010. Ergonomics of human system interaction -Part 210: Human-centred design for interactive systems (formerly known as 13407). International Standardization Organization (ISO). Switzerland
[9] K. Väänänen-Vainio-Mattila, V. Roto, et al. Now let's do it in practice: user experience evaluation methods in product development. CHI '08 extended abstracts on Human factors in computing systems, ACM.
[10] L. Arhippainen.and M. Tähti (2003). Empirical Evaluation of User Experience in two Adaptive Mobile Application Prototypes. Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia, Norrköping, Sweden.
[11] A. P. O. S. Vermeeren, E. L. Law et al. (2010). User experience evaluation methods: current state and development needs Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, ACM, 2010, 521-530
[12] M. Isomursu, M. Tähti et al. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications International Journal of Human-Computer Studies, 2007, 65, 404 – 418
[13] R. L. Mandryk, K. M. Inkpen and T. W. Calvert (2006). Using psychophysiological techniques to measure user experience with entertainment technologies Behaviour & Information Technology, 2006, 25, 141-158
[14] P. M. Sanderson and C. Fisher, C. Exploratory Sequential Data Analysis: Foundations Human–Computer Interaction, 1994, 9, 251-317

This page intentionally left blank

7th International Workshop on Artificial Intelligence Techniques for Ambient Intelligence (AITAmI'12)

Asier Aztiria (Mondragon Unibertsitatea, Spain)

Juan C. Augusto (University of Ulster, UK)

Diane J. Cook (Washington State University, USA)

This page intentionally left blank

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

Introduction to the Proceedings of AITAmI'12

Imagine a future where human environments respond to human preferences and needs. In this world, devices equipped with simple intelligence and the abilities to sense, communicate, and act will be unremarkable features of our world. In sum, we will come to view simple software intelligence as an ambient feature of our environment. This workshop provides the opportunity to understand latest developments and take action to shape the future of the area by gathering researchers in a variety of AI subfields together with representatives of commercial interests to explore the technology and applications for ambient intelligence.

Based on this vision, the workshop on Artificial Intelligence Techniques for Ambient Intelligence (AITAmI) was launched in 2006 co-located with the European Conference on Artificial Intelligence. Since 2009, this workshop is colocated with the International Conference on Intelligent Environments (IE). The previous editions of the workshop indicated a significant interest in research related to use of Artificial Intelligence techniques for Ambient Intelligence environments.

In this edition the focus is on context and situation understanding. This is a key concept which supports accurate diagnosis and reasoning for the decision making processes that sustain Ambient Intelligence applications. Papers presented in this workshop deal with topics such as object recognition for context awareness, activity recognizers, intelligent navigation robots, adaptation of agents' knowledge, reasoning abilities of context aware systems.

The workshop organizers deeply thanks the Program Committee and attendees for their participation and interesting discussions.

Asier Aztiria, Juan C. Augusto, and Diane J. Cook
Co-chairs AITAmI'12

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-82

Contextual Reasoning in Context-Aware Systems

Hedda R. SCHMIDTKE [1]
Carnegie Mellon University, Rwanda

Abstract. Adding reasoning abilities to context-aware systems has been a focus of research in pervasive computing for several years and a broad range of approaches has been suggested. In particular, the well-known trade-off between expressivity and inferential power has been discussed as a major concern, as dimensions of context include well-known hard domains, such as spatial, temporal, and causal knowledge. In practice however, experiments report acceptable run-times and complexity of the used reasoning mechanism does not seem to be an issue at all. Two questions are addressed in this paper: why this is the case and whether these positive results will scale up as context-aware systems are leaving their experimental environments, are extended and modified by application developers, and employed in everyday life by millions of end-users.

The paper presents an analysis of results from pervasive computing, qualitative spatial and temporal reasoning, and logic-based contextual reasoning. The goal was to carve out the theoretical core of fast contextual reasoning reported from experimental context-aware systems, and to discuss how these good properties can be made to scale up. The findings suggest that partial order reasoning is the core of tractable contextual reasoning. Examples illustrate the surprisingly high expressiveness and inferential power, and serve to emphasize the interrelations between tractable reasoning in pervasive computing, qualitative spatial and temporal reasoning, and logic-based contextual reasoning.

Keywords. context, context-awareness, contextual reasoning, context model, granularity

Introduction

Adding reasoning abilities to context-aware systems has been a focus of research in pervasive computing for several years and a broad range of approaches has been suggested. In particular, the well-known trade-off between expressivity and inferential power has been a major concern. Context-aware systems are expected to produce a reaction within milliseconds of detecting a context, they run on light-weight systems such as mobile phones or sensor nodes, and dimensions of context include well-known hard domains, such as spatial, temporal, and causal knowledge. Recent theoretical studies [35,7] consequently suggest that we need to abandon the idea of using an integrated contextual reasoning framework. On the experimental side, however, issues with run-time have rarely been reported, even when using semi-decidable [30] or even undecidable [41] logics.

[1] Address: Hedda R. Schmidtke, Carnegie Mellon University in Rwanda, Telecom House, Boulevard de l'Umuganda, Kacyiru-Kigali, Rwanda; E-mail: schmidtke@cmu.edu.

Contextual reasoning, as a field within the research area of large-scale knowledge representation and reasoning systems [8,5], studies reasoning mechanisms enabled to reproduce the human capability to separate information that is relevant for a task from irrelevant information. Three types of contextual reasoning can be distinguished [5]: *partial reasoning* focusses on relevant parts, *approximate reasoning* leaves out unnecessary detail, *perspectival reasoning* can move the point of view on a problem.

A wealth of research exists concerning formal logic frameworks for reasoning about spatial and temporal structures, including natural language constructs, such as *here*, *now*, or *yesterday*, whose meaning depends on context [11,28,13,3,1]. Especially, the seminal research by Gabbay [12] should be highlighted here, forming a fundament on which non-classical formal logic and data structures for AI can be compared and made to work together.

Spatial and temporal reasoning has been a focus of research also in the area of constraint satisfaction methods. Qualitative spatial reasoning (QSR) is an active research area [9,10] studying qualitative spatial calculi and heuristics that make reasoning about spatial relations, such as the mereotopological region-connection-calculus (RCC) of [29] or the cardinal direction calculus of [25], tractable. Newer results in this area particularly important for this paper shed light on the complexity for combining calculi [34,17,33,32,27].

The paper aims to carve out the theoretical core of a class of particularly fast contextual reasoning mechanisms and to hint at possible approaches for retaining these properties as systems scale up. The rest of the paper is organized as follows. After a brief introduction of contextual reasoning approaches in context-aware computing, the theoretical framework of partial orders and lattice structures underlying hierarchical contextual reasoning mechanisms is discussed. The versatility of this approach is explored in a section on applications. Finally, the possibilities to realize partial reasoning and reasoning with a limited level of detail are discussed. We conclude that the hierarchical contextual reasoning frameworks proposed in the literature will scale well when context-aware systems enter everyday life.

1. Reasoning in Context-Aware Systems

From the earliest location-aware systems [36] to recent pervasive computing systems [19], automatic reasoning capabilities of the system have been a core component. The key idea of being able to react to a context is an essential component of intelligence. Representations that are *cued* by a perceptual trigger have been argued to have come early in the evolution of intelligence, and seem to require less computational power than *detached* representations that can be manipulated independently from any external stimulus [16]. Similarly, context-aware applications can be designed efficiently for the purpose of immediately triggering or adapting behavior in response to sensory input [19]. The objective of adequate representations for such immediate reactive procedures has been pursued in the field of context modeling.

Among the first and most efficient context-modeling approaches were highly efficient tree-structures [36] and directed acyclic graphs (DAG) [23,43] for representing hierarchies of contexts, such as containment hierarchies of locations or class hierarchies of types of situations, possibly supplemented by additional numerical or coordinate information [23,22,4], which can be processed using numerical constraint processing or value comparison.

On the other end of the spectrum between fast inference and high expressiveness lie logic-based approaches that use standard ontology languages [35,42] or rule-based formalisms [26,31]. The knowledge necessary for operating context-aware systems can be distinguished from that for other computing systems as consisting of particularly many statements about individuals for describing the environment, e.g. locating a certain room in a certain building, a certain sensor in a certain room, etc. Class hierarchies for contextual domains such as space or time, are rather flat, comprising e.g. regions and points as possible classes, however, for handling applications and devices the sub-class relation is relevant.

For domains like space and time, relations, such as the well-known 13 temporal interval relations of [2] or the eight mereotopological relations between regions from [29] are highly relevant and modeled in many existing systems. The use of the language SROIQ for OWL 2 [18], featuring means for specifying relations as transitive, reflexive, irreflexive, symmetric, asymmetric, or mutually disjoint, is therefore an important step for context-aware computing [35]. A recent survey [7], however, emphasizes the growing concern that the limited reasoning performance of expressive, integrated reasoning mechanisms could lead to a lack of scalability. The question then arises whether the more light-weight hierarchy-based approaches face similar concerns, and whether their expressiveness is sufficient to enable interesting applications beyond simple location hierarchies.

In line with these observations, we identified in [39] a system of six partial ordering relations that describe context hierarchies in the hierarchical context model of [21]. In [38], we studied a system of 54 compound spatial relations including mereological relations, cardinal direction relations (such as to the North of, to the North-West of, etc.), and size comparisons (smaller than, same size), and found similarly good reasoning properties in a framework of constraint satisfaction. This paper further explains these findings and illustrates why complexity does not increase, even if arbitrary additional partial ordering relations are allowed.

2. Theoretical Foundations

We briefly outline the basic framework of pre-orders, partial orders, and lattice structures that underlies hierarchical reasoning in a first order language and demonstrate the theoretical foundations of its versatility. Here and elsewhere, we omit brackets, in particular outer brackets, if no ambiguity can arise; however, we enclose all atomic formulae in square brackets to support visual separation between terms and formulae.

2.1. Pre-Orders, Partial Orders and Equivalence Relations

A pre-order is a relation $\sqsubseteq_m$ that is reflexive (A1) and transitive (A2).

$$\forall x : [x \sqsubseteq_m x] \tag{A1}$$

$$\forall x, y, z : [x \sqsubseteq_m y] \wedge [y \sqsubseteq_m z] \rightarrow [x \sqsubseteq_m z] \tag{A2}$$

A pre-order $\sqsubseteq$ (A3), (A4) that is antisymmetric (A5) is called partial order.

$$\forall x : [x \sqsubseteq x] \tag{A3}$$

$$\forall x,y,z : [x \sqsubseteq y] \wedge [y \sqsubseteq z] \rightarrow [x \sqsubseteq z] \tag{A4}$$

$$\forall x,y : [x \sqsubseteq y] \wedge [y \sqsubseteq x] \rightarrow [x = y] \tag{A5}$$

Every pre-order $\sqsubseteq_m$ gives rise to an equivalence relation $=_m$, that is, a pre-order that is symmetric (1).

$$\forall x,y : [x =_m y] \overset{def}{\Leftrightarrow} [x \sqsubseteq_m y] \wedge [y \sqsubseteq_m x] \tag{D1}$$

$$\forall x,y : [x =_m y] \rightarrow [y =_m x] \tag{1}$$

$$\forall x : [x =_m x] \tag{2}$$

$$\forall x,y,z : [x =_m y] \wedge [y =_m z] \rightarrow [x =_m z] \tag{3}$$

$$\forall x,y : [x \sqsubseteq_m y] \wedge [y \sqsubseteq_m x] \rightarrow [x =_m y] \tag{4}$$

Moreover, a pre-order $\sqsubseteq_m$ behaves like a partial order on the equivalence classes of its corresponding equivalence relation$=_m$ (4).

An important property of pre-orders is that they have sub-relations that are also pre-orders. In particular, we can define a pre-order $\sqsubseteq_{m,e}$ from a pre-order $\sqsubseteq_m$ and a given element e as the relation that compares for two elements x and y elements x', which are in $\sqsubseteq_m$ both below e and below x, with y. More formally: $\sqsubseteq_{m,e}$ holds for x and y, iff for all x' that are below both x and e, x' is below y in $\sqsubseteq_m$.

$$\forall x,y : [x \sqsubseteq_{m,e} y] \overset{def}{\Leftrightarrow} \forall x' : [x' \sqsubseteq_m x] \wedge [x' \sqsubseteq_m e] \rightarrow [x' \sqsubseteq_m y] \tag{D2}$$

$$\forall x : [x \sqsubseteq_{m,e} x] \tag{5}$$

$$\forall x,y,z : [x \sqsubseteq_{m,e} y] \wedge [y \sqsubseteq_{m,e} z] \rightarrow [x \sqsubseteq_{m,e} z] \tag{6}$$

To show that $\sqsubseteq_{m,e}$ is a pre-order, we need to show that it is reflexive and transitive. It is clear that $\sqsubseteq_{m,e}$ is reflexive (5), since $\forall x' : [x' \sqsubseteq_m x] \wedge [x' \sqsubseteq_m e] \rightarrow [x' \sqsubseteq_m x]$ is a tautology. For transitivity, assume $[x \sqsubseteq_{m,e} y]$ and $[y \sqsubseteq_{m,e} z]$ hold, and consider any arbitrary x' that is below x and e in $\sqsubseteq_m$. From the assumption $[x \sqsubseteq_{m,e} y]$, we know that x' must be below y in $\sqsubseteq_m$, by transitivity of $\sqsubseteq_m$ and (D2). Any such x' is an element that is both below y and below e. By the assumption $[y \sqsubseteq_{m,e} z]$, we know any such x' must be below z, again by transitivity of $\sqsubseteq_m$ and (D2), and therefore $[x \sqsubseteq_{m,e} z]$.

Since $\sqsubseteq_{m,e}$ is a pre-order, an equivalence relation $\equiv_{m,e}$ can be defined following (D1).

$$\forall x,y : [x \sqsubseteq_m y] \rightarrow [x \sqsubseteq_{m,e} y] \tag{7}$$

We can show that $\sqsubseteq_{m,e}$ extends $\sqsubseteq_m$, that is, it behaves like $\sqsubseteq_m$ for all pairs of elements in $\sqsubseteq_m$ (7). Assume $[x \sqsubseteq_m y]$ holds. With respect to (D2), we only need to check elements x' below x in $\sqsubseteq_m$ that are below e – if there are no such x', $[x \sqsubseteq_{m,e} y]$ holds trivially. For any given such x', it holds by transitivity of $\sqsubseteq_m$ that $[x' \sqsubseteq_m y]$, and thus $[x \sqsubseteq_{m,e} y]$.

It thus follows, that a single pre-order (or partial order relation) $\sqsubseteq$ suffices to give rise to a range of pre-orders as extensions. In particular, this entails that a framework for reasoning with context hierarchies need not encode the dimensions of context into the given logical language. Rather system developers can define required relations as needed using the construction (D2). The widely used hierarchical context modeling approach is more versatile than assumed.

2.2. Lattice Structures

If a partial order $\sqsubseteq$ has a unique *least upper bound* (A6) and *greatest lower bound* (A7) for any two elements x and y, it can be extended into a lattice structure.

$$\forall x, y : \exists z : [x \sqsubseteq z] \wedge [y \sqsubseteq z] \wedge \forall z' : [x \sqsubseteq z'] \wedge [y \sqsubseteq z'] \rightarrow [z \sqsubseteq z'] \tag{A6}$$

$$\forall x, y : \exists z : [z \sqsubseteq x] \wedge [z \sqsubseteq y] \wedge \forall z' : [z' \sqsubseteq x] \wedge [z' \sqsubseteq y] \rightarrow [z' \sqsubseteq z] \tag{A7}$$

A lattice structure for $\sqsubseteq$ comprises two binary operations $\sqcup$ (*join*), yielding the least upper bound (D3) of two elements, and $\sqcap$ (*meet*), yielding the greatest lower bound (D4). Furthermore, we characterize two elements $\top$ (top) and $\bot$ (bottom) as the upper bound and lower bound of $\sqsubseteq$ to obtain a bounded lattice (A8).

$$\forall x, y, z : [z = x \sqcup y] \overset{def}{\Leftrightarrow} [x \sqsubseteq z] \wedge [y \sqsubseteq z] \wedge \forall z' : [x \sqsubseteq z'] \wedge [y \sqsubseteq z'] \rightarrow [z \sqsubseteq z'] \tag{D3}$$

$$\forall x, y, z : [z = x \sqcap y] \overset{def}{\Leftrightarrow} [z \sqsubseteq x] \wedge [z \sqsubseteq y] \wedge \forall z' : [z' \sqsubseteq x] \wedge [z' \sqsubseteq y] \rightarrow [z' \sqsubseteq z] \tag{D4}$$

$$\forall x : [x \sqsubseteq \top] \wedge [\bot \sqsubseteq x] \tag{A8}$$

With the operations $\sqcup$ and $\sqcap$ we can define further notions. The relation $\bigcirc$ (overlap) holds between two elements x and y iff the meet of x and y is not the bottom element. Two elements underlap ($\cup$) iff the join of x and y is not the top element.

$$\forall x, y : [x \bigcirc y] \overset{def}{\Leftrightarrow} \neg[x \sqcap y \sqsubseteq \bot] \tag{D5}$$

$$\forall x, y : [x \cup y] \overset{def}{\Leftrightarrow} \neg[\top \sqsubseteq x \sqcup y] \tag{D6}$$

2.3. Linearizations

We call a pre-order $\leqslant_m$ *linear extension* or *linearization* of a pre-order $\sqsubseteq_m$, if and only if $\leqslant_m$ extends $\sqsubseteq_m$ (A11) and is linear (A12):

$$\forall x : [x \leqslant_m x] \tag{A9}$$

$$\forall x, y, z : [x \leqslant_m y] \wedge [y \leqslant_m z] \rightarrow [x \leqslant_m z] \tag{A10}$$

$$\forall x, y : [x \sqsubseteq_m y] \rightarrow [x \leqslant_m y] \tag{A11}$$

$$\forall x, y : [x \leqslant_m y] \vee [y \leqslant_m x] \tag{A12}$$

$$\forall x, y : [x \equiv_m y] \overset{def}{\Leftrightarrow} [x \leqslant_m y] \wedge [y \leqslant_m x] \tag{D7}$$

We again obtain a relation $\equiv_m$ (D7), as equivalence relation for $\leqslant_m$.

Table 1. Pre-orders and their linear extensions in QSR

relation system	relation	symbol	context	type
spatial mereology	part-of	$\sqsubseteq_{m_p}$	m_p	pre-order
	smaller-than	$\leqslant_{m_l}$	m_l	linear extension of part-of
cardinal directions	directly-to-the-North-of	$\sqsubseteq_{n_p}$	n_p	pre-order
	more-to-the-North-of	$\leqslant_{n_l}$	n_l	linear extension of directly-to-the-North-of
	directly-to-the-East-of	$\sqsubseteq_{e_p}$	e_p	pre-order (locally[2])
	more-to-the-East-of	$\leqslant_{e_l}$	e_l	linear extension of directly-to-the-East-of
temporal intervals	during-or-equal	$\sqsubseteq_{i_p}$	i_p	pre-order
	shorter-duration-than	$\leqslant_{i_l}$	i_l	linear extension of during-or-equal
temporal order	before-or-same-time (branching time)	$\sqsubseteq_{t_p}$	t_p	pre-order
	before-or-same-time (linear time)	$\leqslant_{t_l}$	t_l	linear extension of branching before-or-same-time
classes	sub-class	$\sqsubseteq_{c_p}$	c_p	pre-order
	smaller-cardinality	$\leqslant_{c_l}$	c_l	linear extension of sub-class

Many relations that are interesting for reasoning are pre-orders, partial orders, strict partial orders, or linearizations. Table 1 lists a range of relations that have been discussed in the area of qualitative reasoning [2,14,6,25,24,29,9,15] and that can be expressed in terms of pre-orders and their linear extensions.

3. Applications

We can now illustrate how the above framework can be used for modeling context hierarchies. We show examples how specific pre-orders $\sqsubseteq_m$ can be derived as extensions of a partial order $\sqsubseteq$ and how linear extensions $\leqslant_m$ can be defined from pre-orders $\sqsubseteq_m$.

3.1. Modeling Context Hierarchies

We can introduce arbitrary sets of partial ordering relations, such as the six relations of [39], or the detailed modeling of the spatial domain discussed in [38], by deriving extensions $\sqsubseteq_m$ from $\sqsubseteq$ using elements m representing the domains. Here and in the following, we use Greek letters ξ, χ, κ to indicate schema variables ranging over contexts.

$$[\xi \sqsubseteq_\kappa \chi] \overset{def}{\Leftrightarrow} [\xi \sqcap \kappa \sqsubseteq \chi] \tag{D8}$$

$$[\xi \bigcirc_\kappa \chi] \overset{def}{\Leftrightarrow} \neg[\xi \sqcap \chi \sqsubseteq_\kappa \bot] \tag{D9}$$

$$[\xi =_\kappa \chi] \overset{def}{\Leftrightarrow} [\xi \sqsubseteq_\kappa \chi] \wedge [\chi \sqsubseteq_\kappa \xi] \tag{D10}$$

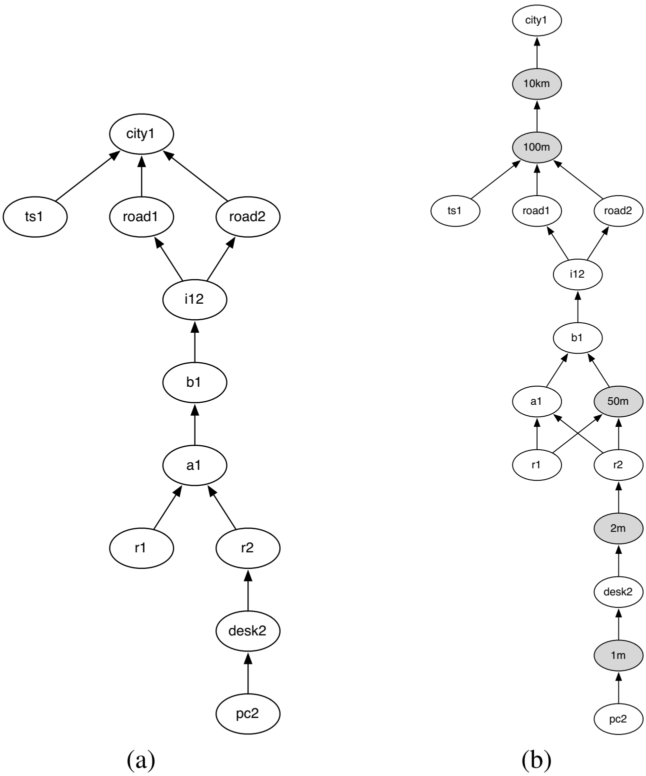


Figure 1. A location hierarchy (a) and size hierarchy (b) specifying the locations (a: relation part-of) and respective sizes (b: relation smaller-than) of a PC (pc2) on a desk (desk2) in a room (r2). In turn, r2 and room r1 are in an office space (a1) in a building (b1) on the intersection i12 of roads road1 and road2. The roads road1 and road2 and the train station ts1 are in a city city1. Arrows indicate the relation described by a spatial containment relation $\sqsubseteq_{m_p}$, not the relation $\sqsubseteq$ itself. The hierarchy of the sizes of locations $\leqslant_{m_p}$ in (b) was obtained by extending the location hierarchy of (a) with additional information about sizes.

Any relation $\sqsubseteq_\kappa$ thus gives rise to relations $\bigcirc_\kappa$ (D9) and $=_\kappa$ (D10).

In general, we retain as a consequence of (D8), that for arbitrary x and y: if x is an m-sub-context of y and m' is a sub-context of m, then x is also an m'-sub-context of y.

$$\forall x, y, m, m' : [x \sqsubseteq_m y] \wedge [m' \sqsubseteq m] \rightarrow [x \sqsubseteq_{m'} y] \quad (8)$$

Hierarchical context models store knowledge given in the form $[a \sqsubseteq_m b]$ in hierarchical data structures that facilitate transitive inference. An example is shown in Figure 1a.

3.2. Modeling Linearizations in Context Hierarchies

Using sub-relations as in (8), we can also define linear relations $\leqslant_m$ that extend partial order relations $\sqsubseteq_m$. However, the linearity constraint (A12) demands that a complete linearization exists, which would in practice be hard to enforce. Moreover, in order to infer something from the linearity constraint a disjunction has to be checked. In the worst case, a complete linearization would have to be generated, with every pair of contexts, that is n^2 disjunctions, to be checked. This would lead to an explosion of computational effort that might only rarely be useful.

Similarly as in [38], we will therefore neglect the disjunctive restriction with respect to reasoning. The argument for this is practical rather than formal: linearizations, such as the size constraints, facilitate reasoning. Following this argument, we can assume that a benefit encourages the reasoning agent to collect this information so that linearity of the $\leqslant_m$ orders is approximated but not required.

We define the following schema exemplarily for the pair of the mereological pre-order $\sqsubseteq_{m_p}$ (part-of) and a linear extension $\leqslant_{m_l}$ (smaller-than) indicated by contexts m_p and m_l so that $\leqslant_{m_l}$ is an extension of $\sqsubseteq_{m_p}$ (D13):

$$[\xi \sqsubseteq_{m_p} \chi] \overset{def}{\Leftrightarrow} [\xi \sqcap m_p \sqsubseteq \chi] \tag{D11}$$

$$[\xi \leqslant_{m_l} \chi] \overset{def}{\Leftrightarrow} [\xi \sqcap m_l \sqsubseteq \chi] \tag{D12}$$

$$[m_p \sqsubseteq m_l] \tag{D13}$$

$$[\xi \equiv_{m_l} \chi] \overset{def}{\Leftrightarrow} [\xi \leqslant_{m_l} \chi] \wedge [\chi \leqslant_{m_l} \xi] \tag{D14}$$

Equivalence with respect to $\leqslant_{m_l}$ is expressed with $\equiv_{m_l}$ (D14). Like knowledge given in the form $[a \sqsubseteq_m b]$, knowledge of the type $[a \leqslant_m b]$ can be stored in hierarchical data structures that facilitate transitive inference. An example is shown in Fig. 1b. The hierarchy for $[a \leqslant_m b]$ can be obtained from that of $[a \sqsubseteq_m b]$ by adding size information. In actual applications, linear ordering can also be encoded numerically.

Large relation systems, such as the 54 combined spatial relations introduced in [38], can then be defined. There, the three spatial pre-orders listed in Table 1, each with linearizations and equivalence relations are used to represent, and reason about, spatial layouts. Relations, such as north-western part ($x\, M_P N_L E_L\, y$), can be defined as combinations of relations. We obtain the inferences from [38] as valid schemata, for instance (9).

$$[\xi\ M_P N_L E_L\ \chi] \overset{def}{\Leftrightarrow} [\xi \sqsubseteq_{m_p} \chi] \wedge [\xi \leqslant_{n_l} \chi] \wedge [\xi \leqslant_{e_l} \chi] \tag{D15}$$

$$[\xi\ M_P N_Q E_L\ \chi] \overset{def}{\Leftrightarrow} [\xi \sqsubseteq_{m_p} \chi] \wedge [\xi \equiv_{n_l} \chi] \wedge [\xi \leqslant_{e_l} \chi] \tag{D16}$$

$$[a\ M_P N_L E_L\ b] \wedge [b\ M_P N_Q E_L\ c] \rightarrow [a\ M_P N_L E_L\ c] \tag{9}$$

Class hierarchies are an important tool for inferring and specifying properties of types of contexts in a generic way. Since the subclass relation is a partial order over the domain of classes, it can be defined as a pre-order $\sqsubseteq_{c_p}$. This allows interesting constructions: for instance, we can define that a context is a class of locations iff it is only in the spatial and sub-class domain and actually overlaps the spatial domain, i.e. has a non-empty spatial extent (D17). If we want to ensure that locations are always given the corresponding type, we can define a relation *subLoc* as an external interface for the user, hiding the spatial mereological $\sqsubseteq_{m_p}$ and sub-class $\sqsubseteq_{c_p}$ as internal operators.

$$[\xi \sqsubseteq_{c_p} location] \overset{def}{\Leftrightarrow} [\xi \sqsubseteq m_p \sqcup c_p] \wedge [\xi \bigcirc m_p] \tag{D17}$$

$$[\xi\ subLoc\ \chi] \overset{def}{\Leftrightarrow} [\xi \sqsubseteq_{m_p} \chi] \wedge [\xi \sqsubseteq_{c_p} location] \tag{D18}$$

Using this, a developer can enter that road intersections are sub-locations of roads (D19). Applying (D18), this would be expanded to mean: the sub-class relation holds between *roadIntersection* and *location* and that the spatial containment relation holds between between *roadIntersection* and *road* (10).

$$[\mathit{roadIntersection\ subLoc\ road}] \tag{D19}$$

$$[\mathit{roadIntersection} \sqsubseteq_{m_p} \mathit{road}] \wedge [\mathit{roadIntersection} \sqsubseteq_{c_p} \mathit{location}] \tag{10}$$

As contexts in this framework are portions of reality not individuals, we have defined classes of locations as classes and as locations.

The is-a-relation between an individual and a class is usually treated in a different manner than the sub-class relation. We can identify individuals as classes of cardinality 1 (D20), based on the comparison of cardinality as a linearization $\leqslant_{c_l}$ of the sub-class relation $\sqsubseteq_{c_p}$.

$$[\xi\ \mathit{instanceOf}\ \chi] \overset{def}{\Leftrightarrow} [\xi \sqsubseteq_{c_p} \chi] \wedge [\xi \equiv_{c_p} \mathit{card1}] \tag{D20}$$

$$[\mathit{roadIntersection12\ instanceOf\ roadIntersection}] \tag{D21}$$

With this definition we can specify that *roadIntersection12* is a sub-class of *roadIntersection* with cardinality 1 (D21).

4. Partial and Approximate Reasoning

Linearizations, ordering the whole domain, are a suitable tool for stratifying a domain into levels of granularity and thus for enforcing that reasoning does not go beyond a given level of detail for a certain task. This holds true even if the total ordering of a domain cannot be enforced (cf. Figure 1b). Size information, even if incomplete, can be used to delimit reasoning and allow for approximate representation.

Size-based granularity is a representational tool that can be used to handle differences in sizes and to model uncertainty [37]. The key idea is to formally specify what is a small, and therefore irrelevant, detail in a context [40]. Size-based granularity is built upon the concept of a grain-size for a context region: parts of an object that are smaller than the grain-size can be disregarded as unimportant details in the context. This concept forms the basis for level of detail reasoning or *approximate reasoning* in the terminology of [5]. Similarly, parts of an object outside the context region are irrelevant. This concept forms the basis for realizing *partial reasoning*[5] in context-aware systems.

We can realize both types of restricted reasoning in the framework. Instead of asking whether $\mathit{CKB} \cup \{\neg\phi(\mathit{ans})\}$ is inconsistent for a contextual knowledge base *CKB* and query $\phi(\mathit{ans})$, we can instead ask whether

$$\mathit{CKB} \cup \{[g_m \leqslant_{ml} \mathit{ans}], [\mathit{ans} \sqsubseteq_{m_p} c]\} \cup \{\neg\phi(\mathit{ans})\},$$

is inconsistent, that is: whether query $\phi(\mathit{ans})$ can be solved for *ans* larger than g for $\leqslant_{ml}$ and below context c for $\sqsubseteq_{m_p}$. In practice, this can be implemented directly into the reasoning mechanism. In a graph-based representation of the context hierarchy, for instance, by pruning the hierarchy below the point g_m and above the point c. The algorithm reasoning about $\phi(ans)$ is then limited to the resulting hierarchy between g_m and c. It follows that additions outside of this local hierarchy would not impede the reasoning process inside, and that growing hierarchical context models can retain their good runtime, a step towards realizing the idea of granularity as advocated by Hobbs [20], that is, towards a tool to make reasoning and representation tractable and scalable.

5. Outlook and Conclusion

Adding reasoning abilities to context-aware systems has been a focus of research in pervasive computing for several years and a broad range of approaches has been suggested. The well-known trade-off between expressivity and inferential power has raised considerable concerns in this area, as dimensions of context include well-known hard domains, such as spatial, temporal, and causal knowledge. We suggested that hierarchical context models, which have been used in numerous practical applications, can be described by pre-order relations and their linearizations. We showed that this framework, for which fast reasoning procedures exist, is expressive, and proposed a mechanism for limiting level of detail and relevant context.

We conjecture that the proposed means can ensure that reasoning with hierarchical context models will scale well. Future works need to include experimental validation especially regarding large-scale context-aware systems. Moreover, heuristics for calculating appropriate level detail and context of reasoning need to be determined.

References

[1] M. Aiello and J. van Benthem. A modal walk through space. *Journal of applied non classical logics*, 12(3-4), 2002.

[2] J. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23:123–154, 1984.

[3] C. Areces, P. Blackburn, and M. Marx. Hybrid logics: Characterization, interpolation and complexity. *J. Symb. Log.*, 66(3):977–1010, 2001.

[4] M. Beigl, T. Zimmer, and C. Decker. A location model for communicating and processing of context. *Personal and Ubiquitous Computing*, 6(5/6):341–357, 2002.

[5] M. Benerecetti, P. Bouquet, and C. Ghidini. Contextual reasoning distilled. *Journal of Experimental and Theoretical Artificial Intelligence*, 12(3):279–305, 2000.

[6] B. Bennett and A. Galton. A unifying semantics for time and events. *Artificial Intelligence*, 153:13–48, 2004.

[7] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A. Ranganathan, and D. Riboni. A survey of context modelling and reasoning techniques. *Pervasive and Mobile Computing*, 6(2):161–180, 2010.

[8] P. Brézillon. Context in problem solving: a survey. *The Knowledge Engineering Review*, 14(1):1–34, 1999.

[9] A. G. Cohn and S. M. Hazarika. Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae*, 46(1-2):1–29, 2001.

[10] A. G. Cohn and J. Renz. Qualitative spatial representation and reasoning. In F. van Hermelen, V. Lifschitz, and B. Porter, editors, *Handbook of Knowledge Representation*, pages 551–596. Elsevier, 2008.

[11] G. Forbes. Indexicals. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume IV, pages 463–490. D. Reidel, 1989.

[12] D. Gabbay. *Labelled Deductive Systems; Principles and Applications.* Oxford University Press, 1996.

[13] D. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Many-dimensional modal logics: theory and applications*. Elsevier, 2003.

[14] A. Galton. *Qualitative Spatial Change*. Oxford University Press, 2000.
[15] A. Galton and J. C. Augusto. Stratified causal theories for reasoning about deterministic devices and protocols. In *TIME*, pages 52–54, 2002.
[16] P. Gärdenfors. The detachment of thought. In C. Erneling and D. Johnson, editors, *The mind as a scientific subject: between brain and culture*, pages 323–341. Oxford University Press, 2005.
[17] A. Gerevini and J. Renz. Combining topological and size information for spatial reasoning. *Artif. Intell.*, 137(1-2):1–42, 2002.
[18] B. C. Grau, I. Horrocks, B. Motik, B. Parsia, P. F. Patel-Schneider, and U. Sattler. Owl 2: The next step for owl. *J. Web Sem.*, 6(4):309–322, 2008.
[19] K. Henricksen and J. Indulska. Developing context-aware pervasive computing applications: Models and approach. *Pervasive and Mobile Computing*, 2:37–64, 2006.
[20] J. Hobbs. Granularity. In A. K. Josh, editor, *Ninth International Joint Conference on Artificial Intelligence*, pages 432–435, Los Angeles, 1985. Morgan Kaufmann.
[21] S. Jang and W. Woo. ubi-UCAM: a unified context-aware application model. In P. Blackburn, C. Ghidini, R. M. Turner, and F. Giunchiglia, editors, *International Conference on Modeling and Using Context*, pages 178–189, 2003.
[22] C. Jiang and P. Steenkiste. A hybrid location model with a computable location identifier for ubiquitous computing. In G. Borriello and L. E. Holmquist, editors, *Ubiquitous Computing*, pages 246–263, Gothenburg, Sweden, 2002. Springer.
[23] U. Leonhardt. *Supporting Location Awareness in Open Distributed Systems*. PhD thesis, Imperial College, London, UK, 1998.
[24] G. Ligozat. Generalized intervals: A guided tour. In *Proceedings of the ECAI-98 Workshop on Spatial and Temporal Reasoning*, Brighton, UK, 1998.
[25] G. Ligozat. Reasoning about cardinal directions. *Journal of Visual Languages and Computing*, 9:23–44, 1998.
[26] S. W. Loke. Representing and reasoning with situations for context-aware pervasive computing: a logic programming perspective. *The Knowledge Engineering Review*, 19(3):213–233, 2005.
[27] B. Nebel and J. Renz. A fixed-parameter tractable algorithm for spatio-temporal calendar management. In *IJCAI*, pages 879–884, 2009.
[28] H. J. Ohlbach and D. M. Gabbay. Calendar logic. *Journal of Applied Non-Classical Logics*, 8(4), 1998.
[29] D. Randell, Z. Cui, and A. Cohn. A spatial logic based on region and connection. In *Knowledge Representation and Reasoning*, pages 165–176. Morgan Kaufmann, 1992.
[30] A. Ranganathan and R. Campbell. An infrastructure for context-awareness based on first-order logic. *Personal and Ubiquitous Computing*, 7(6):353–364, 2003.
[31] P. Reignier, O. Brdiczka, D. Vaufreydaz, J. L. Crowley, and J. Maisonnasse. Context-aware environments: from specification to implementation. *Expert Systems*, 24(5):305–320, 2007.
[32] J. Renz. Qualitative spatial and temporal reasoning: Efficient algorithms for everyone. In *Twentieth International Joint Conference on Artificial Intelligence*, pages 526–531, 2007.
[33] J. Renz and G. Ligozat. Weak composition for qualitative spatial and temporal reasoning. In *CP*, pages 534–548, 2005.
[34] J. Renz and B. Nebel. Efficient methods for qualitative spatial reasoning. *J. Artif. Intell. Res. (JAIR)*, 15:289–318, 2001.
[35] D. Riboni and C. Bettini. Owl 2 modeling and reasoning with complex human activities. *Pervasive and Mobile Computing*, 7(3):379–395, 2011.
[36] B. N. Schilit and M. M. Theimer. Disseminating active map information to mobile hosts. *IEEE Network*, 8(5):22–32, 1994.
[37] H. R. Schmidtke. Aggregations and constituents: Geometric specification of multi-granular objects. *Journal of Visual Languages and Computing*, 16(4):289–309, 2005.
[38] H. R. Schmidtke and M. Beigl. Distributed spatial reasoning for wireless sensor networks. In *CONTEXT*, pages 264–277, 2011.
[39] H. R. Schmidtke, D. Hong, and W. Woo. Reasoning about models of context: A context-oriented logical language for knowledge-based context-aware applications. *Revue d'Intelligence Artificielle*, 22(5):589–608, 2008.
[40] H. R. Schmidtke and W. Woo. A size-based qualitative approach to the representation of spatial granularity. In M. M. Veloso, editor, *Twentieth International Joint Conference on Artificial Intelligence*, pages 563–568, 2007.

[41] T. Strang, C. Linnhoff-Popien, and K. Frank. CoOL: A context ontology language to enable contextual interoperability. In J.-B. Stefani, I. M. Demeure, and D. Hagimont, editors, *Distributed Applications and Interoperable Systems*, pages 236–247. Springer, 2003.

[42] J. Ye, L. Coyle, S. Dobson, and P. Nixon. Ontology-based models in pervasive computing systems. *The Knowledge Engineering Review*, 22:315–347, 2007.

[43] J. Ye, L. Coyle, S. Dobson, and P. Nixon. Using situation lattices to model and reason about context. In A. Kofod-Petersen, J. Cassens, D. B. Leake, and S. Schulz, editors, *4th International Workshop on Modeling and Reasoning in Context (MRC)*, pages 1–12. Roskilde University, 2007.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-94

Real Time Activity Recognition Using a Cell Phone's Accelerometer and Wi-Fi

Enrique A. GARCIA[1] and Ramón BRENA[2]
ITESM Campus Monterrey

Abstract. In this paper, we present an implementation of a real time activity recognizer running on a cellphone. First, simple activity recognition from accelerometer data is performed and then, this information is fused with data from Wi-Fi Access Points to classify the activity being performed by the user. The training set consisted of 8 activities performed in an academic environment and the classification accuracy was 89.7% using a supervised learning approach.

Keywords. activity recognition, accelerometer, Wi-Fi, cell phone.

Introduction

Human activity recognition is an important task for ambient intelligence systems [1]. Being able to recognize the state of a person can provide us with valuable information that can be used as input for other systems. For example, in healthcare, fall detection can be used to alert the medical staff in case of an accident; in security, abnormal behavior can be detected and thus used to prevent a burglary or other criminal activities.

In recent years simple human activity recognition has been achieved successfully, however complex activity recognition is still challenging and is an active area of research. In [2] they pose the following challenges regarding the nature of human activities: *Recognizing concurrent activities, recognizing interleaved activities, ambiguity of interpretation* and *multiple residents*.

In this work we present an implementation of a real time activity recognizer running on a cellphone. First, the *simple activity* is recognized from the cellphone's tri-axial accelerometer and then this information is fused with data gathered from Wi-Fi Access Points to allow a better understanding of the user's context.

The number of Wi-Fi Access Points around the world has increased significantly in the last years. They are installed in many places such as restaurants, hotels, schools, parks, airports, etc. Since every Access Point has a unique identifier namely, the BSSID (Basic Service Set Identifier), it is possible to use this information along with the signal strength for localization and tracking purposes [3,4,5].

The objective of this work is not to infer the spatial location of the user but to determine the user's context by means of a supervised learning approach. For example, if *walking* is detected as a simple activity and the Wi-Fi yields information that tells us that the user is in a library, then we can classify the whole activity as *looking for a book*

[1] E-mail: A00927248@itesm.mx
[2] E-mail: ramon.brena@itesm.mx

and another application can use this information maybe to suggest related titles. If the simple activity is *resting* the whole activity may be classified as *reading a book*. In this case the user's cell phone could block unimportant incoming calls to avoid interruptions. The role of the Access Points is to aid in the discrimination process providing approximate location information, i.e., we can use a fixed set of *primitive activities* as the basis and combine them with 'location' information to generate *contextualized activities.*

This paper is organized as follows. Section 1 presents several recent works in simple and complex activity recognition. Section 2 describes the methodology which includes: data collection, training, and classification. In section 3 we describe the details of the experiment and its results. Section 4 presents different approaches for human activity recognition. Finally, in section 5 conclusions and future work are presented.

1. Related Work

In this section we present a survey of works that tackle the problem of recognizing simple activities (*walking*, *running*, *sitting down*, *falling*, etc.) and complex activities (*making coffee*, *cleaning the house*, *making a drink*, *having dinner*, etc.).

1.1. Simple Activities Recognition

Generally, simple activities do not depend on the context, i.e., they can exist by themselves and they last only a few seconds. Examples of this type of activities are: *running*, *walking*, *resting*, *sitting*, etc.

Brezmes, Gorricho and Cotrina [6] implemented a real time activity recognizer on a mobile phone. They achieved accuracies ranging from 70% to 90% for several activities. Mannini and Sabatini [7] used five bi-axial accelerometers located at the hip, wrist, arm, ankle, and thigh and they reported accuracies between 93% and 98.5% for seven different activities (*sitting*, *lying*, *standing*, *walking, stair climbing*, *running* and *cycling*). Karantonis, *et al*. [8] presented an implementation of a real time activity classifier capable of computing the metabolic energy expenditure. Ravi, *et al*. [9] made a comparison of base-level classifiers and meta-level classifiers and concluded that combining classifiers using Plurality Voting turned out to be the best choice for the recognition of simple activities. Mi Zhang [10] proposed a Bag-of-Features approach which builds activity models using histograms of primitive symbols. Recently, the Bag-of-Features approach has gained significant interest.

1.2. Complex Activities Recognition

Complex activities are composed of a collection of simple activities and may consider information from the context, time, and interactions between other persons and objects. The recognition of these activities generally requires more sensors and a fixed infrastructure (video cameras, RFID tags, several accelerometers, magnetic sensors, etc.).

Tao Gu, *et al.* [11] built activity models by mining a set of Emerging Patterns from a sequential activity trace and used them to recognize sequential, interleaved, and concurrent activities achieving accuracies of 90.96%, 87.98% and 78.58%, respectively. Tam Huynh, *et al.* [12] used topic models to recognize activities such as: *dinner*, *commuting*, *lunch* and *office work*. They automatically extract activity patterns from sensor data (3D accelerometer, clock, binary tilt switches, temperature sensor, and two light sensors) to enable the recognition of daily routines as a composition of such activity patterns. Experimental results obtained by Tam Huynh, *et al.* [13] suggest that the recognition of complex activities can be achieved with the same algorithms of simple activities. The complex activities they recognized were *preparing for work*, *going shopping* and *doing housework*. Tian, *et al.* [14] use accelerometer and GPS information to automatically send updates to a micro-blogging website. They used Hidden Markov Models for the activity recognition and increased the accuracy by constraining the context using GPS location data.

The approach we use lies between simple and complex activities in the sense that we first recognize the simple activity and then we add Wi-Fi information to contextualize it. This work differs from the previously mentioned in the following aspects. First, we take advantage of existing infrastructure (Wi-Fi Access Points) so our approach does not require the addition of sensors to the environment like RFID tags, video cameras, etc. Second, aside from the presence of in range Access Points, we do not need a fixed configuration of the environment. Since we are just reading the BSSID and signal strength to classify the activities, we do not need to configure each Access Point neither know their location. Finally, we focused in using sensors that are commonly available in most smartphones so the user is freed from having to wear several sensors attached to his/her body.

2. Method

In this section we describe the process of data collection, training, and the activity classification phase. From now on, by *primitive* or *simple activity* we mean the type of activities described in section 1.1 such as *running*, *walking*, *resting*, *sitting,* etc. and by *contextualized activity* we mean an activity that is composed by a *simple activity* plus the information gathered from Wi-Fi Access Points.

2.1. Data Collection

An Android 2.2 [15] application running on a LG Optimus Me cell phone was used to collect the data from the accelerometer and Wi-Fi.

The cellphone was placed in the user's belt and the data collection consisted of 8 activities: *1)reading in bedroom A, 2)watching television, 3)reading in bedroom B, 4)sitting in the lobby, 5)reading in the library (first floor), 6)looking for a book in the library (first floor), 7)reading in the library (second floor), 8)looking for a book in the library (second floor).* Activities 1-4 were performed in an apartment building while activities 5-8 were performed in a library. The data collection process was performed by two participants under supervision. The test set was collected independently in a different day from the training set.

From the tri-axial accelerometer sensor, we read the acceleration values from each of the axes (*x,y,z*) and classify the *primitive activity* being performed as one of *walking, running* or *resting.* There are several approaches for recognizing primitive activities (see section 1.1). For this work we used a nearest neighbor approach [6].

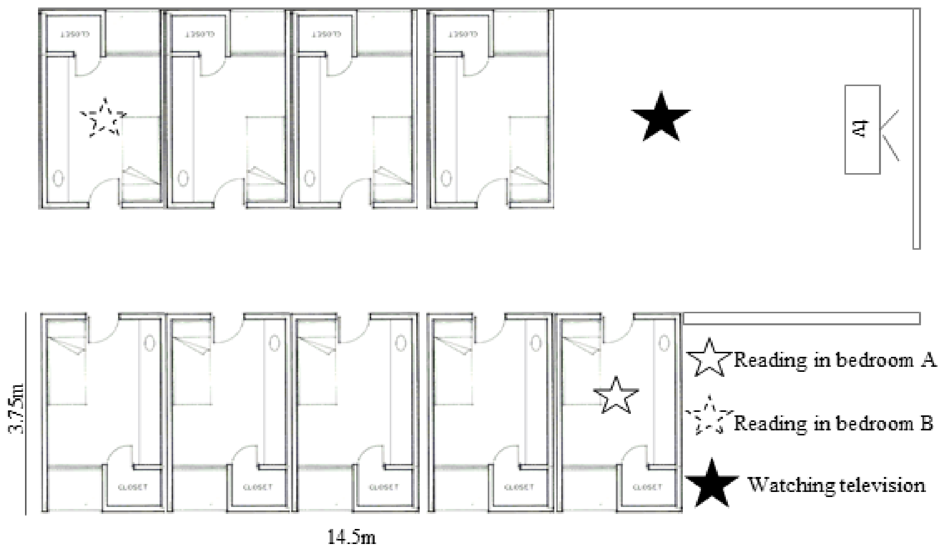


Figure 1. Layout of the apartments building 3rd floor.

From the Wi-Fi sensor, we collected data from the in range Wireless Access Points. Specifically, we collected their BSSID (Basic Service Set Identifier) and signal strength. We selected activities that are very close to each other. Figure 1 shows the layout of the 3rd floor of the apartments building. The lobby is located below the room marked with *Reading in bedroom A* but in the 1st floor.

2.2. Training

In the training phase we generate the instances that will be used as the training set. An instance based learning algorithm (K-nearest neighbors) [16] was used for the *contextualized activities* recognition. Each activity instance has 3 attributes (Table 1).

Figure 2 shows the process of generating one training instance. First, the application performs two scans to collect the data of the in range Access Points. A delay of 500*ms* is set between the scans. Then, every second the *primitive activity* being performed by the user is recognized and stored in a vector *V*. This is done during 5 seconds, i.e., at the end of the 5 seconds, the vector *V* will contain 5 ids' (one for each detected *primitive activity*). Finally the application performs two more scans to gather information from the Access Points. The reason of doing several scans is because in [17] they observed that sometimes one or more Access Points may not be detected because limited sensitivity of the hardware and/or long beacon interval of some Access Points.

Now the instance is created and its *primitive activity* is set to Mode(*V*), i.e., the *primitive activity* that was dominant across the 5 seconds period. For every Access Point found during the scans a pair <*bssid,strength*> is added to *L*, where *bssid* is the Access Point identifier and *strength* is the mean of the signal strength from the 4 scans.

Table 1. Instance Attributes.

Name	Description
Id	Unique identifier of the instance (for debugging purposes)
Class	A number from 1 to 8 to identify which *contextualized activity* the instance belongs to.
Primitive activity	Identifies the *simple activity* associated with this *contextualized activity*. 0 for walking, 1 for running and 2 for resting
List of Access Points	A list *L* in which each element is a pair *<bssid,strength>*

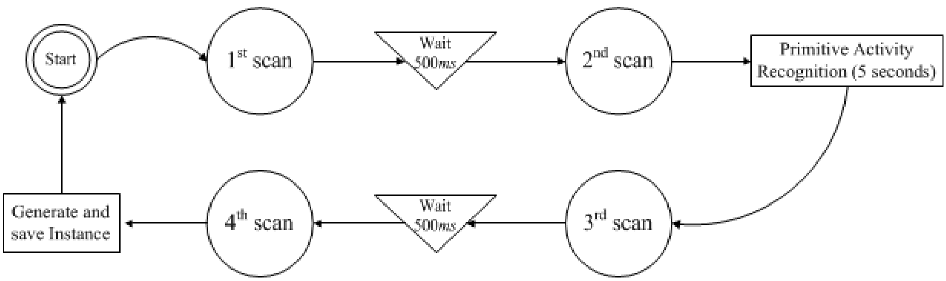


Figure 2. Steps to generate one training instance

2.3. Classification

The K-Nearest Neighbors (K-NN) method was used for the classification. The method consists of computing the distance between the query instance and every other instance from the training set and selecting the *k* nearest instances. For this work we used the Euclidean distance:

$$\sqrt{\sum_{i=1}^{n}\left(b_i - c_i\right)^2} \quad (1)$$

where n is the number of attributes, b is the value of the i^{th} attribute of the query instance and c is the value of the i^{th} attribute of the training instance. For the experiments we used $k = 3$. The differences $(b_i - c_i)$ for $i=1..3$ were computed as follows:

- Primitive activity id. Set to 0 if both *b* and *c* have the same primitive activity.

$$dif(b_1, c_1) = \begin{cases} 0 \; if \; P(b) = P(c) \\ 1 \; otherwise \end{cases} \tag{2}$$

where the function P returns the primitive activity associated with the specified instance.

- Ratio of same Access Points. The extreme cases are when both instances share the same Access Points (in this case the distance is 0) and when they do not have any common Access Point (in this case the distance is 1).

$$dif(b_2, c_2) = 1 - (same / total) \tag{3}$$

where $total = cardinality\{L(b) \cup L(c)\}$ and the function L returns the list of Access Points of the specified instance. Similarly, the variable *same* is computed as $same = cardinality\{L(b) \cap L(c)\}$. Eq. (3) is known as the Jaccard distance.

- Difference of the signal strength's standard deviation. This is defined as:

$$dif(b_3, c_3) = 1 - (1/1 + \alpha) \tag{4}$$

where $\alpha = abs(SD(a,b) - SD(b,a))$ and $SD(p1,p2)$ is a function that returns the standard deviation of the signal strength of all Access Points of *p1* that are also in *p2*.

For the real time classification, a query instance is crated in the same way a training instance is created (see Figure 2) and then it is classified using K-NN algorithm.

3. Experiments

In this section we describe the details of the experiments and then we present the results of the classifications.

3.1. Experiment Description

For this experiment we collected a total of 741 instances for the training set and 243 instances for the test set. Tables 2 and 3 show the number of instances per activity and the average number of detected Access Points. The average number of detected Access Points is important because in [17] they observed that the number of received Access Points strongly affects accuracy of proximity classification. They improved their results by performing three scans and feeding the algorithm with the scan that has the highest number of detected Access Points.

Table 2. Number of training instances and their respective average number of detected Access Points

Composed Activity	**Number of training instances**	**Average number of detected Access Points**
1) Reading in bedroom A	102	2.2
2) Watching television	104	4.0
3) Reading in bedroom B	68	1.5
4) Sitting in the lobby	91	5.1
5) Reading in the library (first floor)	107	4.2
6) Looking for a book in library (first floor)	78	8.5
7) Reading in the library (second floor)	103	4.9
8) Looking for a book in library (second floor)	88	10.8

Table 3. Number of test instances and their respective average number of detected Access Points

Composed Activity	**Number of test instances**	**Average number of detected Access Points**
1) Reading in bedroom A	29	2.3
2) Watching television	28	3.7
3) Reading in bedroom B	27	1.5
4) Sitting in the lobby	32	4.2
5) Reading in the library (first floor)	36	6.0
6) Looking for a book in library (first floor)	30	7.0
7) Reading in the library (second floor)	31	9.9
8) Looking for a book in library (second floor)	30	8.4

The time taken to complete the process shown in Figure 2 for each generated instance takes about 7 seconds, i.e., 5 seconds for the *primitive activity* recognition and approximately 2 seconds for the 4 Wi-Fi scans. Each scan takes less than 1 second but it varies depending on the number of visible Access Points, hardware and other physical issues.

Table 4. Confusion Matrix

Activity	**Classified as**							
	Reading in bedroom A	Watching television	Reading in bedroom B	Sitting in the lobby	Reading in library 1f	Looking for book in library 1f	Reading in library 2f	Looking for book in library 2f
Reading in bedroom A	**22**	5	0	2	0	0	0	0
Watching television	0	**27**	1	0	0	0	0	0
Reading in bedroom B	5	3	**19**	0	0	0	0	0
Sitting in the lobby	0	0	0	**32**	0	0	0	0
Reading in library 1f	0	0	0	0	**32**	0	4	0
Looking for book in 1f	0	0	0	0	0	**26**	0	4
Reading in library 2f	0	0	0	0	0	0	**31**	0
Looking for book in 2f	0	0	0	0	0	1	0	**29**

3.2. Results and Discussion

The overall classification accuracy using holdout validation was 89.7%. Table 4 shows the confusion matrix. From this table we can see, e.g., that 1 of the 28 instances of *watching television* activity was misclassified as *reading in bedroom B* and the remaining 27 instances were correctly classified. The main diagonal shows the number of correctly classified instances. We performed 10-fold cross validation over the entire data set (984 instances = training + test instances) and the resulting accuracy was 90.3%.

It can be seen that there is a relation between the average number of detected Access Points per activity (Table 1, 2) and the accuracy of the classification. For example, *reading in bedroom B* has the lowest average of detected Access Points (just 1), and in the confusion matrix it can be seen that 8 of its instances were misclassified. In contrast, the activities with high number of detected Access Points had fewer misclassifications. Then, based on these results it appears that in order to achieve good classification accuracies for this experiment's configuration, the activity must have at least an average of 3 detected Access Points.

The major source of error was between activities that are physically close to each other. Given that the library is very far from the departments building there were no misclassifications between these activities.

4. Approaches for Human Activity Recognition

Table 5 shows that due to its simplicity and accuracy accelerometers are the main sensors used in activity recognition. For *complex activities* more sensors (RFID, clock, temperature sensor, light sensors, etc.) are required. As mentioned in section 1 this work uses an approach that lies between simple and complex activities because we first recognize a simple activity and then we add Wi-Fi information to contextualize it.

Table 5. Different approaches for human activity recognition

Reference	Activity Type	Sensors	Approach	No. Activities	Accuracy [%]
[6]	Simple	1 tri-axial accelerometer	KNN	6	70-90
[7]	Simple	5 bi-axial accelerometers	HMM	7	93-98.5
[8]	Simple	1 tri-axial accelerometer	Decision Tree	12	90.8
[10]	Simple	1 tri-axial accelerometer, 1 tri-axial gyroscope	Bag of Features	9	92.7
[18]	Simple	Video sequences	Bayesian Classifier	5	1.5 error rate
Our approach	Contextualized	1 tri-axial accelerometer, Wi-Fi	KNN	8	89.7
[11]	Complex	3 iMote2 sets, 2 RFID wristband readers	Emerging Patterns	26	78.58-90.96

[12]	Complex	3D accelerometer, real time clock, 9 binary tilt switches, temperature sensor, 2 light sensors	Topic Models	34	72.7
[13]	Complex	2D accelerometer, 9 binary tilt switches	K-means, SVM, Nearest Neighbor, HMM	3	80.6-91.8

5. Conclusions and Future Work

In this work we presented a way to combine data from an accelerometer and Wi-Fi to recognize different types of activities. Using simple supervised learning algorithms we achieved good results in the classifications accuracy, subject to the number of detected Access Points and the proximity between the activities.

We chose activities that are common in academic environments, however they may not generalize well for a wider range of activities involving other configurations, e.g., office activities, sport activities, home activities, etc.

For this work the *primitive activities* were chosen by getting the mode activity from the 5 second period. Just taking the mode imposes a limit on the number of activities that can be detected within a given area since they may overlap.

The next step of this research is to devise a way of generating a *contextualized activity* that can include more than 1 *simple activity*, i.e. instead of setting the primitive activity attribute just as the mode we want to use the complete sequence of *simple activities* and take the sequence order into account. For the next phase of this research, we will include a wider range of activities in order to tell whether this method is suitable for various situations and configurations.

References

[1] D. J. Cook, J. C. Augusto, and V. R. Jakkula. Ambient intelligence: Technologies, applications, and opportunities. In: Pervasive and Mobile Computing, vol. 5(4) pp. 277–298 (2009)

[2] E. Kim, S. Helal, and D. Cook. Human Activity Recognition and Pattern Discovery. IEEE Pervasive Computing, vol. 9(1) pp. 48–53 (2010)

[3] J. Krumm, and E. Horvitz. LOCADIO: Inferring Motion and Location from Wi-Fi Signal Strengths. In: Mobiquitous, pp. 4–13. (2004)

[4] J. Correa, E. Katz, P. Collins, and M. Griss. Room-Level WiFi LocationnTracking. Carnegie Mellon Silicon Valley, CyLab Mobility Research Center technical report MRC-TR-2008-02 (2008)

[5] G. Zaruba, M. Huber, F. Kamangar, and I. Chlamtac. Monte Carlo Sampling Based In-home Location Tracking With Minimal RF Infrastructure Requirements. In: Proceedings of IEEE GLOBECOM, pp. 3624–3629 (2004)

[6] T. Brezmes, J. L. Gorricho, and J. Cotrina. Activity Recognition from Accelerometer Data on a Mobile Phone. In: Proceedings of the IWANN '09, pp. 796–799. Springer, (2009)

[7] A. Mannini, and A. M. Sabatini. Machine learning methods for classifying human physical activity from on body accelerometers. In: Sensors, vol. 10(2) pp. 1154–1175 (2010)

[8] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Transactions on Information Technology in Biomedicine, vol. 10 pp. 156–167 (2006)
[9] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. Activity Recognition from Accelerometer Data. In: AAAI, pp. 1541–1546 (2005)
[10] M. Zhang, and A. Sawchuk. Motion Primitive-Based Human Activity Recognition Using a Bag-of-Features Approach. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium (IHI '12). ACM, New York, NY, USA, pp. 631-640 (2012)
[11] T. Gu, Z. Wu, X. Tao, H. Pung, and J. Lu. epSICAR: An Emerging Patterns based Approach to Sequential, Interleaved and Concurrent Activity Recognition. In: 7th Conf. on Pervasive Computing and Communications. IEEE P., New York (2009)
[12] T. Huynh, M. Fritz, and B. Schiele. Discovery of activity patterns using topic models. In: Proceedings of the International Conference on Ubiquitous Computing. ACM, New York, pp. 10–19 (2008)
[13] T. Huynh, U. Blanke, and B. Schiele. Scalable recognition of daily activities with wearable sensors. In: 3rd International Symposium on Location- and Context-Awareness (LoCA), pp 50–67 (2007)
[14] H. Tian, P. Lei, L. Xingjuan, and X. Shusong. Wearable activity recognition for automatic microblog updates. In: IEEE/ASME International Conference on Advanced Intelligent Mechatronics, pp. 1720–1723 (2009)
[15] Android 2.2 Platform, http://developer.android.com/sdk/android-2.2-highlights.html
[16] D. Aha, D. Kibler, and M. K. Albert. Instance-Based Learning Algorithms. Machine Learning, vol. 6 pp. 37–66 (1991)
[17] A. Carlotto, M. Parodi, C. Bonamico, F. Lavagetto, and M. Valla. Proximity classification for mobile devices using wi-fi environment similarity. In: ACM International Workshop. (2008)
[18] P. C. Ribeiro, and J. Santos-Victor. Human activity recognition from video: Modeling, feature selection and classification architecture, In: Proc. Int. Workshop Human Activity Recognition and Modeling, pp. 61–70 (2005)

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-104

Structured Light and Shape Descriptors for Automatic 3D Object Classification

Daniel GARCIA [a,1], Manuel ORNELAS [a], Raul SANTIAGO [a], Hector PUGA [a] and Martin CARPIO [a]
[a] *Instituto Tecnologico de Leon*

Abstract. This paper presents a methodology for classification of 3D objects that can be useful to solve some problems in context-awareness, automatic surveillance and ambient assisted living.
This methodology uses light line images projected onto the object being studied. The use of structured light allows to get the three-dimensional object information. The images, that were captured by a CCD camera, were improved with tools of digital image processing and were characterized by three compactness measures. The classification was performed by some algorithms like Naive Bayes, Neural Nets and an associative classifier.

Keywords. Structured light, computer vision, pattern classification, intelligent environments

Introduction

Computer vision can solve some problems around intelligent environments like 3D object recognition for context-awareness, care and tracking of elderly, where it is necessary to get and process visual information from the environment to perform some actions.

One of the current tasks in the field of computer vision is the recognition of 3D objects, as well as getting their shape. Vision systems based in two dimensions lose information of the 3D object when they try to make the projection of $\mathbb{R}^3 \mapsto \mathbb{R}^2$ or $\mathbb{R}^3 \mapsto \mathbb{R}^1$, this causes that a classification module can not distinguish objects with similar 2D shape, but having differences in their 3D shape.

In this paper we propose a cutting edge solution for the classification of 3D objects using structured light, simple descriptors and pattern recognition algorithms.

The structured light is an efficient technique to get 3D object information through simple 2D images. This information is obtained by projecting a laser line on the object being studied and the image is captured by a CCD camera. The acquired images are segmented by the method of Otsu [1] and improved by a morphological closing operation [2]. From the result of the segmentation, it is performed the process to characterize images using some compactness measures [3]. The main features of the images are processed using different algorithms for classification.

[1]Daniel Garcia is a postgraduate student in the Instituto Tecnologico de Leon, Av. Tecnologico s/n, Fracc. Ind. Julian de Obregon, Leon, Gto, Mex.; E-mail: danga23@gmail.com. Other authors are researchers in the Instituto Tecnologico de Leon

The projection of laser light lines on 3D objects is a technique normally used for the reconstruction and modelling of the same objects [4][5][6][7]. However, the structured light has also been used in the recognition of objects, as shown in [8][9][10]. A similar approach of light lines with a color code has also allowed to perform 3D recognition [11][12][13]. Moreover, different similarity measures have been widely used in pattern recognition applications [14][15][16].

For context-awareness, object classification is a necessary step for human-robot and robot-environment interaction [17], scene description [18] and, for a robotic agent is useful to avoid obstacles and pick up an object [19]. In other way, object recognition is a critical step in automatic surveillance and ambient assisted living [20][21][22][23].
This article is organized as follows: section 1 presents the theoretical concepts that were used. Section 2 shows the methodology employed. Section 3 shows the results and Section 4 presents the conclusions of this work.

1. Theoretical Foundations

In computer vision, a three-dimensional scene is viewed by one, two or more cameras to produce monochrome or color images. The acquired images can be segmented to obtain from them interesting features such as edges or regions. From the features, underlying properties are obtained by the corresponding description process. After that, it is possible to achieved the three-dimensional scene structure that is required by the application of interest [24].

1.1. Segmentation

Image segmentation is one of the most important steps leading to the analysis of processed image data. The main goal of segmentation is to divide an image into parts that have a strong correlation with objects or areas of the real world that are contained in the image [25]. The segmentation process terminates when the objects of interest have been isolated [26].

One technique of segmentation consists on binarize the image, which includes selecting an optimum threshold to perform the binarization. In this work, we used the Otsu method to solve this problem [1]. The methodology consists in calculating the threshold value so that the dispersion of the distribution of gray levels within each segment has to be as small as possible but at the same time the dispersion has to be as high as possible between different segments.

1.2. Description

Once the image has been segmented, according to Gonzalez [26], there are two options for the representation of regions: 1) A region can be represented in terms of external features (boundaries) or 2) a region can be represented in terms of internal features. The shape of an object can be a region that is represented in terms of its internal points.

The analysis of the shape of an object is an important field of computer vision area. The shape descriptors are computational tools that are used to analyze information that is obtained from the shape of an image [27].

1.2.1. Compactness

Compactness plays an important role in classification and shape analysis. The measure of compactness is an intrinsic property of objects that can be computed by $(perimeter^2)/area$. Therefore, the measure of compactness is invariant under geometric transformations such as: translation, rotation and scaling [28]. The expression can be described as P2A [3].

The following are some compactness measures that are used in this work[3].

1. Shape compactness associated with the old ratio

$$C = P^2/A \tag{1}$$

where P is the perimeter and A is the area of the shape.

2. The normalized discrete compactness

$$C_{DN} = \frac{C_D - C_{Dmin}}{C_{Dmax} - C_{Dmin}} \tag{2}$$

where C_D is the perimeter of contact of a digital region. If the pixel cell is a square, then C_D is given by:

$$C_D = \frac{2n - P}{2} \tag{3}$$

where n is the number of region pixels, and P is the perimeter of the digital region. C_{Dmin} and C_{Dmax} are the bottom and the top limit of the contact perimeter of a shape composed of n number of region cells, respectively. These parameters are computed by:

$$C_{Dmin} = n - 1 \tag{4}$$

and

$$C_{Dmax} = \frac{4n - 4\sqrt{n}}{2} \tag{5}$$

3. According to Peura

$$C = \frac{P_{circle}}{P_{shape}} \tag{6}$$

where P_{shape} is the perimeter of the digital shape with area A and P_{circle} is the perimeter of a circle with the same area (A) in the continuous plane.

1.3. Algorithms for Pattern Classification

Pattern Recognition's, which is a scientific discipline, goal is the classification of objects in a number of categories or classes [29]. Classification is a process that assigns a label to an object according to some representation of their properties [2]. The pattern recognition algorithms used in this study were:

1. Naive Bayes. The Naive Bayes learning method involves a learning step in which the various $P(C_j)$ and $P(x_i|C_j)$ terms are estimated, based on the frequencies of the training data. The set of those estimates corresponds to the learned hypothesis. This hypothesis is then used to classify each new input vector $X_i = \{x_1, x_2, \ldots, x_n\}$ according to Eq.(7) [30].

$$Assigned\,class = arg\max_{v_j \in V} P(C_j) \prod_i P(x_i|C_j) \tag{7}$$

where $P(C_j)$ is the probability of occurrence of the training patterns of the class C_j and $P(x_i|C_j)$ is the probability of occurrence of individual attributes of an input vector X_i, given the class C_j.

2. Neural Networks - Backpropagation. The backpropagation method consists in applying the steepest descent method to minimize the error produced by neural net output [31].
   The training of a network using backpropagation consists of three stages: (a) feedforward the input training pattern throughout the neural network; (b) back-propagation analysis of the error; and (c) updating the weights.
3. Associative Pattern Classifier (APC). This algorithm is based on the combination of two models of associative memories, the Lernmatrix and the Linear Associator. It combines the principles of matrix algebra that are used by Linear Associator and classification criteria used by the Lernmatrix [32].
   At this time this algorithm works only for two classes, but we can improve its accuracy by using dichotomies between the classes that we need to classify.

2. Image Acquisition, Processing and Description

The following describes the methodology undertaken to achieve the goal of this work:

1. It is performed the projection of a line of laser light on the objects of study. Therefore the entire surface of each object was covered by 10 uniform displacements of the line. The different images of the distorted line were captured by a CCD camera and stored on a computer (Figure 1). The object was displaced six times within the reference plane area to obtain a set of 10 lines per each displacement. Each set of 10 lines represents an element (x_i) that belongs to the class (C_j), then we have six items for each class.

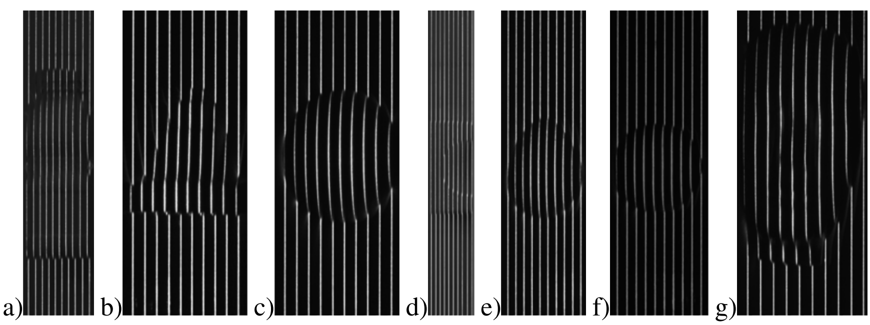


Figure 1. Images of light lines projected onto a) bottle, b) bell, c) sphere, d) concave figure, e) egg, f) mandarin and g) face. These are arbitrary objects that can be found in real intelligent environments

2. As it can be seen in Figure 1, those objects have different sizes, so it's necesary to normalize the images. That was implemented to eliminate noise produced by the straight line that appears at the bottom and top of the deformed line (Figure 2).

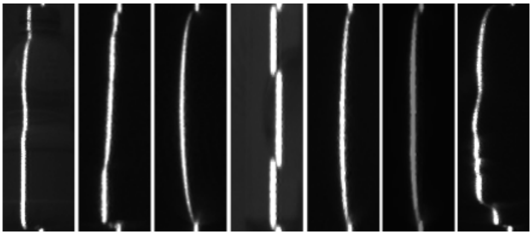


Figure 2. Normalized Images of light lines.

3. Each image was segmented by Otsu method. Figure 3 shows the segmented image.

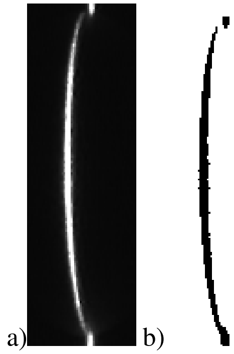


Figure 3. Segmented line by the Otsu method. a) Original image and b) Segmented image.

4. Once the images were segmented, a morphological closing operation was applied, to improve the segmented images (Figure 4) [2], using the structuring element that is illustrated in Figure 5.

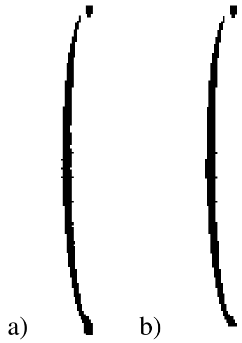


Figure 4. Improved segmented image by a morphological closing operation. a) Segmented image and b) Improved image.

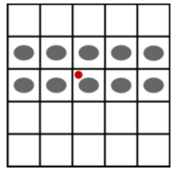


Figure 5. Structuring element that is used to improve the segmented image.

5. After the images have been improved, the images were characterized by the compactness measures to generate the training set.
   Each disturbed line was described by the three compactness measures and the feature vector of every object was formed with the concatenation of each vector of every set of 10 lines, resulting in a vector with 30 components for each object.
6. At this time, there are only six feature vectors for each object or class. Since seven objects were analyzed, the instance contains just 42 vectors, which are named ”original vectors”. This number of vectors is so small, for this classification process, therefore it is necessary to create new feature vectors, based on those that we have already generated.
   These new vectors have been generated by increasing and decreasing, over each of the components of an original vector. For every original vector the first increase and decrease was of $\pm 0.25\%$ of the value for each component, which generates two new vectors. By using a step of 0.25%, 100 new vectors were generated for each original vector.
7. Finally, some classification test were made using different classifier algorithms, like Naive Bayes, neural networks and APC.

3. Experimental Results

The confusion matrix was computed for each implemented algorithm (Table 1), where TC_i and FC_i represent all the vectors or objects that were classified correctly and incorrectly for each class. From this matrix we calculated the accuracy of every algorithm by the Eq. 8 [33]. The k fold cross validation was implemented to compare the performance of each algorithm [34], it was proposed a $k = 10$.

Table 2 shows the results that were obtained on every k-fold for every algorithm for the instance of vectors. Figure 6 illustrates a comparative chart of the results that were obtained for each algorithm.

From the results it can be inferred that Naive Bayes algorithm has the best accuracy.

Table 1. Example of confusion matrix.

		Predicted class		
		C_1	C_2	C_3
True Class	C_1	TC_1	FC_2	FC_3
	C_2	FC_1	TC_2	FC_3
	C_3	FC_1	FC_2	TC_3

$$Accuracy = \frac{TC_1 + TC_2 + TC_3}{C_{1+}C_2 + C_3} \quad (8)$$

Table 2. Results for compactness measures.

k-fold	Naive Bayes	Neural Net	APCDichotomies
1	96.9%	81.43%	73.1%
2	95.95%	63.33%	68.81%
3	97.62%	57.86%	68.33%
4	97.62%	81.67%	70.0%
5	96.9%	93.57%	72.14%
6	95.24%	88.57%	71.67%
7	95.71%	84.52%	72.62%
8	94.29%	70.24%	70.0%
9	95.71%	79.05%	68.81%
10	96.9%	62.38%	67.62%

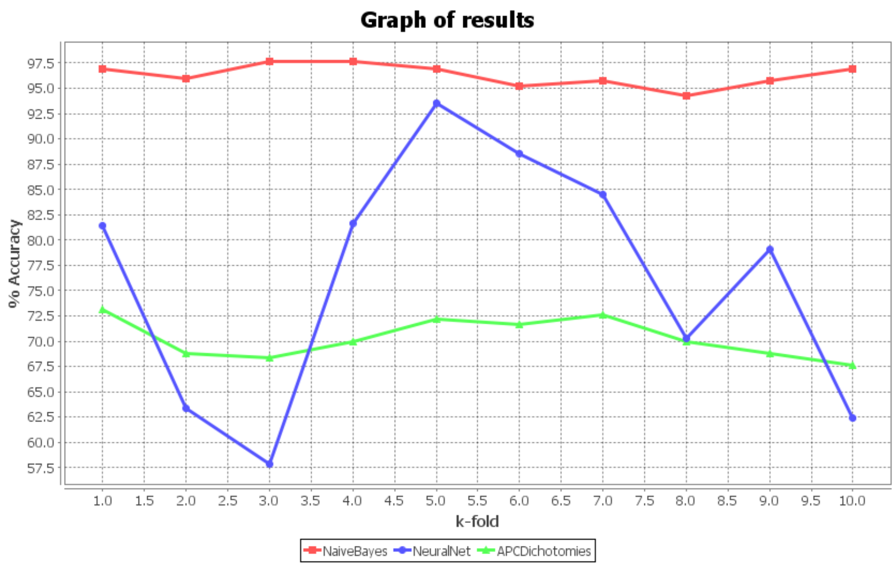


Figure 6. Graph of results for compactness measures.

4. Conclusions

A methodology for classification of 3D objects, by characterizing lines of laser light projected onto them, was presented. With structured light the shape of 3D object can be analized in a 2D space, this prevents loss of information. In this work, some tools of digital image processing were used for segmentation and characterization of images of light lines. These light lines were acquired under controlled laboratory conditions. Some classification algorithms have been implemented, and the best results were obtained with Naive Bayes algorithm. Based on those results, it is possible to apply this methodology on real intelligent environments.

5. Acknowledgements

The authors thank CONACYT and DGEST, that provide financial support for this research.

References

[1] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11:285–296, 1975.

[2] L.G. Shapiro and G.C. Stockman. *Computer visión*. Prentice Hall, 2001.

[3] R.S. Montero and E. Bribiesca. State of the art of compactness and circularity measures. In *International Mathematical Forum*, volume 4, pages 1305–1335, 2009.

[4] Dueñas Juan Espinal. Modelado 3d de objetos utilizando proyección de una línea de luz y redes neuronales. Master's thesis, Instituto Tecnológico de León, 2010.

[5] SY Chen, YF Li, and J. Zhang. Vision processing for realtime 3-d data acquisition based on coded structured light. *Image Processing, IEEE Transactions on*, 17(2):167–176, 2008.

[6] J.A. Muñoz-Rodriguez, R. Rodriguez-Vera, and M. Servin. Direct object shape detection based on skeleton extraction of a light line. *Optical Engineering*, 39:2463–2471, 2000.

[7] Song Zhang. *High-resolution, Real-time 3-D Shape Measurement*. PhD thesis, Stony Brook University, 2005.

[8] G. Arnold, T. Klausutis, and K. Sturtz. Three-dimensional laser radar recognition approaches. *Computer Vision Beyond the Visible Spectrum*, pages 71–114, 2005.

[9] F. Tsalakanidou, F. Forster, S. Malassiotis, and M.G. Strintzis. Real-time acquisition of depth and color images using structured light and its application to 3d face recognition. *Real-Time Imaging*, 11(5-6):358–369, 2005.

[10] M. Unel, O. Soldea, E. Ozgur, and A. Bassa. 3D object recognition using invariants of 2D projection curves. *Pattern Analysis & Applications*, 13(4):451–468, 2010.

[11] M. Buendia, R. Salvador, R. Cibrian, M. Laguia, and JM Sotoca. Determination of the object surface function by structured light: application to the study of spinal deformities. *Physics in Medicine and Biology*, 44:75, 1999.

[12] EM Petriu, Z. Sakr, HJW Spoelder, and A. Moica. Object recognition using pseudo-random color encoded structured light. In *Instrumentation and Measurement Technology Conference, 2000. IMTC 2000. Proceedings of the 17th IEEE*, volume 3, pages 1237–1241. IEEE, 2002.

[13] E. Schubert. Fast 3d object recognition using multiple color coded illumination. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, volume 4, pages 3057–3060. IEEE, 1997.

[14] M. Rohrschneider, G. Scheuermann, S. Hoehme, and D. Drasdo. Shape characterization of extracted and simulated tumor samples using topological and geometric measures. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 6271–6277. IEEE, 2007.

[15] L. Zongmin, K. Hou, L. Yujie, D. Luhong, and L. Hua. The shape recognition based on structure moment invariants. *International Journal of Information Technology*, 12(2), 2006.

[16] Z. Huang and J. Leng. Analysis of hu's moment invariants on image scaling and rotation. In *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, volume 7, pages V7–476. IEEE, 2010.

[17] C.C. Kemp, C.D. Anderson, H. Nguyen, A.J. Trevor, and Z. Xu. A point-and-click interface for the real world: laser designation of objects for mobile manipulation. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 241–248. ACM, 2008.

[18] T.M. Strat and M.A. Fischler. Context-based vision: recognizing objects using information from both 2 d and 3 d imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10):1050–1065, 1991.
[19] H. Celik, A. Hanjalic, and E.A. Hendriks. A framework for unsupervised training of object detectors from unlabeled surveillance video. *Journal of Ambient Intelligence and Smart Environments*, 3(3):213–235, 2011.
[20] H. Aghajan, J.C. Augusto, C. Wu, P. McCullagh, and J.A. Walkden. Distributed vision-based accident management for assisted living. In *Proceedings of the 5th international conference on Smart homes and health telematics*, pages 196–205. Springer-Verlag, 2007.
[21] A. Fod, A. Howard, and MAJ Mataric. A laser-based people tracker. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, volume 3, pages 3024–3029. IEEE, 2002.
[22] T. Kleinberger, M. Becker, E. Ras, A. Holzinger, and P. Müller. Ambient intelligence in assisted living: enable elderly people to handle future interfaces. *Universal Access in Human-Computer Interaction. Ambient Interaction*, pages 103–112, 2007.
[23] W. Yan, C. Weber, and S. Wermter. A hybrid probabilistic neural model for person tracking based on a ceiling-mounted camera. *Journal of Ambient Intelligence and Smart Environments*, 3(3):237–252, 2011.
[24] G.P. Martínsanz and J.M.C. García. *Visión Por Computador: Imágenes Digitales y Aplicaciones*. Ra-Ma, Libreria y Editorial Microninfomatica, 2001.
[25] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. Thompson Learning, 2008.
[26] R.C. Gonzalez and R.E. Woods. *Digital image processing*. Prentice Hall, 2002.
[27] C.A. Martinez-Ortiz. *2D and 3D Shape Descriptors*. PhD thesis, University of Exeter, 2010.
[28] E. Bribiesca. Measuring 2-d shape compactness using the contact perimeter. *Computers & Mathematics with Applications*, 33(11):1–9, 1997.
[29] S. Theodoridis and K. Koutroumbas. *Pattern recognition*. Elsevier/Academic Press, 2009.
[30] T.M. Mitchell. *Machine learning*. McGraw-Hill, Marzo 1997.
[31] M. Friedman and A. Kandel. *Introduction to pattern recognition*. World Scientific, 1999.
[32] R.S. Montero. Clasificador híbrido de patrones basado en la lernmatrix de steinbuch y el linear associator de anderson-kohonen. Master's thesis, Instituto Politócnico Nacional - Centro de Invectigación en Computación, Junio 2003.
[33] T. Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874, 2006.
[34] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint Conference on artificial intelligence*, volume 14, pages 1137–1145. Lawrence Erlbaum Associates LTD, 1995.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-113

ALDRO Learning and Mixed Decision Support Method for Mobile Robot

Fernando Orduña C., Alma Valle R., Alán Dennis V., Javier Rojas D. [a] and Flavio Muñoz B., Alberto Ramírez , Iván Luna, Guadalupe Vásquez C., Karina Hinojosa T.[b]
[a] *Computer Science, ITESCA-SEKAI*
[b] *Advance Research Center CETA-ITESCA*

Abstract. This research is aimed to get an intelligent navigation robot, whose functionality is based on the interaction of two agents, a reactive agent and a deliberative agent, where both agents take decisions jointly to achieve a goal. The deliberative agent makes a decision based on the acquired knowledge and determines which way leads to the best result. The reactive agent receives the decision made by the deliberative agent and transforms it into a physical interaction with the environment using line following techniques. This method was implemented successfully into a robot and the tests results were as expected.

Keywords. Q-Learning, Intelligent Agents.

Introduction

Nowadays, self-learning in robotics is a very common topic [1], although not wondering whether it is possible for a robot to learn on its own, but trying to achieve the most optimal method with the least human intervention. This can be achieved by using reinforcement learning techniques[2], such as Q-Learning, which has been a subject of study in several research projects, such as those of Smart & Kaelbling [3] and Gaskett [4].

In order to learn the environment in research [3], they use a joystick to control the robot, guiding it through the states and actions and making it learn passively with human intervention, so that afterwards it can be left alone and function without any help, using no more than the knowledge previously obtained. As opposed to this, our robot applies line following techniques [5], using a set of sensors to “see” on its own the environment and walk throughout it, learning the states and rewards through stickers located in specific places all around the environment.

The main objectives pursued by line followers is to achieve the best record time, considering tracking of complex trajectories, this aim could be achieved with the newest electronic technology and the help of advanced computational algorithms.

The current aim in this robotic field is the research in getting lower travel times and high transfer speeds. Other research filed is the use of assigning tasks and decision making.

Some examples of the use of robot followers of lines are: [6], the implementation of a proportional + integral + derivative control (PID), with the purpose of error correction to track lines and thus decrease the steady-state error in tracking trajectories.

Another example is [7], where a proposal is made to analyze and calculate the best location to implement sensors on the robot, using a genetic algorithm.

The work in [8], involves a robot with dynamic configuration, as it has the particularity to reconfigure online, without having to stop their activities. The reconfiguration is the ability to detect both black and white lines, without specify to this parameter.

Comparing the works already mentioned in this article, this research shows a line follower robot capable of finding the optimal path in the shortest time possible, using some ability to move it.

This research proposes the ALDRO Method, a combination of reinforcement learning and line following techniques, to help the robot decipher the environment to apply afterwards the knowledge obtained, all of this to get to the final reward in the most optimal manner. This is obtained by a reactive agent (line follower) and a deliberative agent (Q-learning), that way we get a robot that learns everything on its own, without any human intervention.

1. The Robot

Figure 1 depicts the robot's architecture that consists of three different sections. The first one is where the sensors are and read the signals from the environment. The second one is where the agents [9] are positioned and do all the decision making process. Finally, the third one is the one of the mechanics that rotates the motors as needed.

The architecture proposed works with both Reactive Agent (***ra***) and Deliberative Agent (***da***) together. The set of infrared (IR) sensors, of the line following function, detects if an intersection has been found. Then, it sends a message to the ***da*** which indicates the type of intersection. Both agents process that information along with the data read from the stickers --which are located throughout the environment-- to know where the robot is and where it has to go. The ***ra*** communicates with the L293D driver to rotate the motors and move to the desired direction.

The knowledge process section of the robot architecture is the area where the agents are. The agents are the brain of the line follower robot and decide where it has to go to get to the goal. The ***da*** uses a reinforcement learning algorithm to decide the best possible option. The ***ra*** executes that decision. The communication between these agents is what makes the learning of the environment and the movement of the robot possible.

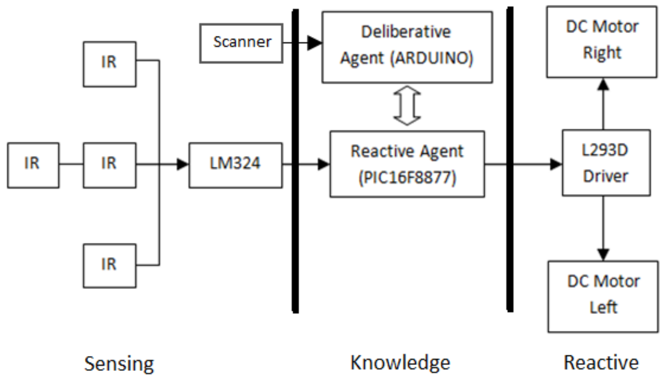


Figure 1. Architecture.

2. Communication Protocol

The interaction between both agents is denoted in the following figure:

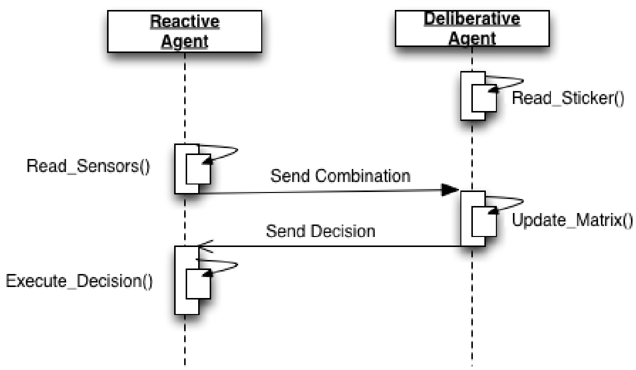


Figure 2. Short caption.

Figure 2 shows the communication process between both agents when the robot is exploring the environment. The robot follows the line until the ***da*** reads a sticker, which determines the current state in the environment. When the sensors of the ***ra*** detect an intersection, the robot automatically stops and waits for the decision made by the ***da***, which sends a message that contains information about the intersection to the ***ra***.

The sensors of the ***ra*** are continuously reading signals from the environment.

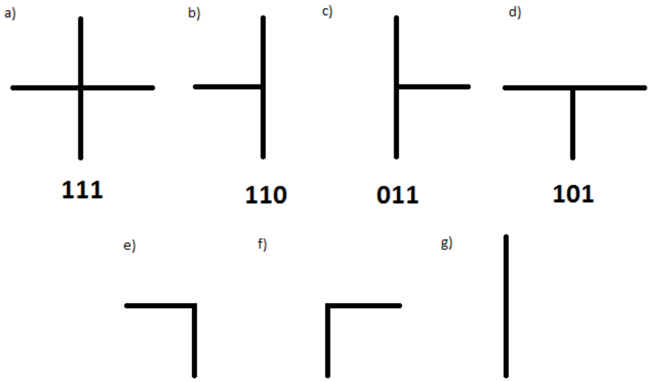


Figure 3. ***ra*** sensors reading combination.

According to figure 3, the ***ra*** reads its sensors and sends a binary number to the ***da*** to indicate the kind of intersection found (a, b, c or d). As long as the reading of the sensors doesn't indicate an intersection, the ***ra*** executes the line following techniques to keep going throughout the environment (e, f or g). The sent combination is interpreted by the ***da*** to develop the structure of the environment; based in it, the ***da*** will randomly choose one of the possible options.

Table 1 shows all the possible decisions the ***da*** can make and the binary number associated to them. The ***da*** has to decide the next direction for the robot to take. This decision is sent to the ***ra*** as a binary number, as shown in table 1.

Table 1. Sent combinations by the ***da***

Possible Decision	Combination Sent to *ra*
Turn Left	000
Turn Right	001
Forward	010
Turn 180º	011
Reached Goal	100

3. Getting the Environment

The Environment used for the learning and experimentation is depicted in figure 4, as follows: The environment has states that are represented by letters and are connected by edges. Each state means a possible decision to be made.

Figure 5 shows four possible decisions to be made in the learning method (***lm***) used, but, internally in the microprocessor, a set of rules has to be implemented to get the mixed method proposed. This is implemented to improve the efficiency of the ***ra***.

Each state has a set of stickers that the ***ra*** will scan to get information about the environment. The stickers are represented by the numbers on each dot. The black circles mean a state. Then, the decision points are labeled as a, b, c, d and the goal as e. The ***ra*** follows a direction in which it will find the sticker always on its right side. The arrow indicates the direction where the ***ra*** is going. The behavior of the ***ra*** is addressed by the sticker according to the direction matrix.

The direction matrix indicates the location of the robot and the directions it can take.

To fill the matrix first the robot has to explore the environment.

The matrix represents all possible directions of each sticker and the current state where the ***ra*** is. The -1 value in a certain field means that there is no line to follow in that direction in the environment. The letters represent the next state, where the ***ra*** is going. e.g., situated on the sticker 0 the possible options are C (going left) and B (going forward).

The italic and bold letters in the matrix are special situations. This means that the ***ra*** found a dead end and it has to go back to the state where it was.

Table 2. Direction Matrix

ID	Left	Right	Forward	Back	Current State
0	C	-1	B	***A***	A
1	B	***A***	-1	C	A
2	-1	C	***A***	B	A
3	-1	B	D	A	C
4	A	D	-1	B	C
5	B	D	A	D	C
6	D	***B***	A	C	B
7	***B***	D	C	A	B
8	A	C	***B***	D	B
9	C	-1	E	B	D
10	E	B	-1	C	D
11	-1	C	B	E	D
12	-1	-1	-1	D	E
13	C	A	D	***B***	B

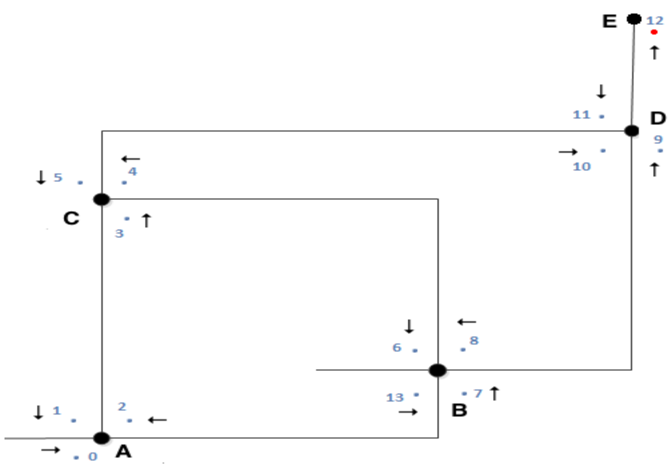


Figure 4. Environment.

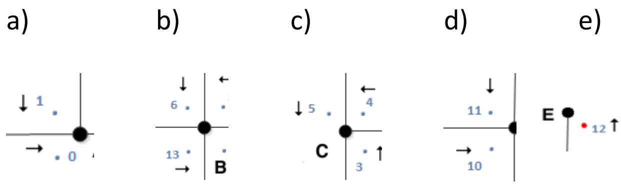


Figure 5. Environment Intersections.

4. Interpreting the Environment

Once the environment is fully known, the next task is to interpret it. The interpretation is done by using the direction matrix. The aim of the interpretation is to build the matrix R, where it reflects the relation between nodes and edges.

The matrix R is a quadratic matrix built by taking into account the number of different states that appear in the direction matrix. The next thing to do is to fill R with the possible actions in each state. The first step in the filling process is to select a state, that becomes the current state in the direction matrix, and to obtain all possible options to take, e.g., located in current state A, the possible options to take are C, B and A, and so on for all the remaining states. In the intersections of the current state and its possible options to go will be a value of 0, e.g. the cell {A,C},{A,B} and {A,A} will have a value of 0. Once done the previous task for all the states, the empty cells are filled with a value of -1. These cells are not considered in the ***lm***. The goal state is identified when there are no possible options to take except to go back, and it takes a value greater than 0, e.g. {D,E} will have a maximum reward value. R is used to construct the knowledge of the ***ra***. To build R the process is depicted as:

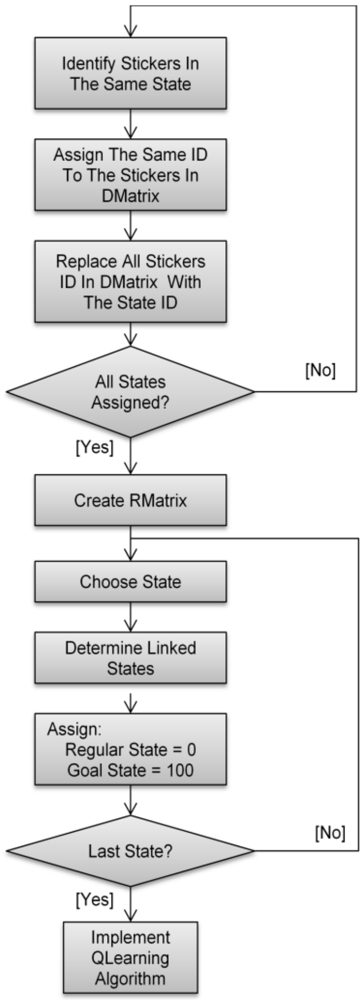


Figure 6. Creating R Matrix.

Once finished the filling process of the R matrix, an exact representation of the environment is built. This representation is shown in the following table.

Table 3. R Matrix

Actions / States	A	B	C	D	E
A	0	0	0	-1	-1
B	0	0	0	0	-1
C	0	0	-1	0	-1
D	-1	0	0	-1	100
E	-1	-1	-1	0	100

5. Learning from the environment

The technique proposed it's a mixed technique considering both line following robot strategy and Q-learning. This technique allows us to have a more efficient navigation ***ra***. The aim of the line follower ***ra*** is to arrive to the end in the best time possible, but this is not considered in our proposal. We consider more relevant the action of going ahead and following the line while an intersection is not found. When the ***ra*** is situated in a dead end it turns 180° to return to the previous state. This is an embedded decision taken by the ***ra***. The ***ra*** reads an input signal from the environment that it has to interpret. Once the combination is interpreted, the ***ra*** has to evaluate it to make a decision. The decision is taken according to the signal gotten from the environment that indicates the combination where the ***ra*** is positioned. The ***ra*** executes the decision taken and keeps following the line until a different combination is found.

When the ***ra*** gets a signal that indicates an intersection, it has to receive a decision from the ***da***. In figure 3 the different intersections are shown. The signal of the environment indicates the different directions the ***ra*** can take. The communication between the ***ra*** and the ***da*** is essential in this part of the process. The ***da*** has to send the decision to the ***ra*** when it gets to an intersection. The signal of the environment indicates where the ***da*** is located so it can make a decision, which is taken based on the Q-learning algorithm and the decision matrix. The ***da*** sends the taken decision to the ***ra***, and it evaluates it to move.

6. Conclusions

The aim of the ALDRO method is to communicate two agents to improve the decision making process for a robot. The proposal describes in depth the decision method used by the deliberative agent having as backbone the reinforcement learning method and the reactive agent which has an embedded decision process to accurately follow the line. The implemented architecture includes an Arduino Uno board for the deliberative agent and a PIC16F8877 microprocessor for the reactive agent. A communication protocol was built and successfully implemented. Our method explores and analyzes the environment to build knowledge about it and find the goal, and then the deliberative agent determines the best option to get there. The combination of the Q-learning algorithm with the line follower improves the robot's movement through the environment. The tests results were the expected; being that the robot obtained the knowledge about the whole environment and arrived to the goal through the best possible option. This was possible because of the good synchronization and communication between the two agents.

References

[1] Nelson Muñoz and Carlos Andrade, "Diseño de un Robot Móvil Orientado a la Enseñanza y la Investigación", Ingeniería y Desarrollo, Universidad del Norte, vol. 19, pp. 114-127, 2006.

[2] Richard S. Sutton and Andrew G. Barto, "Reinforcement Learning: An Introduction", A Bradford Book, 322 p., 1998.

[3] William D. Smart and Leslie Pack Kaelbling, “Effective Reinforcement Learning for Mobile Robots”, ICRA, pp. 3404-3410, 2002.

[4] Chris Gaskett, “Q-Learning for Robot Control”, Ph.D. thesis, Doctor of Philosophy, The Australian National University, 2002.

[5] Andhra Pradesh, “Design and Implementation of Double Line Follower Robot”, International Journal of Engineering Science and Thecnology, vol. 3, no. 6, pp. 47-53, 2011.

[6] Juing-Huei Su, Chyi-Shyong Lee, Hsin-Hsiung Huang, Sheng-Hsiung Chuang & Chih-Yuan Lin, "An intelligent line-following robot project for introductory robot courses", World Transactions on Engineering and Technology Education, vol. 8, no. 4, pp. 455-461, 2010.

[7] Carlos Restrepo Patiño, Jason Milina Varga, Carlos Andrés Torres P. "Algoritmo genético para la ubicación óptima de sensores en un robot seguidor de linea", Scientia et Technica Año XV , no 41, 2009.

[8] Ashish K. Thakre, "Reconfiguration of Mobile Robot", International Journal of Electronics and Computer Science Engineering, vol. 1, no. 2 pp. 161-165, 2012.

[9] Mario González, “Inteligencia Artificial en Robots Móviles para Competencia para su Aplicación en el Diseño de un Agente Inteligente en un Robot Seguidor de Trayectoria”, Bachelor thesis, Instituto Politecnico Nacional, 2009.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-122

Generation of instances for the revision of beliefs

Rodolfo AGUIRRE [a,1], Pedro BELLO [a,2], Ana Patricia CERVANTES [a,3] and Meliza CONTRERAS [a,4]

[a] *Benemerita Universidad Autonoma de Puebla, Faculty of Computers Sciences, Av. San Claudio y 14 Sur, Puebla, Mexico*

Abstract. Given a set of agents with valid previous knowledge bases obtained by a generator of beliefs, we wish to know how new knowledge affects each agent. To model the new knowledge, boolean logic is used, expressed by 2CNF clauses, to reduce the complexity. Upon recieving new knowledge, one or more agents may find it inconsistent with their previous knowledge base, so a mechanism is applied which removes knowledge by using a contraction operation,described by the AGM model in order to ensure consistency of the knowledge base . The goal is to determinate if that contradicting knowledge significantly affects the set of beliefs of each agent.
Furthermore, a problem is modeled in which, given a set of agents(production operators) and their knowledge base(skills in the industry), clauses represent new activities are added when a new process is required and the model determines which is the most suitable agent to perform the activities, through from an evaluation mechanism of the inconsistencies that was generated.
We have developed software which allows to support make decisions over the more capable workers to develop an activity based on their previous skills (knowledge).

Keywords. Knowledge Base, AGM model, 2SAT, Consistency,machine learning techniques

Introduction

In recent years, many formalisms have been proposed in the Artificial Intelligence literature to model common-sense reasoning.

So, the revision and transformation of knowledge is widely recognized as a key problem in knowledge representation and reasoning. Reasons for the importance of this topic are the facts that intelligent systems are gradually developed and refined, and that often the environment of an intelligent system is not static but changes over time [4,9].

Belief revision studies reasoning with changing information. Traditionally, belief revision techniques have been expressed using classical logic. Recently, the study of knowledge revision has increased since it can be applied to several areas of knowledge.

[1]E-mail:r.a.g._192@hotmail.com
[2]E-mail: pbello@cs.buap.mx
[3]E-mail: cervantes.patty@gmail.com
[4]E-mail: mcontreras@cs.buap.mx

Belief revision is a system that contains a corpus of beliefs which can be changed to accommodate new knowledge that may be inconsistent with any previous beliefs. Assuming the new belief is correct, as many of the previous ones should be removed as necessary so that the new belief can be incorporated into a consistent corpus. This process of adding beliefs corresponds to a non-monotonic logic [2,8].

The AGM (Alchourrón, Gärdenfors and Makinson) model addresses the problem of belief revision using the tools of mathematical logic [6,13]. These works are considered the foundation for studying the problem of knowledge exchange. According to the AGM framework, knowledge K is represented by propositional logic theories and new information is represented by the same logic formulas.

One way to represent and check the consistency of a knowledge base is modeled by using the 2SAT problem, which has been shown to be solvable in polynomial time [11].So the knowledge of a set of agents is modeled through this formalism for the purposes of evaluating their consistency when new knowledge is added with the aim of determining which agent is the most affected due to contradiction with the previous knowledge .

1. Preliminaries

During the 1970's from artificial intelligence and information technology the concept of "default reasoning" was introduced and defined by Raymond Reiter. This kind of logic sustains that in the absence of any contrary information, it is plausible to conclude X. It is a form of reasoning that takes into account the limitations of the agent and the commonness of things, which is pretty close to the way that everyday reasoning works. Indeed, it is due to this kind of reasoning that we can act in the world.

Well, the notion of plausible or default reasoning led to a vast area now known as non-monotonic logic or common sense, as well as circumscription logic(McCarthy), modal logic (McDermott and Doyle) and autoepistemic logic (Moore and Konolige) [4].

Non-monotonic logic is that form of reasoning under which a conclusion may be recast, retracted or defeated by an increase in information that modifies its premise. For example, the type of inference of everyday life in which people formulate tentative conclusions, reserving the right to withdraw them in light of new information. This logic satisfies the issue considering the defeatable nature of typical inferences of human common sense reasoning. Considering this type of reasoning, a formal and systematic study of cognitive processes that are present in the manipulation of knowledge structures emerges, by which an intelligent agent can draw conclusions in different ways, without having complete information to do so [5].

Before formalizing changes in beliefs, we must consider several issues: every execution of a dynamic model of beliefs must choose a language to represent them. Whatever the chosen language, the question arises of how to represent the corpus (base) of information as well as the operations for the concepts of minimum and maximum length change of the corpus of information. This implies an epistemic theory which considers the changes in knowledge and beliefs of a rational agent. In our case, we use the criteria of rationality to determine the behavior of changes in beliefs; criteria include the minimum change of preexisting beliefs, the primacy of new information and consistency. Thus for belief revision based on the AGM model using these criteria of rationality, three basic operations are used: expansion, contraction and review [6,3].

Expansion is the operation that models the process of adding new knowledge to the corpus. This can be thought of as the expression of the learning process and is symbolized by the $+$ operator, so it is defined as, $F + p = C(F \cup p)$, where F is the knowledge base, p is the new belief and C is the function that check new knowledge base.

Contraction is the operation that causes a new belief to remove part of the corpus of knowledge, meaning that the agent in question must stop having a certain position on this belief. This becomes complicated when there are other beliefs that would need to be abandoned based on the abandonment of the initial belief, so in the end, only the absolutely necessary beliefs would remain. This is symbolized by the operator $-$ and is defined as $F - p = C(F - p)$ where F is the corpus, the new belief p and C is the function that check new knowledge base.

Revision consists of modifying the set of beliefs when a new belief is incorporated into the previous set so that logical consistency is conserved. If the set of beliefs is already consistent with the new information, then the review coincides with expansion, but if new knowledge is inconsistent with any previous beliefs, the operation of review must determine the resulting set of beliefs which keeps only the part of the original which would obtain a consistent result, so the original set of beliefs must be modified by eliminating as many beliefs as necessary to ensure that the resulting set, which includes the new belief, is consistent, and is defined as $F * p$ where F is the set of beliefs or knowledge base and p is the new belief.

To address the problems of belief revision, it is useful to consider the model using propositional logic to verify the consistency of the knowledge base in order to analyze results from adding new beliefs which are considered valid, so it is necessary to define the concepts of propositional logic involved as follows: a formula is said to be in conjunctive normal form (CNF) if it is composed of a conjunction of disjunctive clauses and will be true if all its clauses are [1,7].

A clause is a disjunction of literals, so that each literal stands for any formula composed of a single proposition symbol x (positive literal) or its negation $-x$ (negative literal) or a constant $\bot$ o $\top$.

So any formula F can be translated into an implication digraph (EF), which is a directed graph whose construction is done by taking each of the clauses (x_i, x_j) of the formula, where vertices of the graph are the x_i and $-x_i$. Here, there is a vertex for each variable and another for its negation. For each clause, two edges are generated by applying the following formula: $(-x_i, x_j)$ and $(-x_j, x_i)$. The implication digraph is widely used to ensure if a formula is satisfiable or not [12].

The Satisfiability Problem(SAT) is posed as follows: given a set of variables and a constraint in conjunctive normal form, a truth assignment that satisfies the constraint must be found. In our case, we worked on CNF for 2SAT problem, which means the formula consists of clauses consisting of two literals [12].

To solve the 2SAT, the implication digraph is built and the strongly connected components of the digraph are calculated. It is said that the problem is solvable if and only if no variable and its negation belong to the same strongly connected component. There is a theorem that supports this formalism [10]: F is unsatisfiable if and only if a variable x exists such that there exist trajectories x a $-x$ and $-x$ to x in EF.

2. Generator of instances

In artificial intelligence, there are several problems where an initial knowledge base is considered. That is the case in belief revision, which can be considered a propositional theory.

In this case, a problem is modeled to determinate the consistency of the knowledge of a group of agents whose initial knowledge base is made up of the same variables (same context). Each agent learns within the same context, so after adding a set of clauses, the affect of this new knowledge on each agent is determined.

It is therefore necessary to apply operators of expansion and contraction of knowledge to evaluate the consistency of the resulting knowledge base according to the following strategy:

2.1. Knowledge Evaluation Strategy

First of all, it is necessary to model the knowledge base by 2-CNF to prove the validity of the knowledge of each agent. Afterwards, inconsistencies are searched for in each knowledge base. In order for new learning to take place, the base knowledge must be satisfiable, meaning that at least the positive or negative literal of each variable must be consistent.

So, the same new knowledge is gradually added to each agent's knowledge base, meaning that all agents must adjust to the new knowledge based on their previous beliefs. Each new clause represents a new belief, so each agent must analyze if the new knowledge affects their knowledge base or not. If an agent's knowledge base is not affected, it will increase, with the new knowledge, representing the assimilation of the new knowledge without contradictions.

Otherwise, if an agent's knowledge base is unsatisfiable, then that agent will have to remove the previous knowledge that generates contradictions with the new knowledge. In this case, an exhaustive search will be done to remove an indefinite number of clauses.

This process will execute as many times as necessary until the knowledge base of each agent is not fed.

2.2. Evaluation Strategy of Generator

Input:

- A set of n agents with m clauses (x_i, x_j) that make up the satisfiable knowledge base F_i, where $i = 1 \dots n$.
- The new knowledge C_j, with $j = 1 \dots T$, where T is the number of new beliefs to be added.

For each new belief C_j For each base F_i

1. Obtain the extended formula EF_i using equation (1) below:

$$EF = \{(-x_1 \vee x_2) \wedge (-x_2 \vee x_1) \wedge (-x_i \vee x_j) \\ \wedge (-x_j \vee x_i) \dots (-x_m \vee x_n) \wedge (-x_n \vee x_m)\} \tag{1}$$

2. Create the linked list L to store the implication graph of EF_i.

3. Calculate the consistency sets TX for each literal.
   $TX[x_i] = x_i, L[x_i] \cup L[L[x_i]]$ for each $L[x_i]$ that does not belong to the set. x_i is said to be inconsistent if in all of $TX[x_i]$ there is both a variable x_j and its negation $-xj$.
4. Verify the consistency of the knowledge base F_i. If in the calculation of the set TX, some x_i, $TX[x_i]$ is inconsistent and $T[-x_i]$ is also inconsistent, then the base F_i is unsatisfiable. Otherwise, the base F_i is satisfiable.
   (a) If the base F_i is unsatisfiable, we evaluate the new knowledge base $F_i * C_j$. If the result is unsatisfiable, then we apply the contraction process $F_i - C_j$ on the knowledge base.

Table 1. Initial knowledge base of the five agents

A_1	A_2	A_3	A_4	A_5
(x_1, x_2)	$(-x_1, -x_2)$	$(-x_1, x_2)$	$(-x_1, -x_4)$	$(-x_1, -x_2)$
(x_1, x_3)	$(x_1, -x_3)$	$(x_1, -x_3)$	$(-x_4, -x_3)$	$(-x_1, -x_3)$
(x_1, x_4)	$(-x_1, -x_4)$	(x_1, x_4)	$(-x_1, -x_2)$	$(-x_1, -x_4)$
(x_2, x_3)	$(-x_2, x_3)$	$(-x_2, x_3)$	(x_4, x_1)	$(-x_2, -x_3)$
(x_2, x_4)	(x_2, x_4)	(x_2, x_4)	$(-x_4, x_1)$	$(-x_2, -x_4)$
(x_3, x_4)	$(-x_3, -x_4)$	$(x_3, -x_4)$	(x_1, x_2)	$(-x_3, -x_4)$

Table 2. Number of inconsistencies when new knowledge is added

Clauses	A_1	A_2	A_3	A_4	A_5
$(x_1, -x_2)$	1	4	3	3	1
$(-x_3, x_4)$	2	4	4	4	2
$(x_2, -x_4)$	3	8	4	4	3
$(-x_1, x_4)$	3	4	4	7	4
$(x_2, -x_3)$	3	4	4	4	4
$(-x_1, x_3)$	4	4	4	8	4
$(-x_2, x_4)$	4	4	4	4	4
Total Incons	20	32	27	34	22

2.3. Using the model in Intelligent Factory Automation

The Strategy in the previous section described was applied to a group of five agents that represents a group of operators of industrial processes with an initial consistent knowledge base, whose clauses appears in table 1, these clauses represent the skills that have developed over time operators. When the implantation of a new process (group activities) required, it is necessary to determine which of the agents is the most suitable to adapt.

Table 2 shows the number of inconsistencies generated when new process is added to each agent, as well as the sum of all inconsistencies. When the process finishes the knowledge base so the total number of clauses is shown in Fig. 1 and the final knowledge base of agents is showed in table 3.

To guarantee the consistency of the knowledge base of each agent, a certain number of inconsistencies was obtained, as show in Fig. 2, which depicts the increase of inconsistencies with respect to the new knowledge added.

Figure 1. Final knowledge base for each agent

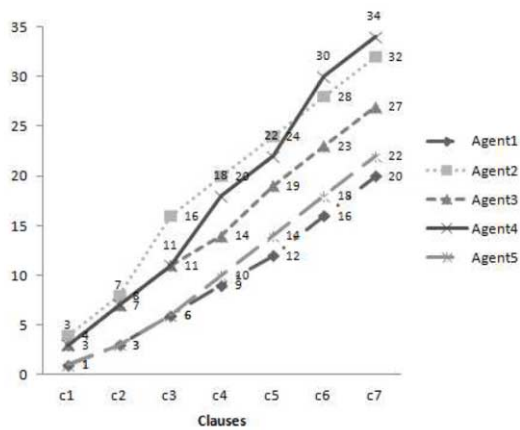


Figure 2. Cumulative frequency of inconsistencies

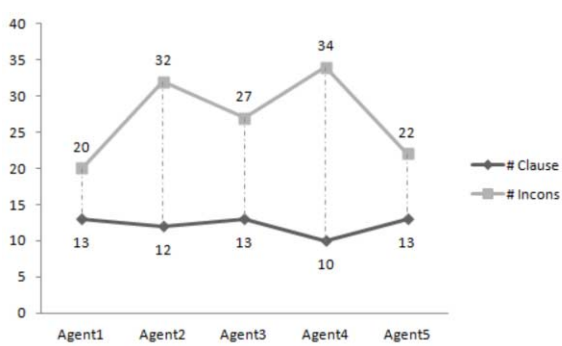


Figure 3. Clauses vs inconsistencies

Table 3. Final knowledge base of the five agents

A_1	A_2	A_3	A_4	A_5
(x_1, x_2)	$(-x_1, -x2)$	$(-x_1, x_2)$	(x_4, x_1)	$(-x_1, -x_2)$
(x_1, x_3)	$(x_1, -x_3)$	$(x_1, -x_3)$	$(-x_4, x_1)$	$(-x_1, -x_3)$
(x_1, x_4)	$(-x_1, -x_4)$	(x_1, x_4)	(x_1, x_2)	$(-x_1, -x_4)$
(x_2, x_3)	$(-x_2, x_3)$	$(-x_2, x_3)$	$(x_1, -x_2)$	$(-x_2, -x_3)$
(x_2, x_4)	$(-x_3, -x_4)$	(x_2, x_4)	$(-x_3, x_4)$	$(-x_2, -x_4)$
(x_3, x_4)	$(x_1, -x_2)$	$(x_3, -x_4)$	$(x_2, -x_4)$	$(-x_3, -x_4)$
$(x_1 - x_2)$	$(-x_3, x_4)$	$(x_1, -x_2)$	$(-x_1, x_4)$	$(x_1, -x_2)$
$(-x_3, x_4)$	$(x_2, -x_4)$	$(-x_3, x_4)$	$(x_2, -x_3)$	$(-x_3, x_4)$
$(x_2, -x_4)$	$(-x_1, x_4)$	$(x_2, -x_4)$	$(-x_1, x_3)$	$(x_2, -x_4)$
$(-x_1, x_4)$	$(x_2, -x_3)$	$(-x_1, x_4)$	$(-x_2, x_4)$	$(-x_1, x_4)$
$(x_2, -x_3)$	$(-x_1, x_3)$	$(x_2, -x_3)$		$(x_2, -x_3)$
$(-x_1, x_3)$	$(-x2, x4)$	$(-x_1, x_3)$		$(-x_1, x_3)$
$(-x_2, x_4)$		$(-x_2, x_4)$		$(-x_2, x_4)$

Finally, Fig. 3 shows the relationship between the clauses that make the new knowledge bases of the agents with respect to the inconsistencies generated in them.

From the above results it was determined that the agents A_1, A_3 and A_5 are the most feasible to adapt to new processes, but these candidates, the candidate A_1 is ideal because the number of inconsistencies generated by adding new processes is the smaller, so these results can be considered as a suggestion to staff responsible for the performance of workers.

3. Conclusion

A strategy was developed to evaluate the knowledge bases of a group of agents, with the objective of guaranteeing their consistency, considering the same context.

This process resulted in determining a relationship between the number of inconsistencies and the number of beliefs that were preserved in the knowledge base, that is, the agents with fewer clauses have a greater number of inconsistencies. Initial knowledge bases are useful for modelling problems involving behavior analysis of a population and guaranteeing knowledge consistency.

The final knowledge base also shows that agents with an initial monotonous knowledge base or with few contradictions do not have big changes, whereas the knowledge base of agents with a great number of contradictions tends to become reduced.

We have a simple method based on the elimination of the clause that generates the fewest inconsistencies by adding new knowledge p, this thanks to the calculation of set TX and the implication generated by the implication graph.

This type of strategy can be applied in situations where the agents involved share the same context but do not interact with one another directly. Particularly, it can be used to model social problems where the knowledge base represents a set of conflicts that an individual has, and therefore applying a new therapy or treatment (new knowledge) implies determining if it will allow for a change in behavior or not. This strategy also serves to identify the conflicting knowledge.

This type of model also has applications in disease diagnosis, in administrative decision-making or consumer habit systems, logical reasoning games and the definition of rules for smart appliances as support for updating their status based on your previous settings.

References

[1] G. De Ita,P. Bello, M. Contreras, M.: New Polynomial Classes for #2SAT Established Via Graph-Topological Structure. Jour. Engineering Letters, 15(2),250-258, 2007.

[2] H. Arló Costa, A. Pedersen: Social norms, rational choice and belief change. In Erik J. Olsson and Sebastian Enqvist, editors, Belief Revision meets Philosophy of Science, volume 21 of Logic, Epistemology, and the Unity of Science. Springer,2010.

[3] A. Pedersen, H. Arló-Costa :Belief Revision. In L. Horsten and R. Pettigrew, editors, Continuum Companion to Philosophical Logic. Continuum Press, 2011.

[4] A. Baltag,S. Smets: Belief revision as a truth-tracking process. TARK XIII Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge,ACM, New York,2011.

[5] A. Baltag,S. Smets: Learning by Questions and Answers: From Belief- Revision Cycles to Doxastic Fixed Points,LNAI Lecture Notes in Computer Science.5514, to 139 (2009).

[6] E. Ferme : Revisión de Creencias, Revista Iberoamericana de Inteligencia Artificial. 11, 17 to 39 (2007).

[7] C. Guillen,C., A. López, A., G. De Ita: Diseño de Algoritmos Combinatorios para # SAT y su Aplicación al Razonamiento Proposicional. Technical report, 2005.

[8] H. Katsuno, A.O. Mendelzon : On the Difference between Updating a Knowledge Base and Revising it. In Proceeding KB-91,387-395(1991).

[9] F. Liu: Reasoning about Preference Dynamics, Springer-Verlag, Series: Synthese Library, Vol. 354,2011

[10] C. Papadimitriou : Computacional Complexity Addison Wesley Publishing Company San Francisco,1994.

[11] D. Roth , On the hardness of approximate reasoning.Artificial Intelligence. 82, 273-302,1996.

[12] M. Wahlström: A Tighter Bound for Counting Max-Weight Solutions to 2SAT Instances. LNCS, vol 5018, 202-213 Springer-Verlag,2008.

[13] Y. Shoham, K. Leyton-Brown:Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations, Cambridge University Press.2009

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-130

A Comparative Study of Intelligent Bio-inspired Algorithms Applied to Minimizing Cyclic Instability in Intelligent Environments

Leoncio ROMERO [a,1], Victor ZAMUDIO [a,2], Rosario BALTAZAR [a],
Marco SOTELO [a], Carlos LINO [a], Efren MEZURA [b] and Vic CALLAGHAN [c]

[a] *Division of Research and Postgraduate Studies, Leon Institute of Technology, Leon, Guanajuato, Mexico*

[b] *Laboratorio Nacional de Informatica Avanzada, Xalapa, Veracruz, Mexico*

[c] *School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, United Kingdom*

Abstract. Cyclic instability is a problem that, despite being shown to affect intelligent environments and in general any rule-based system, the strategies available to prevent it are still limited and mostly focused on centralized approaches. These approaches are based on topological properties of the ***Interaction Network*** (IN) associated, and locking a set of agents. In this paper we present a comparative study of the performance of different optimization techniques when solving the problem of cyclic instability in synthetic scenarios. Instead of using the Interaction Network of the System (which can be computationally expensive, specially in very dense systems). We introduced the concept of ***Average Change of the System*** (ACS) in order to measure the oscillatory behavior of the system and an optimization strategy. In particular Particle Swarm Optimization (PSO), Micro-Particle Swarm Optimization (μ-PSO), Bee Swarm Optimization (BSO), Artificial Immune Systems (AIS) and Genetic Algorithms (GA) were considered. The results found are very promising, as they can successfully prevent unwanted oscillations. Additionally, some of these strategies could be implemented using parallel and distributed processing, in order to be used in real-time scenarios.

Keywords. Cyclic Instability, Ambient Intelligence, Locking

Introduction

Intelligent Environments are systems that are affected by errors, as any other computer system. Among the problems affecting rule-based multiagent ambient intelligence sce-

[1]Corresponding Author: Leoncio Romero, Avenida Tecnologico S/N Fracc. Industrial Julian de Obregon, Leon, Gto., Mexico; E-mail: leoncior@acm.org

[2]Corresponding Author: Victor Zamudio, Avenida Tecnologico S/N Fracc. Industrial Julian de Obregon, Leon, Gto., Mexico; E-mail: vic.zamudio@ieee.org

narios we find cyclical instability [1]. This behavior is characterized by the presence of unexpected fluctuations caused by the interaction of the rules governing the different agents [1].

In the literature there are several approaches to the problem of cyclic instability [2], [1], showing good results. However, these approaches require a large number of calculations, therefore the possibility to be applied to real-time scenarios are very limited. Bio-inspired algorithms have been successfully applied to the problem of cyclic instability, using the ***Game of Life as a scenario*** [3].

In this paper we present two functions to measure the oscillatory behavior of the system. Several optimization techniques were applied to these functions, in order to compare their performance preventing oscillatory behavior. Among the optimization techniques used we can mention *Particle Swarm Optimization* (PSO), *Micro Particle Swarm Optimization* (μ-PSO), *Bee Swarm Optimization* (BSO), *Immune Artificial Systems* (AIS), and *Genetic Algorithms* (GA). These algorithms were applied to well known scenarios where the strategy INPRES has successfully prevented oscillatory behavior [1]. In our experiments the number of locked agents was used to measure the performance of the algorithms applied. The strategy with the minimum number of agents locked was considered the best.

1. Cyclic Instability in Intelligent Environments

Rule-based systems play an important role in Ambient Intelligence. In particular, multi-agent systems can provide different functionalities to the final user, generating complex interactions between a large number of connections in the system. Under some conditions (in particular the presence of feedback in the rules), unwanted oscillations of the agents can affect the expected performance of the system. These changes over time can even cause interference with other devices or unwanted behavior [1,4–6].

The state of the system $s\left(t\right)$ (see eq. 1 is defined as the logarithm base 10 of the decimal representation of the binary state of the agents involved. An agent can have two states on (1) and off (0).

$$s\left(t\right)=\log\left(s\right) \tag{1}$$

where:
$s\left(t\right)$ is the state of the system at time t.
s is base-10 representation of the binary vector of the agents.

2. Measuring the cyclical instability

In order to measure the oscillatory behaviour of the system, two functions are used: Average Cumulative Oscillation (ACO) and Average Change of the System (ACS).

Average Cumulative Oscillation (ACO) [3] measure the difference between $s\left(t\right)$ and $s\left(t+1\right)$ on the assumption that if a system is unstable difference will always exist among the states which will cause the value of the ACO will increase as increases the number of generations of the system, however if the system is stable this value will tend to 0.

$$o = \frac{\sum_{t=1}^{n-1} |S(t) - S(t+1)|}{n-1} \tag{2}$$

where:
o: average cumulative oscillation
n : number of generations (total time of testing)
$S(t)$: state of system at time t
$S(t+1)$: tate of system at time $t+1$

In this paper we introduce another function, called Average Change of the System (ACS). This equation means that an unstable system will remain constantly changing ie. the state $s(t)$) is always different from the state $s(t+1)$ this implies that if a system is unstable the value of ACS obtained is close to 1, while if the system is stable this value is approximated to 0.

$$p = \frac{\sum_{i=1}^{n-1} x_i}{n-1} \tag{3}$$

where:
p: average change in system
n : number of generations of scenario to test (total time of test)
$x_i = \begin{cases} 1\ si\ S(t) \neq S(t+1) \\ 0\ in\ other\ case \end{cases}$

with $S(t)$ being the state of the system in time t and $S(t+1)$ being the state of system in time $t+1$

In the case of a stable system, these two equations will show a flat line. Due to the previous, it is possible to use them as objective functions in a minimization algorithm.

3. Optimization Algorithms

3.1. Particle Swarm Optimization

Particle Swarm Optimization (PSO) [7, 8] algorithm was proposed by Kennedy and Everhart. It is based on the choreography of a flock of birds [7–12]. The basic PSO algorithm [8] uses two equations. The first one 4 finds the velocity, describes the size and direction of the step that will be taken by the particles and is based on the knowledge achieved until that moment.

$$v_i = wv_i + c_1 r_1 (lBest_i - x_i) + c_2 r_2 (gBest - x_i) \tag{4}$$

where:
v_i is the velocity of the i-th particle.

$i = 1, 2 \ldots, N$ and N is the number of the population.
w is the environment adjustment factor
c_1 is the memory factor of neighborhood
c_2 is memory factor
r_1 and r_2 are random numbers in range $[0, 1]$
$lBest$ is the best local particle founded for the i-th particle
$gBest$ is the best general particle founded until that moment for all particles
The equation 5 updates the current position of the particle to the new position using the result of the velocity equation.

$$x_i = x_i + v_i \tag{5}$$

where x_i is the position of the i-th particle.

3.2. Binary PSO

Binary PSO [12, 13] was design to work in binary spaces. Binary PSO select the $lBest$ and $gBest$ particles in the same way as PSO. The main difference between binary PSO and normal PSO are the equations that are used to update the particle velocity and position. The equation for updating the velocity is based on probabilities in the range $[0, 1]$. For that mapping is established for all real values of velocity to the range $[0, 1]$. The normalization function 6 used is a sigmoid funcion.

$$v_{ij}(t) = sigmoid(v_{ij}(t)) = \frac{1}{1 + e^{-v_{ij}(t)}} \tag{6}$$

and equation 7 is used to update the new particle position.

$$x_{ij}(t+1) = \begin{cases} 1 \, if \, r_{ij} < sigmoid(v_{ij}(t+1)) \\ 0 \, in \, other \, case \end{cases} \tag{7}$$

where r_{ij} is a random vector with uniform values in the range $[0, 1]$.

3.3. Micro PSO

Micro-PSO (μ-PSO) algorithm [14, 15] is a modification made to the original PSO algorithm in order to work with small populations. PSO and μ-PSO are very similar, but μ-PSO has more exploratory power. In order to avoid local optimum (exploring the configuration space), μ-PSO includes the concepts operators of replacement and mutation [15, 16].

3.4. Bee Swarm Optimization

This algorithm is based on PSO and Bee algorithm [13] and uses a local search step to to improve the performance. This algorithm was proposed by Sotelo [13]. The local search is made around $gBest$ in each iteration of the algorithm.

Another variant of the algorithms consists in applying the local search around $lBest$ after comparing it with a bee.

In the future we will refer to BSO algorithm with the $gBest$ enchancer as BSO1 while BSO algorithm with the $lBest$ enhancer will be known as BSO2.

3.5. Artificial Immune System

The Artificial Immune System (AIS) [17] is a metaheuristic based on the Immune System behavior of living things [13], particularly of mammals [17].

One of the main functions of the immune system is to keep the body healthy. A variety of microorganisms (called pathogens) could invade the body, which could be harmful. Antigens are molecules that are expressed on the surface of pathogens that can be recognized by the immune system and are also able to initiate the immune response to eliminate them [17].

Artificial immune systems have various types of models, in this work we use the one that implements the clonal selection algorithm that emulates the process by which the immune system in the presence of a specific antigen, stimulates only those lymphocytes that are more similar, then they are cloned and mutated [17].

3.6. Genetic Algorithm

Genetic algorithms (GAs) [18] proposed by John Hollan based on the theory of evolution by Darwin [18–20]. This technique is based on the selection mechanisms that nature uses, according to which the fittest individuals in a population are those who survive, to adapt more easily to changes in their environment.

A fairly comprehensive definition of a genetic algorithm is proposed by John Koza [21]:

"It is a highly parallel mathematical algorithm that transforms a set of individual mathematical objects with respect to time using operations patterned according to the Darwinian principle of reproduction and survival of the fittest and after naturally have arisen from a series of genetic operations from which highlights the sexual recombination. Each of these mathematical objects is usually a string of characters (letters or numbers) of fixed length that fits the model of chains of chromosomes and is associated with a certain mathematical function that reflects their ability".

The GA seeks solutions in the space of a function through simple evolution. In general, the individual fitness of a population tends to reproduce and survive to the next generation, thus improving the next generation. Either way, inferior individuals can, with a certain probability, survive and reproduce.

3.6.1. Clones and Scouts

In order to increase the performance of the GAs the concept of clones and explorers is considered [22]. A clone is an individual whose fitness is equal to the best individual fitness. When it reaches a certain percentage of clones a percentage of the worst individuals in the population is then mutated. Mutated individuals are named scouts. The application of clones and explorers is in addition to the mutation carried out by the GA generation.

4. Experimental Results

In order to test the performance of the previous algorithms 5 diferent well known topologies were used: 1) iDorm [1], 2) Strong coupling [6], 3) Week coupling [6] 4) non-coupled cycles [1], and 5) cycles coupled in two points [1]. In iDorm topology the maximum allowed percentage of locked agents was 30% while in other was 20%.

If a solution generated by any of the algorithms was good with respect to the metric used but the percentage of locked agents exceeded the maximum allowable then the solution is penalized by increasing the value of the metric obtained by a constant value.

In our experiments we set as a parameter, 3000 functions calls as a measure of success of the algorithms i.e. the system has 3000 opportunities to find a better solution. If after 3000 functions calls a better solution is not found, the system has failed.

The best solution not only minimizes the value of the metric used but also minimizes the number of locked agents. In the experiments the percentage of agents that can be blocked is set also as a parameter. This is important because if this percentage grows the system can be disabled.

Interaction Networks test instances [1] are showed on Table 1.

Table 1. Benchmark

Instance	# of Agents	ACO	ACS
1 (iDorm)	4	0.07183234371149662	1.0
2 (Strong coupling)	7	0.26938367763858284	1.0
3 (Week coupling)	7	0.1362409906735542	1.0
4 (Non-coupled)	64	0.35847554949972144	1.0
5 (Coupled in 2 points)	64	1.1840427653926249	1.0

A summary of the parameters used in our experiments is shown in Table 2.

Based on the previous parameters Tables 3 and 4 shows the results obtained in the case of the function ACO, for each algorithm.

Table 5 shows the results obtained when using the function ACS.

All algorithms were tested using well know topologies and rules, and where the strategy INPRES had been previously successfully applied [1, 6]. In order to have a fair comparison, we considered only the number of agents locked (as is the only parameter considered by INPRES).

Table 6 shows the number of locked agents, when the ACO is applied (see Tables 3 and 4.

Table 7 shows the number of locked agents corresponding to the results obtained with the ACS (shown in Table 5).

Table 2. Algorithms Parameters

Algorithm	Parameter	Value
PSO BSO	Particles	45
	w	1
	c_1	0.3
	c_2	0.7
μ-PSO	Particles	6
	w	1
	c_1	0.3
	c_2	0.7
	Replacement generation	100
	Number of restart particles	2
	Mutation Rate	0.1
AIS	Antibodies	45
	Antibodies to select	20
	New Antibodies	20
	Beta Factor	2
GA	Chromosomes	30
	Mutation percentage	0.15
	Elitism	0.2
	Clones percentage	0.3
	Scouts percentage	0.8

Table 3. ACO Results (A)

Instance	Average Cumulative Oscillation		
	PSO	**BSO1**	**BSO2**
1 (iDorm)	8.628730269198046E-4	0.0	0.0
2 (Strong coupling)	0.0032242818868545857	0.0	0.0
3 (Week coupling)	0.003153206791536531	0.0	0.0
4 (Non-coupled)	0.002508031766729257	3.8389318144437694E-4	7.7867850781205E-9
5 (Coupled in 2 points)	3.022378356262937E-4	5.27630744376312E-15	0.0

In order to show how the oscillations are succesfully removed, in Figures 1 to 5 the evolution of the system is shown. In figure 1a the oscillatory behavior of instance 1 (iDorm) is showed and in figure 1b the instabilities are successfully removed. For the instance 2 (Strong coupling) the evolution of the system is shown in figure 2. For the instance 3 (Weak coupling) the behavior is whown Figure 3. In figure 4a the oscillatory behavior of the instance 4 (Non-coupled) is showed, and behavior without oscillation is showed in figure 4b. The oscillatory behavior of the instance 5 (Coupled in 2 points) is showed in figure 5a, and behavior without oscillation is showed in figure 5b.

Table 8 shows the results obtained by the algorithms considered in this paper. In order to determine whether an algorithm outperforms another one, the Wilcoxon test was applied for both the number of agents locked and the oscillatory functions ACO and ACS.

Table 4. ACO Results (B)

Instancia	**Average Cumulative Oscillation**		
	AIS	**GA**	**μPSO**
1 (iDorm)	0.0	0.0173	0.0
2 (Strong coupling)	0.0	0.0	0.0
3 (Weak coupling)	0.0	0.00253	0.0
4 (Non-coupled)	7.63419134025501E-6	0.0	0.008354842172152571
5 (Coupled in 2 points)	0.0	0.0	3.2806731364235696E-4

Table 5. ACS Results

Instance	**Average Change of the System**					
	PSO	**BSO1**	**BSO2**	**AIS**	**GA**	**μ-PSO**
1 (iDorm)	0.0	0.0	0.0	0.0	0.0	0.0
2 (Strong coupling)	0.02	0.0	0.0	0.0	0.0	0.0
3 (Weak coupling)	0.03	0.0	0.0	0.0	0.0	0.0
4 (Non-coupled)	0.03	0.0	0.0	0.03	0.0	0.03
5 (Coupled in 2 points)	0.07	0.0	0.0	0.0	0.0	0.07

Table 6. ACO Locked Agents

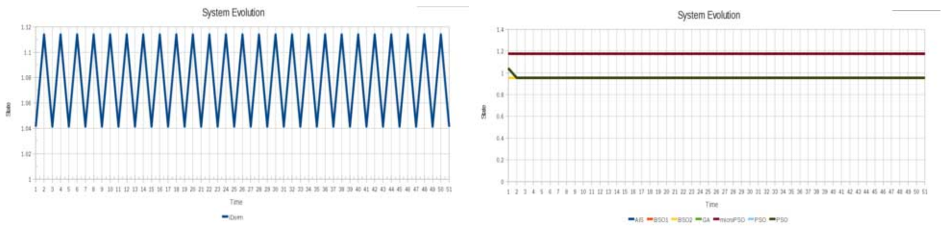
Instance	**Number of Locked Agents**							
	Permited	**INPRESS**	**PSO**	**BSO1**	**BSO2**	**AIS**	**GA**	**μ-PSO**
1 (iDorm)	1	1	1	1	1	1	1	1
2 (Strong coupling)	1	3	1	1	1	1	1	1
3 (Weak coupling)	1	1	1	1	1	1	1	1
4 (Non-coupled)	12	16	6	9	12	6	5	7
5 (Coupled in 2 points)	12	28	7	12	9	9	5	5

Table 7. ACS Locked Agents

Instance	**Number of locked agents**							
	Permited	**INPRESS**	**PSO**	**BSO1**	**BSO2**	**AIS**	**GA**	**μ-PSO**
1 (iDorm)	1	1	1	1	1	1	1	1
2 (Strong coupling)	1	3	1	1	1	1	1	1
3 (Weak coupling)	1	1	1	1	1	1	1	1
4 (Non-coupled)	12	16	9	12	12	9	3	12
5 (Coupled in 2 points)	12	28	5	9	7	6	2	6

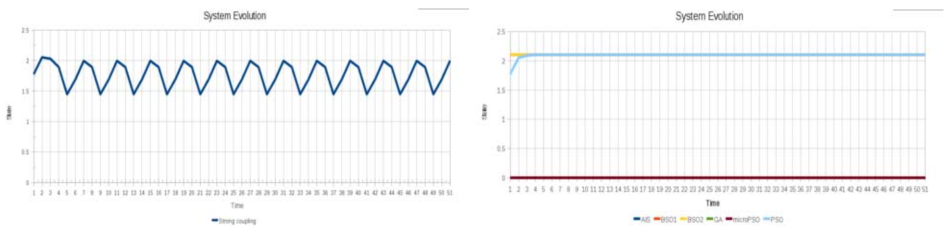
From Table 8 it was found that GAs were able to obtain smaller for ACO and ACS but if we take into account the number of locked agents the algorithms PSO and μ-PSO were the best.

All the algorithms used were able to prevent cyclic behaviour and showed low values (good performance) in the different parameters involved. However the most important parameter is the number of locked agents. In this case, μ-PSO showed the best performance in the experiments performed.



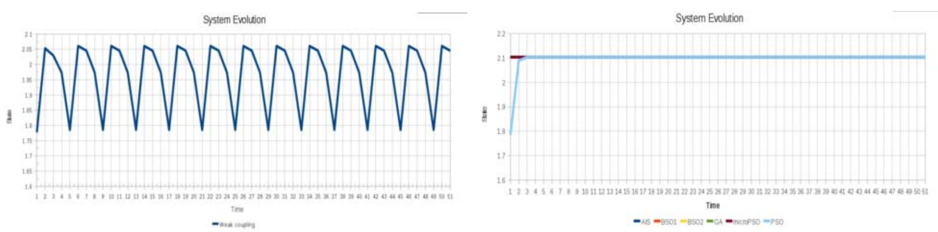
(a) Oscillatory behavior of the system (b) Instabilities are successfully removed

Figure 1. Evolution of the system for the case of 1 (iDorm)



(a) Oscillatory behavior of the system (b) Instabilities are successfully removed

Figure 2. Evolution of the system for the case of 2 (Strong coupling)

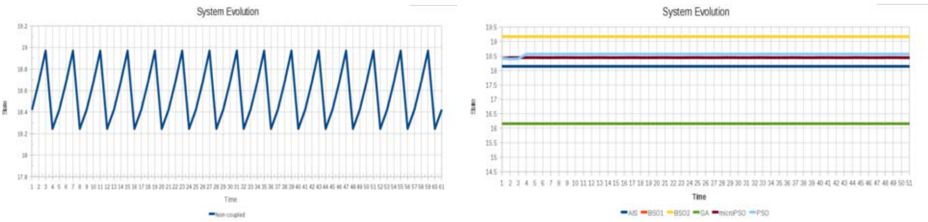


(a) Oscillatory behavior of the system (b) Instabilities are successfully removed

Figure 3. Evolution of the system for the case of 3 (Weak coupling)

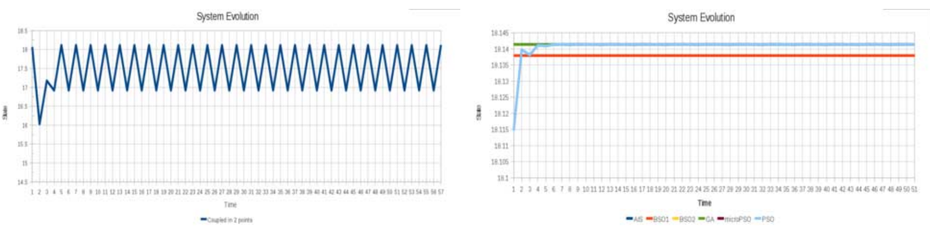
5. Conclusions and Future Work

From our experiments we found that bio-inspired algorithms can successfully minimize the oscillations when applied to different well-known test instances. These results are very encouraging, as bio-inspired algorithms could be applied to real-time scenarios where rules of interactions could be changing on time, and where the agents could be nomadic (with the corresponding changes of the rules). Additionally, this approach prevents having a large number of agents locked, disabling the system. The metrics used in our experiments *Average Cumulative Oscillations* ACO and *Average Change of the System* ACS showed good results. However it was found that for ACO when the relation of change between states is very small the system can have oscillation values close to 0 or significantly below than the original oscillation value and cyclic instability was still



(a) Oscillatory behavior of the system (b) Instabilities are successfully removed

Figure 4. Evolution of the system for the case of 4 (Non-coupled)



(a) Oscillatory behavior of the system (b) Instabilities are successfully removed

Figure 5. Evolution of the system for the case of 5 (Coupled in 2 points)

Table 8. Algorithms Compared Using the Wilcoxon Test

Algorithm	Number of Algorithms (Metric AAO)				Number of Algorithms (Metric ACS)			
	by Value		by # of Agents		by Value		by # of Agents	
	Overcome	Not Overcome	Overcome	Not Overcome	Overcome	Not Overcome	Overcome	Not Overcome
PSO	0	5	4	1	1	4	5	0
BSO (1)	1	4	2	3	3	2	1	4
BSO (2)	2	3	2	3	4	1	2	3
μPSO	4	1	5	0	0	5	4	1
AIS	4	1	3	2	2	3	3	2
GA	5	0	0	5	5	0	0	5

present. In the case of ACS we found that if you have oscillations close to 1 the instability is present, but if the oscillations are equal or close to 0 the system is stable.

From the experiments performed it was found that PSO and μ-PSO are the most likely to be implemented in real-time scenarios. Additionally, a parallel implementation of these algorithms could improve the results obtained in these experiments. More research is needed in these directions, in particular more complex scenarios and dynamic conditions (rules and nomadic agents), and additional optimization techniques. We hope to report our results in future publications.

Acknowledgments

The authors want to thank Jorge Soria for their comments and suggestions to his work. Leoncio Romero acknowledges the support of the National Council for Science and Technology CONACyT. Additionally, Efren Mezura acknowledges the support from CONACyT through project No. 79809.

References

[1] Victor Manuel Zamudio. *Understanding and Preventing Periodic Behavior in Ambient Intelligence*. PhD thesis, University of Essex, October 2008.
[2] Jaafar Gaber. Action selection algorithms for autonomous system in pervasive environment: A computational approach. *ACM Transactions on Autonomous and Adaptative Systems*, 6, February 2011.
[3] Leoncio Alberto Romero, Victor Zamudio, Rosario Baltazar, Marco Sotelo, and Vic Callaghan. A comparison between pso and mimic as strategies for minimizing cyclic instabilities in ambient intelligence. In *5th. International Symposium on Ubiquitous Computing and Ambient Intelligence UCAmI*, 2011.
[4] V. Zamudio and V. Callaghan. Facilitating the ambient intelligent vision: A theorem, representation and solution for instability in rule-based multi-agent systems. *Special Section on Agent Based System Challenges for Ubiquitous and Pervasive Computing. International Transactions on Systems Science and Applications.*, 4(2):108–121, May 2008.
[5] Victor Zamudio and Vic Callaghan. Understanding and avoiding interaction based instability in pervasive computing environments. *International Journal of Pervasive Computing and Communications*, 5:163–186, 2009.
[6] V. Zamudio, R. Baltazar, and M. Casillas. c-INPRES: Coupling analysis towards locking optimization in ambient intelligence. In *The 6th International Conference on Intelligent Environments IE10*, Monash University (Sunway campus), Kuala Lumpur, Malaysia., July 2010.
[7] Anthony Carlise and Gerry Dozier. An off-the-shelf pso. 2001.
[8] Russel C. Eberhart and Yuhui Shi. Particle swarm optimization: Developments, applications and resources. *IEEE*, pages 82–86, 2001.
[9] Carlos A. Coello and Maximo Salazar. Mopso: A proposal for multiple objetive particle swarm optimization. *Evolutionary Computation, IEEE*, pages 1051–1056, 2002.
[10] Swagatam Das, Amit Konar, and Uday K. Chakraborty. Improving particle swarm optimization with differentially perturbed velocity. *GECCO*, pages 177–184, 2005.
[11] K.E. Parsopoulos and M. N. Vrahatis. Initializing the particle swarm optimizer using the nonlinear simplex method. *GECCO*, 2005.
[12] Tamer M. Khali, Hosam K. M. Youssef, and M. M. Abdel Aziz. A binary particle swarm optimization for optimal placement and sising of capacitor banks in radial distribution feeders with distored substation voltages. *AIML 06 International Conference*, Jun 2006.
[13] Marco A. Sotelo. Aplicacion de metaheuristicas en el knapsack problem. Master's thesis, Leon Institute of Technology, Guanajuato, 2010.
[14] Juan C. Fuentes Cabrera and Carlos A. Coello Coello. Handling constraints in particle swarm optimization using a small population size. In *6th. Mexican International Conference on Artificial Intelligence*, Aguascalientes, Mexico, November 2007.
[15] Francisco Viveros Jiménez, Efrén Mezura Montes, and Alexander Gelbukh. Empirical analysis of a micro-evolutionary algorithm for numerical optimization. 2011.
[16] Efrén Mezura Montes, Omar Cetina Domínguez, and Betania Hernández Ocaña. Nuevas heurísticas inspiradas en la naturaleza para optimización numérica. 2010.
[17] Nareli Cruz Cortés. Sistema inmune articial para solucionar problemas de optimización. Octuber 2004.
[18] Jhon Holland. Adaptation in natural and artificial systems. *MIT Press*, 1992.
[19] Christorpher R. Houck, Jeffery A. Joines, and Michael G. Kay. A genetic algorithm for function optimization: A matlab implementation. 1995.
[20] Carlos A. Coello Coello. Introducción a la computación evolutiva. January 2004.
[21] Jhon R. Koza. Genetic programming: On the programming of computers by means of natural selection. *MIT Press*, 1992.

[22] Jorge A. Soria-Alcaraz, J. Martin Carpio-Valadez, and Hugo Terashima-Marin. Academic timetabling desing using hyper-heuristics. *Soft Computing for Intell. Control and Mob. Robot.*, 318:43–56, 2010.

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-142

Optimization of a fuzzy contrast method for a pattern recognition system

Miguel Angel LOPEZ [a], Arnulfo ALANIS [a], Bogart Yahil MARQUEZ[a], Karina ROMERO[a], Maria del Rosario BALTAZAR[b], Victor Manuel ZAMUDIO[b]

[a]*Department of Systems and Computer Tijuana Institute of Technology Tijuana, Baja California, México.*
[b]*Research Division, Technology Institute of Leon, Tijuana BC, México*

Abstract— This article develops a method of fuzzy fingerprint pre-processing to apply in a pattern recognition system, in which a neural network is developed with fuzzy response integration for fingerprint recognition. There are also shown neural network training results using Biometrics Laboratory database at the Bologna University and of using a genetic algorithm, as well as the chromosome used in the genetic algorithm development of membership functions optimization of fingerprint pre-processing fuzzy method.

Keywords. Pattern Recognition System, Hybrid Intelligent System, Biometric Fingerprints.

Introduction

Using a combination of neural networks and biometrics makes it a much more efficient project for daily life, one of the most common and acceptable is the fingerprint.

Fingerprints are only arbitrary shaped rugosity in the skin over the fingertip. They are formed by grooves: mountains and valleys. The "mountains" are called papillary crests and "valleys" interpapillary grooves. The crests are the glands that produce sweat, which contains oil that slides into the grooves, where it is stored. As we touch, oil stored in the grooves passes to the surface. Figure 1 shows an image of a fingerprint.

Figure 1. Fingerprint

Fingerprints are formed in the sixth week of intrauterine life and do not vary in their characteristics over the individual lifetime.

They are unique and irreplaceable even in identical twins because their design is not strictly determined by genetic code, but for small variable in growth factor concentrations and hormones located within the tissues. Note that in the same individual the print of each of your fingers is different.

Fingerprints are taken from index fingers because unlike the thumbs are less prone to injuries that leave scars impairing identification. Biometric identification is the identity verification of a person based on their body characteristics or behavior, for example using our hand, eye iris, voice or face on facial recognition.

Although biometric studies are not perfect, they are a very powerful tool to identify people. Of all existing biometric identification systems, fingerprints are the only legally recognized as reliable proof of identity. This system in addition to being effective, it is easy to apply and authentication is obtained rapidly.

Appling pre-processing to the images the result becomes more effective. In this paper we propose to work with features extraction of the image based on the contrast, in order to obtain a higher recognition rate.

The image processing is given by a set of operations performed on the images to perform quantitative measurements to describe them, that is, to extract features that enhance, improve or detail the image [1].

In this research experiments we used the biometric laboratory database at the Bologna University [1]. The characteristics of the database are:

- Fingerprints were acquired using a low-cost optical sensor.
- Image size is 300 x 300 pixels in .tif format.
- Fingerprints are mainly from students between 20 to 30 years of age (approximately 50% men).
- Four fingers were scanned for each volunteer (the index and middle fingers of both hands).
- The database has 8 samples from 10 people making a total of 80 images [1], Figure 2.

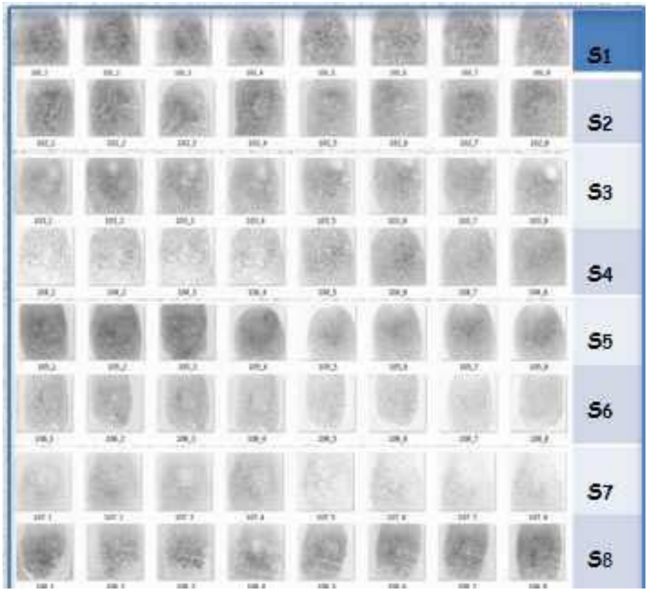


Figure 2. Fingerprint database sample from the laboratory at the Bologna University.

In recent years many researchers have applied the fuzzy set theory to develop new techniques to improve image contrast.

In this work it was developed a system in which can be improve the rate of recognition of individuals through fingerprint biometrics measurement since fingerprint is considered the most efficient, comfortable and easy to apply.

This work was based on the average of fingerprint contrast that we took as reference, fingerprints that have been worked previously and have always been recognized.

The average comes from the 8 samples of an individual, as shown in Figure 3.

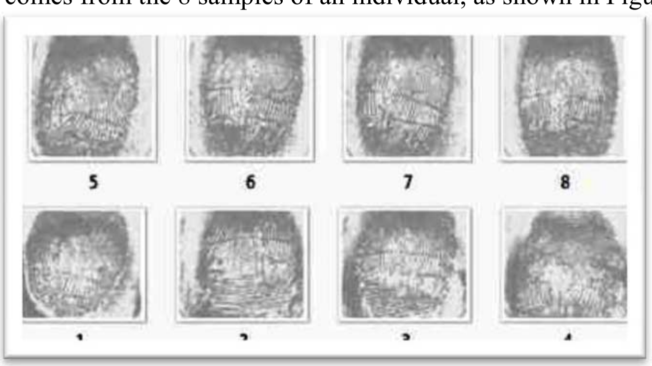


Figure 3. Set of fingerprint images used to get the gray contrast average.

The gray tone range to consider improving the contrast of the image must remain in the range of values of each pixel of the image of the fingerprint of the input variables membership functions of fuzzy inference system not higher than the average contrast in the range of [-5.5], that is, if the value of the pixel in the image of the fingerprint is analyzed within a range of [0 to 255] ±5 values the average value of average contrast of all images used for reference then the fuzzy inference system adjusts it to remain within the average range of contrast otherwise remains with the original pixel value.

Definition of the chromosome was made as real with a length of 18, where the first three bits belong to the first input membership function, the following three belong to the second input membership function, and the next three to the third input membership function, the following nine belong to the output membership functions. As shown in Figure 4 [6.7].

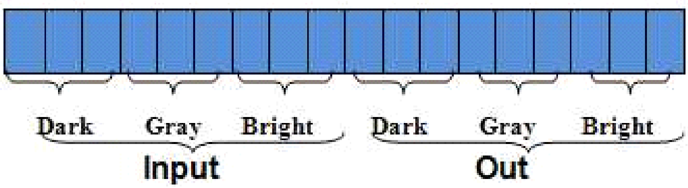


Figure 4. Chromosome definition.

The genetic algorithm process is as follows:

- Start
- Variables definition
- Generate population
- Call image
- Set the image contrast in the range of average established
- Yes- apply cross selection
- No- go to step 4
- Save image, fuzzy inference system (FIS) and runtime.

The inference system of diffuse contrast has the following rules to adjust the contrast of the image [5]:

a). if (input is dark) the (output is bright)
b). if (input is gray) the (output is gray)
c). if (input is bright) the (output is bright)

The picture brightness varies to 0 (black) 255 (white) fuzzy inference system has an input variable and output variable, with three triangular type input membership functions [2,3] as shown in Figure 5.

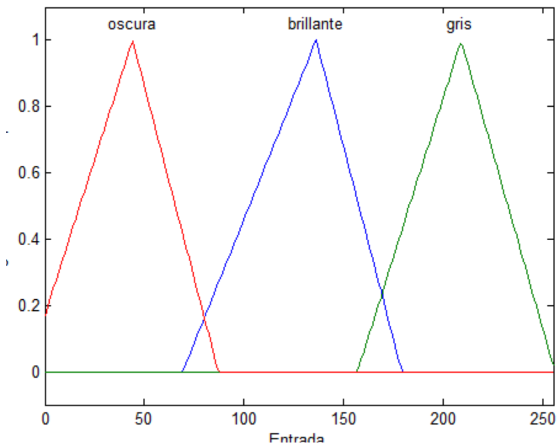


Figure 5. Input membership functions

Output membership functions of the inference system of fuzzy contrast are shown in Figure 6.

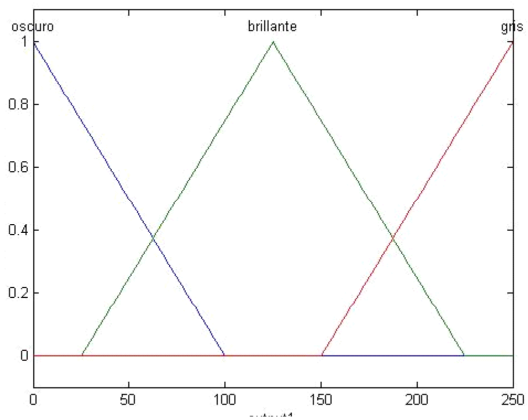


Figure 6. Output membership functions.

The optimization system consists of two parts: a local architecture and an overall architecture, the first can be seen in Figure 7 which consists of checking the contrast of the images to decide whether to modify the image contrast.

- **Local system architecture.**

If the image is within the average contrast range no change is made and then stored in the database, also the fuzzy inference system is stored, FIS, so as the runtime.

If the image is not within the average contrast range the genetic algorithm is responsible of moving the image contrast using the fuzzy contrast inference system's rules until it can adjust the contrast of the image within the average contrast.

The local architecture is to have a prior adjustment of image contrast to make the checking of the results with the pattern recognition system achieve better recognition results for this case study is used fingerprints.

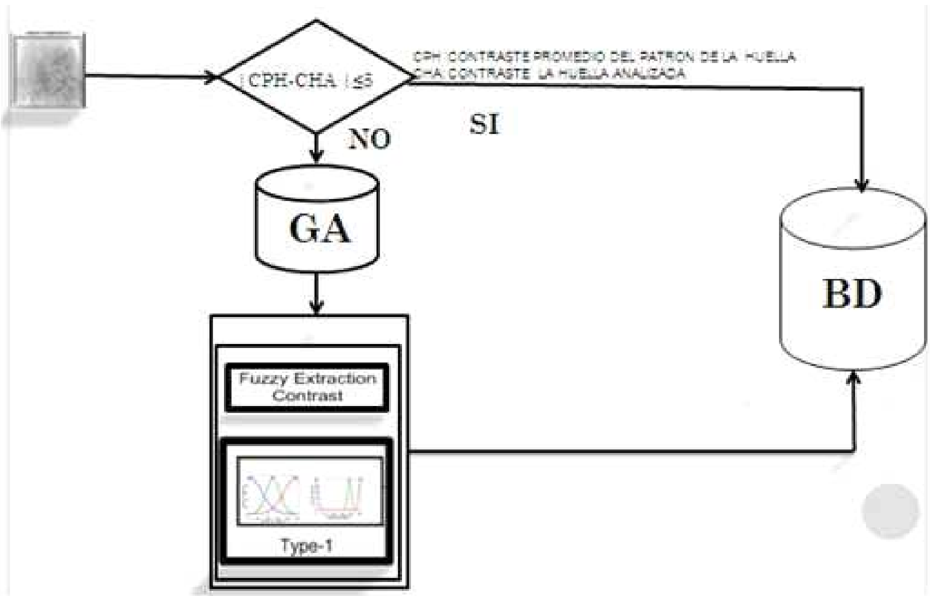


Figure 7. Local system architecture.

- **Global system architecture**

The pattern recognition system smart hybrid in its overall architecture works as follows. The results obtained with the local architecture to get the best contrast of fingerprints are stored in a database and the image obtained fuzzy inference system to achieve the contrast of the image.

By accessing the images obtained from the database are checked against the pattern recognition system, complete in its overall architecture, when the mark is recognized by the system is saved along with its system of fuzzy inference contrast.

In cases where fingerprints are not recognized are wont to utilize a genetic algorithm to move the parameters of fuzzy inference system contrast. To further verify the results of modification of the contrast of the fingerprint to check if it is recognized by the pattern recognition system completely.

The global architecture presents the complete process of the system which consists of obtaining a higher percentage of identification using footprints as a biometric measure. Neural networks were used for identification and genetic algorithm for membership functions optimization in the inference system used for image contrast.

For experiments with the pattern recognition system was used the laboratory's database at the Bologna University, which consists of 8 samples from each of 10 individuals making a total of 80 images, we consider the images contrast of individual 8 for being a group of fingerprint images which have been tested in previous experiments and the pattern recognition system always identified them [2,3].

The fingerprint images go into a modular artificial neural network consisting of three modules of 100 * 100 pixels each, the neural network output of each module enters an responses integrator of a fuzzy inference system and a response is obtain to see if the image is identified, if the image is identified is stored in the identified fingerprints database, otherwise the genetic algorithm is responsible for conducting the work, which will move the fuzzy inference system membership functions until the image is identified to store it, and the process is repeated with each of the fingerprint images used [9,10,11,12].

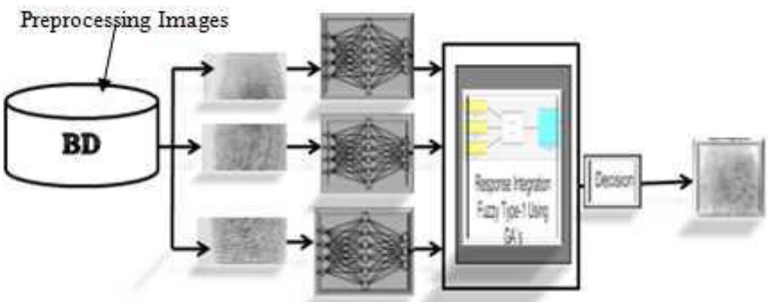


Figure 8. Global system architecture.

- **Results and simulation**

Network is trained with 80% of the remaining images are used to identify, results of neural network training and genetic algorithms application are shown.

Two tests were also conducted to verify the effectiveness of the system, which were:
False acceptance is to use images not used before, for this we use the database that has 100 people taking 8 samples for a total of 800 images, images of 640 x 480 pixels in .tif format. 10 people and their 8 samples were taken and images were normalized to a size of 300 x 300 for the new 80 images data base and for the false rejection test noise was applied

Figure 8 shows the fuzzy contrast inference system simulation applying blur motion noise to the left and blur radial noise to the right at 20 distance pixel level.

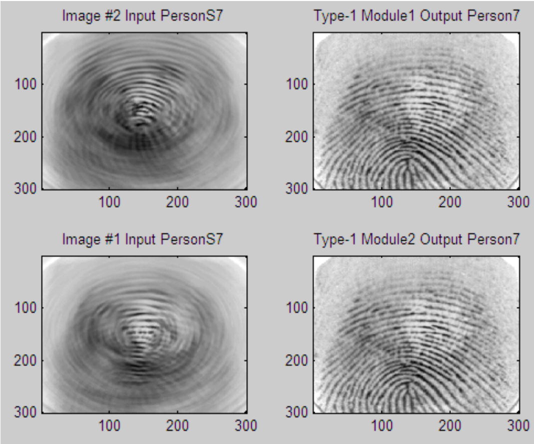


Figure 9. Results of fingertips using motion blur at 20 distance pixel

Table 1 shows a sample of the best results obtained from the artificial neural network without pre-processing of the fingerprint image in which blur motion was applied with a noise level at 10 distance pixel.

Table 1. Blur Motion Noise 10 distance pixel.

Training	Images not identified	Identification %
1	15/60	75%
2	12/60	80%
3	10/60	83.3%
4	18/60	70%
5	16/60	73.3%
6	12/60	80%
7	12/60	80%
8	12/60	80%
9	19/60	68.3%
10	16/60	73.3%

Table 2 shows results of working with pre-processing the image within an average contrast range of [-7,7].

Table 2. Blur Motion with 10 distance pixel.

Training	Established range	Images not identified	Identification %
1	-7 , 7	10/60	83.3%
2	-7, 7	10/60	83.3%
3	-7, 7	10/60	83.3%
4	-7 , 7	9/60	85%
5	-7, 7	7/60	88.3%
6	-7 , 7	7/60	88.3%
7	-7 , 7	7/60	88.3%
8	-7 , 7	9/60	85%
9	-7 , 7	7/60	88.3%
10	-7 , 7	7/60	88.3%

Table 3 shows results of pre-processing the image within an average contrast ratio of [-5.5].

Table 3. Blur Motion Noise level at 10 distance pixel.

Training	Established range Average contrast	Images not identified	Identification %
1	-5 , 5	13/60	78.3%
2	-5 , 5	10/60	83.3%
3	-5 , 5	10/60	83.3%
4	-5 , 5	10/60	83.3%
5	-5 , 5	10/60	83.3%
6	-5 , 5	11/60	81.7%
7	-5 , 5	11/60	81.7%
8	-5 , 5	11/60	81.7%
9	-5 , 5	9/60	85%
10	-5 , 5	10/60	83.3%

Table 4 shows a comparison with pre-processing method and without preprocessing using blur motion noise levels of 10, 20 and 30 pixels distance, results are:

Table 4. Comparison Results blur motion noise

Method	System without pre-processing	System with pre-processing
Motion blur 30	47.44% of identification	55.05% of identification
Motion blur 20	70.09% of identification	82.31% of identification
Motion blur 10	75.3% of identification	84.39% of identification
Radial blur 30	55.09% of identification	63.05% of identification
Radial blur 20	70.96% of identification	79.33% of identification
Radial blur 10	69.65% of identification	81.10% of identification
False acceptance	61.53% of identification	78.96% of identification

False acceptance experiments were performed using blur motion and blur radial noise, a contrast ratio average [-5.5] [-7.7], [-10.10], [-10.4] [10.7], with results shown in Table 5, where we can see that there is a better rate of detection in all cases when applying pre-processing method [9,10,11,12].

Table 5. Comparison Results Blur Motion and blur radial noise

Method	System without pre-processing	System with pre-processing
Motion blur 30	**47.44%** of identification	**55.05%** of identification
Motion blur 20	**70.09%** of identification	**82.31%** of identification
Motion blur 10	**75.3%** of identification	**84.39%** of identification
Radial blur 30	**55.09%** of identification	**63.05%** of identification
Radial blur 20	**70.96%** of identification	**79.33%** of identification
Radial blur 10	**69.65%** of identification	**81.10%** of identification
False acceptance	**61.53%** of identification	**78.96%** of identification

- **Conclusion**

Based on experiments using the method of diffuse contrast pre-processing, increases the percentage of identification of a pattern recognition system.

In the experiments were used different motion blur noise levels of 10, 20 and 30 pixels distance to generate a database that is used as a validation test of false acceptance, that is, images were tested without pre-processing and with pre-processing, resulting in a higher percentage of fingerprint identification using pre-processing. Experiments were also made expanding the search range of the average contrast of the image in a range of [-7,7] getting a better rate of detection.

The experiments carried was found to the value range in which there is a more efficient fingerprint recognition is when driving between [0,10] of the average contrast of the image of the fingerprint reference, we conclude that the exploration of the range of values is to move values from one side in this case values above average contrast value.

Some of the objectives we want to get into this line of research is to obtain a relationship between the appropriate image contrast and pattern recognition system so that the image can be recognized by the system, and therefore able to suggest which is the contrast must have the appropriate image to be recognized.

References

[1] http://biolab.csr.unibo.it/Home.asp, Biometric System Laboratory University of Bologna, Marzo del 2010.

[2] M. Lopez, P. Melin Member, IEEE, and Oscar Castillo Member IEEE, A New Approach for Fuzzy Extraction based on Pixel's Brightness, ISSN:1098-7576.

[3] M. Lopez, P. Melin Member, IEEE, and Oscar Castillo Member IEEE, Comparative Study of Extraction Features Type-1 and Type-2 Fuzzy Logic approaches for Pattern Recognition System based on the mean pixels, pp. 171-188, DOI: 10.1007/978-3-642-15111-8_11.

[4] Miguel Lopez, Patricia Melin, Oscar Castillo: Comparative Study of Feature Extraction Methods of Fuzzy Logic Type 1 and Type-2 for Pattern Recognition System Based on the Mean Pixels. Soft Computing for Recognition Based on Biometrics 2010: ISSN: 171-188.

[5] Elías Todorovich,Nelson Acosta, Gerardo Acosta, XVII Congreso Argentino de Control Automático, 11-13 de Septiembre 2000, Bs. As. Argentina, pp 315-330..

[6] Francisco Herrera, Ten lectures on genetic fuzzy systems, , 1997. pp. 1–69, ISBN 80-227-0978-6

[7] Wang, Wendong, and Susan M. Bridges, "Genetic Algorithm Optimization of Membership Functions for Mining Fuzzy Association Rules," Proceedings of the 7th International Conference on Fuzzy Theory & Technology, Atlantic City, NJ, February 27 – March 3, 2000, pp.131-134.

[8] T.van, K.Cheung, M.Ha, A Fingerprint Recognizer Using Fuzzy Evolutionary Programming, University of Canberra.

[9] Miguel López, Patricia Melin, Oscar Castillo: Pattern Recognition in Blur Motion Noisy Images using Fuzzy Methods for Response Integration in Ensemble Neural Networks. IFSA/EUSFLAT Conf. 2009: 809-814.

[10] Miguel Lopez, Patricia Melin, Oscar Castillo: Comparative Study of Fuzzy Methods for Response Integration in Ensemble Neural Networks for Pattern Recognition. Bio-inspired Hybrid Intelligent Systems for Image Analysis and Pattern Recognition 2009: 123-140.

[11] Miguel Lopez, Patricia Melin: Response integration in Ensemble Neural Networks using interval type-2 Fuzzy logic. IJCNN 2008: 1503-1508.

[12] Miguel Lopez, Patricia Melin, Oscar Castillo: Optimization of Response Integration with Fuzzy Logic in Ensemble Neural Networks Using Genetic Algorithms. Soft Computing for Hybrid Intelligent Systems 2008: 129-150.

This page intentionally left blank

1st Workshop on Large Scale Intelligent Environments (WOLSIE'12)

James P. Dooley (University of Essex, UK)

Matthew H. Ball (University of Essex, UK)

Mohammad Al-Mulla (Kuwait University, Kuwait)

This page intentionally left blank

J.A. Botía et al. (Eds.)
IOS Press, 2012

Introduction to the Proceedings of the 2012 Workshop on Large Scale Intelligent Environments (WOLSIE'12)

As technology permeates into every aspect of our lives and becomes embedded in the environments that we encounter on a daily basis, the dreams for a world of ubiquitous computing are being realised. Augmented by the deployment of artificial intelligence, the technology rich spaces that surround us are given an ambient intelligence where software samples the real world, reasons with available information and pro-actively takes action. Much work has been done in this area as reflected by the literature, but as research progresses and the technologies available to us advances, the scale of the environments we envision increases to that of (for example) intelligent office blocks, university campuses, and cities.

This workshop aims to create a forum for intelligent environment researchers to meet and discuss not only the challenges related to scaling-up intelligent environments research "beyond four walls" but also the opportunities that may arise allowing us to create technology-rich intelligent buildings, neighbourhoods and cities that operate in a cheaper, greener, safer and more secure way. What's more, the workshop seeks to stimulate discussion around the more humanistic and social factors resulting from the deployment of ubiquitous computing and ambient intelligence on a large scale.

The papers included in this workshop discuss a series of ongoing projects that are working towards the realisation of real-world large-scale intelligent environments and each provides a valuable insight into the possible future world that we will inhabit.

We would like to thank all of the organising committee members and international program committee who helped to organise the workshop. Also, we'd like to thank all of the authors who submitted papers and presented at WOLSIE, and the organisers of IE'12.

James Dooley, Matthew Ball, Mohammed R. Al-Mulla,
Víctor González and Rolando Menchaca-Mendez
(Co-Charis WOLSIE'12)

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-158

Beyond Four Walls: Towards Large-Scale Intelligent Environments

James DOOLEY[a,1], Matthew BALL[a] and Mohammed R. AL-MULLA[b]

[a] *Intelligent Environments Group, University of Essex, UK.*

{jpdool, mhball}@essex.ac.uk

[b] *Computer Science Department, Kuwait University, Kuwait.*

mrhalm@sci.kuniv.edu.kw

Abstract. As Intelligent Environments research continues in its various guises, there is an increase in the scale at which investigation is being conducted. Projects are looking beyond single-room and apartment sized living-labs that have provided us with initial testbeds for early research and are envisioning grander designs on the scale of entire buildings, campus or towns. But now we must ask if the knowledge we have gained previously can scale upwards, or are new methods and models required to break free from the confines of controlled labs and into real-world deployments where multi-user is the norm not the exception. In the discourse of this paper, we describe what Large-Scale Intelligent Environments (LSIEs) are and identify some of the challenges that relate to realizing them. We highlight the importance of security, standards and infrastructure so that human users can roam confidently from place-to-place whilst enjoying a seamless continuity of experience.

Keywords. Large-Scale, Intelligent Environment, Ambient Intelligence.

Introduction

As technology permeates into every aspect of our lives and becomes embedded in the environments that we encounter on a daily basis, the dreams for a world of Ubiquitous Computing (UC) are being realized [1]. Augmented by the deployment of computational / artificial intelligence, the technology rich spaces that surround us are given an Ambient Intelligence (AmI) [2] [3] where software samples the real world, reasons with available information and pro-actively takes action. Such spaces are known as Intelligent Environments (IEs). Thus far, research has specifically focused on creating IEs from the familiar places around us in which a series of interconnected computational devices can be embedded.For example, IEs have been created in: *homes, vehicles, offices, classrooms, shops, libraries, and museums*.

[1] Corresponding Author.

Much work has been done in the IE area as reflected by the literature, but as the field matures and we delve further, the scale of the environments we envision increases. A trend is emerging where the focus of research is being placed on investigating the challenges related to an increase in deployment size; a community is forming whose intent is to take IE research *"beyond four walls"* and realize the creation of Large-Scale Intelligent Environments (LSIEs) as real-world deployments that surpass early simulation and proof-of-concept work. This has a special relevance in the realisation of real-world applications and is one of the next steps required on the road to widespread adoption of IEs as instances of UC and AmI.

In this paper we first provide an overview of the conventional IE research scale - so called *"living-labs"*. We then provide an operational definition and description to measure what constitutes a Large-Scale Intelligent Environment (LSIE). It follows to explore the various requirements imposed upon them and describe the various forms in which they can manifest. We continue to examine some of the research community projects that are already identifying the shift from small-scale to large-scale.

1. Towards Intelligent Environments

We classify an IE as a logical space that contains a ubiquitous computing deployment (i.e. a set of interconnected computational devices that are embedded in the space itself) and is endowed with an AmI that perceives and affects the real-world through sensors and actuators. This is typically realized by embodying computational / artificial intelligence within agent-based software [4] [5]. Our future world is envisioned to be a set of these IEs through which we roam – each person visiting just a subset of the overall IEs in existence.

The remainder of this section provides some examples of UC and IE works including a case study on one IE in particular – the University of Essex *"iSpace"*. Being rich in interconnected computing devices, sensors and actuators; *"technology rich"* UC environments are example precursors to IEs - lacking only a quality of intelligence that is achieved through the deployment of suitably endowed software, such as intelligent embedded-agents. This subtle difference between UC and IEs has a large impact on the experience of occupants: while a UC is largely reactive to explicit command and control by users, an IE is more proactive and will take actions independently based on decision-making processes.

While not an exhaustive or detailed exploration of the living-labs that exist (or have existed), it is clear that there is a well-established international interest in UC / IE research from hobbyists and students through to academic and industrial researchers.

1.1. Ubiquitous Computing

Circa 2000, the Cisco *"Internet House"* was constructed on a full building scale, but its purpose was to showcase a home with always-on Internet connectivity and appliance automation (where the home and its appliances could be controlled over the internet). Similarly, the Philips *"HomeLab"* [6] was a fully functional apartment whose purpose was aimed more at user experience evaluation through the use of monitoring technologies (such as cameras and microphones). The greater extent of technology deployment in the MIT *"Placelab"* also took place in a dedicated apartment scale space and focused on the space construction, technology deployment and user experience. The Stanford *"iRoom"* [7] and National Institute for Science and Technology (NIST) *"smart space"* [8] have investigated the deployment of ubiquitous computing in the office / meeting room context. The Fraunhofer inHaus-Center run two labs called the *"SmartHome"* and *"SmartBuilding"* for research into many different areas of innovation including user interfaces [9], an area of research also investigated by the *"iRoom"* at LIMSI [10]. The *"Ambient Kitchen"* at Newcastle University [11] examines the application of UC within the specific context of a domestic kitchen, while the *"Classroom 2000"* [12] and *"SCALE-UP"* [13] projects have explored the application of UC in an education setting.

More recently, the Media Interaction Lab at the University of Applied Sciences Upper Austria have engaged in projects centered around the use of technology in collaborative and office based activities. Of particular interest here are the "*Nice Discussion Room"* [14] and the "*Active-Office"* [15]. Similarly, the *"Future Meeting Room"* [16] at Edinburgh Napier University has focused on the deployment of multi-touch technologies and large screen displays built into tables and walls.

Facilities such as the Duke University *"Smart Home"* have been used primarily for student projects, while the more recent emergence of community-lead *"hackerspaces"* around the world have promoted public participation in technology-oriented projects.

1.2. Intelligent Environments

Researchers have deployed intelligence into numerous spaces; the MIT *"Intelligent Room"* [17] was created in order to investigate new modes of Human-Computer Interation (HCI) and relied on the deployment of embedded cameras, microphones and displays. Here, AI was employed to enable the room to interpret and react to users through gestures (from video) and speech (from microphones). This work also incorporated person tracking (from video) for context awareness that reportedly enhanced the operation of other sub-systems such as speech recognition.

At the University of Colorado, the *"Adaptive Home"* used a centralised neural-network based controller that monitored approximately 75 sensors (light, temperature, sound, motion, door/window state, etc.) and then took appropriate action on related actuators in the home [18]. Over the lifetime of this lab, many experiments were conducted and results published regularly. Such a rich publication history also exists for the Georgia Institute of Technology *"Aware Home"* that explores a huge diversity of subject areas including sociological applications such as assisted-living and home-care [19]. The *"PEIS home"* at the Orebro University further extends the capability of environment manipulation that lies within control of software intelligence by deploying and integrating mobile robots into its infrastructure [20]. Elegantly, some labs (such as the iRoom at the German University in Cairo [21] and the MavHome at Washington State University [22]) are used to experiment with populations of software *"agents"* that provide the AmI (this is especially interesting when considering emergent behaviour from populations of agents that compete or collaborate).

The University of Essex *"iDorm"* mimicked single room student accommodation and was deployed with an extensive array of sensors and actuators that were integrated onto a network in a distributed / grid architecture [23]. Many investigations were carried out across several projects and PhD investigations throughout its life, including the use of neural networks, fuzzy logic, embedded agents and genetic algorithms [4] [24] [25]. The iDorm evolved into a larger 3-room space and was succeeded by the iClassroom [26] and the iSpace…

1.3. Case Study: the University of Essex "iSpace"

The *"Intelligent Space"* (known as *"iSpace"*) is a fully furnished apartment that is designed to emulate a typical two-bedroom home and contains a spacious kitchen / living area, bathroom, master bedroom and study / bedroom as shown in Figure 1. The vision became part of the design for a new computer-science and electronic-engineering building on the University of Essex campus in Colchester, UK. This afforded its construction to include unique features not otherwise available as identified by the previous iDorm work – a capacious false ceiling, false walls and an adjoining control room; all of which are hidden from view and can be used to deploy electronic equipment from embedded systems and sensors up to full size computers and control / automation electronics. This is enabled by the pervasive availability of mains power sockets, serial busses, Ethernet sockets, and wireless networks.

Figure 1. The University of Essex iSpace.

Within the iSpace, a myriad of heterogeneous technology is deployed; An Ethernet / Wi-Fi backbone harbours a distributed computing system that interconnects sensors, actuators and computational devices. For security and privacy reasons, this network is self-contained within the iSpace and is protected by a gateway / firewall that provides an uplink to the University network (and in turn, the Internet). The gateway also hosts certain services such as a DHCP server for dynamic device configuration inside the space and port forwarding for external access to certain resources. Figure 2 shows the general architecture of technology deployment in the iSpace, which builds upon the approach used in the iDorm by adding more technologies and increasing the scale of deployment.

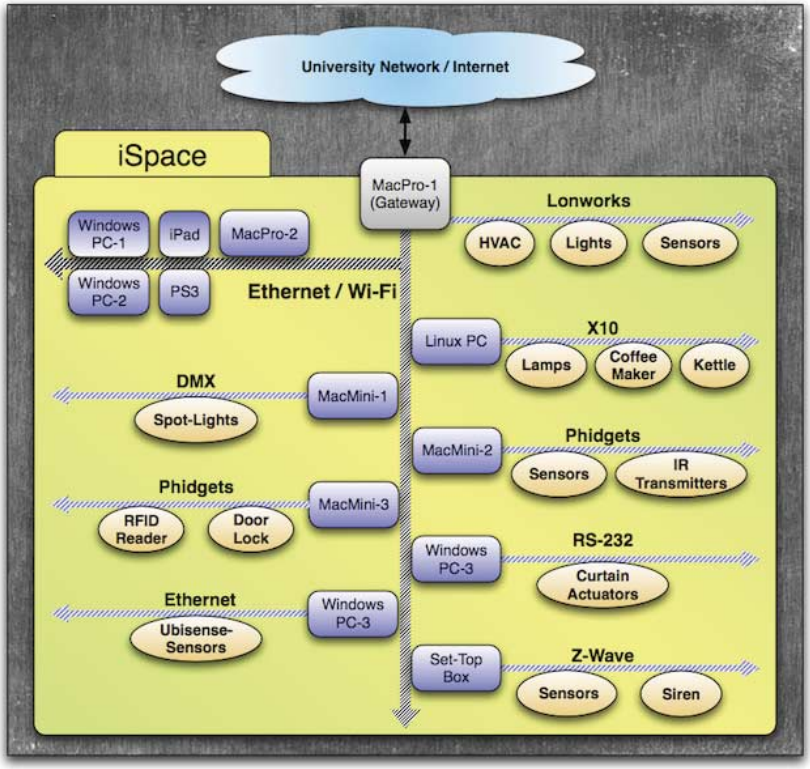


Figure 2. Technology deployed in the iSpace.

To facilitate the discovery and interaction of software components on the network, a middleware layer is deployed that sits on top of the IP network and presents a common *"view"* to the software layer (sometimes called an *"overlay network"*). Each *"peer"* on the network exports its resources through this middleware so that software agents can discover and make use of them. For example, a light-control agent may need to discover and monitor light-sensors and light-switches then discover and manipulate light devices. Thus the heterogeneous network of devices is rendered homogeneous to the software layer for use among different models of distributed computing (service-oriented, multi-agent oriented, component oriented, etc.). Universal Plug and Play (UPnP) has been used as the middleware layer in much of the research that has occurred, but researchers are free to use or develop the middleware that is appropriate to their projects under investigation (in some cases, the middleware itself was the subject of investigation [27] [28]). At the time of writing there are over 100 UPnP devices on the iSpace network, most logically map to some sensor or actuator (such as those shown in Figure 2.) while some are virtual groups (for example, each room has a virtual UPnP device that encompasses a group of individual lights) or higher level sub-systems (such as a user context system that is backed by a Ubisense real-time location tracking system composed of its own gigabit network and 12 sensors). UPnP provides the network with several features that are valuable in such a dynamic and real-time environment. In particular, discovery and subscription allow agents to discover and subscribe to UPnP devices in real-time. This permits an agent to receive asynchronous events when the network topology changes (devices arrive or leave) or when context changes occur (light levels / temperatures change, lights are switched / dimmed from other software sources, etc.).

It is worth noting that software agents can directly interact with all the deployed technologies permitting them to sense, reason and act without human intervention. Thus the iSpace can be placed under AI control (Fuzzy-Logic, Neural Networks, Genetic Algorithms and rule-based systems are most commonly used), where one agent is usually responsible for a particular function (for example security) or sub-system (such as lighting). This kind of relationship between AI and space forms a collective AmI and is a pre-requisite of truly intelligent environments that are required to exhibit user-centric environment adaptation.

2. What are Large Scale Intelligent Environments?

From the literature it is easy to see that a lot of work has been carried out in the IE area, much of it motivated from unique perspectives and investigations. Whilst this work has been proceeding, the need for experimentation has been limited in size to proof of concept work. However, now that the field is maturing and the world is becoming increasingly deployed with networked technology, new opportunities are arising which will allow larger-scale experimentation and real-world deployments to be explored.

The concept of expanding beyond four walls is a defining characteristic of a Large-Scale Intelligent Environment (LSIE); the scale of a LSIE could range from that of a multi-floor building up to an entire town or city, or even up to a country wide or global scale. There are a number of ways in which the vision of LSIEs could be achieved, they could be monolithic or composite systems occupying the physical world but could also exist in virtual reality, or even cross both the physical and virtual worlds using mixed reality. These various forms of LSIE are outlined in the next sub-sections.

2.1. Monolithic Form

In a monolithic LSIE, the whole systems is viewed and governed as a single entity, even though in reality it would most likely contain several sub-spaces. For example, a multi-floor office building could be governed as a single (monolithic) LSIE even though the physical space is separated into multiple floors, each with multiple offices that have their own individual requirements of the LSIE system. This kind of architecture will require extra effort to manage as the technology will cross borders of governance and responsibility, but is well suited to centralised control and therefore may be of use in specific cases despite a probable cost overhead (both for installation and maintenance).

2.2. Composite Form

LSIEs can be realised through the composition of multiple geographically or organisationally separated IEs that are electronically joined to form a larger compound. For example, an organisationally separated building scale LSIE could be decomposed into several individual and independent IEs (in turn of monolithic, composite, virtual or mixed-reality form). This could be useful in situations where each individual component IE requires its own governance (such as an apartment block, or office building). As the links between constituents are electronic, geographic separation could be global. For example, international offices of the same company could transcend their physical separation and be brought together. Similarly, *ad hoc* LSIEs can be formed to host events such as conferences and workshops without the need for participants to leave their parent countries / regions.

2.3. Virtual-reality Form

Entire LSIEs can be simulated / emulated virtually using 3D modelling software or games engines. This form of LSIE has the benefit that it is not restricted by things like the size of the physical space or the cost of equipping the environment with an extensive set of sensors and actuators. What's more, in the virtual world, time can be sped up or slowed making experiments more dynamic [29]. This form offers a safe and cheap way of developing / prototyping / training soft components – an especially useful function for AmI. Other applications of this form include games and social applications in which large numbers of users and non-player characters can interact.

2.4. Mixed-reality Form

As well as entire LSIEs existing in virtual space, they can also be made to span across both the virtual and physical worlds. This form of LSIE has interesting applications in the areas such as eLearning and teleconferencing, in which real-world objects can be linked to virtual representations and manipulated in real time. For example, a lecturer could be giving a talk at the University of Essex in the UK, while a student in another country attends the lecture virtually and is able to interact with the lecturer in real time by raising his hand to ask a question [30].

3. Other Considerations

For LSIEs to become commonplace in the real world, there is a large number of challenges, considerations and requirements that must first be addressed.

A paramount challenge is that of user acceptance, for which many personal concerns of the user and societal-factors must be taken into account. For example, the issue of balancing user-control and system autonomy is a challenge cited by many researchers of IEs [31]. For small-scale IEs, such as a single smart home, this may be a question of user preference (how much control they are willing to delegate to the system), however, for a LSIE that includes multiple different environments or contexts the challenge becomes much more complex. What's more, steps must be taken to deal with the sheer amount information to be processed and work to be done in controlling future LSIE systems, which may simply overload the user. Conversely, a LSIE will also have to deal with an overload of users. The majority of IE research to date has only been focused around single or small groups of users. A city-wide LSIE, however, will have to deal with hundreds of thousands of users, all with individual preferences and requiring different services to be provided constantly and reliably.

Privacy and security are also major challenges that must be addressed in realising the LSIE vision. With an abundance of sensors embedded deeply into many if not all of the environments that we inhabit, many of our personal actions and personal information can easily be detected, shared, and possibly misused by others. In real-world LSIE deployments, every step possible should be taken to prevent this and safeguard personal information. A further security concern, is the potential for malicious use (or abuse) of the system given that LSIE are intrinsically designed to be easily accessible by everyone anytime and via any of a myriad of devices.

4. Community Interest

4.1. Pervasive Computing at Scale (PeCS)

There has been recent attention surrounding the increase in scale of Pervasive Computing (PerCom) applications. Although there is a slightly different focus in the PerCom versus the IE field, they share many commonalities. For instance, much of the PeCS work identifies the increase in embedded / mobile computing with always on network capabilities [32] and examines how systems must scale in order to provide constant and consistent functionality to users that operate with multiple computing devices as they travel through the physical world. The diversity of work examined in this area is reflected by the international attention gained by activities such as the 2011 *"Workshop on Pervasive Computing at Scale"* that was sponsored by the US National Science Foundation. Such is the symbiotic relationship between the IE and PerCom communities that as one field advances, the other inescapably progresses.

4.2. Large-Scale Middleware

As ubiquitous computing deployments, there is a fundamental need for robust and reliable middleware that can securely integrate the various distributed resources of any IE. However, existing approaches have been shown to suffer in terms of both performance and reliability when scaled up. This shortcoming has been identified, investigated and resolved through the work discussed in [27] - co-funded by British Telecom (BT) and the Engineering and Physical Sciences Research Council (EPSRC). The advances made through this work improved the scalability and performance time of distributed middleware for use in IEs by three orders of magnitude and included complex object descriptions that could be queried in a distributed manor reducing the network overhead associated with other approaches.

4.3. ScaleUp

The ScaleUp research project is a collaboration between the University of Essex and King Abdulaziz University. It is focused specifically upon investigating the needs of migrating from current-scale IE research up to LSIEs. The work covers the individual areas of formal methods, infrastructure / middleware, video processing & distribution, intelligent agents and virtual learning environments. The project intends to deploy building, campus and global scale LSIEs through its scope.

4.4. Workshop On Large Scale Intelligent Environments (WOLSIE)

The 2012 Workshop On Large Scale Intelligent Environments (WOLSIE) has been created in response to the inevitable demand for the next phases in development of the IE vision towards real-world adoption. It has been born within the annual Intelligent Environments conference and engages the community of researchers that has emerged through the eight years of conference activity. This will offer a timely exploration of the technical and social LSIE challenges ready for future exploitation in the real world – a world that has been recently primed for the adoption of IE technology by the increasingly widespread infiltration of research that endeavours to create technology-rich buildings, neighbourhoods and cities that operate in a cheaper, greener, safer and more secure way.

5. Conclusions

As the state of the art improves and research in the area of Intelligent Environments progresses, the scale at which those environments manifest increases. We are now at the point when scalability of initial works is being considered and explored by a large community of researchers who share a common motivation to realise larger, multi-user Intelligent Environments in the real-world up to City-scale. The fruit of these investigations will yield a very exciting time as we see not only research, but also practical deployments with real users *“in the wild”*.

References

[1] M. Weiser. 1991. ***“The Computer for the Twenty-First Century”***. Scientific American, pp. 94-10.

[2] W. Weber, J. Rabaey and E. Aarts. 2005. ***“Ambient Intelligence”***. Springer-Verlag.

[3] R. José, H. Rodrigues & N. Otero. 2010. ***"Ambient Intelligence: Beyond the Inspiring Vision"***, Journal of Universal Computer Science (J.UCS), 2010, pp.1480-1499.

[4] H. Hagras, V. Callaghan, M. Colley, G. Clarke, A. Pounds-Cornish and H. Duman. 2004. ***“Creating an Ambient-Intelligence Environment Using Embedded Agents”***, IEEE Intelligent Systems, vol.19, no.6.

[5] D.J. Cook, J.C. Augusto & V.R. Jakkula. 2009. ***“Ambient intelligence: Technologies, applications, and opportunities”***, Pervasive and Mobile Computing, Vol. 5, No. 4, pp. 277-298.

[6] B. de Ruyter & E. Aarts. 2004. ***"Ambient intelligence: visualizing the future"***. In proc. of the Working Conference on Advanced Visual interfaces (AVI '04), pp. 203-208.

[7] B. Johanson, O. Fox & T. Winograd. 2002. ***"The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms"***, IEEE Pervasive Computing, vol.1, pp.67-74.

[8] A. Fillinger, I. Hamchi, S. Degré, L. Diduch, T. Rose, J. Fiscus & V. Stanford. 2009. ***"Middleware and Metrology for the Pervasive Future"***. IEEE Pervasive Computing Mobile and Ubiquitous Systems. vol. 8, no. 3, pp 74-83.

[9] C. Ressel, J. Ziegler & E. Naroska. 2006. ***"An approach towards personalized user interfaces for ambient intelligent home environments"***, 2nd IET International Conference on Intelligent Environments, 2006. (IE'06), vol.1, pp. 247 – 255.

[10] Y. Bellik, & C. Jacquet. 2008. ***"From the Intelligent Room to Ambient Intelligence",*** 1st Digiteo Annual Forum, Poster n°2008-01, October 2, 2008 – Supélec.

[11] P. Olivier, G. Xu, A. Monk & J. Hoey. 2009. ***"Ambient kitchen: designing situated services using a high fidelity prototyping environment"***, In proc. of the 2nd International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '09).

[12] G.D. Abowd, C.G. Atkeson, J. Brotherton, T. Enqvist, P. Gulley, and J. LeMon. 1998. ***"Investigating the capture, integration and access problem of ubiquitous computing in an educational setting".*** In proc. of the 1998 conference on Human Factors in Computing Systems (CHI'98), p.p. 440-447.

[13] R. Beichner, J. Saul, D. Abbott, J. Morse, D. Deardorff, R. Allain, S. Bonham, M. Dancy & J. Risley. 2006. ***"Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) project"***, In E. F. Redish and P. J. Cooney (Eds.), PER-Based Reform in University Physics. College Park, MD: American Association of Physics Teachers.

[14] M. Haller, J. Leitner, T. Seifried, J. Wallace, S. Scott, C. Richter, P. Brandl, A. Gokcezade & S. Hunter. 2010. ***"The NiCE Discussion Room: Integrating Paper and Digital Media to Support Co-Located Group Meetings"***, In proc. of the 8th SIGCHI conference on Human factors in computing systems, 2010.

[15] K. Probst, J. Leitner, F. Perteneder, M. Haller, A. Schrempf & J. Glöckl, 2012. ***"Active Office: Towards an Activity-Promoting Office Workplace Design"***, in proc. of the 30th international conference extended abstracts on Human factors in computing systems, 2012.

[16] O. Mival & D. Benyon. 2011. ***"Outputs from the Future Interactions Network"***, Interactive Experience: The 25th BCS Conference on Human-Computer Interaction. July 2011.

[17] M.H. Coen. 1998. ***"Design principles for intelligent environments"***, In proc. of the 1998 Innovative applications of artificial intelligence. American Association for Artificial Intelligence, pp. 547-554.

[18] M.C. Mozer. 1998. ***"The neural network house: An environment that adapts to its inhabitants"***, In proc. of the American Association for Artificial Intelligence Spring Symposium on Intelligent Environments (pp. 110-114). Menlo, Park, CA: AAAI Press.

[19] C.D. Kidd, R. Orr, G.D. Abowd, C.G. Atkeson, I.A. Essa, B. MacIntyre, E.D. Mynatt, T. Starner & W. Newstetter. 1999. ***"The Aware Home: A Living Laboratory for Ubiquitous Computing Research"***, In proc. of the 2nd international Workshop on Cooperative Buildings, integrating information, Organization, and Architecture. Lecture Notes In Computer Science, vol. 1670. Springer-Verlag.

[20] M. Broxvall, B.S. Seo & W.Y. Kwon. 2007. ***"The PEIS Kernel: A Middleware for Ubiquitous Robotics"***, In proc. of the IROS-07 Workshop on Ubiquitous Robotic Space Design and Applications.

[21] B. El-Desouky, & H. Hagras. 2009. ***"An Adaptive Type-2 Fuzzy Logic Based Agent for Multi-Occupant Ambient Intelligent Environments"***, In proc. of the 5th International Conference on Intelligent Environments, Barcelona, Spain, July 2009.

[22] D. J. Cook, M. Youngblood, E. Heierman, K. Gopalratnam, S. Rao, A. Litvin & F. Khawaja. 2003. ***"MavHome: An Agent-Based Smart Home"***, In proc. of the IEEE International Conference on Pervasive Computing and Communications, pp. 521-524, 2003.

[23] A. Pounds-Cornish, A. Holmes. 2002. "***The iDorm - A Practical Deployment of Grid Technology"***, 2nd IEEE International Symposium on Cluster Computing and the Grid (CCGRID'02), pp.470.

[24] E. Tawil & H. Hagras. 2004. ***"A Novel Multi-Objective Multi-Constraint Genetic Algorithms Approach for Co-ordinating Embedded Agents"***, In proc. of the IEEE International Conference on Systems, Man, and Cybernetics, 10-13 October 2004, The Hague, The Netherlands.

[25] F. Rivera-Illingworth, V. Callaghan, H. Hagras. 2005. ***"A Neural Network Agent Based Approach to Activity Detection in AmI Environments"***, International Workshop on Intelligent Environments 2005.

[26] J. Dooley, V. Callaghan, H. Hagras, M. Gardner, M. Ghanbaria & D. Al-Ghazzawi. 2011. ***"The Intelligent Classroom : Beyond Four Walls"***, Intelligent Campus Workshop (IC'11), July 2011.

[27] J. Dooley. 2011. ***"An Information Centric Architecture for Large Scale Description and Discovery of Ubiquitous Computing Objects"***, PhD Thesis, University of Essex.

[28] J. Dooley, C. Wagner, H. Hagras, G. Pruvost. 2011. ***"FollowMe: The Persistent GUI"***, In 1st International Workshop on Situated Computing for Pervasive Environments (SCOPE'11).

[29] J. Dooley, M. Davies, M. Ball, V. Callaghan, H. Hagras, M. Colley & M. Gardner. 2010. ***"Decloaking Big Brother: Demonstrating Intelligent Environments"***, In proc. of the 6th International Conference on Intelligent Environments (IE'10). Kuala Lumpur, July 2010.

[30] T. Zhang, V. Callaghan, R. Shen & M. Davies. 2011. ***"Virtual Classrooms: Making the Invisible, Visible"***, Intelligent Campus Workshop (IC'11), July 2011.

[31] M. Ball & V. Callaghan. 2012. ***"Managing Control, Convenience and Autonomy"*** in Special Issue on Agent-Based Approaches to Ambient Intelligence, AISE book series, IOS Press.

[32] D. Cook & S. Das. ***"Pervasive computing at scale: Transforming the state of the art"***. Journal of Pervasive and Mobile Computing, 8(1):22-35, 2012.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-169

Smart Home in a Box: A Large Scale Smart Home Deployment

Aaron S. CRANDALL [a] and Diane J. COOK [a,1]
[a] *School of Electrical Engineering and Computer Science, Washington State University, USA*

Abstract. Smart home technologies hold promise for many aspects of daily life. Research and development of these systems has matured for elder care, energy efficiency, and home safety applications. The focus of most implementations has been on single living spaces for a small number of individuals. New low-power wireless systems, inexpensive computing power, and widely available network access has reached the point where large scale ubiquitous computing technologies have become much more feasible.

The smart environments research community has few projects to explore issues and techniques for deploying large scale ubiquitous systems. This work summarizes some of the existing works and introduces the Smart Home in a Box (SHiB) Project. The upcoming SHiB Project targets building 100 smart homes in several kinds of living spaces for gathering longitudinal data from a significant number of residents. The resulting data set will provide opportunities to answer open questions in the areas of transfer learning, active learning, digital asset migration, middleware architectures, activity detection and discovery, human factors, smart home installation, and others.

Keywords. smart homes, large scale ubiquitous computing, Smart Home in a Box, crowd sourcing, active learning, transfer learning

Introduction

Since the foundation of ubiquitous computing was laid out in the early 1990s [1], the promise of ambient, smart, or calm technologies has driven work in the areas of engineering, computer science, mobile applications and health care. The application of these high density sensor platforms to behavior modeling [2], activity detection [3,4], gerontechnology [5] and adaptive environments [6] has pushed a boom in both the research and commercialization fields. Many of these tools have reached maturity on small scales, but work on large scale implementations poses new hurdles to overcome.

Part of the ubiquitous computing vision is the ability to move from one space to another, and drawing upon locally available computing resources to suit your needs. Having an intelligent environment in a home or office that learns an individual's preferences has been explored [7,8,9], but the ability to move those preferences between spaces is still

[1]Corresponding Author: Diane J. Cook, Washington State University P.O. Box 642752, Pullman, WA, USA; E-mail: cook@eecs.wsu.edu.

an oft-discussed, though rarely tested goal. There is still plenty of work to be done in all areas of having spaces conform to people, especially in the case of multi-site situations.

This work summarizes some of the existing small and large-scale smart environment projects to date, and introduces the Washington State University Center for Advanced Studies in Adaptive Systems' (CASAS) upcoming Smart Home in a Box (SHiB) project. The SHiB Project is projected to include 100 new smart homes deployed at a wide range of living spaces over the next three years. The kinds of spaces SHiB will focus on are private homes and elder care facilities.

The primary goal of SHiB is to provide a massive data set with high quality annotation. The project is a multi-disciplinary, with participation from researchers in the fields of computer science, psychology, health care, and other engineering fields. The hope is to have a test bed sufficient to address issues in the areas of large-scale deployment, activity detection, transfer learning, individual preference mobility, human factors, active learning, and more.

1. Grand Challenges in Large Scale Smart Homes

There are a number of open grand challenges to be addressed in the deployment of large scale smart home systems. This work proposes upcoming tools to address some of these, and other groups are investigating solutions to others. Listed here is a partial list of challenges, in no particular order of importance.

1. A smart home system completely deployable in a home in under three hours by a non-specialist.
2. Strategies to mitigate the visible impact of the technology on the look and feel of the home, especially in a retrofitting situation.
3. Smart homes able to model residents in an unknown space and of an unknown quantity "out of the box" to a reasonable accuracy.
4. Subtle and calm prompting systems that adapt to the current home conditions.
5. A smart home user interface that may be accessed ubiquitously throughout the home and without inconveniencing most residents.
6. Methods to harness and integrate information from disparate sources: environmental sensors, smart phones, wearables, etc.

2. Related Work

To date there has been a plethora of small-scale smart home installations, such as the Aware Home [10], the iDorm [11], House_n [12], the Gator Tech Smart home [13], the Neural Network House [14], the MavHome [15], and the CASAS project [16,17]. These projects encompass one or a handful of smart environments. They have been used to understand human behavior and provide context-aware services.

Historically, there has been few multi-site, large-scale smart environment research projects. Even a single smart environment represents a significant investment in equipment, engineering time and maintenance. Keeping and maintaining numerous sites is a daunting task for most research groups. To date there have been two well known large scale smart environment projects: TigerPlace at the University of Missouri [18] and the ORCATECH Living Laboratory at Oregon Health and Science University [19].

TigerPlace is a monolithic smart home system installed in a long term elder care facility, with the sensors installed in 17 apartments, plus shared living spaces. This facility has provided numerous opportunities for developing novel approaches to elder care and behavior analysis over the years.

The ORCATECH Living Laboratory (OLL) is a distributed in-home smart environment project. To date it has gathered data from over 400 homes across the United States. The project has spanned numerous health care, sensor technology development and behavior analysis projects. Additionally, OLL has published information about the engineering and logistical problems overcome during its operation [20,21]. The sheer scale and distributed nature of the project have required innovative and strict approaches necessitated by managing so many installed testbeds and data sources.

2.1. WSU CASAS Horizon House

A more recent large-scale smart environments project is the Washington State University's CASAS Horizon House installation. This project is installed at Seattle's Continuing Care Retirement Center (CCRC), called Horizon House, where volunteer older adults have the CASAS sensor technology installed in their apartments. The project is targeting longitudinal data collection of both sensor data and biannual psychological evaluations to track individual residents' cognitive capabilities. To date, the facility has 16 running apartments with a year's worth of data each.

3. CASAS Smart Home in a Box

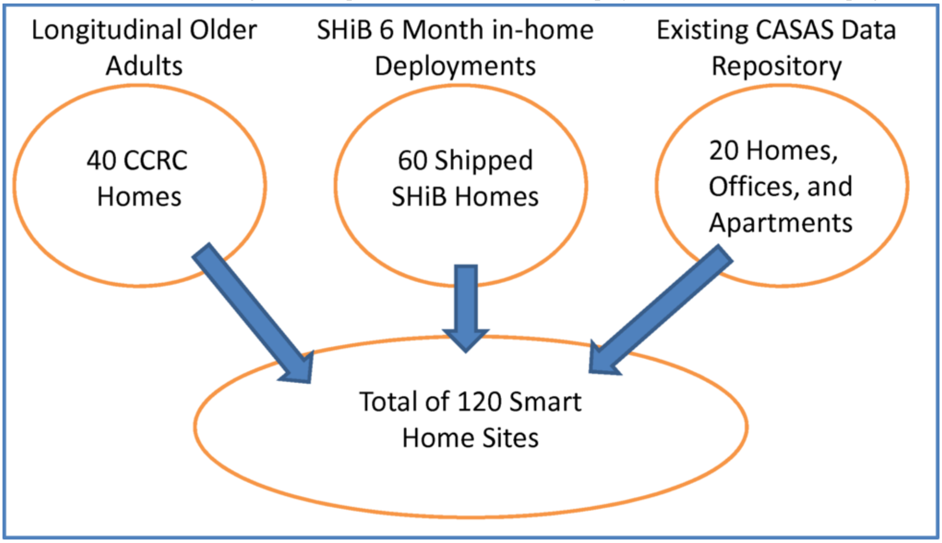
Smart Home in a Box (SHiB) is an upcoming project initiated by the CASAS research group. The primary goal of this project is to provide a large-scale smart environment data and tool infrastructure. The project will increase the number of four currently operational CASAS smart home testbeds and the 16 long-term suites at Horizon House to a much larger collection. A second goal is to provide longitudinal data that monitors behavioral patterns over months or years and allows researchers to thus examine behavioral cycles and trends. Finally, we want to include collection of smart environment data specifically targeted towards older adults with cognitive decline.

3.1. SHiB Planned Deployments

To target the older adults, the project will expand the installation at Horizon House to 25 suites, plus another 15 suites at other CCRCs around the region. These CCRC installations will run for upwards of five years each. This provides a set of longitudinal data on older adults matched with biannual psychological evaluations to give personal physical and cognitive health trends for analysis.

For the large-scale in-home deployments, SHiB will create 20 "CASAS in A Box" kits and make them available for volunteer participants to install in their own homes. These SHiB kits will run in the volunteer's home for at least six months before being shipped back to CASAS to be re-deployed somewhere else. After two years of this process, the SHiB Project plans on gathering data from 60 private homes. In total, the project

Figure 3.1. Planned categories and quantities of Smart Homes deployed for the CASAS SHiB project.



will have collected smart home data from 100 homes. All of this data will be integrated into the CASAS data repository [22], taking the total CASAS database up to 120 data sets, as shown in Figure 3.1.

The net result of CASAS' SHiB effort will be a massive data collection. These smart homes generate an average of 5,000 sensor events each day. When smart phones and power meter readings are incorporated, the volume increases dramatically. The data collection will offer unprecedented opportunities for researchers to study human health and behavior. It will also offer a large step toward the goal of scaling pervasive computing to larger population groups and settings.

As a research project, the SHiB addresses issues in a number of fields. These include how to deploy a smart home without a specialist on hand, managing large scale data gathering, dealing with annotation of exceptionally large data sets, making a resident's data available and understandable to themselves, plus core algorithms research in the areas of transfer learning, trend analysis, active learning, and activity detection.

3.2. CASAS in a Box Design

The first component of the project is the design of the "CASAS in A Box" kits. Each kit consists of a standard set of parts, plus a few optional ones. The core of the infrastructure is the computing and communications infrastructure, as shown in Figure 3.2. This box contains the power sources, networking and a Dream Plug computer to handle the computation, data collection and storage for the home.

Each of the installations will include 30–40 sensors. These are Control4 [23] and Card Access [24] ZigBee-based wireless sensors that have a proven battery life of 6 months to a year and reliable performance. This sensor package, as shown in Figure 3.3, provides motion/lighting detectors, door open/close, and temperature sensors. Each home will also optionally include a TED power meter, object shake sensors and a smart home with data collection software to capture acceleration and direction parame-

Figure 3.2. CASAS smart home components for power, networking, computation and middleware stored in a box.

ters while the phone is in the resident's possession. All of the sensors will be attached using 3M Command Strips, which allow objects to be easily attached and removed from walls without damage.

The current CASAS smart homes upload their data collection at 15 minute intervals if they have access to an Internet connection. The data is copied using an encrypted connection, checked for transmission errors and loaded into the CASAS database if the files were copied correctly. If the smart home server fails to connect to the CASAS central system it spools the new data until an uplink can be established. Because the CASAS sensor platform is event driven and the repetitive nature of the data compresses efficiently, a year's worth of data takes up around 60MB of disk. With higher resolution data from smart phones and power metering this will balloon to several hundred megabytes, which is still easily kept locally on the in-home server.

To handle the case of homes without an available Internet connection, the in-home server will be expected to keep all gathered data locally until it is shipped back to CASAS at the end of the deployment. Once the in-home server connects to the main data repository, it will upload all new data for long term storage. In the interests of protecting the data over the six month deployment, the servers will use a RAID 1 disk configuration to reduce the chance of data loss if a single disk fails.

3.3. SHiB Kit Installation

One of the key components of a smart environment is installation. Traditionally, this process is completed by specialists and may require numerous tweaks to get the system running. Since having specialists on hand to perform every install does not scale, the

SHiB Project is researching methods and tools to overcome this requirement. The goal is for laymen volunteers to install the system straight from the box without a specialist's intervention. How to achieve this is a notable project unto itself.

CASAS has built a team of experimental psychology and engineering researchers with the aim of developing the best set of equipment and instructions for the kit. Since the layout of the resident's home is unknown, the kit and instructions need to be flexible and clear to achieve the best results. There is an inherent trade off of sensor density with ease of installation, and finding the required sensor capabilities for driving the data gathering goals of the SHiB Project is currently under review.

Figure 3.3. CASAS in A Box sensor components, including motion / lighting sensors attached to the ceiling (upper left), door / temperature sensors (upper right), object shake sensors (lower left), smart phone (lower middle), and TED power meter (lower right).

3.4. Sensor Placement

Sensor placement for smart environments, especially when done by laymen, is difficult. Being able to determine which sensors are required for a given space and to communicate how to place them effectively is a crucial process. There has been very good work into where to place sensors to meet a variety of constraints [25], but also doing this in a previously unknown space to meet amorphous goals is less explored. The SHiB Project hopes to build new tools and approaches to overcome these issues, making the kits more effective data gathering platforms.

3.5. Human Factors in Smart Environments

SHiB hopes to address a number of human factors issues with smart environments. Having a large scale deployment and hands-on experience feedback from 60 homes will give significant results on several fronts:

1. Technology perception, especially regarding intrusiveness and expectations
2. Installation problems and understanding of complex instructions
3. Sensor labeling, both for installation and how that influences perceptions
4. In-home impacts on behavior, both conscious and unconscious
5. Privacy and control over personal data

To date, much of the data gathered about living and interacting with smart environments has been anecdotal. There exists some recent work on privacy in sensor rich environments [26,27], and CASAS researchers hope to leverage the large scale of the SHiB Project to formally evaluate these factors to help guide future projects.

3.6. Data Annotation

During the course of the previous CASAS smart homes, it became apparent that the value of the smart home data is increased dramatically if it is annotated. These annotations are historically done by human hand for features such as activities being performed, the identity of residents, and environmental information. Experience shows that it takes roughly 1.5 hours of labor to annotate one 24 hour period in a home. Since this hand-annotation process is expensive, the SHiB Project will use a four stage process to annotate the vast data collected:

1. Self reporting
2. External observers
3. Crowd sourced information
4. Transfer and Active Learning algorithms

3.6.1. Self Reporting

Initially, the residents will be asked to complete a form that indicates when they perform key activities. In the CCRCs with the older adults, this process will be completed with the help of the research team. SHiB participants in private homes may ask CASAS engineers for clarification and direction, but most will be expected to complete it on their own. The targeted activities include: Cook, Eat, Wash dishes, Bathe, Dress, Groom, Make phone calls, Sleep in bed, Sleep out of bed, Work, Watch TV, Read, Leave home, Enter home, Entertain guests, Exercise, and Take medicine. Optionally, the residents can specify exact dates and times when activities occurred.

3.6.2. External Observers

This stage is the classic human annotation phase for the data. People hired by the SHiB Project will view the data from homes using a graphical interface, and annotate activities by hand. Their primary guidance will be information provided by the residents about their behaviors and homes. This stage is expected to label 10 days for each home and is a requirement for many supervised and semi-supervised algorithms used for classifying events automatically.

3.6.3. Crowd Sourced Information

For this approach, people will be invited to annotate a series of sensor events in a visualization on a web page. Small segments of data and detailed instructions will allow visitors to vote on which activity is being performed. The goal is to provide reasonably accurate labels for "interesting" series of events in the data set, as well as driving the cost of annotation down. How to present the data, get involvement, and resolve inconsistencies between volunteers are open research questions.

3.6.4. Transfer and Active Learning Algorithms

There are also approaches for using transfer learning and semi-supervised algorithms for "out of the box" annotation being evaluated. Transfer learning leverages labeled data from other smart environments to begin classifying unlabeled at the new site. With the new data sets generated and labeled on the scale of the SHiB Project, these approaches can be truly evaluated for a wide variety of environments.

When using semi-supervised classifiers, the initial 10 day annotation from self reporting and external observers is used to automatically classify new data from the same site. These tools can also be boosted by an initial transfer learning stage with data from other homes. The new automatically determined labels are then used to help classification as the home runs, essentially learning more about itself to get better at recognizing the residents and their activities.

Another approach being explored is using Active Learning algorithms. The goal is to query the resident as an expert label provider. This is expected to be done through the smart phone or other interface delivered with the SHiB kit. As the classification tools operate, they will have the opportunity to ask the resident about a series of events and what activity it should be classified as. This process can happen "out of the box" or at any time later to augment an existing model provided by a transfer learning approach.

3.6.5. Digital Asset Migration

The CASAS SHiB team has begun implementing approaches to digital asset migration between varied smart home facilities. By having a wide range of living and working spaces to draw upon the migration of learned personal preferences can be well explored. We already have data sets for individuals that include both their behavior and environmental controls in homes, offices and offices within a home. Being able to identify a resident and bring their assets across from one space to another is a key step in Ubiquitous Computing. The SHiB project will enable the expansion of this work to many sites with many individuals.

4. Summary

Large-scale smart environment technology research projects are a necessary step towards realizing the ubiquitous computing concept. These projects are resource and time intensive, but offer a wealth of research opportunities. The upcoming CASAS Smart Home in a Box project is designed to support a wide range of research targets, both immediately practical, and to evaluate open ended theories.

The field of intelligent environments needs to draw upon its existing successes and begin exploring the possibilities presented by these large-scale systems. SHiB provides a solid platform with significant tools to take these next steps towards making smart technologies available to large numbers of people.

References

[1] Mark Weiser. The computer for the 21^{st} century. *Scientific American*, pages 94–104, 1991.

[2] Hui-Huang Hsu and Chien-Chen Chen. RFID-based human behavior modeling and anomaly detection for elderly care. *Mobile Information Systems*, 6(4):341–354, December 2010.

[3] Usman Naeem and John Bigham. Recognising activities of daily life through the usage of everyday objects around the home. In *International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth, pages 1–4, April 2009.

[4] Geetika Singla, Diane J. Cook, and Maureen Schmitter-Edgecombe. Recognizing independent and joint activities among multiple residents in smart environments. *Journal of Ambient Intelligence and Humanized Computing*, 1(1):57–63, 2010.

[5] G. Demiris and B. K. Hensel. Technologies for an aging society: a systematic review of "smart home" applications. *Yearbook of Medical Informatics*, pages 33–40, 2008.

[6] Parisa Rashidi and Diane J. Cook. Adapting to resident preferences in smart environments. In *Proceedings of the AAAI Workshop on Advances in Preference Handling*, pages 78–84, 2008.

[7] N. Kushwaha, Minkoo Kim, Dong Yoon Kim, and We-Duke Cho. An intelligent agent for ubiquitous computing environments: smart home ut-agent. In *Proceedings of the IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems*, pages 157–159, May 2004.

[8] Scott Davidoff, Min Lee, Charles Yiu, John Zimmerman, and Anind Dey. Principles of smart home control. In Paul Dourish and Adrian Friday, editors, *Ubiquitous Computing*, volume 4206 of *Lecture Notes in Computer Science*, pages 19–34. Springer Berlin / Heidelberg, 2006.

[9] George Demiris, Brian K. Hensel, Marjorie Skubic, and Marilyn Rantz. Senior residents' perceived need of and preferences for "smart home" sensor technologies. In *International Journal of Technology Assessment in Health Care*, volume 24, pages 120–124. Cambridge University Press, 2008.

[10] Elizabeth D. Mynatt, Anne-Sophie Melenhorst, Arthur D. Fisk, and Wendy A. Rogers. Aware technologies for aging in place: understanding user needs and attitudes. *Pervasive Computing*, 3(2):36–41, April 2004.

[11] Faiyaz Doctor, Hani Hagras, and Victor Callaghan. A fuzzy embedded agent-based approach for realizing ambient intelligence in intelligent inhabited environments. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 35(1):55–65, January 2005.

[12] A. Helal, E. Kim, and S. Hossain. Scalable approaches to activity recognition research. In *Pervasive*, volume 7040. Springer Berlin / Heidelberg, May 2010.

[13] Sumi Helal, William Mann, Hicham El-Zabadani, Jeffrey King, Youssef Kaddoura, and Erwin Jansen. The Gator Tech Smart House: a programmable pervasive space. *Computer*, 38(3):50–60, March 2005.

[14] Michael C. Mozer. The neural network house: An environment that adapts to its inhabitants. In *Proceedings of the AAAI Spring Symposium on Intelligent Environments*, pages 110–114, Menlo Park, CA, 1998. AAAI Press.

[15] Diane J. Cook, Michael Youngblood, Edwin O. Heierman III, Karthik Gopalratnam, Sira Rao, Andrey Litvin, and Farhan Khawaja. MavHome: an agent-based smart home. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications*, PerCom '03, pages 521–524, March 2003.

[16] Aaron S. Crandall. *Behaviometrics for Multiple Residents in a Smart Environment*. PhD thesis, Washington State University, 2011.

[17] Diane Cook. Learning setting-generalized activity models for smart spaces. *IEEE Intelligent Systems*, 27(1):32–38, January–Feburary 2012.

[18] Marilyn Rantz, Myra Aud, Greg Alexander, Debra Oliver, De Minner, Marge Skubic, James Keller, Zhihai He, Mihail Popescu, George Demiris, and Steve Miller. Tiger place: An innovative educational and research environment. Technical Report FS-08-02, AAAI Fall Symposium, Washington, DC, November 2008.

[19] OHSU ORCATECH. Oregon Center for Aging & Technology (ORCATECH) at Oregon Health & Science University. http://www.orcatech.org/.

[20] Jeffrey Kaye, Tamara Hayes, Tracy Zitzelberger, Jon Yeargers, Misha Pavel, Holly Jimison, Nicole Larimer, Jessica Payne-Murphy, Eric Earl, Kathy Wild, Linda Boise, Devin Williams, Jay Lundell, and Eric Dishman. Deploying wide-scale in-home assessment technology. In *Technology and Aging*, volume 21 of *Assistive Technology Research Series*, pages 19–26. IOS Press, 2008.

[21] Tamara L. Hayes, Pavel Misha, and Jeffery A. Kaye. Gathering the evidence: Supporting large-scale research deployments. *Intel Technical Journal*, 13(3):148–167, 2009.

[22] WSU CASAS. CASAS public smart home datasets repository. http://casas.wsu.edu/datasets.html.

[23] Control4. Control4 home automation and smart home control web site. http://www.control4.com.

[24] Card Access. Card access incorporated web site. http://www.cardaccess-inc.com.

[25] Andreas Krause. *Optimizing Sensing: Theory and Applications*. PhD thesis, Carnegie Mellon University, December 2008.

[26] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. Exploring privacy concerns about personal sensing. In *Proceedings of the International Conference on Pervasive Computing*, PerCom '09, pages 176–183, Berlin, Heidelberg, 2009. Springer-Verlag.

[27] George Demiris, Debra Perker Oliver, Jarod Geiger, Marjorie Skubic, and Marilyn Rantz. Older adults' privacy considerations for vision based recognition methods of eldercare applications. *Technology and Health Care*, 17(1):41–48, 2009.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-179

Towards FollowMe User Profiles for Macro Intelligent Environments

Luke WHITTINGTON [a], James DOOLEY [a,b]. Martin HENSON [a] and Abdullah AL-MALAISE AL-GHAMDI [a,b]

[a] *Department of Computer Science, University of Essex, UK. Email: lwhittl@essex.ac.uk*

[b] *Faculty of Computing and Information Technology, King Abdulaziz University, KSA*

Abstract. We envision an Ambient Intelligent Environment as an environment with technology embedded within the framework of that environment to help enhance an users experience in that environment. Existing implementations , while working effectively, are themselves an expensive and time consuming investment. Applying the same expertise to an environment on a monolithic scale is very inefficient, and thus, will require a different approach. In this paper, we present this problem, propose theoretical solutions that would solve this problem, with the guise of experimentally verifying and comparing these approaches, as well as a formal method to model the entire scenario.

Keywords. ambient intelligent environments, formal methods, mobile devices, cloud server, bi-graphs

1. Introduction

Intelligent environments are a wonderful example of ubiquitous computing[1], but ask two different researchers what an Intelligent Environment is and you'll probably get two very different answers. In our vision, Intelligent Environments tend to be common spaces (such as the home or the office) which contain a plethora of embedded computer devices that help enrich user experiences. These devices are generally interconnected by a network infrastructure (most likely a Local Area Network (LAN)), and controlled by a group of intelligent agents. Environments have a set of users, with each user profile containing their own set of preferences and applications.

Applications have a specific meaning in this context, for example: each light in the environment has a Universal Plug and Play (uPnP) wrapper which exposes functionality of the light across the network. This means that the light power can be changed, as well the dimming level of the light. Having control over individual lights, while beneficial, isn't exactly desirable. Ideally we would have control over *groups* of lights (e.g. *living room* lights, *kitchen* lights and so on). It follows that upon an event triggering (e.g. a DVD player powering on), this could trigger a rule in the intelligent agent, causing the living room lights to be dimmed and the curtains to automatically close. This would be called the "movie application" - users can create these applications as they need and/or desire.

These environments are able to recognise human occupants, reason with context and program itself to meet user(s) needs by learning from their behaviour [2]. The University of Essex has a purpose built Intelligent Environment, the iSpace, which is a fully functioning apartment (complete with bedrooms, kitchen, bathrooms etc.), which has been augmented by intelligent agents and a plethora of sensors. The iSpace contains false walls and false ceilings, allowing devices to be embedded directly into the framework of the apartment. This is used as a model deployment and is a template to shape new spaces such as the iClassroom and an office based scenario.

ScaleUp is a research project, led by the Intelligent Environments Group (IEG) at The University of Essex. The aim of the ScaleUp project is achieving large scale Intelligent Environments. This project has already been outlined in a previous paper published by the IEG [3]. The existing implementations tend to be the size of a single apartment, which in their current format present scaling issues when attempting to create monolithic sized environments. It is natural progression that these environments need to scale up so it is vital that these scalability issues are solved. In this paper, we introduce the problem in more depth (Section 3) and then move onto the proposed solutions, either a centralised user-profile respiratory, a quite literal "mobile" user profile or a hybridisation of the two.

2. Initial Survey of existing Work

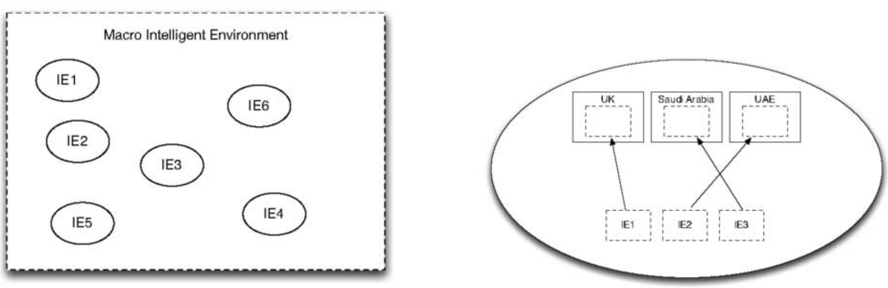
During the initial survey of existing work on large scale intelligent environments, two issues become apparent very quickly; while there has been a significant amount of publication on the topic of intelligent environments[4][5], there is significant fragmentation in the language used to describe these environments (even to the extent whereby the environment itself has different naming conventions[6]). This fragmentation causes difficulties ensuring the different research groups are talking about the same subject. In order to keep consistency with the IEG at Essex, we shall be using the language expressed in their published papers[7,8].

Some of the earliest examples of a purpose built intelligent environments are the "Intelligent room" built in Bristol, UK [9] and the iSpace built at The University of Essex [10]. Since then, the field has made massive strides of progress, with other spaces being built; examples including workplace environments: the "smart lab" at the University of Deusto and recreational purposes[4] and home place environments: the Phillips "HomeLab"[11], which is a fully functional apartment similar to IEG's iSpace. The Cisco "Internet House", while larger than an apartment, was built to show an environment with an "always-on" internet connection and appliances which could be controlled via the internet. The theory that these other environments propose would allow them to fit into the model this paper proposes; creating potential opportunities for collaboration between research groups, which would hopefully lead towards the unification of research in this area.

The other issues is that of all the major publications read, rather than how these environments interact and communicate between each other, the focus was on the *internals* of an intelligent environment[12]. Naturally, there is a lot of novel research still required for the internals of intelligent environments, ranging from the automation of these environments (through the use of embedded agents)[13,14], to the Human-Computer Interfacing[15,16]; it is vital this research continues as the model presented in this paper, means that each environment is self-governed and thus will still require a high level of individual intelligence. M. Habib published a paper [17] on bringing together geographically separated intelligent environments. However, this paper uses a virtual world in an attempt to bring together these geographically separated environments (similar to the concepts in [7,18]). We are aiming more at bringing together these environments in the physical world; though it would be feasible to say these ideas could be implemented in the virtual world as well. At the time of writing, this is the only paper to solely focus on merging smaller scale environments into a larger one. However, it has been noted that there is a need to start scaling up existing implementations of pervasive computing [19]. The fact that the majority of these publications only briefly touch upon the topic on the communication between intelligent environments and scaling, provides further emphasis on the originality and novelty of the proposed research area.

3. Problem Definition

Most computer systems authenticate a user at initial login session [20] and Intelligent Environments conform to this paradigm; requiring users to explicitly *"login"*, using some form of contextual credentials in order to access the assets within that environment. In their current format, scaling up the existing implementations presents issues; keeping track of multiple users in a large area, ensuring each user has continuity of experience and in particular the *"humanistic"* problem of continually being forced to re-authenticate oneself via entry of username and password. This is the equivalent of having to dig your front door key out of the bottom of your bag each time you wish to enter your house, it would ideal for the house to simply *"know"* who to grant access to. To solve these scalability problems, the project envisions creating a set of independent spaces that are treated as a traversable set rather than a monolithic environment.



(a) (b)

Figure 1. (a) shows a large scale IE actually consists of a network of smaller, individual IEs. (b) shows that these individual IEs have no requirement to be geographically co-located to be considered within a Macro IE

Figure 1.a shows the set of environments that are geographically co-located but this need not be the case; the environments can be distal or proximal. The connections between each environment is electronic so there is no requirement for them to be on the same campus, territory or even country. You'll notice in figure 1.a that the border is dotted; this represents a "region" [21], a concept which is discovered in detail in section 5. It is entirely feasible to create these Large Scale Intelligent Environments (as shown in

Figure 1.a) in Essex, Saudi Arabia and the United Arab Emirates, and as long as they can communicate electronically then they may exist as a Macro Intelligent Environment (Figure 1.b illustrates that the ”regions” represented in Figure 1.a can be geographically sparse).

Users will have a set of environments that they are familiar with (there exists a mutual acceptance of data between the user and the space), and can roam freely between each environment (assuming they have the authorisation to do so). Solving this problem is significant because the core aspect of an ”Intelligent World” is the ability to roam freely between these individual ”islands” (individual environments), while enjoying technological transparency and continuity of experience[22,23,24]. In order to ensure the user experience is consistent across multiple environments, each user must have a profile that is mobile and can follow them around. In this way, the continuity of user experience can be maintained.

As a consequence, it is important that the users profile follows them between environments. When a user roams between multiple, individual environments, how does the environment obtain the relevant information (environment preferences, authentication and authorisation details, available applications) about that user?

Another advantage with using a set of traversable, individual environments is it allows ”fences” to be created within the Macro Intelligent Environment. Using a university as an example, the entire campus would be the Macro Intelligent Environment, M. Each department within that university could contain 2-3 of their own individual environments, IE. A student who studies Mathematics probably wouldn’t gain any advantage of using an environment in the English department, so it logically follows that the Mathematics student can only access the environments associated with his department; but you still want to keep that technological transparency and continuity of experience between the environments he has authorisation to use.

This creates ”virtual” large scale intelligent environments, within an existing large scale intelligent environment. The beauty of this solution is that it can keep going down as many layers as is required. These ”virtual” large scale intelligent environments augment the work started in [7]. The internals of these environments are not perfect either; the technology currently used to expose the assets within the environments has a large margin to improve[25]. While these are important problems that need to be addressed (especially if these environments are to be used in real time computing), they are out of the scope of this paper.

So far, we have talked about Macro Intelligent Environments from a *management* (top down) point of view, but there is another vital perspective that needs to be addressed; the user’s perspective. The model proposed for Macro Intelligent Environments is very dynamic; individual environments can be introduced and removed with relative ease. With the ability to create departments, as mentioned above, it also means that each user’s perspective of the environment may differ. This also introduces some interesting insights into the way security management would work for MIE; the traditional role based security or user based security models may not fit, forcing an entirely new model to be created.

4. Proposed Solutions

Generally, in computer science, the word macro is used to define an input pattern (usually abbreviated) that will create a larger, more complex output[26]. However, the word *macro* is from the greek ($\mu\alpha\kappa\rho\sigma$) for "big" or "far" and it is an amalgamation of these definitions we used when creating "Macro Intelligent Environments". A Macro Intelligent Environment is a composition of several smaller Intelligent Environment instances that are integrated to form the whole. This means each member of a Macro Intelligent Environment is autonomous and self governing, providing an abstraction to the Macro Intelligent Environment. This is in contrast to the Monolithic Intelligent Environment that would attempt to manage all the low level details for all the spaces from the top down. It is possible to create an abstract outline for all the proposed solutions:

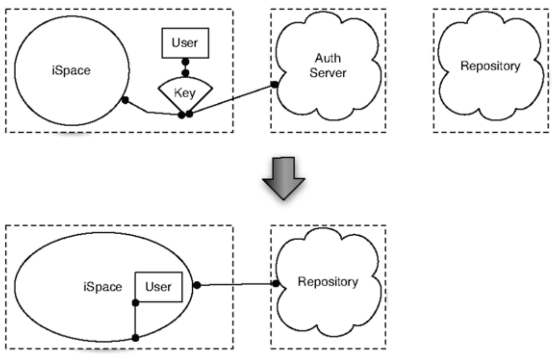


Figure 2. The key is used by the user to grant access to the Intelligent Environment

The similarities between all the solutions are such that once a user has gained access to the space, their user profile is acquired from *somewhere*. Although not explicitly mentioned, security is a vital role in Macro Intelligent Environments, thus we have used a "key" to represent this security layer. This key could be any appropriate security solution: something you know (username/password), something you have (RFID tag) or who someone is (biometrics)[27,28]. There will need to be another server which contains these authorisation and authentication. Figure 2 shows that the user requires some form of key to gain access to the space (and the use of that key disappears once the user has gained entry, as shown in the reaction rule in Figure 2). Figure 2 looks very similar to Figure 3 presented in *4.1* - however one fundamental difference is the repository isn't explicitly defined.

In order to create our Macro Intelligent Environments, there are a plethora of suitable solutions available, this paper proposes three solutions to the problem:

4.1. Cloud Server

Cloud computing is a marketing term used for the delivery and/or consumption of computing services over a network (usually the internet). With cloud computing the end user isn't aware of the implementation of the server but is merely concerned with the delivery of the service. With this solution, the user information will be stored on the server, and upon successful user authentication, the environment will attempt to retrieve the user's profile from the cloud server. This method still requires the user to carry around a trusted device / token of some description but it would only be used to gain entry to the environments. As shown in Figure 3, the notion of some form of key is still required to gain entry to the environment, but upon successful entry, the reaction rule shows that the key and authentication server play no further role. Each individual component in Figure 3 is shown to be contained within separate "regions"[21] but it is possible for them to be contained within the same region.

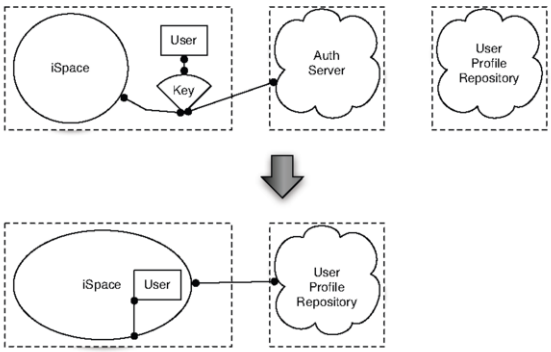


Figure 3. User 1 enters the iSpace and it immediately communicates with the Centralised User Repository.

Take this example: Jenny is a lecturer at the University of Essex, she is currently working in the iSpace, preparing her next lecture. The time approaches for said lecture so she leaves the iSpace; everything she has been working on is being uploaded and stored in the cloud as she walks to the iClassroom. When she arrives at the iClassroom, she gains entry by waving her RFID tag over the reader, prompting her profile be downloaded from the cloud, which recognises her status as the teacher and automatically adjusts the lights in the classroom for teaching mode. She pulls up the lecture slides she was previously working on in the iSpace and is almost immediately prepared for the students to arrive in the classroom.

This scenario immediately presents some potential issues that will need to be resolved; if a connection cannot be established to the server, is the user not granted access? Does the environment load a blank profile locally and attempt to synchronise at given intervals? The advantage of using formal methods to model this is we can anticipate for this sort of problem, and hopefully come up with a suitable solution during the modelling stage (rather than hacking around at code in the last minute!).

This description of the cloud server sounds like a simple file host, serving and storing files when requested but this would be severely underusing the power of server side technology available. We envision the server actually providing services to the user, which would be available in most, if not all, of the environments that user has access to. These services would be synonymous with the services provided by the space itself; the user would be unaware of what was providing the services, just that the services are available for consumption. One such function is a messaging service; take the previous example of our lecturer, Jenny. Fellow lecturer Ingrid wishes to contact Jenny immediately but doesn't know where she is. The server would know whether Jenny was currently active in an Intelligent Environment, and providing Ingrid was also in a similar environment, she could send a message to Jenny via the server; acting as a simple routing host. This is similar to how instant messaging traditionally works but illustrates how the cloud server could be used as more than just a file repository. This leads onto the concept of ones entire presence following them around the Macro Intelligent Environment, their preferences, their documents, communications...the possibilities are endless.

4.2. Trusted Device

A trusted device is a device that the user would carry around with them the majority of the time. The most obvious example would be a smartphone, and seeing how smartphones allow the deployment of powerful applications, it makes them suitable to play this role. This trusted device would contain all the relevant information about the user on the device itself; upon successful user authentication, the environment would then retrieve the information from the device. This method has the advantage of not required an internet connection (or a complex network); thus being ideal for remote locations where a strong internet connection isn't available (e.g the International Space Station).

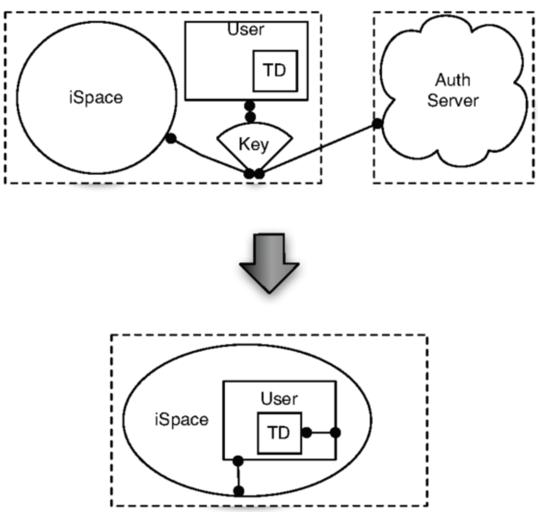


Figure 4. User 1 enters the iSpace and it immediately communicates with the Trusted Device to learn about the user.

Figure 4 shows the user still requires a key to gain entry to the IE (subsequently also requiring the server architecture for authentication); however, with the advances in mobile technology (particularly Near Field Communications) it may be possible to combine the trusted device and key. Once the user has been granted access to the space then no interaction occurs externally from the intelligent environment (as shown in the reaction rule of Figure 4).

We cut back to our lecturer Jenny. Again, she is working in the iSpace, ready for her lesson in the iClassroom. This time when she leaves the iSpace for the classroom, the information is stored on her trusted device (which in this case is her smartphone). She is *"logged out"* of the iSpace when her trusted device is out of range of the environment (either the environments local area network or some location system, such as ubisense). On the way to the classroom, Jenny makes changes on the trusted device that affect her profile. When she comes into range of the classroom, the trusted device automatically connects to the network, sends a handshake message, if accepted, she will be logged into the space and granted access. As before, the information will be pulled off the device (including the new changes) and she will be ready to teach.

As before, this presents some potential issues that will need to be resolved. In an ideal situation, every member would have the exact same trusted device but in reality the fragmentation of smartphone devices is somewhat prevalent in the consumer market so it is of utmost importance that we keep the core logic written in a way that is supported by the majority of devices available, thus resulting in only having to develop a unique presentation layer for the different devices.

4.3. Hybridisation of A. and B.

Subsequently, comparing the two previous approaches to determine the most appropriate situation in which to implement each solution, could lead to a series of scenarios where one solution outshines the other and vice versa. Following on, the idea of a hybrid solution could solve most of the potential issues raised by each individual solution.

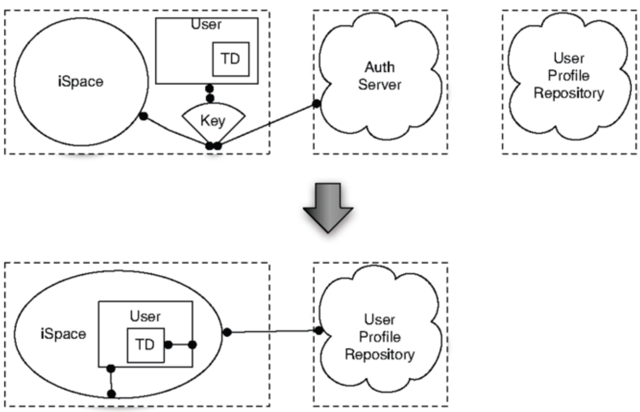


Figure 5. User 1 enters the iSpace and it immediately communicates with the Trusted Device to learn about the user, then sends this information to the Centralised User Repository.

As with Figures 2-4, the user requires a key to gain entry to the intelligent Environment. Figure 5's reaction rule shows that once the user is authenticated that the space has an active connection to both the trusted device and the Centralised User Repository.

The hybridisation of the previous two solutions provides a more complete system as it inevitably cuts out some of the problems presented. The idea behind this approach is to use the trusted device to store the user's profile and the server to also store the user's profile (as well as other files that may be too large to fit on a trusted device). When the user authenticates, the profile is pulled from the trusted device and any subsequent changes are saved to both the device and the server.

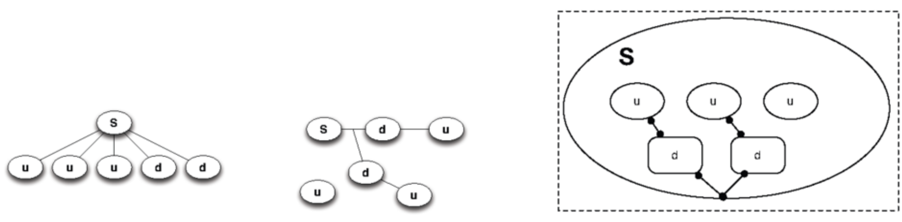
5. Introduction of Formal Methods

When designing a system, it is typically good practice to at least sketch a design first. For example, when designing a C++ program, you may use UML diagrams to show the class inheritance structure of the program. This gives you an insight into the properties of the system you're designing; how all the components interact, and wherein any potential problems may arise. The advantage of formal models is they usually give substantial evidence that these properties will hold.

Usually, Formal Methods are only used in the development of high-integrity systems (due to the associated cost with using Formal Methods)[29]. With Macro Intelligent Environment's vast size, it isn't only a case of designing a framework but essentially it is an entire eco-system being created; this requires extensive planning and modelling to ensure the proposed solution is going to work. Obviously, with a system of this size, not every problem can be anticipated at the design stage but it worth the time taken to do so.

Fragmentation of research into Intelligent Environments was mentioned earlier; different research groups have very differing opinions on what an Intelligent Environment actually compromises. Due to this, there hasn't been much research into creating a formal model for Intelligent Environments. This raises questions on whether the research will be into the formal methods of Macro Intelligent Environments or into the middleware; but that is synonymous with asking whether programming is about UML or software development. It is the entire model that we are trying to capture, at both the design stage and the implementation stage too.

The concept of using Bi-graphs as a formal framework for description, design and analysis is discussed in [30], though the ideas have been around for longer. As the name suggests, Bi-Graphs consist of two graphs; a *place graph* and a *link graph*. These two graphs share nodes. the place graph is restricted to a tree (no cycles), whereas a link graph tends to be a hyper-graph (a link can connect more than two objects)[31]. By representing both components of the bi-graph in one picture, we can get an impression of object locations and connections simultaneously[7].



(a) (b)

Figure 6. This figure shows the underlying bigraphs ((a) and then the result of representing both the graphs in one picture ((b).

Bi-graphs offer a simple understanding of newly introduced concepts and show possibilities for specifying behaviours at more than one level of abstraction. All the figures shown in this paper are themselves Bi-graphs. While Bi-Graphs are a high level, visual formal model, it is possible to break them down into their algebraic form (however this is not within the scope of this paper). To give an example of their use, figure 7 shows

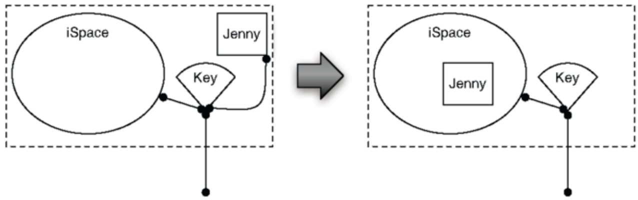


Figure 7. Jenny uses the keypad to gain entry into the iSpace. The key persist as it contains a link outside of the region.

our lecturer Jenny transitioning from outside an Intelligent Environment to inside, using a persistent key (which in this case, is a keypad on the wall next to the door).

You'll notice the dotted line in Figure 7, this represents a region. It allows the Bi-Graph to give a notion of locality for individual components (this case, Jenny approaching the iSpace and using the keypad on the wall of the iSpace to gain entry). Figure 8 shows that you may have multiple regions contained within the same bi-graph, which in turn means that the place graph will contain a forest of tree graphs, one for each region.

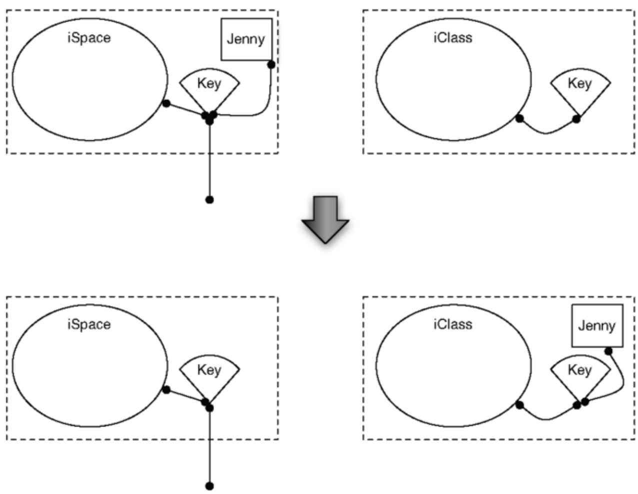


Figure 8. Jenny leaves the iSpace, walks over to the iClass and gets ready to use her one-time-use key to gain entry to the iClass.

Figure 8 shows the scenario explained *4.1* where Jenny is preparing a lecture in the iSpace, then physically travels to the iClass; which is geographically located elsewhere from the iSpace. We have only scratched the surface on using bi-graphs as a formal method of modelling Intelligent Environments but there is far more use than modelling the simple descriptions contained within this paper. It ranges from using a trace graph to predict how a system will react depending on which order reaction rules take place to using non-determinism to capture the complex details of a reaction rule. The use of bi-graphs as a formal model for Intelligent Environments is introduced in further detail in [32].

6. Conclusion and future work

It is the natural progression for these Intelligent Environments to start increasing in size, and this paper presents the one of the major problems with expanding the existing implementations. The solutions presented are only *theoretical solutions*, and while initial experimentation work has started (in the form of a paper published in the 2012 IEEE International Conference on Fuzzy Systems[14]), there is still much work to undertake implementing these solutions and experimentally verifying the outcomes. An in depth comparison of the approaches will also need to be undertaken, in order to fully appreciate with solution works best, and in what scenario. One interesting outcome will be whether the hybrid approach is a *"one-size fits all"* solution or whether the other approaches outshine it in certain situations. The connection between the theoretical practice of formal methods in Intelligent Environments and the practical implementation based off these models is a vaguely explored area, yet is proving to be an exciting one. We are hoping to use this work as a springboard to create interest and a community around this topic in order to further explore it.

Acknowledgments

This work has been undertaken as part of the ScaleUp project, which is funded by King Abdulaziz University, KSA

References

[1] L. Roalter, A. Moller, S. Diewald, and M. Kranz. Developing intelligent environments: A development tool chain for creation, testing and simulation of smart and intelligent environments. In *7th Internation Conference on Intelligent Environments (IE'11)*, Nottingham, 2011.

[2] W. Weber, J. Rabaey, and E. Aarts. *Ambient Intelligence*. The Netherlands: Springer-Verlag Berlin Heidelberg, 2005.

[3] J. Dooley, M. Henson, V. Callaghan, H. Hagras, D. Al-Ghazzawi, A. Malibari, M. Al-Haddad, and A. Al-Ghamdi. A formal model for space based ubiquitous computing. In *7th International Conference on Intelligent Environments (IE11)*, Nottingham, 2011.

[4] D. Lopez-de Ipina, A. Almeida, U. Aguilera, I. Larizgoitia, X. Laiseca, P. Orduna, A. Barbier, and J.I. Vazquez. Dynamic discovery and semantic reasoning for next generation intelligent environments. In *4th International Conference on Intelligent Environments (IE08)*, Seattle, USA, 2008.

[5] R. Bose and A. Helal. Distributed mechanisms for enabling virtual sensors in service oriented intelligent environments. In *4th International Conference on Intelligent Environments (IE08)*, 2008.

[6] A. Kameas. Towards the next generation of ambient intelligent environments. In *In proceedings of 19th IEEE International Workshop on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE)*, pages 1–6, 2010.

[7] J. Dooley, M. Davies, M. Ball, V. Callaghan, H. Hagras, M. Colley, and M. Gardner. Decloaking big brother : Demonstrating intelligent environments. In *6th International Conference on Intelligent Environments (IE10)*, Kuala Lumpur, 2010.

[8] M. Ball, V. Callaghan, and M. Gardner. An adjustable-autonomy agent for intelligent environments. In *6th International Conference on Intelligent Environments (IE'10)*, Kuala Lumpur, 2010.

[9] M. Cohen. Towards interactive environments: The intelligent room. In *Proceedings of the 1997 Conference on Human Computer Interaction*, Bristol, U.K., 1997.

[10] H. Hagras, V. Callaghan, M. Colley, G. Clarke, A. Pounds-Cornish, and H. Duman. Creating an ambient-intelligence environment using embedded agents. *IEEE Intelligent Systems*, 19(6), 2004.

[11] B. de Ruyter and E. Aarts. Ambient intelligence: visualizing the future. In *In proceedings of the Working Conference on Advanced Visual interfaces (AVI04)*, pages 203–208, 2004.

[12] P. Bull, P. R. Benyon, and P. R. Limb. Residental gateways. *BT Technology Journal*, 20(2):73–81, 2002.

[13] Y. Jin, R. Wang, H. Huang, and L. Sun. Agent-oriented architeceutre for ubiquitous computing in smart hyperspace. *Wireless Sensor Network*, 2(1):74–84, 2010.

[14] A. Bilgin, J. Dooley, L. Whittington, H. Hagras, M. Henson, C. Wagner, A. Malibari, A. Al-Ghamdi, M. Alhaddad, and D. Alghazzawi. Dynamic profile-selection for zslices based type-2 fuzzy agent controlling multi-user ambient intelligent environments. In *International Conference on Fuzzy System*, Brisbane, Austrailia, 2012.

[15] S.T. Shivappa, M.M. Trivedi, and B.D. Roa. Audiovisual information fusion in human-computer interfaces and intelligent environments: A survey. *Proceedings of the IEEE*, 98(10):1692–1715, 2010.

[16] C. Ressel, J. Ziegler, and E. Naroska. An approach towards personalized user interfaces for ambient intelligent home environments. In *2nd International Conference on Intelligent Environments (IE'06)*, Athens, Greece, 2006.

[17] M. K. Habib. Collaborative and distributed intelligent environment merging virtual and physical realities. In *In proceedings of 5th IEEE International conference on Digital Ecosystems and Technologies (DEST)*, pages 340–344, Daejeon, Korea, 2011.

[18] M. Jianhua, L.T. Yang, B.O. Apduhan, H. Runhe, L. Barolli, and M. Takizawa. A walkthrough from smart spaces to smart hyperspaces towards a smart world with ubiquitous intelligence. In *Proceedings 11th International Conference on Parallel and Distributed Systems*, Fuduoka, Japan, 2005.

[19] D. Cook and S. Das. Pervasive computing at scale: Transforming the state of the art. *Pervasuve and Mobile Computing*, 8(1):22 – 35, 2012.

[20] K. Niinuma, P. Unsang, and A.K. Jain. Soft biometric traits for continuous user authentication. *IEEE Transactions on Information Forensics and Security*, 5(4):771–780, 2010.

[21] T. Hildebrandt and J. Winther. Bigraphs and (reactive) xml - an xml-centric model of computation. Technical report, The IT University of Copenhagen, 2004.

[22] J. Dooley, C. Wagner, H. Hagras, and G. Pruvost. Followme: The persisent gui. In *1st International Workshop on SituatedComputing for Pervasive Environments (SCOPE-2011) at 6th International Symposium on Parallel Computing in Electrical Engineering (PARELEC-2011)*, Bedfordshire, UK., 2011.

[23] A. Ranganathan, C. Shankar, and R. Campbell. Application polymorphism for autonomic ubiquitous computing. *Multiagent Grid Systems*, 1(2):109–129, 2005.

[24] S. Pardha, M. Tungare, J. Holman, and M. Perez-Quinones. Continuous user interfaces for seamless task migration. In Julie A. Jacko, editor, *In Proceedings of the 13th International Conference on Human-Computer Interaction. Part III: Ubiquitous and Intelligent Interaction*, pages 77–85. Springer-Verlag, Berlin, Heidelberg, 2009.

[25] J. Dooley. *An Information Centric Architecture for Large Scale Description and Discovery of Ubiquitous Computing Objects*. Thesis, (phd), University of Essex, 2010.

[26] Irwin D. Greenwald and Maureen Kane. The share 709 system: Programming and modification. *journal of the ACM (New York, NY, USA: ACM)*, 6(2):128–133, 1959.

[27] Federal Financial Institutions Examinations Council. Authentication in an internet banking environment. `http://www.ffiec.gov/pdf/authentication_guidance.pdf`, 2012.

[28] S. Creese, M. Goldsmith, B. Roscoe, and I. Zakiuddin. Authentication for pervasive computing. *Proc. Security in Pervasuve Computing 2003*, 2803:116–129, 2004.

[29] M Archer, C. Heitmeyer, and E. Riccobene. Proving invariants of i/o automata with tame. *Automated Software Engineering*, 9:201–232, 2002.

[30] R. Milner. *The Space and Motion of Communicating Agents*. Cambridge University Press, 2009.

[31] Denmark IT University of Copenhagen. A brief introduction to bigraphs. `http://www.itu.dk/research/pls/wiki/index.php/A_Brief_Introduction_To_Bigraphs`, 2012.

[32] M. Henson, J. Dooley, L. Whittington, and A. Al-Ghamdi. Towards simple and effective formal methods for intelligent environments. In *8th International Conference on Intelligent Environmnets (IE'12)*, Guanajuato, Mexico, 2012.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-191

Collaborative Agents Framework for the Internet of Things

Priscila Angulo-Lopez[a,1] and Guillermo Jimenez-Perez[a]
[a]*Computer Science Department, ITESM, Mexico.*

Abstract. The trend towards implementing applications supporting the “Internet of Things” (IoT) concept is allowing the discovery of amazing new opportunities that current technologies could bring to us. Previous efforts have shown the challenges yet to be solved. Research for increasing the capabilities of the IoT is being conducted under various paradigms and analyzing different approaches for finding useful ways to implement valuable applications under the IoT concept. This paper proposes a framework based on a MultiAgent Systems Paradigm, which is an accepted form for modeling and implementing applications that are strongly related real-world concepts. The idea behind the proposal is to simplify the development of applications in the IoT which could perform complex tasks in a distributed environment.

Keywords. Internet of Things; IoT; MultiAgent Systems; framework; agents

Introduction

The Internet started as a network of computers only oriented to work realization, this concept and use have been changing in the past years, in part, caused by its expanded boundaries (e.g., the bandwidth, the device processing power, the storage capacity) and prices reduction. The Internet has become more accessible, common and expanded. The next step in its evolution, is marked by the trend of shifting away from the use of personal computers as the primary access tool towards mobile devices, like smart phones, tablets, notebooks, etc., which are also fitted with sensors and actuators with the ability to perceive and sense the real world, and even more, physical objects could be tagged to be sensed by all this devices. This new environment with the ability of having an intelligent participation of all kind of devices is increasing the reachability of the Internet and taking us to the new Internet, the “Internet of Things” (IoT) [1].

The original concept of the Internet of Things was coined in 1998 by Kevin Ashton of the MIT Auto-ID Center. The concept served to describe the attachment of RFID tags to objects in order to use them as pointers to Internet databases, which contained information about the objects with the tag added [2]. By now, an accepted definition of the IoT, is “Things having identities and virtual personalities operating in smart spaces using intelligent interfaces to connect and communicate within social, environmental, and user contexts” [3], that means a network of all kind of devices with sensing, actuating and connecting capabilities, devices that could be anywhere and serve for

[1] Corresponding Author. Priscila Angulo-Lopez , Computer Science Department, ITESM, Monterrey, Mexico; Email: A00813460@itesm.mx

creating a digital interpretation of our world. It is possible to observe how the original concept has changed and now it also includes requirements for intelligence, which opens a new world of possibilities for the development of applications under this concept. As it sounds very exciting, it also reveals all the challenges that need to be affronted in order to reach this vision in a safe and productive way.

Few years ago, in 2005, the ITU (International Telecommunications Union) presented a report, explaining the feasibly and truly taking place of the IoT [4]. One of the main contributing factors was the widespread use of gadgets which have become an integral part of everyday life for millions of people, and which is promoting the main growing of this technology. These gadgets are allowing new forms of communication, not only between people, but also between people and things and between things themselves. This envisioned network is more heterogeneous, more space expanded and more interconnected. Fig. 1 shows how a network of things (sensors and actuators), distributed along the space, could support a variety of applications, like for Vehicular control, Power supply control, Security vigilance and so on.

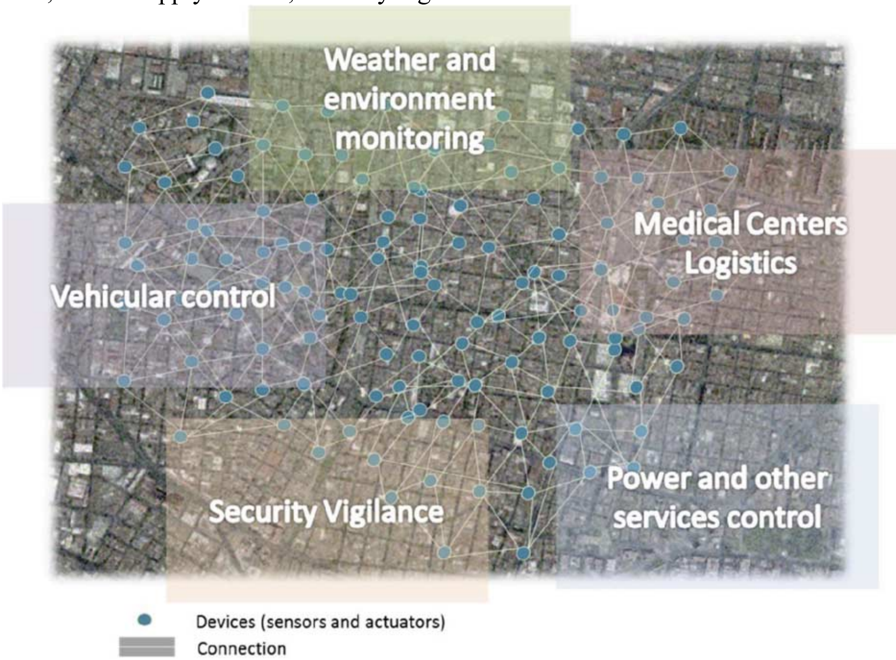


Figure 1. Applications of large distributed things

The IoT applications have special requirements to work properly. One of them is the mechanism for the identification and discovery of devices. Another is the acquisition of measurements and data from the devices, its processing and translation to a context of useful information, actuation over the environment, and finally there is an obvious need for security. There are, likewise, the requirements related to the scalability of this new kind of network, affected by a broad variety of devices that could be connected, its amount and the volume of data that needs to be transported and processed. There is not yet an approach that addresses all the requirements, therefore here we propose one that could fulfill them.

1. Related work for accomplishing IoT requirements

One of the proposals is to adapt the Service Oriented Paradigm to the IoT, thus resulting in an architecture which suggests "abstract things as services", which allow a loosely coupled relationship between things and improves the reusability level [5]. In this model, the things are not accessed directly by other things, instead, there is a middleware which manages this process, and this middleware has the responsibilities of receiving the requests and then find all the "services" available which could supply the request, select one service for solving the request, retrieve the response and deliver it to the requester. This middleware works with a Knowledge Base of the services and provides a better way to find the candidates that satisfy a request. The authors of this model explain that the graph of possible IoT services is not exact, but it is approximately optimal, also the discovery of matching constraints for services is probabilistic, the main issue is to find the services that have the more required characteristics for solving the request. The core of this solution relies on the Knowledge Base, composed by three ontologies: Device Ontology (repository of devices), Domain Ontology (models of real world entities through their physical concepts) and Estimation Ontology (web of relationships between models, devices and physical concepts). This proposal considers some of the current and future necessities for the IoT. However, a centralized approach like this could represent a disadvantage when the environment is completely distributed, which is the case of the IoT.

A distinct approach for building the IoT is abstracting the devices involved as Smart Objects [6]. In this approach, a smart object includes abilities for sensing, processing, and act. Every smart object is autonomous and includes a piece of application logic, providing the capabilities of non-centralized processing. On this approach the objects collaborate with the nearest objects in order to provide services to persons. There are three types of smart objects interacting in this model:

1. Activity aware objects. These objects understand the world in terms of their use for performing a certain activity. Its representation is formed by the accumulation of activities. This kind of objects doesn´t has interaction with other objects but is able to log data. The benefits of these objects is that they are able to monitoring themselves, they are aware of their history, status and other properties that could facilitate their maintenance and provide useful information to persons.
2. Policy aware objects. These objects understand the activities in which are involved and also know the policies that must be followed for their correct use, its representation consist in a set of rules. This type of objects is interactive and can provide context sensitive information, principally alerts about trespassing use policies.
3. Process aware objects. These objects understand the organizational processes which they perform and are able to relate the real world activities to these processes, their representation is a workflow model with context information, their level of interaction is the highest one because they are in charge of providing context guidance about the tasks, deadlines and decisions to the persons involved in the process.

On this approach the services provided by the devices became critical for the realization of the activities, even if the approach sounds more oriented to the development of business activities, it can also be applied to other environments where some degree of intelligence could be delegated to the devices.

Similarly to the last one, there is another proposal which depicts a smart environment composed of "smart artifacts", these artifacts are common objects (like simple furniture) dotted with sensors, a processor, network capabilities and actuator devices. In this scenario the artifacts are aware of themselves and also can be aware of other artifacts who belong to the same environment or have the same context. Herein every artifact has its own logic and goals, but also they collaborate in order to accomplish a goal given by the user, every artifact shares a database with their peers in order to request the development of a subgoal to one of them and be able to achieve a main goal [7].

2. A Framewok for the IoT

In order to propose a Framework for the IoT it is first necessary to conceptualize what a framework is. Some definitions are listed next:

- "Software frameworks are a form a software reuse that primarily promotes the reuse of entire architectures within a narrowly defined application domain. […] An important distinctive feature of software frameworks is that they do not simply supply individual components but also provide the interconnections among the components they supply. […] Frameworks can be seen as generative devices, namely as starting points for the construction of individual applications in the framework domain. The process whereby a concrete application is derived from the framework is often called framework instantiation. […] A framework should be constituted of three types of constructs: abstract interfaces, concrete components and domain-specific design patterns. […] Abstract interfaces consist of declarations of sets of related operations for which no implementation is provided. […] The framework components can be of two types: core components and default components. The core components encapsulate behaviors that are common to all applications in the framework domain. […] Default components represent default implementations for some of the abstract interfaces in the framework. They encapsulate behaviors that are not intrinsic to the framework domain. […]" [8, pp. 29-35].
- A Framework is a way to facilitate working with complex technologies, forces the implementation of consistent code, provides a structure and facilitates reuse [9].
- "A Framework should hide implementation details of the underlying system, to make it abstract enough – only properties interesting in the application domain should be left visible in the fixed part and the variable part. The variable part must be variable enough to allow for useful instantiations, but the variability should be quite limited, to make sure that components built for the framework will work in all. […] It is clear that any framework must have a clearly defined application domain. Otherwise, the variable part must be very large, causing the framework to be hard to create, understand, and test" [10].

There is also a motivation for proposing a framework and no other kind of approach; a framework facilitates the implementation of real world applications, the components of a framework tend to keep a strong adherence to the real-world concepts and enable encapsulation. The use of frameworks allows defining components that cooperate to provide main functions and it is also open to plug new components if it is necessary. Working under this approach also provides a way for dealing with the security issues and error handling. The frameworks could provide a system wide vision and define expected standards for the system [11, pp. 125-126].

Given the definitions and the motivation, a Framework for the IoT must undertake the duties related to define an application domain, give the component structure for developing applications in the framework domain and prove to be reusable. The expected application domain consists of any environment wherein there exits multiple sensors and actuators that could be grouped by a processing unit with the required logic and abilities for interaction, measuring and acting. Every one of these groups will have the capabilities for interacting with another unit within a context an also learn about his actions and their results. The component structure of the framework will be presented later.

The paradigm utilized for our proposal of a Framework for the IoT is the MultiAgent Systems Paradigm. Like Wooldridge did envision, the trend of computing is shifting to represent the real world, and "software Agents" are an excellent way to represent the "things" depicted by the IoT [12]. An intelligent Agent is self-aware, environment-aware and capable of ma king choices to interact with the world and change it in order to accomplish their goals, and also collaborate with other Agents to realize major tasks. An Intelligent Agent has some characteristics that are suitable for the requirements of the IoT [12] [13]:

- Perception. Ability to perceive and describe the state of the environment.
- Reactivity. Ability to react to the changes in the environment.
- Proactiveness. Ability to take the initiative in order to accomplish its goal.
- Social Ability. Interact with other agents, including humans.
- Autonomy. Operate by own decision and control its actions.
- Learn. Ability to gain knowledge based on the results of its actions.

Even if a centralized system has the advantage of controlling the environment, know its capabilities, and assure the delegation and accomplishment of tasks, it is clear that it will be hard to maintain when the amount of devices connected reach a large number, and there is also the risk that a failure on the centralized system could bring down all the operations related to the connected devices. Consequently it is valid to consider an approach that keep the distributed nature of the IoT, allow more scalability, and also provides control and security capabilities to the environment where it works. A framework based on collaborative agents to develop applications for the IoT could achieve these competencies [14, pp. 253-268]. The relationships between the environment and the agents in an application based on the IoT would be like the shown in Fig.2, in which it is possible to observe how in an environment could exist multiple contexts given by the human perception. The context gives the meaning of a defined situation and those contexts could vary according to the situation and the requirements of the application. Every agent in the environment would be associated to a single context at a time, but the association could be changed to other context. Every Agent will have the abilities for taking measures or sensing the environment and perform actions on it. Agents would also be able to interact between them, but this will be limited by their communication capabilities, their collaborative behavior and the requirements for accomplishing the given goals.

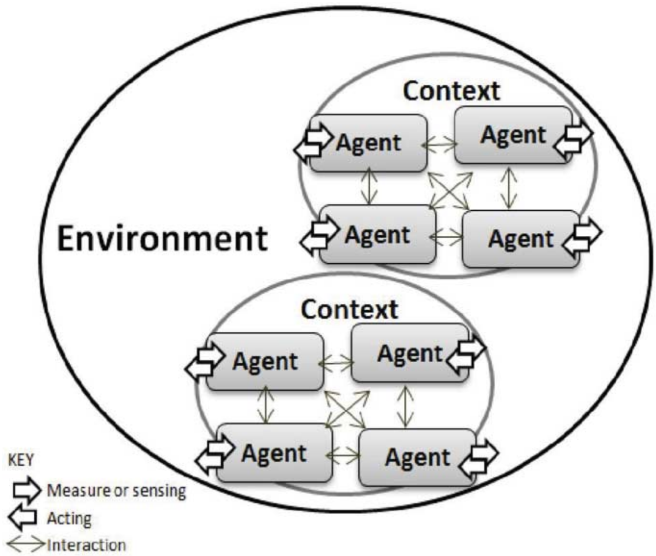


Figure 2. Agents interaction within the environment.

3. Brief description of the Agent based Framework Components

Fig. 3 shows the components of the proposed framework which covers the aspects presented in previous sections. The component behaviors and interactions will be described next:

- Knowledge. Herein is located all the knowledge gained and given to the agent. This component has two subcomponents identified:
    - Context. This subcomponent will have the data related to the real-world concepts to which the agent is associated to. This component is essential to find the peers of the agent and enable the interactions and collaboration among them. It is also important for the interpretation of the data retrieved from the environment and the other peers.
    - Actions learning. The agent will have a memory of how his actions affected the environment, how did a task get accomplished; the state of the environment before and after a task realization, and also, there will be memory about the interactions realized and the results of them.
- Discovering. The capabilities for identifying the devices that compose the agent itself and the capabilities for recognizing a peer will be implemented in this component.

  - Peers discovering. The capabilities for detecting other peers will be configured in this component. Every agent will have a context specification describing its context concepts assigned (which are provided by the Knowledge component). This information will allow the identification of other peers representing concepts related to his context.
  - Device plugging. An agent's resources will be its connected devices. The agent must register the drivers and meta-data related to the devices, their use policies and constraints. All this information and configuration will be in this subcomponent.
- Security assurance. Security verification will occur when a device is plugged, when a peer is discovered or when a decision is taken, before performing it. This component implements the logic that will protect the integrity of the agent and assure his safe operation and interactions.
- Decision Making. In this component relies the responsibilities and goals of the agent. Based on the sensing of the environment and the knowledge acquired, the agent must be able to take decisions in order to accomplish its goals and perform its tasks.
- Data processing. In this component the data retrieved from the environment will be interpreted and transformed to be kept as knowledge and also to provide it to the Decision Making component. Respectively, the decision made in the Decision Making component will be transformed to instructions for its actual realization in the environment.
- Actions. This component will implement the ways of interaction with the environment. In this component could also be defined the means for presenting data to human users. The identified interactions are the next:
  - Interactions with peers. This subcomponent must address the collaborative requirements, including the controlling elements for shared tasks and goals. In the proposed scenario won't exist a central controller in the environment, the management of collaboration will have to rely on every agent. An option could be that for every shared goal one of the agents will assume the leadership role and responsibilities to ensure the achievement.
  - Actions performance. This subcomponent will implement the actions over the environment.
  - Environment measures. This subcomponent will implement the functionality for measure gathering.

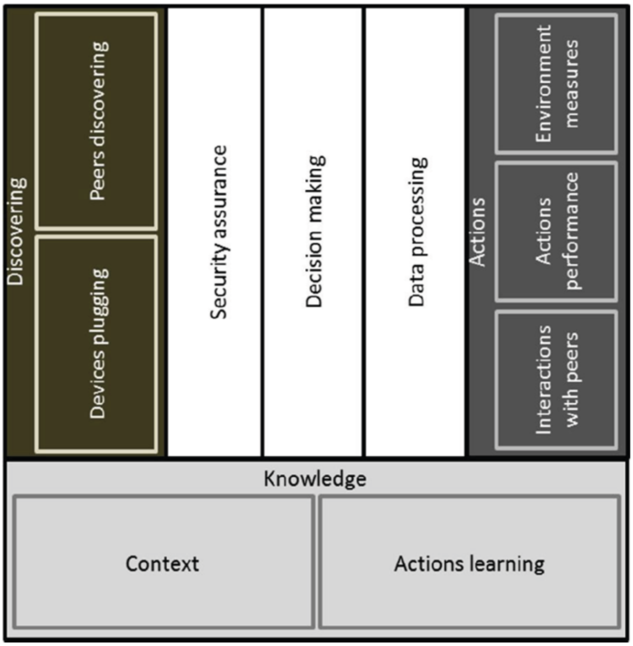


Figure 3. Framework based on Collaborative Agents for the IoT.

4. Conclusions and Future Work

The advent of possibilities for developing new applications that involve the capacities of the available devices, network technologies and processing capabilities present an opportunity for researching ways for development that could serve as starting point for reaching the desired vision of the IoT. The proposed framework includes in its components the capability to support identified requirements for an IoT application. There is still the need for detailing every component and give them the intelligent characteristics that will be needed and useful in a smart and changing environment. Given that, the future work will consist on defining and providing the strengths to every component, validate them, integrate them, implement the framework, and finally prove its utility and advantages.

Acknowledgment

This work is being supported by the Monterrey Tech CAT-145 "Autonomous Agents in Ambient Intelligence" research chair.

References

[1] L. Coetzee y J. Eksteen, «The Internet of Things – Promise for the Future? An Introduction,» de *IST-Africa*, Gaborone, Botswana, 2011.

[2] A. Serbanati, C. M. Medaglia y U. B. Ceipidor, «Building Blocks of the Internet of Things: State of the Art and Beyond,» de *Deploying RFID - Challenges, Solutions, and Open Issues*, InTech, 2011, pp. 351-366.

[3] European Commission, Information Society and Media Directorate General, «Internet of Things: an early reality of the Future Internet,» European Commission, Prague, 2009.

[4] International Telecommunication Union, «ITU,» November 2005. [In line]. Available: http://www.itu.int/osg/spu/publications/internetofthings/InternetofThings_summary.pdf. [Visited on: January 2012].

[5] S. Hachem, T. Teixeira y V. Issarny, «Ontologies for the Internet of Things,» de *8th Middleware Doctoral Symposium*, Lisbon, Portugal, 2011.

[6] G. Kortuem, F. Kawsar, V. Sundramoorthy y D. Fitton, «Smart Objects as Building Blocks for the Internet of Things,» *IEEE Internet Computing,* vol. 14, nº 1, pp. 44-51, 2010.

[7] S. W. Loke, «Building Intelligent Environments By Adding Smart Artifacts to Spaces: a Peer-to-Peer Architecture,» de *Sixth International Conference on Intelligent Environments*, Kuala Lumpur, Malaysia, 2010.

[8] A. Pasetti, Software frameworks and embedded control systems, New York, NY: Springer, 2001.

[9] M. Clifton, «What is a Framework?,» 03 November 2003. Available at: http://www.codeproject.com/Articles/5381/What-Is-A-Framework. [Accesed on: 01 April 2012].

[10] R. Land, «How can frameworks facilitate component reuse?,» Artech House, 2002.

[11] M. Norris, R. E. Davis y A. Pengelly, Component-based Network Systems Engineering, Norwood, MA: Artech House, 2000.

[12] M. J. Wooldridge, An Introduction to MultiAgent Systems, New York, NY: John Wiley & Sons, 2009.

[13] T. Sánchez López, A. Brintrup, M.-A. Isenberg y J. Mansfeld, «Resource Management in the Internet of Things: Clustering, Synchronisation and Software Agents,» de *Architecting the Internet of Things*, New York, NY, Springer, 2011, pp. 159-193.

[14] I. Gorton, Essential Software Architecture, Springer, 2006.

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-200

Intelligent Signaling for Prevention of Intersection Collisions in Urban Zones

Juan Antonio GUERRERO-IBAÑEZ[a,1], Carlos FLORES-CORTES [b], Juan Manuel RAMIREZ-ALCARAZ [c], Hector VIZCAINO-ANAYA [d], Tomás MENDOZA-ROBLES [e], Alvaro ANGUIANO-MANCILLA [f] and Emmanuel PEÑA-CARDENAS [g]
[a] *Telematics Faculty, University of Colima, México*

Abstract. Safety is one of the most important concerns in the Intelligent Transportation Systems (ITS). One of the most relevant research areas in safety is the development of Collision Warning Systems (CWS). One goal of the research groups is to take the initiative to decrease the number of collisions occurring on the intersections. In urban zones, traffic lights are utilized to control the traffic flow on some intersections. However, it is not possible to place traffic lights on all intersections due to the enormous economical costs. This paper describes ISAPCI (*Intelligent Signaling Approach for Prevent Collisions at Intersections*), an ITS based on vehicular ad-hoc networks (VANET) to mitigate collisions at intersections, and to regulate the traffic flow at intersections that are not controlled by traffic lights.

Keywords. Algorithm, collisions, intersections, VANET, V2I, V2V.

Introduction

In recent years, the number of vehicles has been increasing drastically due to rapid urbanization. The technical report of the *United Nations Population Fund* showed that for the first time, more than half of the world's population, around 3.3 billion people live in urban areas. By 2030, this number is expected to swell to almost 5 billion [1]. In this sense, critical traffic problems such as accidents and traffic congestion require the development of new transportation control systems. To date, different *Intelligent Transportation Systems (ITS)* are being developed to improve the safety, security and efficiency of the transportation systems and to foster the development of novel vehicular applications [2].

Vehicular Ad-Hoc Network (VANET) is a very important componet of the *ITS, and o*ne of the most significant opportunities of research in this area is the development of safety applications. Researchers have defined two types of safety applications: passive and active applications. Passive safety applications, work inside the vehicle protecting passengers against injury in the event of an accident. Active safety applications are designed to prevent accidents and work as pre-crash application [3]. Intersection collision avoidance is amongst the top five components of safety applications [4].

[1] Corresponding Author: Juan Antonio Guerrero-Ibañez, Telematics Faculty, University of Colima, Av. Universidad #333, Colonia las víboras, Colima, Col., Mexico; E-mail: antonio_guerrero@ucol.mx.

Complex intersections are controlled by traffic lights, where drivers cross when they have the right of way. However, implementing traffic lights on all intersections is not the best solution because of the high economical cost and because in zones with high rates of traffic, traffic congestion problems may arise. At intersections where there not exist traffic lights, a waiting driver has to judge when it is safe to cross the intersection based on the speed and distance of approaching vehicles. One of the main problems of this type of intersection occurs when the driver's view at a given junction or intersection is obstructed as shown in figure 1.

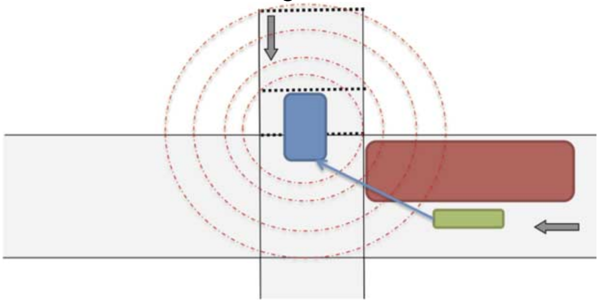


Figure 1. A possible collision scenario.

In this scenario the truck (red rectangle) is obstructing the view of the driver in the blue car, and as a consequence, the driver is not aware of the approaching motorcycle (green rectangle) on the other side of the road.

In this paper we present a method of early warning that helps to automatically manage traffic flow in intersections where there are not traffic lights installed. For our case of study we designed an algorithm based on zones that uses exchange of warning messages between nearby vehicles to alert drivers of possible collisions. We implemented a fast prototype based on mobile mini-robots and propose two possible solutions: i) the Vehicle-to-Vehicle (V2V) approach and ii) the Vehicle-to-Infrastructure (V2I) approach.

The remainder of this paper is organized as follows. Section1 describes the related work. In section 2 we briefly describe our case study. The proposed solutions are described in section 3. Preliminary results of our work are presented in section 4. Finally, section 5 summarizes our work and provides directions for future research.

1. Related work

Different efforts have been made to decrease the number of accidents, however with today's technology it is still very difficult to achieve this objective. In the last years, research and development work has been carried out to overcome road safety problems by developing driver assistance systems that use technologies such as sensor networks, *radio-frequency identification* (RFID) readers, video cameras, among others to model the traffic state that surrounds the vehicle in order to warn the driver in case of a potential dangerous situation [6], [7].

Collision-warning systems focus on providing warning information to drivers based on contextual information of other vehicles such as speed, direction, and distance among others. Some of the related work in this area relies on the use of digital maps, which are used as support for determining a vehicle's position. There exist a tendency for fitting Global Positioning Systems (GPS) and navigation systems in vehicles; however, the information of these systems may not always be available or it may be out of date. Thus, ed an alternative way for managing traffic at intersections to prevent collisions is necessary.

One of the most comprehensive works in this area is the cooperative collision warning system developed in the context of the PATH project [8]. Other papers e that investigate the propagation of messages in a multihop mode or utilize opportunistic relay messaging in vehicular networks are available in [9] – [11].

2. Case study

For this research work we envisage a scenario where vehicles travel across an urban area where intersections such as cross streets, side streets, alleys, freeway entrances and roundabouts among others are present. To effectively prevent accidents a vehicular safety application must be capable of dealing with all types of intersections. To accomplish this it is required: i) a mechanism for gathering traffic information such as number of lanes, direction of each lane, number of vehicles waiting on each lane, destination lane or route of each vehicle, etc., and ii) an algorithm to manage traffic flow i.e. determine how many vehicles on a given lane should cross the intersection on a given moment.

As seen, a solution to this problem is not trivial. Firstly, because of the number of sensors, algorithms, devices and infrastructure necessary for gathering all traffic information and secondly, because of the safety, security and privacy issues that must be considered.

This paper focuses on investigating a case study where vehicles must be capable of automatically driving across intersections as the one depicted in figure 2. On this type of intersections two uni-directional lanes converge. This type of intersection is also referred as a T intersection. When arriving at an intersection the algorithm must manage traffic by indicating each vehicle its right-of-way. The right-of-way is alternated between lanes allowing one vehicle at a time to cross the intersection. We investigate: i) an infrastructure-less approach, where vehicles communicate directly with each other to determine who has the right-of-way and ii) an infrastructure-based approach, where an intermediary device installed at a given intersection informs vehicles of its right-of way.

3. The intelligent signaling approach

The Intelligent Signaling Approach for Preventing Collisions at Intersections (ISAPCI) proposed in this paper automatically manages the right-of-way of vehicles at intersections.

The proposed mechanism allows drivers to be aware of approaching vehicles that are out of their current line of sight (LOS). With this feature, drivers are prevented of possible collisions when arriving at an intersection. Our algorithm actively warns drivers as they approach to an intersection. Thus, instead of stopping at an intersection to see who has the right-way, the mechanism automatically informs vehicles if they must stop or continue.

Our solution controls traffic flow at intersections by sending messages to warn drivers of the possibility of a collision when another vehicle in a neighboring lane is approaching to the intersection at a same time.

The proposed algorithm is based on the concept of zones as shown in the figure 2. These zones are delimited by a set of sensors, in our case these sensors are RFID tags located on the side of the road. It is assumed that vehicles are equipped with a RFID reader and a wireless communication device, which allows exchange of information with other vehicles or traffic control base stations. The algorithm defines four zones. The way in which these zones are defined and utilized is detailed in the following subsections.

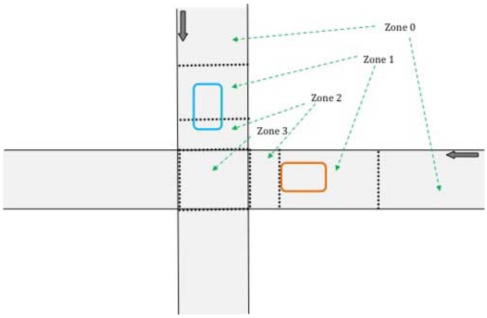


Figure 2. The concept of zones for the proposed approach.

Communication between vehicles is achieved utilizing a wireless network architecture. Two operational modes are considered: V2V and V2I.

3.1. Frame format

In our approach, a series of messages are exchanged between vehicles and/or base stations depending on the communication mode. A warning message consists of several fields, which are summarized as follows. The *Type* field refers to the type of message (request/response/warning, etc). The *CarID* field refers to the source address of the node; this identifier is unique for each vehicle. The *ZoneID* field indicates the zone identifier where the vehicle is located and the *Priority* field represents field stores the priority of the road where the vehicle is moving. The *DestID* field contains the destination address for the frame, depending on the communication mode this field can contain a broadcast address or a unicast address. The *Data* field contains the actual information of the frame. Finally, the *Direction* field is utilized for storing the direction of movement of the vehicle.

3.2. V2V algorithm

In the case of the V2V architecture, approaching vehicles to an intersection exchange warning messages and based on the type of received message the driver decides either to stop the vehicle or continue driving to cross the intersection. Figure 3 shows the pseudocode for the V2V algorithm.

```
// Pseudo code V2V
BEGIN:
C1 reads the RFID tag
 If (zoneID = 0) then C1 is in listening mode
   If (zoneID = 1) then
```

```
      C1 broadcasts a request message and switches to listening mode
      If (C1 receives a response message) Then
        If ((ZoneIDreceived > ZoneID_C1) or (ZoneIDreceived = ZoneID_ C1 and
          Priorityreceived > Priority_ C1))
              C1 stops and sends a notification message to the other vehicle
        EndIf
      Else If (C1 does not receives a response message) and (ZoneID=2)
              C1  crosses the intersection.
      EndIf
      If (zoneID=3)
         C1 broadcasts a warning message
      EndIf
END
```

Figure 3. Pseudo code for V2V architecture.

The algorithm works as follows. When a vehicle C1 reads the first RFID tag, C1 updates theits ZoneID field with the value of 0.. At zone 0 vehicles passively listen for incoming messages from nearby vehicles. If a message is received while at this zone, the vehicle will stop when it arrives to zone 2, otherwise it crosses the intersection. When the vehicle is in zone 1, it starts advertising its presence periodically announcing other vehicles its proximity to the intersection. If the priority of a received message is the same as its own priority, it means that both vehicles are traveling on a same lane and therefore the information is discarded. If the priority of the received message is higher , then, the other vehicle has the right of crossing the intersection.. C1 waits for the crossing confirmation, which is sent when the other car has crossed zone 3, at this point C1 waits or reduces its speed, and also continues broadcasting messages to other vehicles indicating that they must wait as one vehicle is crossing the intersection. Once C1 receives the crossing confirmation it can continue its trip crossing the intersection. If C1 does not receive any message while in zone 2 then it indicates that there are not vehicles nearby and it can cross the intersection saeftly.

3.3. V2I algorithm

In the case of V2I, an embedded system installed at each intersection (e.g. on the stop sign) controls traffic flow.. Figure 4 shows the pseudo code for the embedded system and for the vehicle.

```
// Pseudo code V2I for embedded system
BEGIN:
TC is in listening mode
 if(TC receives a request message) then

  save data and address in queue Q

  if (length of Q == 1) then
    send a response message with the GO

    code

  else
    send a response message with the
    STOP code

  end if
   else
```

```
// Pseudo code V2V for vehicle
BEGIN:
C1 reads the RFID tag
If (zoneID = 0) then
  C1 is in listening mode
  If (zoneID = 1) then
      C1 broadcasts a request message
         and switches to listening mode
  end if
  If (C1 receives a response message)
     Then
   If (Data == stop)
      C1 stops and waits for another
      response message
   end if
   If (Data == go)
       C1 does not stops and continues.
   Else
       If (C1 does not receives a

        response message) and (ZoneID=2)
```

(a)	(b)

<pre> if (length of Q > 0) then
 	
   read Q[0]
   send a response message to the
    address of Q[0]
   switch to listening mode
  end if
  if (TC receives a confirmation
   message) then
    update queue Q
  end if
END</pre>
	<pre> C1 sends a warning message
	
	        and passes the intersection.
	    If (zoneID=3)
	       C1 sends TC a confirmation
	       message
	    Endf if
	   End if
	END</pre>

Figure 4. Pseudo code for V2I mode.

The traffic controller (TC) starts, in the listening mode. If TC receives a request message from a vehicle (C1) then TC stores the CarID in a queue Q. If the length of Q is equal to one then TC sends a response message to C1. This message contains the GO code that indicates that the vehicle must continue. If the length of Q is greater than one then TC sends a response message with the STOP code to C1. If TC does not receive any message then TC reads the first element of the queue and sends a response message with the GO code and switches to listening mode. If TC receives a confirmation request then TC deletes the address of the vehicle and continues in the listening mode.

Figure 4b shows the algorithm that controls each the vehicle. When the vehicle (C1) reads a RFID tag, C1 updates the ZoneID value. At zone 0 vehicles listen for incoming messages to detect nearby TC. If zone 1 is detected then C1 starts the transmission process sending a request message .If C1 receives a response message, and it contains a STOP code then C1 must reduce the speed or stop., If the message contains a GO code then C1 does not stops and continues its trajectory. If C1 does not receives a response message, it waits for the crossing confirmation which is sent when the car that arrived first passes to the zone 3, at this time C1 remains stopped or reduces its speed, but also remains broadcasting.. When the crossing confirmation is received then this car keeps moving. If C1 does not receives any message when it arrives to zone 2 then it indicates that there are not vehicles nearby and can cross the intersection in a safe way.

4. Preliminary results

To evaluate the algorithms described in Section 3, a scenario with mini car-robots was implemented. The car-robots utilized are manufactured by the Moway Company. These robots have several sensors like anti-collision, infrared, light intensity and temperature, among others. They have an expansion bus to connect some modules to expand their functions, e.g. a WiFi interface that allows to control and program them wiresly.

As expected, the algorithms perform well and accomplished their function completely. Some faults, due to sensors precision were observed, e.g. sometimes mini cars were deviated of their road because of the infrared sensor did not detected obstacles; or lost messages synchronization due to bad sensing of dark lines on the road that were an important base to correct performance of the algorithms. However, most of the time, the sensors were working well enough and allowed to observe that in normal conditions the proposed solution resolve the problem with no ambiguity. Finally, figure 5 shows the implemented scenario with the mini-robots.

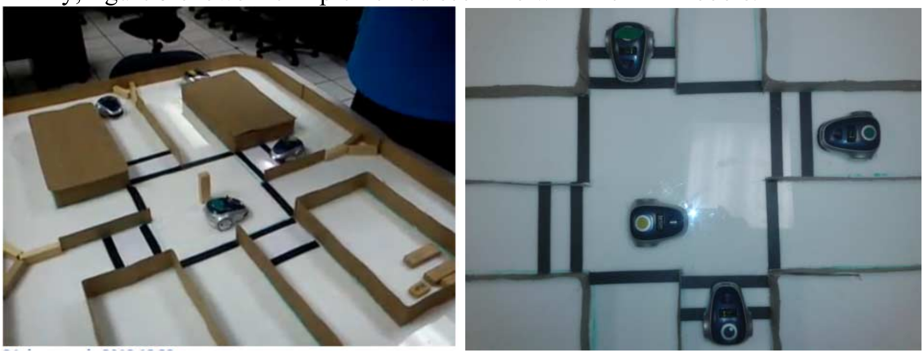


Figure 5. Our fast prototype with four mini-robots.

Mini-robots detect the different zones and broadcast warning messages trying to detect a possible vehicle in the neighboring lane. The radio-frequency devices allow mini-robots the exchange of information with other vehicles or traffic control stations. In the case of V2V, approaching mini-robots to the intersection exchange warning messages and based on the type of received message the mini-robot decides either to stop or continue and cross the intersection. In the case of V2I, a mini-robot acts as a traffic light and controls the traffic in the intersection. Our fast prototype allowed us to verify the efficiency of our proposal. We can observe how through the exchange of message the algorithms for both V2V and V2I modes could prevent possible collisions at intersections.

5. Conclusions

This paper has explored the concept and design of ISAPIC, which is an Intelligent Signaling Approach for Prevention of Intersection Collisions. Our approach is based on the two communication modes of VANET: V2I and V2V. The approach uses the zone concept. In this model, the streets are divided into zones that are identified through RFID tags.

These zones are used to provide more accurate motion state estimation of approaching vehicles into an intersection, and intelligent signaling algorithms are proposed for prevention of possible collisions upon received information from neighboring vehicles. Collision-detection algorithms are then designed based on the position of the vehicles, and warnings messages are given upon different criteria. Our testbed results show that our approach outperforms the referenced approach in successful warning ratio. Our future work will be focused on the evaluation of more complex scenarios as mentioned in the study case section.

Acknowledgment

This work has been supported by the Programa de Mejoramiento al Profesorado (PROMEP) in Mexico with reference PROMEP/103.5/11/1025. (1)

References

[1] Y. Qian and N. Moayeri, "Design of Secure and Application-Oriented VANETs", Vehicular Technology Conference, pp. 2794-2799, Singapore, 2008.

[2] United Nations Population Fundation, "Technical Report: State of World Population 2007: Unleashing the potential of urban growth", New York: Retrieved from http://www.unfpa.org/swp/2007/english/introduction.html, 2007.

[3] S. Yousefi and M. Fathy, "Metrics for performance evaluation of safety applications in vehicular ad hoc networks", Transport. Vilnius: Technika, 2009, vol. 23, No. 4, pp. 291-299.

[4] National Highway Traffic Safety Administration, "Vehicle Safety Communications Project Task 3 Final Report, Identify Intelligent Vehicle Safety Applications Enabled by DSRC", DOT HS S09 S59, Washington, D.C., March 2005.

[5] Car 2 Car Communication Consortium. Use case: Approaching motorcycle warning. Retrieve from http://www.car-to-car.org/, 2011.

[6] Doshi, A., Shinko Yuanhsien, C., & Trivedi, M. M. "A novel active heads-up display for driver assistance". IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics, 39(1), 85–93. doi:10.1109/TSMCB.2008.923527, 2009.

[7] Bertolazzi, E., Biral, F., Da Lio, M., Saroldi, A., & Tango, F. "Supporting drivers in keep- ing safe speed and safe distance: The SASPENCE". Subproject Within the European Framework Programme 6 Integrating Project PReVENT. IEEE Transactions on Intelligent Transporta- tion Systems, 11(3), 525–538. doi:10.1109/ TITS.2009.2035925, 2010.

[8] R. Sengupta, S. Rezaei, S. Shladover, D. Cody, S. Dickey, and H. Krishnan, "Cooperative collision warning systems: Concept definition and experimental implementation", Intelligent Transportation Systems Journal, vol. 11, no. 3, pp 143-155, July 2007.

[9] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading structure for randomness in wireless opportunistic routing," in Proc. SIGCOMM Conf. Appl., Technol., Architectures, Protocols Comput. Commun., 2007, pp. 169–180.

[10] S. Biswas and R. Morris, "ExOR: Opportunistic multi-hop routing for wireless networks," in Proc. SIGCOMM Conf. Appl., Technol., Architec- tures, Protocols Comput. Commun., 2005, pp. 133–144.

[11] A. Balasubramanian, R. Mahajan, A. Venkataramani, B. N. Levine, and J. Zahorjan, "Interactive wifi connectivity for moving vehicles," in Proc. ACM SIGCOMM Conf. Data Commun., 2008, pp. 427–438.

[12] R. Popescu-Zeleti, I. Radusch, & M. Rigani, "Vehicular-2-X communication", Berlin, Germany: Springer. doi:10.1007/978-3-540- 77143-2.

This page intentionally left blank

1st International Workshop on Intelligent Domestic Robots (iDR'12)

Simon Egerton (Monash University, Australia)

Eduardo Morales (National Institute for Astrophysics, Optics and Electronics, México)

This page intentionally left blank

J.A. Botía et al. (Eds.)
IOS Press, 2012

Introduction to the Proceedings of the First International Workshop on Intelligent Domestic Robots (iDR'12)

Welcome to the proceedings of the 1st International Workshop on Intelligent Domestic Robots (iDR'12)

6.5 million robots were in operation in the world in 2007, by 2012 this number has reached over 18.0 million. A growing significant proportion of these robots are domestic service robots which are quickly becoming an everyday part of our lives. This workshop provides an opportunity for the research community to discuss the latest research and future directions in the area. The papers appearing in our first workshop were chosen for their research contributions and to stimulate and provoke discussion. The papers share a common theme, human robot interaction, a key area for critical research input if domestic robots are to work and live alongside us in our homes.

Our keynote speaker, Ricardo Téllez, gives an inspiring address, arguing for the case of strong AI and the need to adopt cognitive approaches if we are ever going to be truly successful in developing robust service robots.

The first paper in the workshop presents a framework for human pose and gesture recognition using a light scanner, the paper takes a very methodical experimental approach and reports a very comprehensive set of experimental results which examines the effectiveness and impact of different classification methods.

The second paper provides a discussion of key issues relating to human robot interaction and deals with appearance, spatial and psychological aspects of robots occupying our living spaces.

The Microsoft Kinect sensor has rapidly found a place within the robotics community since its launch towards the end of 2010. The third paper in this workshop reports on a framework that uses the Kinect to detect static gestures. Our final presentation reports on work carried out to enable robust face detection and tracking in crowded environments, this gives us pause to consider the times our homes, which are otherwise sparsely populated with people, fill with occasion and large numbers of people. What would your domestic service robot make of this?

Lastly the organisers would like to thank all the authors who contributed their work and to the Program Committee for their help to make our first workshop possible.

Simon Egerton (Monash University, Sunway Campus, Malaysia) and
Eduardo Morales (National Institute for Astrophysics, Optics and Electronics, Mexico)

12th May 2012

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-212

Cognitive Service Robots

Ricardo TELLEZ[1]
PAL Robotics, C/Pujades, 77-79 Planta 4º Puerta 4ª
08005 Barcelon, Spain.

Keynote Abstract: Current approaches to intelligent service robots follow an idea borrowed from Good-Old-Fashioned Artificial Intelligence (GOFAI): very complex planning algorithms will produce artificial intelligence. Hence is basically a matter of CPU power to have an operative artificial intelligence for a robot. An example of this belief is the current trend in robotics that uses the cloud as the brain of the robots. It is assumed that the cloud will have CPU power enough to handle object recognition, speech recognition, or grasping. However, current experiments in speech recognition indicate that despite the increase in CPU power along the last years, the speech recognition rate has reached a plateau (not good enough for a real life system).

I will argue that this brute force approach is preventing us from attacking the hard problem of artificial intelligence, that is, make a robot understand. Cloud robotics will not lead to the kind of intelligent robots we need for a human environment. Instead, the solution goes by having robots that are more cognitive (whatever it means), being understanding the main cognitive ability they need. I will define (or not!) what a cognitive robot is, and why this type of cognition is required for robots that have to live on a dynamic human environment.

Brief Biography: Ricardo Téllez, holds a PhD on Artificial Intelligence by the Technical University of Catalonia. His thesis was devoted to the control of complex robots using neural networks, and how those robots can create their own concepts about the world.

He has been working for more than five years at company Pal Robotics developing service robots. He developed the navigation systems of human size humanoid robot Reem-B, making it the first human size humanoid robot able to move autonomously. At present, he develops better navigation systems for humanoid robots in crowded environments, allowing them to move safely for both the robot and people.

He is especially interested on the development of service robots that can really understand their environment.

[1] Corresponding Author: Ricardo Téllez, Pal Robotics, Spain. E-mail: ricardo.tellez@pal-robotics.com

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-213

Real-Time Human Pose and Gesture Recognition for Autonomous Robots Using a Single Structured Light 3D-Scanner

Tim VAN ELTEREN and Tijn VAN DER ZANT
Cognitive Robotics Laboratory, Department of Artificial Intelligence, Faculty of Mathematics and Natural Sciences, University of Groningen, Groningen, The Netherlands

Abstract. We propose a method for real-time human pose and gesture recognition for autonomous robots using a structured light 3D-scanner. Poses are recognized using skeleton representations by performing classification using the Nearest Neighbour algorithm. The whole-body pose recognition approach uses the joint coordinate data from the processed depth images. The quality of the classification is determined by 10-fold cross validation in which the recognition rate is 99.9028%.

Keywords. machine learning, artificial intelligence, pose recognition

Introduction

Robust and realtime tracking of a person's body has applications in many domains such as human-computer interaction interfaces, telepresence and monitoring for security and healthcare. Another challenging and interesting application, which will be the focus of this paper, is the field of Human-Robot Interaction (HRI) for autonomous robots. The introduction of realtime depth aware cameras has made this challenge somewhat easier but the state of the art systems still have their limitations.

The recent introduction of depth cameras using a structured light 3D-scanner approach, such as the Kinect Sensor System[1], brings realtime human pose recognition at consumer prices.

This paper covers the recognition of custom defined poses and gestures by an autonomous robot as part of a multimodal HRI system to autonomously perform tasks within the RoboCup@Home competition[2]. The RoboCup@Home competition [1] is an international benchmark for domestic service robots. It aims to develop service and assistive robot technology with high relevance for future

[1] Microsoft Corp. Redmond WA. Kinect for Xbox 360
[2] `http://www.ai.rug.nl/robocupathome/`

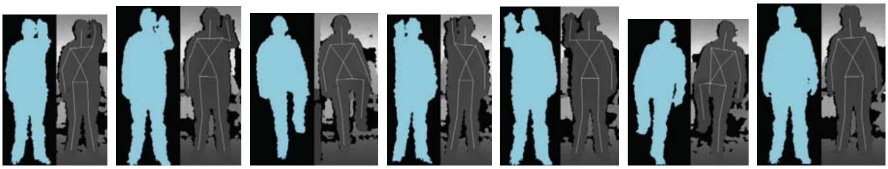


Figure 1. Seven examples of the set of seventeen classes of poses

personal domestic applications [2]. The performance and abilities of the robots in the competition are benchmarked using a series of tests. These tests all take place in a realistic non-standardized home environment that does not contain any artificial markers [3].

Natural interaction methods without the use of artificial markers are of special importance and relevance because of the applications for use of robots in real world domestic environments. The multimodal HRI system consisting of the combination of a speech recognition and markerless gesture recognition system is an integral part of the behavior-based architecture [4] [5] that has been developed at the Cognitive Robotics Laboratory of the University of Groningen[3].

Human pose estimation is an active area of research that has delivered a vast amount of literature surveyed in [6] and [7]. The advances made by [8] in real-time identification and localization of body parts from depth images and the research performed by [9] in 3d model based tracking approaches for human motion capture in uncontrolled environments show an object recognition and respectively a modeled approach. The vision based motion capture and analysis described in [10] performs real-time motion capture using a single time-of-flight camera.

The research performed by the Microsoft Research Cambridge and Xbox Incubation on real-time human pose recognition in parts from single depth images [11] forms the basis for our research on the recognition of human poses for natural HRI with an autonomous robot.

The depth images received from a structured light 3D-scanner such as Kinect are the result of an algorithm that performes dense 3D image acquisition using structured light [12] with a pattern of projected infrared points. The deformation of a speckle pattern projected on the scene, with respect to a reference pattern, reveils information about the distance of the objects and results in a calibrated depth mapping of the scene.

1. Methods

3D joint position data of a test subject is used to classify the pose. The performance of a total of twenty-six machine learning algorithms are used to perform classification on the dataset to perform pose recognition. Ten-fold cross validation is performed to determine the quality of the classification.

[3] `http://www.ai.rug.nl/crl/`

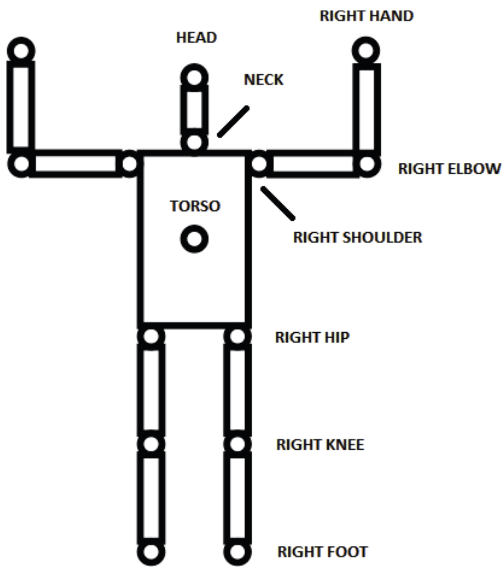


Figure 2. The circles in the skeleton represent the fifteen degrees of freedom or joints used for the classification of poses

1.1. Data

In our research we use the Kinect sensor system with the PrimeSense OpenNI[4] framework and the NITE[5] middleware which gives us access to the 3D-coordinates of the joints position of a calibrated test subject. Using this setup we create a dataset for machine learning that consists of the test subjects in a number of natural poses. The data that the poses consist of are fifteen 3D coordinates as shown in figure 2.

The current implementation uses the position data of fifteen joints. This implementation draws the corresponding skeleton over the depth map of the scene as can be seen in figures one, three and five. The poses used for training are shown as a person segmented on the left and the skeleton overlay over the depth image on the right.

The seventeen classes that are defined and used for classification cover a wide range of natural human poses. The complete set of classes are shown in figure one, three and five. Classification was performed using the WEKA[6] machine learning toolkit. The approaches used and their resulting performance are listed in table 1. Initial training of the classifier was performed with the joint data pose representations of over 26000 poses from four different test subjects.

In a follow up experiment the training of the classifier was performed with the joint data pose representations of 46863 poses from fifteen different test subjects. Building the dataset takes about fifteen minutes per person resulting in a total time to build the larger dataset of about four hours.

---

[4]`http://www.primesense.com/en/openni`
[5]`http://www.primesense.com/en/nite`
[6]`http://www.cs.waikato.ac.nz/~ml/weka/`

Figure 3. Five examples of the set of seventeen classes of poses

In addition to the aformentioned datasets a random subset consisting of 1% and 10% of the dataset consisting of 46863 poses was used for training to determine which algorithms are the most suitable for online implementation based on the execution time of the 10-fold cross validation.

1.2. Machine learning

The data is processed by a number of classification algorithms. The five best performing algorithms are shown in figure 4. A selection of the benchmarked machine learning algorithms are covered in more detail in the following subsections.

1.2.1. 1-Nearest Neighbour

The Nearest Neighbour algorithm [13], a type of instance based learning, uses past data instances, with known output values, to predict an unknown output value of a new data instance. Normalized Euclidean Distance was used as distance measure with a k value of 1.

1.2.2. Random Forest

The Random Forest consists of an ensemble of Tree-based classification algorithms in which the best performing classifier is selected and used for classification of the test data. The default parameters are used.

1.2.3. Random Tree

The Random Tree algorithm is a variant of the REPTree algorithm, Reduced Error Pruning Tree, which in turn is a type of Tree classifier. The default set of parameters were used and already showed sufficient performance.

1.2.4. J48

The J48 algorithm [15], also known as C4.5 algorithm, is an algorithm used to generate a decision tree which in turn is used for classification. Performance of the algorithm is tested using the default parameter set.

1.2.5. Voting Features Interval

Classification by Voting Features Interval [14] is performed in which a concept is represented by a set of feature intervals on each feature dimension separately. Each feature participates in the classification by distributing real-valued votes among classes. The class receiving the highest vote is declared to be the predicted class.

Figure 4. Five examples of the set of seventeen classes of poses

1.3. Ten-Fold Cross Validation

On all the algorithms in Table 1 10-fold cross validation is performed to test the accuracy and quality of the classification. k-fold cross validation is a technique in which the original sample is randomly partitioned into k subsamples. Of the k subsamples, a single subsample is retained as the validation data for testing the model, and the remaining k − 1 subsamples are used as training data.

The cross-validation process is then repeated k times (the folds), with each of the k subsamples used exactly once as the validation data. The k results from the folds are averaged to produce a single estimation. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and validation, and each observation is used for validation exactly once.

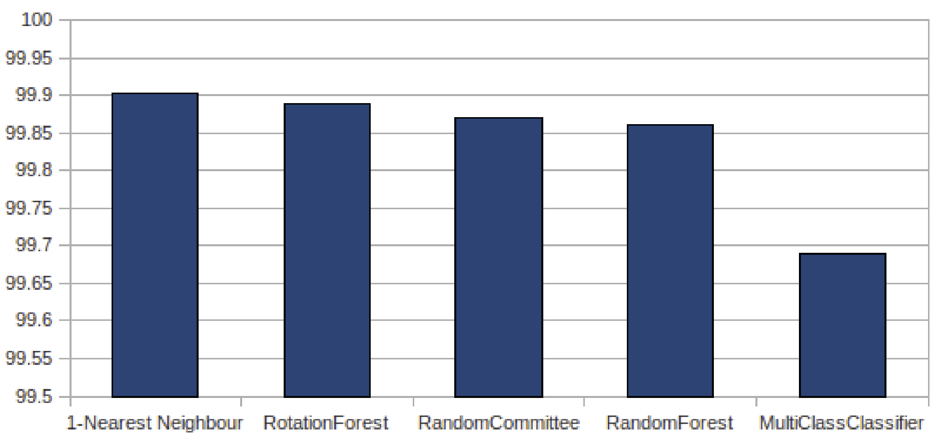


Figure 5. Performance of a selection of best performing machine learning algorithms by their performance and execution time

2. Results

The results from the experiments as shown in table 1 show state of the art performance in classification of the poses. The 1-Nearest Neighbour algorithm with euclidean distance shows the best performance with 99.9028% correct classification of the poses in approx. 165 seconds. The Random Tree algorithm takes some time in building a model but performs fast in 10-fold cross validation of approx.

10 seconds. It still has a high performance with over 98% correctly classified instances. The Voting Features Interval algorithm is fast in building a model, it only takes around .3 seconds, and in 10-fold cross validation which takes around 9 seconds. This however comes with a price: a lower performance of approx. 80 % correctly classified instances.

2.1. Experiments

The classification experiments performed with the twenty-six machine learning algorithms on the full dataset of 26000 poses from four test subjects ordered by their performance are shown in table 1.

Figure 6 shows the five algorithms that perform best using a 10% subset of the datset with respect to their execution time in 10-fold cross validation. These algorithms are the most suitable for online implementation because of their real-time behavior. Though a much smaller dataset is used for training the classifier, the performance of the algorithms, as shown in Table 2, does not drop below 77%.

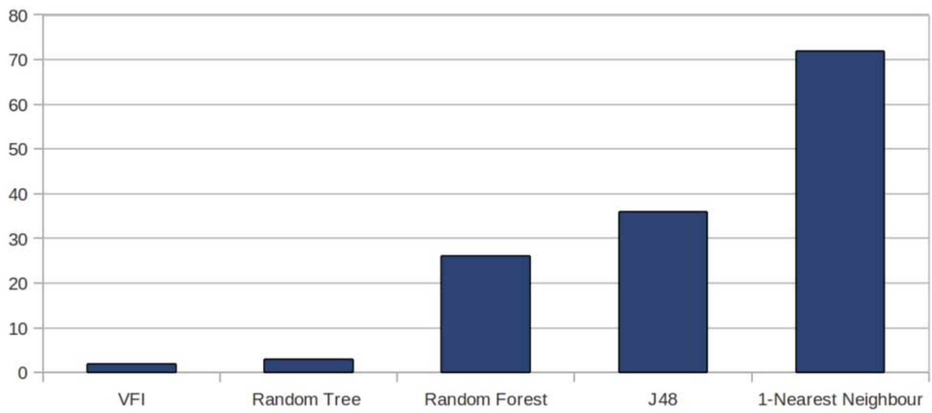


Figure 6. The five best performing algorithms ordered by their execution time

3. Discussion

A method for human pose and gesture recognition for human-robot interaction in autonomous robots is proposed and tested. The Voting Features Interval and Random Forest algorithm are the most suitable for online implementation. The Nearest Neighbour algorithm shows the best performance but this comes with a trade-off: it needs a considerable amount of time to perform matching of the pose which might make it less suitable for online implementation.

Name	Performance	Build time	Cross validation	Error
1-Nearest Neighbour	**99.9028**	**0.01**	**165**	**0.0104**
Rotation Forest	99.8878	281.54	2374	0.0144
Random Committee	99.8691	19.78	205	0.0156
Random Forest	**99.8616**	**15.69**	**159**	**0.0175**
Multi-Class Classifier	99.6896	581.15	3751	0.2275
Classification Using Regression	99.4765	151.1	1623	0.0248
NNge	99.2895	100.66	478	0.0281
Random Sub-Space	99.2858	53.9	557	0.0253
J48	**99.1587**	**21.92**	**195**	**0.0299**
PART	99.1474	107.96	959	0.0305
J48graft	99.0764	28.42	268	0.0315
Random Tree	**98.9680**	**2.05**	**19**	**0.0339**
JRIP Tree	98.8969	195.21	1560	0.0343
REP Tree	98.8140	10.09	86	0.0351
Data-Near-Balanced ND	98.7286	39.4	577	0.0373
SMO (SVM)	98.7212	16.92	172	0.2176
Nested Dichotomies	98.6987	62	1485	0.0376
Class-Balanced ND	98.6352	37.2	429	0.0386
Logit Boost	97.6816	321.45	3608	0.0525
Raced-Incremental Logit Boost	96.9338	49.06	945	0.0525
Bayes Net	93.5161	20.97	216	0.0825
Filtered Classifier	91.4968	20.41	222	0.0866
VFI	**79.4264**	**0.59**	**9**	**0.2070**
Decision Table	74.9654	63.99	814	0.1645
Hyper Pipes	65.1610	0.1	3	0.2260
One-R	35.5345	1.71	15	0.2676

Table 1. Performance comparison of the twenty-six classification algorithms. The columns from left to right show the name of the classification algorithm, the percentage correctly classified instances, the time it takes to build the model in seconds, the time it takes to perform 10-fold cross validation in seconds and the root mean squared error.

Name	Performance	Build time	Cross validation	Error
Voting Features Interval	77.3416	0.05	2	0.2189
Random Tree	94.5381	0.28	3	0.0802
Random Forest	98.8479	2.64	26	0.0433
J48	95.9462	4.18	36	0.0678
1-Nearest Neighbour	99.0186	0.02	72	0.0340

Table 2. The five best performing algorithms on 10% of the dataset ordered by their execution time in seconds. The columns from left to right show the name of the classification algorithm, the percentage correctly classified instances, the time it takes to build the model in seconds, the time it takes to perform 10-fold cross validation in seconds and the root mean squared error.

3.1. Future work

Future work consists of the integration of a system that performs online human pose recognition for human-robot interaction. The first step towards such a system requires the development of an interactive pose training behaviour for the autonomous robot. A possible follow-up will be to apply confidence scores in the classification of poses and create a behaviour that combines the confidence scores from both speech and gesture recognition systems to enhance the performance of the overall human-robot interaction system.

A first step towards benchmarking the performance of the Nearest Neighbour algorithm in online gesture recognition is to implement it using the Approximate Nearest Neighbors algorithm.

References

[1] van der Zant, T. and Wisspeintner, T. (2006). Robocup x: A proposal for a new league where robocup goes real world. In [Bredenfeld et al., 2006], pages 166-172.

[2] Wisspeintner, T., van der Zant, T., Iocchi, L., and Schiffer, S. (2009a). Robocup@home: Developing and benchmarking domestic service robots through scientific competitions.

[3] Wisspeintner, T., van der Zant, T., Iocchi, L., and Schiffer, S. (2009b). Robocup@home: Results in benchmarking domestic service robots. In Proceedings of the XIth RoboCup symposium.

[4] Brooks, R. (1986). "A robust layered control system for a mobile robot". Robotics and Automation, IEEE Journal of.

[5] Arkin, R. C. (1998). Behavior-Based Robotics. MIT Press.

[6] T. Moeslund, A. Hilton, and V. Kruger. A survey of advances in vision-based human motion capture and analysis. CVIU, 2006.

[7] R. Poppe. Vision-based human motion analysis: An overview. CVIU, 108, 2007.

[8] C. Plagemann, V. Ganapathi, D. Koller, and S. Thrun. Real-time identification and localization of body parts from depth images. In proceedings of the ICRA, 2010.

[9] A comparison of 3d model-based tracking approaches for human motion capture in uncontrolled environments. M. Shaheen, J. Gall, R. Strzodka, L. van Gool, and H.-P. Seidel. MPI Informatik, Germany. BIWI, ETH Zurich, Switerserland IBBT, ESAT-PSI, K.U.Leuven, Belgium.

[10] Real Time Motion Capture Using a Single Time-Of-Flight Camera. V. Ganapathi, C. Plagemann, D. Koller, S. Thrun. Stanford University, Computer Science Department, Stanford, CA, USA.

[11] Real-Time Human Pose Recognition in Parts from Single Depth Images, IEEE Conference on Computer Vision and Pattern Recognition (2008) (2011), Volume: 2, Issue: 3, Publisher: Ieee, Pages: 1297-1304, J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, A. Blake, Microsoft Research Cambridge and Xbox Incubation

[12] Fofi, D., Sliwa, T., Voisin, Y., A comparitive study on invisible structured light.

[13] Cover, T., Hart, P., Nearest neighbor pattern classification, Information Theory, IEEE Transactions on, vol.13, no.1, pp.21-27, January 1967

[14] Demiröz, Gülşen and Güvenir, H., Classification by Voting Feature Intervals, Machine Learning: ECML-97, vol. 1224, pp.85-92, 1997.

[15] Quinlan, J. R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-221

Development of Means for Support of Comfortable Conditions for Human-Robot Interaction in Domestic Environments

Andrey RONZHIN[a, 1], Maria PRISCHEPA[b], Viktor BUDKOV[b]
[a] *St. Petersburg State University, Russia*
[b] *Institution of the Russian Academy of Sciences St. Petersburg Institute for Informatics and Automation of RAS, Russia*

Abstract. Humanoid shape and style of communication became more popular and possible for state-of-the-art social robot due to rapid development of the embedded systems and the methods for audiovisual processing of natural modalities. The detection of psychologically comfortable conditions for communication is a one of fundamental problems at the development of user interfaces for domestic robots. In this paper the modern works devoted to the psychological aspects of user-robot interaction, issues of the proximity and appearance of robot are considered. The heterogeneous sensors used for detection and tracking user as well as techniques for registration of audio-visual data used for extraction of user preferences, which are applied in the developed information mobile robot, are presented.

Keywords. Mobile domestic robots, spoken dialogue system, multimodal interfaces, user preferences, proximity, naturalness.

Introduction

The design of appearance and behavior of social robots similar to human style became popular recently. Such robots are able to interact with a person, not only through speech but also with gestures. Naturalness of the interface and effectiveness of the user-robot dialog depends on used modalities (speech, gestures, sensor input etc), and also on the robot appearance (naturalness, degree of detail and plasticity of the robot face, head and other robot parts). The considering of these psychological aspects of human-robot interaction is very important for the development of user interfaces for the domestic service robots.

One of the main tasks in the development of behavior strategies of social robots is the establishment of emotional contact between robot and human [1]. For example, human-like robots are capable to express the basic human emotional states by means of moving neck, eyelids, eyebrows, mouth and nose. Nevertheless, the number of freedom degrees of robot parts is minimized in the majority of cases in order to reduce the production costs [2].

[1] Corresponding Author: Andrey Ronzhin, St. Petersburg State University, 11, Universitetskaya nab., St. Petersburg, 199034, Russia; E-mail: ronzhin@iias.spb.su

Moreover, many developers raise the question of user-robot proximity and corresponding strategies for the choice of optimal robot position toward a user for more effective interaction. In psychology there is a notion of personal human space, which is characterized by four conventional zones [3]: (1) intimate zone (from 15 to 46 cm), people who are permitted to this zone are those with whom person is in a close emotional contact; (2) personal zone (from 46 cm to 1.2 m), where acquaintances are permitted; (3) social zone (from 1.2 to 3.6 m), which is mainly for strangers; (4) public zone (more than 3.6 m), where the communication with a large group of people takes place. In the majority of cases personal zone is most preferable for interaction between robot and human. At the same time the choice of robot-human distance is varied in different applications. Depending on robot assignment it is possible to classify the service robots into personal and public. Personal domestic robots are used for home assistance (e.g. baby-sitter robots), while public robots provide services in public places such as shopping malls or museums.

The designing of dialog models is made taking into account the knowledge domain as well as preliminary experiments on user-robot interaction in circumstances close to the real operation. In many works it was noted that the style of user communication changes with the growing of familiarity with the robot, its functions and the utility degree of provided services [4, 5]. For the personified setup of the communication style with a robot advance learning should be performed as well as accumulation of information about user behavior and preferences during the communication.

For the estimation of user-friendliness of speech dialog interface several measures are used and it requires significant time for the work of user-speakers as well as for a further processing the results of the experiment. The analysis of wide range of metrics, which is used for estimation of the effectiveness of dialog systems, is presented in [6]. The simplest way to estimate the user satisfaction as result of a dialog is based on the measurement of interaction duration and accuracy of speech recognition system. In the PARADISE system a regression model is used to estimate the user satisfaction based on the analysis of successful task fulfillment, quality and dialog effectiveness. This model is used in wide range of the systems [7, 8]. Similar metrics of automated estimation of system behavior were developed in certain other fields: BLEU-estimation for machine translation and ROUGE-estimation for text reviewing.

Experimental investigations of social robots in the real world showed, that, in user's judgment, the most qualitative interaction was one the with lots of dialogs and high speed of speech dialog system, while interaction based on the voice commands is not natural and does not induce users to employ the social norms of communication, which are typical for dialog among humans. The steps to improve human-machine interaction include analysis of user behavior, personified setup of dialog scenarios, and growth of system capabilities, e.g. the capability to detect and to correct user's mistakes.

Further in this paper we provide the analysis of modern works devoted to the issues of social and domestic robots development; also peculiarities of informational mobile robot, developed by the authors, are reviewed. The analysis of hardware and software means, which realize user search and tracking as well as registration of audio-visual data used for extraction user preferences is given.

1. Related Work

In the work [9] the psychological factors that influence on the people communication and that are applicable to the human-machine interaction are described. The most investigated parameter is the interaction distance. Developers pay attention mostly to the analysis of means that would provide maintenance of the psychologically comfortable distance between user and robot in real-time. That is, for example, in the work [10] the combination of two laser range finders and video camera is used for the user detection and tracking. One of the range finders is mounted at 24cm height above the floor and used for user's legs detection. Second one is mounted at 80cm above the floor and used for trunk detection. The video camera is placed at 1.5m above the floor and used for face detection. The decision about user presence is made upon the range finders and the video camera responses in one direction. The changing of robot position is estimated by means of an odometer sensor. As far as the robot moves on a certain territory and the location of static obstacles and such objects as chairs is known in advance, if the user is detected in chair area the assumption of a face location is made given that user may be sitting. For new users the list of their characteristics (height and width of a video-object) is made and entered into a database. In the course of user tracking the information of user's physical characteristics is corrected and entered into the database. Thus, sensors used in this work serve not only for distance estimation, but also for accumulating the personal data of user's preferences and physical characteristics.

Modern social robots are able to interact with a person, not only through speech but also with gestures. For example, the robot Robotinho, developed at the University of Bonn in Germany, has a humanoid form, and can interact with humans through speech, gestures and facial expressions [11]. The robot uses mixed system of dialogue, and is able to determine position of a user and his face, as well as to recognize and synthesize speech. Robotinho can express its emotional state and communicate with many people simultaneously. Since the robot has a humanoid body shape, it can nonverbally communicate with users through gestures during the dialogue, as well as attract users' attention to itself or to the objects of the environment by gestures or gaze direction. The robot detects a user with two laser range finders, and then he finds a human face with two video cameras. When interacting with users it creates a database of users containing user's face images and his/her preferences based on the query history. In future the robot will be able to identify user.

Thus, the appointment of the robot and the possibilities of potential users are necessary to consider at the development of multimodal interfaces for a social robot. Ways of interaction must be easy-to-use and do not require special training of users. Speech and multimodal interfaces are being actively researched and applied in robotic systems now. Despite the fact that user interaction with social robots in most cases takes place in an environment with high noises the speech interfaces are being actively studied and applied research in robotic systems [12, 13, 14].

In the work [2] the influence of robot physical features, its specific means of emotion expressions (e.g. expressive head movements, active eyes movements) on the human social response to a robot is considered. The task of self-disclosure is a very important aspect in communication among people; so in the development of human-like robots capable of expressing emotions naturally there appears a possibility to establish confidential relations between user and robot [15]. The purpose of this investigation was the study of following issues. Whether the robot design specifics may influence the human wish to disclose personal information? How does the number of freedom degrees of the robot's neck influences on the human attitude toward the robot and self-disclosure? Whether the probability that patient will disclose personal information to a robot is higher than that to a doctor? There were advanced two hypotheses: (1) human is more relaxed during communication with a robot if there is robot's neck moving; (2) the more expressive robot's mimics, the more information person discloses. To prove the advanced hypotheses two experiments were performed: in the first experiment the robot used only hand gestures and hand gestures with neck moving; the second experiment was performed for three robot's head configuration: (1) eyes and eyebrows are hidden from a user, (2) only eyes are hidden from a user, and (3) robot's face is completely open for a user. During the first experiment the first hypothesis was proved. Total number of words and emotions of users is increased then the robot used gestures and neck movements. For the second hypothesis only preliminary results were obtained, which showed that it may be proved if there is sufficient amount of statistical data.

How the style of interaction changes during everyday use of social robot is considered in the work [5]. Speech dialogs were analyzed with a purpose of detection of those topics, which are inherent only in inter human communication, and with a purpose of checking: whether users will follow the social norms of communication while interacting with the robot. The given results indicate that approximately half of users followed the minimum set of social norms inherent in inter human communication (such as greeting, farewell and gratitude for rendering help) while interacting with the robot. Since greeting is a social norm for humans, many people started a dialog with the robot with a greeting, e.g. "hi" or "hello" (60%). Certain users at the end of interaction said goodbye to the robot (24%). The majority of users asked a number of questions without waiting system response. And only few users (2%) paid attention to robot's responses or followed the robot speech asking later more specific question. Certain users (8%) established more confidential relations with the robot and addressed it with its name at the beginning of interaction. Also few people emphasized a dialog upon their selves (2%), for example, "I'm alone" or "I'm bored". At the same time if the robot is capable to detect such utterances and respond to them it would help to establish more confident relationship between human and robot.

During interaction with a robot-administrator installed in the building of academic organization 15% of users preferred to use a keyword as a command instead of full sentence as if they were talking to a human [5]. 15% of users thanked robot for granted information even if it had wrongly recognized the request. It is interesting, that people, who used a command type of dialog instead of full sentences turned out to be more polite toward the robot than those, who used full sentences. It is also worth noting, that certain users, which had already interacted with the robot (it wasn't something new to them) let it know about their emotional state, while others were only asking robot about its possibilities and the necessary reference information.

In conclusion of review it is worth noting, that psychological aspects of human-machine interaction are still insufficiently studied. The understanding of why people follow the social norms more or less requires further research. Experience of interaction with robots and degree of robots animation by people are the most influential factors in the effectiveness and friendliness of dialog. The development of software and hardware of registration and analysis of user preferences is, from the technical point of view, one of the most urgent problems in creating of social robots. In the next section we observe the technical means that we used for realization of informational robot (particularly the user detection by means of range finders and audio-visual monitoring of surrounding space).

2. Multi Sensor Approach to User Detection

Detection of the obstacles and finding the users is a one of the major issues that must be solve during the development of mobile social robots. Different types of range finders are used for obstacles detection. Additional sensors should be used for the user search. In our work we applied ultrasonic and infrared range finders, web cameras and microphone arrays that are placed as shown in Figure 1.

The infrared sensors find obstacles at close distances with high speed and precision. The ultrasonic sensors provide a finding of obstacles on larger distance. The location of eight infrared and six ultrasonic sensors on the lower part of the robot was chosen in such way in order to cover all directions around the robot, especially critical areas: front and back of the robot. Data from the infrared and ultrasonic sensors are averaged in a sliding window, whose size is chosen for each sensor independently by analyzing the incoming data, in order to eliminate false detection of an obstacle or a user. The current position of the robot is determined based on readings of odometers.

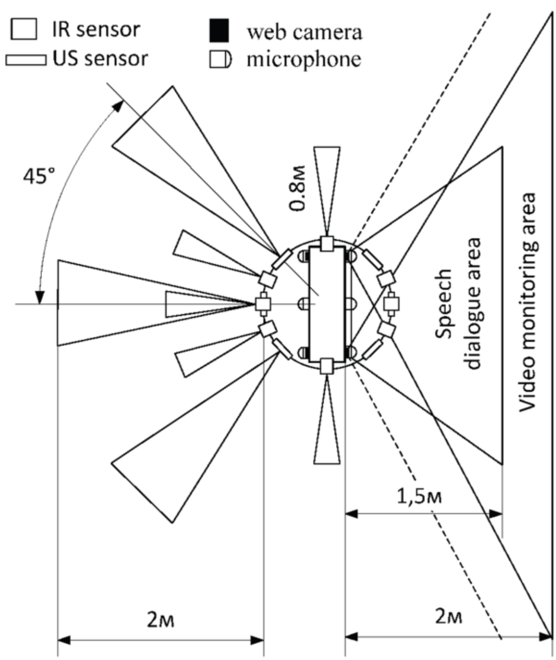


Figure 1. Sensor placement in the information mobile robot.

To find the face of a user and record the video two cameras located on each robot side at the height of the average human height are used. The human face detection and tracking is implemented by using the original algorithm that tracks the movement of natural markers (the center of the upper lip, nose, the point between the eyes, the pupil of the right eye and the pupil of the left eye). This algorithm provides the robustness of the positioning of the head [16]. Also a special method to keep track of control points and automatically restore them in the work area in case of loss of the system is developed. The Intel OpenCV library is used for the basic procedures of video capturing and processing.

The T-shaped microphone arrays installed around touch screen on the each side of the robot are used to record the audio signals. The microphone arrays are used for speech recognition, sound source localization and accumulation of audio data of the users for biometric processing. The applied sound source localization module allows the robot to obtain sounds that outgoing from a point in space corresponding to the mouth of a user, and suppress the sounds coming from other directions. It significantly improves the quality of distant speech recognition. Sound source localization algorithm is based on the GCC method with the PHAT weighting function [17]. The SIRIUS system is implemented for speaker independent continuous speech recognition [18].

During the free movement of the robot the sensor system performs real-time analysis of the range finder data. The robot stops and determinates the location of obstacle, when the range finders detect it appearance on the robot way. The video cameras are used to verify presence of a person in the obstacle direction by searching face in the proposed area. If a person is found in the interaction zone and his/her face is directed on the robot, then the robot considers the person as a user and starts a dialogue. Next section describes more detail the structures of dialogues, which used in different operation modes of the information mobile robot.

3. Hierarchical Dialogue Model

The developed mobile social robot provides reference services to visitors of shopping malls. Therefore, the interaction dialog model corresponds to robot's tasks. The use of the dialogue model in the shopping center was selected because the system-led mode in most cases degrades the naturalness of interaction and imposes strong restrictions on the user's phrase. On the other hand, the modern means of natural language processing have not yet reached the level of completely user-led strategy. In addition, the system should be available to people with varying degrees of training. Also visitors, by psychological factors are afraid to start a dialogue with the robot. The user-led strategy means that a user has a freedom to build a course of dialogue and every phrase. This leads to a significant increase in databases, which requires a speech recognition system to handle voice messages. There is a tradeoff between making the robot's recognition grammars sufficiently large so that people can express themselves somewhat freely versus making the grammars small enough so that the system runs with high accuracy and in real-time [20]. We tuned speech recognition performance by creating the sets of the specialized grammars for different robot contexts and the robot's software dynamically activates a corresponding subset of grammars depending on the context of the robot activity. This allows us to overcome the combined increase in parsing time due to incorporating natural syntactic variability in the recognition grammars.

Instead of keeping a single enormous recognition grammar active, the robot keeps subsets of small grammars active in parallel, given what it currently expects to hear. The key assumptions here are that certain types of utterances are only likely to be said under particular circumstances, and these are circumstances among which the robot is capable of distinguishing [20].

Structuring the dialogue model in accordance with the current mode of operation is shown in Figure 2. Such division has several advantages. First, the system performance is increased due to the fact that each sub dialog has own vocabulary, so the search is performed in the reduced area. Second, it simplifies the procedure of adding new or updating existing sub dialogs. There is no reason for redo of the entire system, that's enough to add a new sub dialog and conditions of its call.

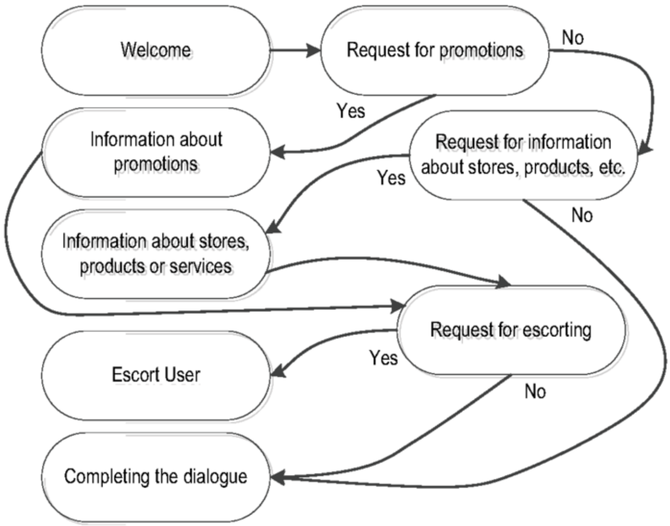


Figure. 2. Structure of sub dialogs.

Let us consider structure of sub dialogs and transitions between them. In the beginning the system greets user and suggests to listen a brief information about promotions. If a user disagrees then the robot suggests user to ask information about shops or goods. In other case objects which held in the current promotion, are marked at the screen and the robot changes the sub dialog to the "Information about promotions", in which data on all the objects and their events are provided. A visitor can select interesting information by clicking on the appropriate store displayed on the map and pronouncing its name, and if user wants the robot changes sub dialogue to “Escort” and escorts him/her to the selected point. After displaying information about all actions or a specific sale in the store, the robot asks the user about necessity of information about other shops, goods and etc. If the user refuses to continue the dialogue, the robot turns to the sub dialog "Completing the dialogue". If user agrees then robot selects sub dialog “Information about shops” in which user can ask to find certain shop, good or a set of shops that sell certain goods category. When a shop is selected, the robot also asks to escort user to it. During the interaction is created a user profile. If the user is already registered in the system, the existing profile is supplemented by new data. More information about it in the next section.

4. Creation and Implementation of User Profile for Social Mobile Robots

Creating a personal profile is produced during robot-user interaction. The height and age category of a user are recorded in the user's profile at the beginning of interaction.

As further interaction the user profile is supplemented the personal user data and preferences, such as preferable interface modalities, services, etc.

The structure of user profile data is shown in Figure 3. The presented structure can be divided into two categories of data: entered by the user and accumulated automatically. The first category consists of the subcategories of "Personal data" and "Contact information". The personal data includes personal characteristics of a user, which also can be used to improve interaction. Information from the category of "Contact information" may be used to deliver various information materials to the user address.

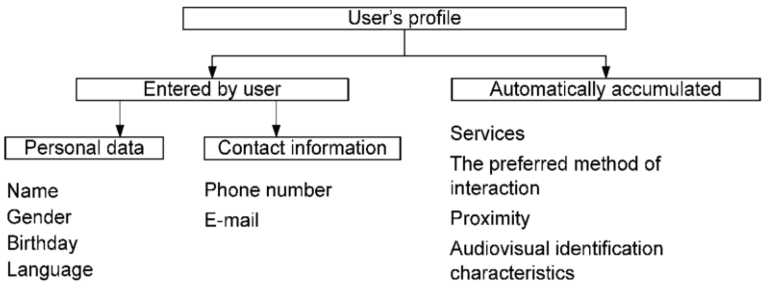


Figure 3. The scheme of the user profile.

The data about selected services, audio-visual user data, including his/her photos, list of natural modalities preferred in the user interface, as well as information on the distance, which a user kept during the time of interaction with the robot are automatically collected during the dialogue of the robot with a user.

If a user is registered in the profile database, the robot automatically adapt the interface, which is preferred by the user and offers a choice of services that are most often used by the user. Also the strategy of dialogue, which is typical for the particular user, is chosen in the automatic mode. The data accumulated during the interaction are used for updating user's profile.

Let us to consider two ways of using the robot: 1) an information kiosk in a mall, and 2) an information kiosk at domestic assistant environments. In the first case, the robot is placed in a constantly changing environment in difficult conditions, because in shopping malls are always a lot of people. The mobile robot should automatically stop in the case of appearance of any unidentified obstacles by range finders. It is also necessary to solve the problem of the user detection. In the case of high noise area, the user can be determined only on the basis of data from the cameras, without the use of audio localization. If the audio noise level does not exceed the permissible, the choice of the nearest user is based on the data of the audio localization too. Since the shopping mall has enough background noise the using a voice interface is not always available, therefore, the use of verbal interaction with the user may only be used at a certain level of noise and most of the time the graphical user interface will be used.

The main tasks of the information mobile robot in a shopping mall are: 1) promotion; 2) inquiry; 3) moving to the base. In the “Inquiry” mode the data required to a user are displayed on the screen and pronounced by speech synthesis module. For example, output of the current location of a user and the robot, the route to the point of interesting, the founding goods and services on the database of a shopping center, search the store by the name or belonging to a category of goods are carried out. The robot also needs to be able to make online ordering of goods in the store, and support the direct communication with a representative of the store. The information about current promotions in shops, goods and service is displayed on screen in the "Promotion" mode. When the robot battery has low charge it can change the current mode to moving to base, but first of all the robot will give information about changing the work mode to the current user.

Use of an information mobile service robot at domestic assistant environments is differed from shopping mall, it usually have less audio noise conditions. The robot should move on a given area, and provide only reference service. Since the number of users in such establishments is significantly smaller than in shopping malls, the working area of interaction can be expanded and search for persons can also be produced in a more wide range. The information, which the robot should provide, includes the structure of the building (map), location of the premises, consulting rooms, and information about doctors, including their contact data. Also the robot can have possibility to accompany user to the required destination by the user's request. Since the number of patients at domestic assistant environments is fixed, then it is possible to make personal profiles for all users during the first interaction with the robot and provide the personified service for each patient in further.

5. Conclusion

The development of means of detection of user presence and user tracking during different stages of service (search, greeting, dialog, escorting) requires solving the range of questions concerning psychological aspects of communication as well as ways of automatic analysis of human behavior. The information service robot described in this article is equipped with sensors of different operation types; this enables to detect an undefined object in time and distinguish a user from obstacles that occur on the robot's way of moving. The history of user behavior is fixing during interaction and it used to extract user's preferences and for personified service during further communication sessions. This work is supported by the Russian Federal Targeted Program (contract No.876) and Saint-Petersburg State University (project # 31.37.103.2011).

References

[1] C.L. Breazeal, *Designing Sociable Robots*. MIT Press, 2002.
[2] J. K. Lee, C. Breazeal, Human Social Response Toward Humanoid Robot's Head and Facial Features, In: *Proceedings of CHI'2010* (2010), 4237-4242.
[3] D. Lambert, *Body Language*. Harper Collins, London, 2004.
[4] M. E. Foster, M. Giuliani, A. Knoll, Comparing Objective and Subjective Measures of Usability in a Human-Robot Dialogue System, In: *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP (2009)*, 879-887.
[5] M. K. Lee, M. Makatchev, How Do People Talk with a Robot? An Analysis of Human-Robot Dialogues in the Real World, In: *Proceedings of CHI'2009 (2009)*, 3769-3774.
[6] N.O. Bernsen, L. Dybkjær, Multimodal Usability: Human-Computer Interaction Series, *Springer 2010 (2010)*, 431.
[7] A. Walker, D.J. Litman, C.A. Kamm, A. Abella, PARADISE: A framework for evaluating spoken dialogue agents, In: *Proceedings of ACL/EACL'1997 ACL Antology* (1997), 1035.
[8] S. Moller, K.-P. Engelbrecht, R. Schileicher, Predicting the quality and usability of spoken dialogue systems, *Speech Communication 2008 (2008)*, 730-744.

[9] R. Mead, A. Atrash, M. J. Mataric, Proxemic feature recognition for interactive robots: automating metrics from social sciences, In: *Proceedings of ICSR 2011, Springer Verlag Berlin Heidelberg 2011 (2011)*, 52-61.
[10] J. Stuckler, S. Behnke, Improving People Awareness Of Service Robots by Semantic Scene Knowledge, In: *Proceedings of RoboCup 2010, Springer Verlag Berlin Heidelberg 2011 (2011)*, 157-168.
[11] M. Nieuwenhuisen, J. Stuckler, S. Behnke, Intuitive Multimodal Interaction for Service Robots, In: *Proceedings of HRI'2010 (2010)*,177 -178.
[12] T. Kollar, S. Tellex, D. Roy, N. Roy, Toward Understanding Natural Language Directions, In: *Proceedings of HRI'2010 (2010)*, 259-266.
[13] R. Cantrell, M. Scheutz, P. Schermerhorn, X. Wu, Robust Spoken Instruction Understanding for HRI. In: *Proceedings of HRI'2010 (2010)*, 275-282.
[14] S. Kriz, G. Anderson, G. Trafton, Robot-Directed Speech: Using Language to Assess First-Time Users' Conceptualizations of a Robot. In: *Proceedings of HRI'2010 (2010)*, 267-274.
[15] A. Powers, S. Kiesler, The advisor robot: tracing people's mental model from a robot's physical attributes, In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, ACM 2006 (2006)*, 218-225.
[16] A. Karpov, A. Ronzhin, ICANDO: Low Cost Multimodal Interface for Hand Disabled People, *Journal on Multimodal User Interfaces* **Vol. 1, No. 2** (2007), 21-29.
[17] O. Maurizio, S. Piergiorgio, B. Alessio, C. Luca. Speaker Localization in CHIL Lectures: Evaluation Criteria and Results, *Machine Learning for Multimodal Interaction, Springer, Berlin* (2006), 476–487.
[18] A. Ronzhin, A. Karpov, I. Kipyatkova, M. Zelezny. Client and Speech Detection System for Intelligent Infokiosk, *Springer-Verlag Berlin Heidelberg, Petr Sojka et al. (Eds.): TSD 2010, LNAI 6231* (2010), 560–567.
[19] Ronzhin, A., Karpov, A., Kipyatkova, I., Zelezny, M.: Client and Speech Detection System for Intelligent Infokiosk, In: *Proceedings of TSD 2010, Springer-Verlag Berlin Heidelberg, LNAI 6231* (2010), 560–567.
[20] M.H. Coen, Design principles for intelligent environments, In: *Proceedings of the 15 National conference on Artificial intelligence (1998)*, 547-554.

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-231

Recognition of Static Gestures using Three-Dimensional Chain Codes using Dominant Direction Vectors

Jose FIGUEROA[b], Jesus SAVAGE[a], Ernesto BRIBIESCA[b] and Enrique SUCCAR[c]
[a] *Universidad Nacional Autonoma de Mexico, Department of Electrical Engineering*
[b] *Universidad Nacional Autonoma de Mexico, Instituto de Investigaciones en Matematicas Aplicadas y Sistemas, Department of Computer Science*
[c] *Instituto Nacional de Astrofisica, Optica y Electronica, Department of Computer Science*

Abstract. This paper presents a novel method for static gesture recognition based on three-dimensional chain codes which are computed from three-dimensional skeletons acquired by three-dimensional vision sensors, such as Microsoft Kinect™. The method has two stages: a digitization stage and a recognition stage. The digitization stage is based on the orthogonal direction change chain code, which represents the changes of direction on segments of a three-dimensional curve which was fitted in a three-dimensional grid; these changes of direction are invariant to rotation and translation. The recognition stage is based in the detection of dominant changes of direction on segments of a three-dimensional curve which was fitted in a three-dimensional grid. The experiments for testing this method of static gesture recognition involved recording the pose of the arms of a subject who was standing at increasing distances and the body was oriented in frontal and three-quarters angles; the poses were matched against a set of reference arm poses which were taken from a frontal body pose at an specific distance. The results show that the generation and matching of three-dimensional chain codes of arm poses captured on varying configurations are reliable enough for being used for gesture recognition. The gesture recognition system will be used on the robots of the Robocup@Home teams Markovito and Pumas from Mexico.

Keywords. Feature extraction, Chain code, Machine vision, Threedimensional Curve, People detection, Pattern recognition, Gesture recognition.

Introduction

Human beings have several means of communicating ideas to their peers: speech, which relies on emitting sounds with the vocal organs; writing, which relies on drawing symbols on a surface; and gesturing, which relies on the presentation or movement of body parts.

In daily life, people use gestures in order to give more information when a person is speaking or for transmitting ideas when the situation does not allow to speak, either because loud noises overwhelm speech loudness, silence is required to perform a task, the interlocutors are too far from each other, or any of the interlocutors is unable of speaking.

The interaction between a human being and a computer usually is done by devices which must be manipulated directly, such as keyboards, mice, touch screens, digitizing tablets, trackballs, joysticks, just to name a few. Another way of interacting with a computer without relying on manipulable input devices is by using voice commands, which are recorded by a microphone and the computer interprets the sound information in order to get commands from it.

The advantage of using gestures is that they work in environments where voice acquisition for verbal commands is complicated or infeasible, as well as gestures provide a more universal way of inputting commands because voice input is too dependent on both the accent and the pronunciation of the person who is speaking.

The scope of this article is focused in presenting a novel way of representing and recognizing human pose data from vision and motion capture systems.

The representation of human pose data depends on the source data which is being used for tracking the human being. For the case of bi-dimensional image data, the ways of acquiring human pose are: binary silhouettes obtained either by background removal, color segmentation or depth segmentation [8,13,10,9,1]; invariant features extracted from image gradient [11], or three-dimensional joint data acquired through motion capture [15,14,6,7] or by probabilistic recognition of the position of the body parts using depth data [12]. According to the source data and the scope of the work, the representation of the pose data has varied formats, such as directed histograms of pixels [8], mathematical models for body parts [13], edge oriented histograms [10], sets of images containing key poses [9], residual vectors [1], SIFT-like features [11], differences of joint angles [15], quaternions of each joint [14], statistical data from acceleration and angular data [6], or vectors of angle features [7]. All these representations of pose data are used as input for algorithms of temporal pattern recognition, such as Markov Models, State Machines, Neural Networks, or Support Vector Machines.

The advantages of this method are the following: 1. the chain codes represent three-dimensional orthogonal direction change, which makes the representation of a three-dimensional curve invariant to translation and rotation; 2. the chain codes can represent any three-dimensional curve in a compact fashion, as a string of chain codes; 3. the computation of the chain codes and the computation of the dominant direction changes use vector math and basic math, so they can be optimized easily for computer architectures which lack of floating-point units.

1. Proposed Approach

The proposed method for static gesture recognition consists of two stages, a digitization stage, where the three-dimensional joint data is converted to a discrete three-dimensional curve, and a recognition stage, where the measure of similarity between two discrete three-dimensional curves is computed in order to know if two discrete three-dimensional curve have similar shapes.

1.1. Orthogonal Direction Change Chain Code

The digitization stage is based on the orthogonal direction change chain code [2,5], which digitizes three-dimensional curves into a set of codes which represent orthogonal direction changes between three constant length segments of a three-dimensional curve (u, v, w) (Equation 2), which is aligned to the vertices of a three-dimensional grid of a constant cell size.

In order to convert a set of three-dimensional lines into a set of constant length segments, the first step consists in aligning the vertices of each line, *p* and *q*, to the corners of the three-dimensional grid, by rounding the values of *p'* and *q'*, according to the smallest distance of each axis and two neighbouring vertices of the grid (Equation 1), getting the vertices *p'* and *q'* as a result.

Most of the time, the length of the line is longer than the size of the cell of the grid, additional points are added to the line by means of linear interpolation, whose amount is the Chebyshev distance of *p'* and *q'*, $D_{Chebyshev}(p',q')=max(|p'-q'|)$. The interpolated points are adjusted to the grid using the Equation 1.

$$v'_x=\begin{cases} g_x & if g_x \le v_x < \frac{g_x+g_{x+1}}{2} < g_{x+1} \\ g_{x+1} & if g_x < \frac{g_x+g_{x+1}}{2} \le v_x \le g_{x+1} \end{cases} \quad (1a)$$

$$v'_y=\begin{cases} g_y & if g_y \le v_y < \frac{g_y+g_{y+1}}{2} < g_{y+1} \\ g_{y+1} & if g_y < \frac{g_y+g_{y+1}}{2} \le v_y \le g_{y+1} \end{cases} \quad (1b)$$

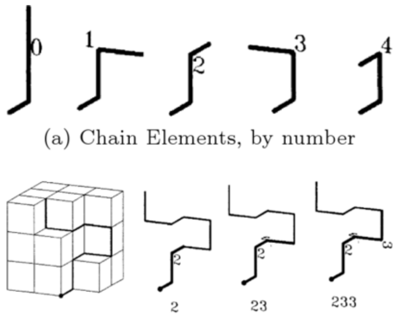
$$v'_z=\begin{cases} g_z & if g_z \le v_z < \frac{g_z+g_{z+1}}{2} < g_{z+1} \\ g_{z+1} & if g_z < \frac{g_z+g_{z+1}}{2} \le v_z \le g_{z+1} \end{cases} \quad (1c)$$

Once all the points are adjusted to the grid, the next step is to split the coordinates of the lines between two consecutive points to make a set of single line segments. And, the final step is to assign chain codes by taking three consecutive single line segments, starting from the first line segment, and apply the rules of orthogonal direction changes to compute the corresponding chain element (Figure 1b).

There are five different orthogonal direction changes for representing any three-dimensional curve (Figure 1) [4]:

- The Chain Element "0" represents a direction change which goes straight through the contiguous straight-line segments following the direction of the last segment.
- The Chain Element "1" represents a direction change to the right.
- The Chain Element "2" represents a direction change upward (stair-case fashion).
- The Chain Element "3" represents a direction change to the left.
- The Chain Element "4" represents a direction change which is going back.

$$chain\,element(u,v,w)=\begin{cases}0 & if \quad w=v\\1 & if \quad w=u\times w\\2 & if \quad w=u\\3 & if \quad w=-(u\times v)\\4 & if \quad w=-u\end{cases} \tag{2}$$



(b) Example of Chain Code Sequence

Figure 1. Orthogonal Direction Change Chain Codes.

1.2. Digitization of Three-dimensional joint data

The three-dimensional joint data is captured by a three-dimensional vision system, such as the Microsoft KinectTM, which acquires the joint data by analysing the depth map captured by the sensor. The joint data is organized in a humanoid skeleton hierarchy.

For purposes of this work, the joints are numbered in this way (Figure 2): *0)* Head (0); *1)* Neck (1); *2)* Left Shoulder (2); *3)* Left Elbow (3); *4)* Left Hand (4); *5)* Right Shoulder (5); *6)* Right Elbow (6); *7)* Right Hand (7); *8)* Torso (8) *9)* Left Hip (9); *10)* Left Knee (A); *11)* Left Foot (B); *12)* Right Hip (C); *13)* Right Knee (D); *14)* Right Hip (E).

The skeleton data is converted to a series of chain codes, which are merged in a tree structure (Figure 3), using the notation described in [3].

1.3. Dominant Direction Changes

The first attempts of recognizing static gestures were based on the work of [4], which describes an algorithm to compute the measure of dissimilarity between two discrete three-dimensional curves (A and B), which are composed of a series of orthogonal direction change chain codes, based on the quantification of the differences between partial common couples contained in two curves. An implementation of the former algorithm was made in C++ and tested with sets of chain codes, the results of the matching had large amounts of false positives, specially between the poses for the arm being extended to the front and the arm being lowered down, whose reference poses which had a matching rate close to 90% between both poses.

Due to these problems, alternative approaches for analyzing the chain codes to determine poses are researched. Through visual analysis of the chain codes, it is pointed out that the chain codes of the three-dimensional curves for each reference pose have different distributions of orthogonal segments. As an example, the chain code for the left arm when is extended to the side has a large proportion of orthogonal segments whose directions are parallel to the positive horizontal axis +X, while the chain code for the right arm when is extended to the side has a large proportion of orthogonal segments whose directions are parallel to the negative horizontal axis -X (Table 1).

Table 1. Dominant Axis for each Orthogonal Direction for static arm poses.

Direction	Dominant Axis
Left	+X
Right	-X
Up	+Y
Down	-Y
Back	+Z
Front	-Z

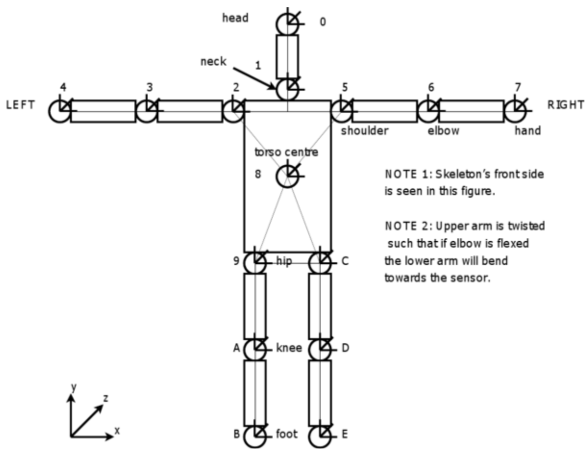
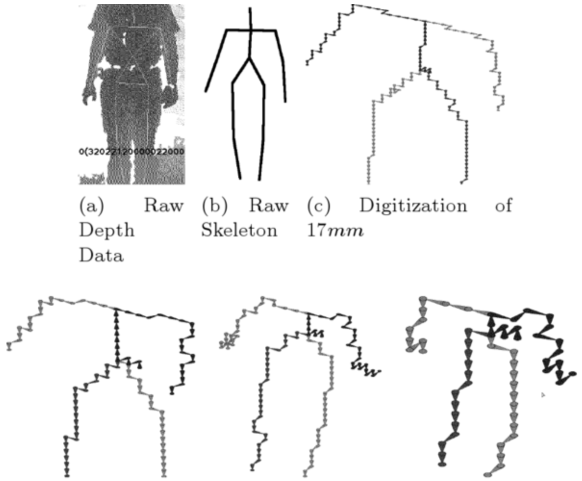


Figure 2. Three-dimensional Joint Data Hierarchical Structure.

In order to determine the Dominant Direction Vectors of a chain code, the chain code needs to be converted to a set of orthogonal segments with a known direction. In order to do that, two direction vectors are given arbitrarily, whose only restriction is that they must be orthogonal to each other. For example, $u = (0, 1, 0)$ and $v = (0, 0, 1)$. The next direction vector, w, is computed by applying the rules for chain codes to each character of the chain code (Equation 2). Each direction vector, including the arbitrary direction vectors used at the beginning, is stored in a list of vectors V. Every time that the direction vector w is stored in the list of vectors, the direction vectors u and v have their values reassigned to $u = v$ and $v = w$. Once all the direction vectors are stored in the list of vectors, the next step is to compute the percentages of each direction, P, by dividing the count of the vectors on each direction contained in the list of vectors, $\{V_{+X}, V_{-X}, V_{+Y}, V_{-Y}, V_{+Z}, V_{-Z}\}$, between the total amount of vectors in the list of vectors, $|V|$ (Equation 3). The Dominant Direction Vectors, D, are obtained from the two largest percentages on P (Equation 4).



(a) Raw Depth Data (b) Raw Skeleton (c) Digitization of $17mm$

(d) Digitization of $29mm$ (e) Digitization of $42mm$ (f) Digitization of $85mm$

Figure 3. Skeleton Processing. Figures 3c-3f shows the trees of chain codes at different resolutions, which are computed from raw skeleton data (Figures 3a, 3b).

$$P=\begin{cases} P_{+X} & \frac{|V_{+X}|}{|V|} \\ P_{-X} & \frac{|V_{-X}|}{|V|} \\ P_{+X} & \frac{|V_{+Y}|}{|V|} \\ P_{-Y} & \frac{|V_{-Y}|}{|V|} \\ P_{+Z} & \frac{|V_{+Z}|}{|V|} \\ P_{-Z} & \frac{|V_{-Z}|}{|V|} \end{cases} \tag{3}$$

$$D=\begin{cases} D_1=max(P) \\ D_2=max(P \setminus D_1) \end{cases} \tag{4}$$

Figure 4. Orthogonal Direction Change Chain Codes.

2. Experiments

The purpose of the experiments is to test the reliability of computing the Dominant Direction Vectors of three-dimensional chain codes for recognizing pose data from body parts, specifically both arms, captured by three-dimensional vision sensors and stored as video files.

2.1. Configuration

The vision sensor is set to a height of 1000*mm* above the ground, later the test subject is set at a certain distance from the vision sensor and is asked to perform a pose with the arm extended, in a determinate body orientation angle. The set of distances are {2000*mm*, 2500*mm*}, the set of poses for each arm is lowered down, extended to the front, and extended to the side; and the orientation angles set are 0°, 45°, -45°, where 0° is the frontal angle camera of the subject and 45° is the subject set on a three-quarters orientation. The height of the vision sensor does not allow to capture poses where the arms are raised above the head.

The joints of the skeleton are used in relative coordinates, using the joint of the torso as reference. The reference size for the resolution of the three-dimensional grid for computing the three-dimensional chain codes is set to 170*mm*, because is the length of the vector which is formed by the coordinates of both head and neck joints. By using this reference size, other resolutions, which are fractions of the reference size, are selected for the experiments, the resolution set is:{17*mm*, 29*mm*, 42*mm*, 85*mm*}; the purpose of testing multiple resolutions for the three-dimensional grid is to figure out at which resolutions the computation of the Dominant Direction Vectors becomes unreliable while keeping close to real-time performance. A subset of the joints of the skeleton is used to determine the pose of each arm, without the forearm: torso, neck, shoulder and elbow.

The tree which describes the skeleton has the structure shown in Figure 4. The tree is formed by four sequences of joints which have the torso centre as its root. This configuration was chosen because the height of the vision sensor does not allow acquiring the joint of the head when a person is at a close distance (<= 1500*mm*).

2.2. Orientation Angle of the Body

An issue with the matching of chain codes from skeletons captured by three-dimensional vision sensors is the orientation angle of the torso of the subject: the visual analysis of skeletons made of orthogonal direction vectors shows that the same pose has varying proportions of orthogonal direction vectors, according to the orientation angle of the subject to the camera, which changes the apparent pose of each limb. For example, when the subject is in front of the camera and has both arms extended to the side, each arm has a large percentage of orthogonal direction vectors on the X axis and a small percentage of orthogonal direction vectors in the Z axis, which would match to a pose where a limb is extended to the side; however, when the subject is in a three-quarters angle, the percentage of orthogonal direction vectors on the X axis decreases and the percentage of orthogonal direction vectors on the Z axis increases, which would match to a pose where the limb is extended to the front or it would match to a pose where the limb is extended to the back.

In order to avoid that, the orientation angle of the subject is computed from the normal vector N of the triangle formed by the joints of the torso, the left shoulder and the right shoulder, *p1*, *p2*, *p3*. From that vector, the orientation angle θ is computed using N_x and N_z (Equation 8). The numerical value of the normal vector of the torso, when the subject is in front of the camera, is 90º, therefore, when the subject is not oriented towards the camera, the complementary angle φ must be computed to rotate each joint of the skeleton, whose coordinates are set relative to the torso joint, in order to have the angle of the normal vector in the numerical value of the frontal camera angle (Equation 9).

$$U = p3 - p1 \tag{5}$$

$$V = p2 - p1 \tag{6}$$

$$N = U \times V \tag{7}$$

$$\theta = \tan\left(\frac{N_x}{N_z}\right) \tag{8}$$

$$\varphi = 90^\circ - \theta \tag{9}$$

As the chain codes for each arm are start at the joint of the torso, a reference set of direction vectors can be added to reduce the ambiguity of the chain codes, in this case the following sequence of direction vectors was used $\{(0,1,0),(0,0,1),(0,1,0),(0,0,1)\}$ which is converted to the chain code sequence *22*. With this sequence, the line between the torso joint and the neck joint can be formed by a set of orthogonal direction vectors with the values (0, 1, 0) if the resolution is high enough.

2.3. Arm Pose Matching

The matching process consists in computing the Dominant Direction Vector of the chain code sequence for each arm, using the algorithm described in the Subsection 1.3. The algorithm was applied with a variation in the way the direction vectors are counted: at high resolutions, the line between the torso joint and the neck joint can be formed by a set of positive vertical orthogonal direction vectors with the values (0, 1, 0), which can be used as a reference to find the origin of the shoulders by finding any orthogonal direction vector whose value is different to (0, 1, 0).

That sequence of consecutive positive vertical orthogonal direction vectors can be ignored in order to limit the proportion of positive vertical orthogonal direction vectors to those vectors which belong to poses where the arm is raised above the level of the shoulders. Once an orthogonal direction vector which is different to a positive vertical vector is found, at the level of each shoulder, the count for computing the Dominant Direction Vectors is started.

For each pose there is a set of Dominant Direction Vector which are expected (Table 2).

2.4. Results

In the testing stage, the subject is set in the distance and angle sets which are mentioned in the section 2.1 and 200 skeleton samples of each arm pose are captured, whose chain codes are computed later and matched against the reference chain codes.

Both accuracy and running time are assessed for each configuration of the arm pose tests. In the matching accuracy tests (Table 3-8), the overall best results were achieved when the resolution became higher (Table 3, 8). As for the average running time of the matching, this value increases with the resolution, the tests performed with the highest resolution show an average time of 0.003 milliseconds, which is good enough for real-time applications. These results can be compared against a set of results from control data, which was obtained from the Dominant Direction Vectors of raw join data captured by the Kinect (Tables 9 and 10).

A problem with the generation of chain codes for the skeleton is that either by the noise on the depth map or by the lack of complete information in the skeleton, it is not possible to generate a root node for the skeleton tree, which is used as reference during the matching process. The lack of a root node reduces the length of the chain code and increases the probability of a mismatch.

Another problem with the generation of the skeleton comes when the subject has the arms extended towards the camera, as the visibility of the elbow is reduced, OpenNI is unable of computing the coordinates of the joint of the elbow correctly, which alters the proportion of orthogonal direction vectors over the -*Z* axis, resulting in a set of Dominant Direction Vectors which does not identify the pose for arms extended towards the camera (Tables 5, 6).

Table 2. Expected Dominant Axis for each Arm Pose in the Testing Stage

Pose	First Dominant Axis	Second Dominant Axis
Left Arm Side	+X	+Y or -Y
Right Arm Side	-X	+Y or -Y
Left Arm Down	+Y	+X
Right Arm Down	-Y	-X
Left Arm Front	+Z	+X
Right Arm Front	-Z	-X

Table 3. Matching Accuracy of Down Pose (2000mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)				Average Running Time (ms)			
		17mm	29mm	42mm	85mm	17mm	29mm	42mm	85mm
Down (0°)	Left	99.50	99.50	99.50	8.50	0.0032	0.0027	0.0031	0.0019
	Right	99.50	99.50	99.50	21.00	0.0033	0.0027	0.0028	0.0020
Down (45°)	Left	99.50	99.50	99.50	15.50	0.0033	0.0026	0.0026	0.0020
	Right	99.00	79.50	99.50	2.00	0.002609	0.002178	0.002024	0.001532
Down (-45°)	Left	96.50	96.50	95.50	96.50	0.002509	0.002055	0.002063	0.001504
	Right	98.00	88.50	99.50	96.00	0.002763	0.002201	0.002094	0.001609

Table 4. Matching Accuracy of Down Pose (2500mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)				Average Running Time (ms)			
		17mm	29mm	42mm	85mm	17mm	29mm	42mm	85mm
Down (0°)	Left	99.50	98.00	99.50	89.00	0.0032	0.0027	0.0024	0.0020
	Right	99.50	99.50	99.50	98.50	0.0036	0.0029	0.0025	0.0020
Down (45°)	Left	99.50	99.50	99.50	99.50	0.0041	0.0027	0.0025	0.0020
	Right	95.00	94.50	94.50	86.00	0.002658	0.002189	0.001855	0.001675
Down (-45°)	Left	99.50	99.50	99.50	99.50	0.002614	0.002142	0.001996	0.001539
	Right	96.50	94.00	95.50	94.00	0.002938	0.002240	0.002022	0.001568

Table 5. Matching Accuracy of Front Pose (2000mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)				Average Running Time (ms)			
		17mm	29mm	42mm	85mm	17mm	29mm	42mm	85mm
Down (0º)	Left	48.00	49.50	48.00	48.00	0.0034	0.0026	0.0025	0.0022
	Right	10.00	9.00	2.00	0.00	0.0031	0.0025	0.0026	0.0020
Down (45º)	Left	95.00	95.00	95.00	94.00	0.0032	0.0027	0.0025	0.0020
	Right	93.00	94.00	91.00	95.00	0.002712	0.002112	0.001945	0.001611
Down (-45º)	Left	94.00	94.00	94.00	94.00	0.002425	0.002004	0.002022	0.001568
	Right	9.00	7.00	10.00	0.00	0.002591	0.002194	0.001996	0.001596

Table 6. Matching Accuracy of Front Pose (2500mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)				Average Running Time (ms)			
		17mm	29mm	42mm	85mm	17mm	29mm	42mm	85mm
Down (0º)	Left	7.00	4.00	4.00	4.50	0.0036	0.0026	0.0024	0.0020
	Right	59.00	58.00	54.50	37.00	0.0026	0.0021	0.0019	0.0018
Down (45º)	Left	80.50	81.50	76.50	78.00	0.0033	0.0026	0.0024	0.0020
	Right	95.50	96.00	95.50	56.00	0.002656	0.002068	0.001906	0.001539
Down (-45º)	Left	74.00	74.00	74.00	74.00	0.002045	0.001673	0.001521	0.001311
	Right	25.50	26.50	21.50	17.00	0.002951	0.002296	0.002073	0.001678

Table 7. Matching Accuracy of Side Pose (2000mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)				Average Running Time (ms)			
		17mm	29mm	42mm	85mm	17mm	29mm	42mm	85mm
Down (0º)	Left	99.50	99.50	99.50	99.50	0.0033	0.0026	0.0025	0.0020
	Right	99.50	99.50	99.50	99.50	0.0036	0.0030	0.0025	0.0021
Down (45º)	Left	99.50	99.50	99.50	99.50	0.0034	0.0028	0.0025	0.0020
	Right	99.50	99.50	99.50	99.50	0.002543	0.002012	0.001858	0.001557
Down (-45º)	Left	99.50	99.50	99.50	99.50	0.002799	0.002337	0.002001	0.001578
	Right	99.00	99.50	99.50	99.50	0.002779	0.002217	0.001947	0.001516

Table 8. Matching Accuracy of Side Pose (2500mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)				Average Running Time (ms)			
		17mm	29mm	42mm	85mm	17mm	29mm	42mm	85mm
Down (0º)	Left	99.50	99.50	99.50	99.50	0.0035	0.0027	0.0026	0.0020
	Right	99.00	99.00	99.00	99.00	0.0035	0.0028	0.0028	0.0021
Down (45º)	Left	99.50	99.50	99.50	99.50	0.0034	0.0027	0.0024	0.0020
	Right	99.50	99.50	99.50	99.50	0.002473	0.002009	0.001911	0.001529
Down (-45º)	Left	99.00	99.00	99.00	99.00	0.002838	0.002281	0.002258	0.001645
	Right	99.50	99.50	99.50	99.50	0.002658	0.002101	0.001876	0.001529

Table 9. Stats of Control Group (2000mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)			Average Running Time (ms)		
		0º	45º	-45º	0º	45º	-45º
Down	Left	99.50	99.50	99.50	0.00682752	0.00801804	0.00938817
	Right	99.50	99.50	99.50	0.00494681	0.00566779	0.00618352
Front	Left	95.00	94.00	94.50	0.00858765	0.00839521	0.00764601
	Right	95.00	88.00	95.00	0.00609884	0.00608858	0.00538813
Side	Left	99.50	99.50	98.00	0.00745101	0.00843626	0.0073612
	Right	98.50	98.50	98.50	0.00538813	0.00630154	0.00556003

Table 10. Stats of Control Group (2500mm Away, Relative Coordinates)

Pose	Arm	Accuracy (%)			Average Running Time (ms)		
		0º	45º	-45º	0º	45º	-45º
Down	Left	99.50	99.50	99.50	0.00714568	0.00681213	0.00861843
	Right	99.50	99.50	99.50	0.00526497	0.00509563	0.00664279
Front	Left	96.50	74.00	84.00	0.00859021	0.00659147	0.00799238
	Right	96.50	73.00	83.50	0.00606036	0.00482879	0.00537273
Side	Left	98.50	99.00	99.50	0.00916751	0.00706357	0.00788206
	Right	99.00	99.50	99.50	0.00611937	0.00560108	0.00540865

3. Conclusions

The tests show that the matching of skeletons generated by three-dimensional vision sensor, by means of a set of chain codes generated by high-resolution three-dimensional resolution grid, using only the chain codes of arm poses when the subject is in front of the camera can be used for gesture recognition. The matching accuracy is reliable enough to use it as input for gesture recognition using probabilistic signal analysis models, such as Hidden Markov Models; and the matching running time is fast enough to use the chain code matching algorithm for real-time applications. This gesture recognition system will be used on the robots of the Mexican Robocup@Home teams Markovito (INAOE) and Pumas (UNAM).

References

[1] S. Aslam, C. F. Barnes, and A. F. Bobick. Video action recognition using residual vector quantization and hidden markov models. In *IPCV'10*, pages 659-666, 2010.

[2] E. Bribiesca. A chain code for representing 3d curves. *Pattern Recognition*, **33** (5):755 - 765, 2000.

[3] E. Bribiesca. A method for representing 3d tree objects using chain coding. *J. Vis. Comun. Image Represent.*, **19**:184-198, April 2008.

[4] E. Bribiesca and W. Aguilar. A measure of shape dissimilarity for 3d curves. 2008.

[5] E. Bribiesca and C. Velarde. A formal language approach for a 3d curve representation. *Computers & Mathematics with Applications*, **42**(12):1571-1584, 2001. [6

[6] E. Guenterberg, H. Ghasemzadeh, V. Loseu, and R. Jafari. Distributed continuous action recognition using a hidden markov model in body sensor networks. In *Proceedings of the 5th IEEE International Conference on Distributed Computing in Sensor Systems*, DCOSS '09, pages 145-158, Berlin, Heidelberg, 2009. Springer-Verlag.

[7] A. K. Koc. Body posture analysis in the context of shopping. Master's thesis, TU Delft, June 2011.

[8] S. Koceski and N. Koceska. Vision-based gesture recognition for human-computer interaction and mobile robot's freight ramp control. In *Proc. 32nd Int Information Technology Interfaces (ITI) Conf*, pages 289-294, 2010.

[9] C.-M. Oh, M. Z. Islam, and C.-W. Lee. Pictorial structures-based upper body tracking and gesture recognition. In *Proc. 17th Korea-Japan Joint Workshop Frontiers of Computer Vision (FCV)*, pages 1-6, 2011.

[10] J.-W. Park, Y.-C. Lee, B.-S. Jo, and C.-W. Lee. Virtual playing ground interface using upper-body gesture recognition. In *Proc. 17th Korea-Japan Joint Workshop Frontiers of Computer Vision (FCV)*, pages 1-5, 2011.

[11] Q. Shi, L. Wang, L. Cheng, and A. Smola. Discriminative human action segmentation and recognition using semi-markov model. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1 -8, june 2008.

[12] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time human pose recognition in parts from single depth images. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 1297 -1304, june 2011.

[13] M. Sigalas, H. Baltzakis, and P. Trahanias. Gesture recognition based on arm tracking for human-robot interaction. In *Proc. IEEE/RSJ Int Intelligent Robots and Systems (IROS) Conf*, pages 5424-5429, 2010.

[14] J. Sung, C. Ponce, B. Selman, and A. Saxena. Human activity detection from RGBD images. Technical report, Department of Computer Science, Cornell University, Ithaca, NY 14850.

[15] J. K. T. Tang, J. C. P. Chan, and H. Leung. Interactive dancing game with real-time recognition of continuous dance moves from 3d human motion capture. In *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*, ICUIMC '11, pages 50:1-50:9, New York, NY, USA, 2011. ACM.

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-242

Human-Robot Interface using Face Detection and Recognition, for the Service Robot, "Donaxi"

Hector VARGAS[a,1], Esperanza MEDINA[a], Daniel MARTINEZ[a], Edson OLMEDO, Gerson BERISTAIN
[a]Universidad Popular Autonoma del Estado de Puebla, Computer Faculty

Abstract. This paper is in the context of algorithms for detection and face recognition, as well as face tracking, integrating them into service robots at home, this paper was made for service robot "Donaxi", to the contest in RoboCup. The proposal was to set up a hybrid algorithm with different functional algorithms for the detection, recognition and face tracking, even after losing face. For detection we used the Haar algorithm, for tracking Kalman filter was used, and recognition strategy was used to segment the image into regions weighted, and discriminate against them through a decision tree for possible recognition. The experimental results show that the system is stable when saturated its field of view with many faces or persons, in this case only discriminate which is closer to the robot. It is noteworthy that followed the rules of test "Who's Who" of the category of the RoboCup at Home, this test is still an open problem for the robotics community in the area of Human-Robot Interface.

Keywords. Hybrid algorithm, face detection, face tracking, face recognition, Kalman filter, Decision trees, Phase correlation, Histogram comparison.

Introduction

One of the great challenges in computer vision is to improve the automated systems for objects detection and tracking or regions of a set of images. Nowadays, detecting human-face is a discussed problem using different types of features. Face tracking is a difficult problem because faces are deformable objects with areas of little texture. Most of the algorithms for detecting face using binary pattern-classification. This means that the content of an image is transformed into features for a face classifier previously trained, decides if the image is a face or not. The process of face detection consists of two steps. The first is building models. The second step is to find a particular region in the image, called area of interest.

---

[1] Corresponding Author: Hector Vargas, Esperanza Medina, Daniel Martinez, Edson Olmedo, Gerson Beristain UPAEP, 21 Sur 1103CP 72410, Puebla, Mexico;
E-mail: {hectorsimon.vargas, aldodaniel.martinez , edson.olmedo, gerson.beristain}@upaep.mx, esperanza. medina@upaep.edu.mx

Some other tracking algorithms use the advantage of correlations between image frames to accelerate tracking process. Fundamentally, the features of the object itself are local information, and the features of an image sequence belong to global information.

There have been many works for object detection and tracking for object detection and tracking. Most of these algorithms have the main idea on estimating the movements of area of interest using probabilistic theories like some popular models and approaches, like Kalman filter.

Baek proposes a face tracking algorithm, this consider the corresponding face vector status, including the central position, the size of the containing rectangle, the average colored area and their first derivatives [1]. For the evaluation of new face candidates we employ a Kalman's estimator. In the tracking phase the algorithm detect if the face is new or not, using the face's previous frame like a template, with this data calculates the new face position.

In parallel, another point of attention has been the faces recognition. In this context is presented problems in the variability of the head rotation, the intensity and angle of light, facial expression to take a picture and others which have complicated the recognition systems [2] [9].

Face recognition methods and algorithms commonly assume [10] that face images are aligned and have an equal posture, but in many practical applications it is impossible that these conditions will be present. One solution that has been presented to reduce these factors is to obtain sub-features of a face, which enables a better approximation to say whether a face is known or not based on the analysis results and related features.

Some work on face recognition [10,11] have shown that an adequate solution to measure the characteristics is the use of decision trees, a probabilistic technique that through certain events and probability weights can make a decision in context.

This paper is organized as follow. In section 1, related algorithms that will be used to built a hybrid system for detecting human-face detection, recognition and tracking. Section 2 describes the detailed information of hybrid system that we propose. The results are presented in section 3 and the conclusions in section 4.

1. Related Algorithms to build a hybrid system

To build a hybrid system for detecting human-face and tracking unrestricted is necessary to use different algorithms. In This paper we propose an algorithm for recognizing and tracking a person.

The goal to this algorithm is to search a face in an image using a recognition process based on the detection of features that encode some information about the face to be detected, once face image detected is cropped, the segmentation is to determine if a face is already known, continues to track the face, so it will trace the journey that the faces will be making and storing a image is taken from the coordinates in which the face is detected, it is able to continue to monitor if the face is not fully facing the camera.

When the face is back facing the camera turns to make the recognition to ensure that the face is the same as it began to follow, Figure 1.

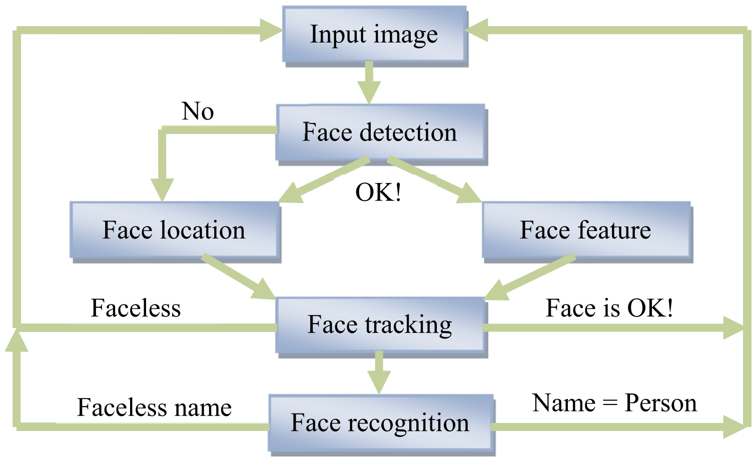


Figure 1: Hybrid system to Detection, Recognition and Tracking Unrestricted.

Importantly, firstly the algorithm must find a face to follow. We should also mention that this algorithm was tested on the service robot "Donaxi@HOME".

1.1. Face Detection

Face detection is the first stage of an automated face recognition system, since a face has to be located before it is recognized. In this case, we use the object detector of OpenCV. This object detector has been initially proposed by Paul Viola and improved by Rainer Lienhart. The algorithm's first step is the creation of a training file. The generation involves an analysis of various images and face views (positive examples), and arbitrary images (negative examples).

This training can be applied to a region of interest in an image this region must be to the same size as used during the training. The working of the training produces a "1" (if the región shows a face) and "0" (not face in image). To search for the object in the whole image one can move the search window across the image and check every location using the classifier.

1.2. Face location

Once a face is detected within the image, the center coordinates of the face to know are stored so the system can map the exact location of the face within the image when the face changes its position; the new coordinates are also stores in order to know the person trajectory.

With the obtained data is generated a line graph as this type of graphic is used to compare values over time.

After we are going to analyze the behavior of the graph and discard those values that are far away from others to ensure that the plotted points show the person trajectory.

To find the direction and extent to which the person moved draw a line from the point where the face was detected for the first time to the last point where the face was detected. The information that we obtain after to detect a face is used in the Kalman filter configuration to face tracking.

1.3. Kalman Filter

In order to keep the person even when the face isn't fully in front of the camera we use the Kalman filter. Kalman filter estimates the state variables of a process with feedback. Calculate the process state at some point and then get information (fed back) of the measure.

If you want to apply the Kalman filter to face tracking, you must provide a characteristic representative thereof, to be taken as the observation of the object. In this case an image is stored each time a face is detected or a picture of where the face was detected last. To calculate this point we will call the center of mass and determines the position of the stored image is necessary to perform a series of operations on the image.

To identify movement of the face within the image is necessary to apply motion detection techniques. This way we continue to face that hovers over the image.

The study of movement of the face within the image is essential to introduce the time variable. A sequence of images is given by the function f (x, y, t), where x, y are the coordinates of the center of the head at a particular time instant, t. Therefore the value of f (x, y, t) represents the intensity of pixel (x, y) within the image t.

In face tracking measures are necessary observation corresponds with the position of center of mass of the face. In this case it is used to determine the position based on projections which use the center of mass, which, as mentioned above is obtained from the center point of the head. Previously eliminates potential noise given only to those parts of the image over a fixed number of pixels above the threshold. In this way we make sure not to consider some isolated pixel face.

1.4. Probabilistic Decision Trees

Decision trees [4] are a widely known formalism for expressing classification knowledge. The traditional approach to constructing a decision tree from a training set of cases described in terms of a collection of attributes is based on successive refinement. In a general context, the decision trees are particularly useful when one or more decisions in sequence should be taken and they are affected by one or more uncertain events to which they are assigned a certain probability weighting which directly affect the value (usually expressed in earnings) for final decision.

In the context of this work, we take the principles of decision trees for use in face recognition, using as a key decision if the face is known or not, and as uncertain events similarities and differences in the detected face real time, against the face stored in the database of known persons. Similarity values and differences were obtained using phase correlation algorithms and comparing histograms.

In order to achieve this aim with better efficiency, face division by subsections was used, based on the aesthetic theories [5] where each subsection is analyzed in an independent way with the similarity and difference algorithms

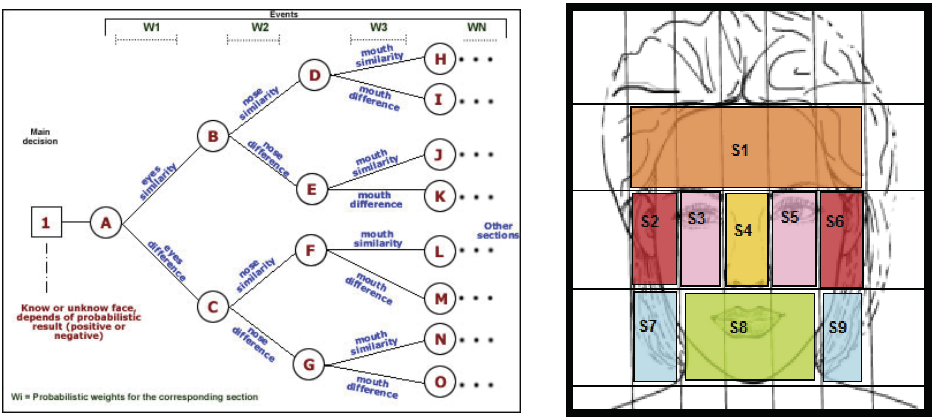


Figure 2. Decision tree and Face division in 9 characteristic subsections.

Each one of these subsections received a probabilistic weight based on how characteristic for the face recognition is, for finally integrate them in a probabilistic decision tree.

1.5. Phase Correlation

It is a method for determinate the displacement of one picture respect to other similar [7]. However phase correlation utilizes a direct way comparison, results are too ambiguous because it's required that the compared face images accomplish the same characteristics when it is used in a direct way. When this technique is combined with the face subdivisions technique, efficiency is incremented and results are more trustable for faces equality comparison.

1.6. Histogram Comparison

An image histogram [8] is a graphical representation of the intensity distribution of an image that quantifies the number of pixels for each intensity value considered. For the face recognition purpose we use the histogram equalization like a way to improve the contrast in an image in order to stretch out the intensity range.

With this first treatment, an image with a more generic tonality was achieved in order to make the recognition process based on the accumulative distribution equation.

$$H^{\wedge\prime}(i) = \sum_{0 \le j < i} H(J)$$

Figure 3. Histogram equalization

The second histogram utilization is a comparison with the saved histograms of known subsections faces in order to be utilized as concept of comparison in the decision tree algorithm for face recognition. To compare two histograms (H1 and H2), first we have to choose a metric (d (H1, H2)) to express how well both histograms match.

For the algorithm, correlation and intersection histograms comparison metrics [8] are utilized.

2. Hybrid Algorithm

2.1. Tracking Face

The hybrid algorithm to track a person works as follows:

1 Detection and face cutting: At the beginning of the process, the system detects the person's face using the haar classifier, once is defined a face is there, it is cut out from the picture and becomes the new interest region.

2 The mass center of the interest region is obtained and stored in a database which will store all coming up mass centers.

3 The stored mass centers are plotted and the Kalman filter is applied to them in order to exclude those items which are far apart.

4 Estimate the position whre the person is located.

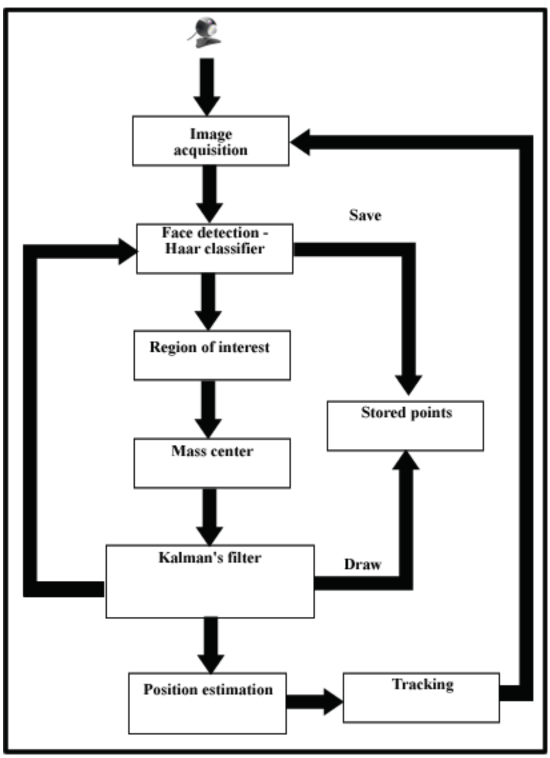


Figure 4. Face tracking hybrid algorithm

2.2. Face Recognition

The hybrid algorithm for face recognition works in three main phases:

1 Detection and cut face: In this first stage, we detect the face of the person using the haar classifier, once defined that there is a face, this is cut out of the picture.

2 Division of the face: Having defined the image available of the face, this is divided into 9 regions of space.

3 Calculation of similarity: Finally for the algorithm, correlation and intersection histograms comparison metrics [8] are Utilized For Each one of 9 Subsections. Both operations are performed with the detected face and the familiar faces and previously stored. The results of these operations are introduced to the decision tree, by which you get a percentage of similarity of the detected face against a familiar face before.

If the percentage of similarity exceeds a minimum value assigned to the face is considered as known, otherwise, the face belongs to a person not known to the algorithm.

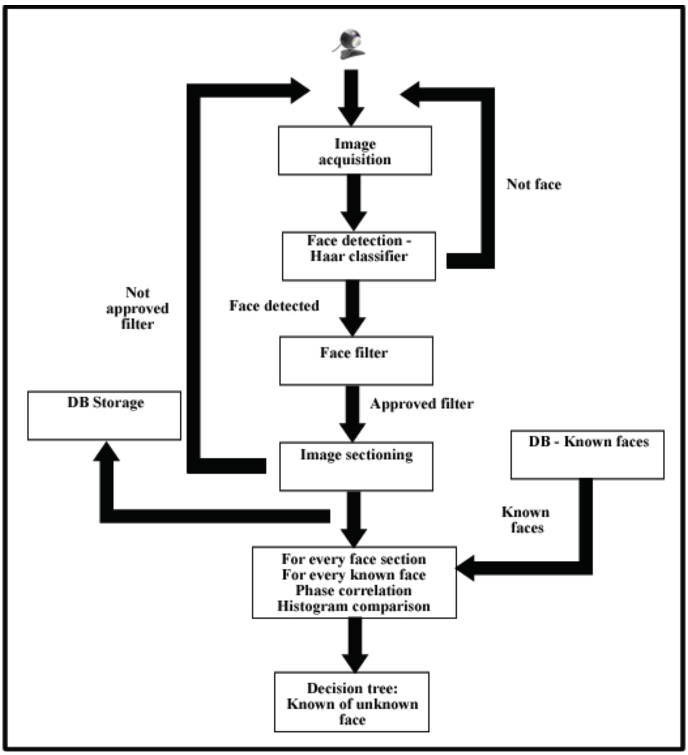


Figure 5. Face recognition hybrid algorithm

3. Results

3.1. Tracking Face

For the implementation of the experiment, we were the participation at least one person to track. When starting up, the person must be placed in front of the camera to detect possible face forward. Once the face is detected locked up his face in a circle and paint the center point of the face and the journey that makes the person as it moves.

Finally, the line graph which would indicate the direction and magnitude at which the person moves. To verify if the plotted course was right, it was proof that a person stands in front of the camera waiting for her face is detected and so get moving. If the program traces the route that matches the person that I do, it is considered that the test was successful.

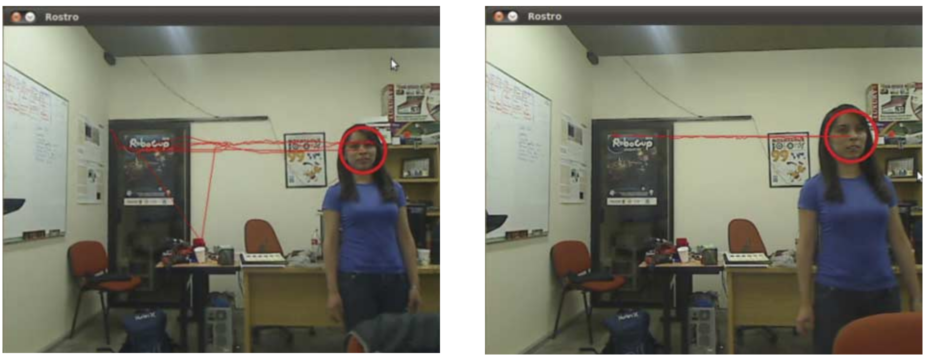


Figure 6. Face Tracking and Kalman filter application. The first image on the left is seen to have many disturbances generated by the same variation in the identification of the face, in the second image on the right is applied Kalman and observed substantially improving the correction of the estimated center position of the head.

3.2. Face Recognition

This face recognition algorithm was planned to solve the problem of vision test posed "Who is Who" for the RoboCup 2012, so for testing and obtaining the same results were a series of tests emulating this specific activity. The evidence that was submitted to the algorithm were to teach four previously (acquaintances) and then start the recognition phase where different people (both known and unknown) were located in front of the camera and then proceeded to run your algorithm procedure to determine whether they were known or unknown to him.

In tests, the system has been exposed to a total of 205 choices of action (situation where you have to decide if a face is known to him or not) of which were obtained the following results:

Face Recognition Tests		
	Quantity	Average (%)
Action decisions	205	100
Correct decisions	178	86.83
Uncorrect decisions	27	13.17

Table 1 Obtained results for face recognition.

These tests were performed by taking a static camera images, and showed positive results with a high percentage of correct recognition, as well as efficient processing speed.

Figure 7. Face recognition real time application.

4. Conclusions

The algorithm was tested in the service robot "Donaxi @ HOME", which participates in the RoboCup since 2009.

The model for face tracking and recognition is proposed in this paper seeks to exploit the ability to use the comparison of characteristics to define whether a face is known or not, based on a current face and the features of a face stored above so we can determine if the face that is always the same. This was determined by face recognition at the beginning and end of follow up.

The combination of different techniques shown here (haar classifier, face split into subregions, phase correlation, comparison of histograms, decision trees and Kalman filter) have unique capabilities and advantages that may develop in part a recognition and tracking faces, but in combination (to generate the proposed hybrid algorithm) can get a better focus and greater efficiency when performing this task.

The results have shown that one can obtain a high efficiency rate with this algorithm, although the results should be viewed with caution because this algorithm has not yet passed the phase of high-level tests (test the algorithm bases data and robust free faces that are available on the web) so what is the next phase of its growth.

5. Acknowledgment

We want to thank to the Mexican Robotics Federation (FMR) and Robotics and Mechatronics Network of Mexico for the support provided to carry out this work. Also at the National Council of Science and Technology (CONACYT) of Mexico for push

forward technological development in the area of service robotics in Mexico.

References

[1] K. B. et al. AI 2005, chapter Multiple Face Tracking Using Kalman Estimator Based Color SSD Algorithm, pages 1229–1232. LNAI 3809. Springer-Verlag Heidelberg, 2005.
[2] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET Database and Evaluation Procedure for Face-Recognition Algorithms," Image and Vision Computing, vol. 16, pp. 295-306, 1998.
[3] Computer Vision Research Progress http://books.google.com.mx/books?id=g9bP-7oBulUC&dq=%22Some+researchers+proposed+different+control+and+noise+models%22&hl=es&source=gbs_navlinks_s.
[4] Yves Kodratoff, Ryszard S. Michalski, "Machine learning: an artificial intelligence approach Volume III", Morgan Kaufmann; 1 edition, ISBN-10: 1558601198, pages 140 – 146, August, 1990
[5] Daniel Maturana, Domingo Mery y Alvaro Soto, "Face Recognition with Decision Tree-based Local Binary Patterns", Department of Computer Science, Ponticia Universidad Católica de Chile, Chile, 2010
[6] Jesus B. Cedeño, "La cara, sus proporciones estéticas", Clinica Central Cira García, La Habana, Cuba.
[7] B. S Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration", IEEE Transactions on Image Processing 5, no. 8 (1996): 1266–1271.
[8] OpenCV dev-team, "OpenCV v2.3 documentation" - Histogram Equalization - Histogram Calculation - Histogram Comparison, August, 2011
[9] Carlos Villegas Quezada, "Reconocimiento de Rostros Utilizando Análisis de Componentes Principales: Limitaciones del Algoritmo", Universidad Iberoamericana, 2005, URL: http://www.bib.uia.mx/tesis/pdf/014620/014620_00.pdf
[10] D. Maturana, D. Mery and Á. Soto. Face Recognition with Decision Tree-Based Local Binary Patterns. Tenth Asian Conference in Computer Vision (ACCV) 2010.

This page intentionally left blank

1st Workshop on Future Intelligent Educational Environments (WOFIEE'12)

Victor Callaghan (University of Essex, UK)
Minjuan Wang (San Diego State University, USA)
Juan Carlos Augusto (University of Ulster, UK)

This page intentionally left blank

J.A. Botía et al. (Eds.)
IOS Press, 2012

Introduction to the Proceedings of WOFIEE'12

As the world moves steadily to become a knowledge-based economy, education and learning have never been more important. Technology is playing an increasingly crucial role in the delivery of education, which in turn is driving research into the search for ever better technological solutions. The age of intelligent environments is bringing such pedagogical advances as smart classrooms, intelligent campuses, immersive and mixed-reality learning, affective learning, mobile learning, intelligent learning clouds and personalised intelligent tutors to revolutionize current learning practices, and to challenge the traditional notion of a university or school.

Based on this context we launched the 1st International Workshop on *Future Intelligent Educational Environments (WOFIEE'12)*. We hope this will provide a forum where to discuss the current state-of-the-art, imagine solutions for current limitations and plan steps within our community which may help to achieve some of the required advances in this area. This event will serve as a forum for researchers and practitioners to discuss the latest intelligent technologies that can support the development of new educational technologies and environments around the world.

Central to this forum are enriched physical and virtual environments such as smart classrooms, virtual / mixed-reality environments, intelligent learning clouds or mobile and augmented-reality systems that can interact with students and teachers at a pedagogical level, so as to bring true innovations to education. We also include the systems that support the learning of practical skills, such as those typified by science and engineering laboratories that are critical to students. We also consider the wider campus infrastructure, which can also impact the cost and effectiveness of education. Examples include smart signage that can guide people around a campus or smart applications for timetabling or managing the environment of the teaching facilities. Education is increasingly global and the cultural dimension is a topic we consider important for this event. Finally, scientific research, engineering Innovation and business advancement are beneficiaries of good education and it's fitting that we are including a paper that considers how entrepreneurship for creating and commercialising future educational environments might be supported.

An additional notable development we are pleased to report is the creation of the *Transactions on Future Intelligent Educational Environments (TOFIEE)* [1] which will provide a rigorous academic forum where significant achievements will be published for global access.

As a final note, we wish to express our sincere thanks to the WOFIEE'12 Program Committee for their thorough reviews and strong support. We are looking

[1] http://icst.org/future-intelligent-educational-environments

forward to meeting you all at this first workshop which we hope is the start of a series of future editions, which will allow us to meet and build a strong community dedicated to introducing exciting innovations to education.

Juan C. Augusto, Vic Callaghan, Minjuan Wang
Co-chairs WOFIEE'12

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-257

The University as an incubator for technology, innovation, and entrepreneurship

Keynote Speaker

Professor James Carlson[1]
Clay Mathematics Institute, Cambridge, Massachusetts

Abstract: One of the great drivers of economic growth in the US is the creation of new companies by entrepreneurs coming from academia. Notable examples are Adobe Systems, Google, and the Internet itself. While some academic environments are wildly successful as incubators for technology and entrepreneurship, not all are. I will discuss individual cases, some large, others small, some from other countries, by means of which one can identify factors for success.

Keywords. Academic Enterprise, Entrepreneurship, Technology Innovation, Business Incubators, Knowledge Transfer.

Brief Biography: James Carlson received his B.S. in mathematics at the University of Idaho in 1967, his Ph.D. in mathematics from Princeton University in 1971. From 1995 to 2002 he was chair of the Mathematics Department at the University of Utah. Since 2003 he has served as president of the Clay Mathematics Institute. He is the author of two books and over forty research articles.

[1] Corresponding Author: James Carlson Email: jxxcarlson@me.com

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-258

Physical Activity and Team Work with Pervasive Games

Fernando MARTÍNEZ-REYES
Facultad de Ingeniería, Universidad Autónoma de Chihuahua
Chihuahua, México
fmartine@uach.mx

Abstract. The availability of computer technology in educational institutions offers opportunities for the implementation of augmented spaces for human-computer interactions. We can observe the increasing adoption of technology such as computers, wireless networks, and electronic whiteboards in schools to support teaching activities. Whereas it is plausible that schools use technology to reinforce learning, we think there are other educational activities being set aside. This article presents two works in progress: The Game of Pirates and The Virtual Maze. The Game of Pirates is a game which aims to encourage physical activity, therefore it addresses childhood obesity. This game asks children to engage in vigorous physical activity in order to obtain the keys that open the treasure chest. The Virtual Maze is a resource that can be used to promote children's teamwork. Two children, the main player and the collaborator, need to communicate, exchange opinions, and take joint decisions in order to traverse the labyrinth and complete the game. Current implementations are being tested in the laboratory and will be soon available to offer children of primary school age an experience with these kinds of mobile computing experiences.

Keywords. Mobile computing, pervasive games, virtual environments, children's physical activity.

Introduction

Computer technology is pervading educational settings. Computers, wireless networks, electronic whiteboards, and web-based "Enciclomedias" [1] can be part of today's classrooms. These types of technology enable teachers to deliver digital content which can influence the teaching-learning process. Whereas the efforts schools are making to improve traditional teaching of science subjects using technology are valuable, we argue that other academic elements such as physical activity and social interaction, important elements in the development of children and young people [2], can also be empowered with technology-augmented spaces. Lack of physical activity and the encouragement of teamwork are two of the social issues that schools worry about the most.

The lack of physical activity in primary schools is one of the greatest international concerns, because the children's health and academic performance has been seen affected [3, 4]. Although schools try to promote, for instance, physical activity by providing training courses in sports, this could not provide the recommended one hour per day for young people [5]. Moreover, schools may not have an adequate level of infrastructure to promote sufficient physical activity in children. Jennings et al. [6], from a study with twelve Mexican primary schools, reported that in primary education children have often only two opportunities to develop physical activity, during the lunch break and the physical education session. They also found that the opportunities a child might have to undertake physical activity during the 30 minutes of the daily school "recess" are limited because there are usually limitations of space and equipment, and even schools' policies prohibiting ball games. Their study indicated, additionally, that there are typically only two sessions a week were children are taught sports; each session lasting 40 minutes, and 17 minutes as the average time a child has to undertake vigorous physical activity.

Another area of concern for primary schools is the children's social integration, key to encourage cooperative and collaborative group work, and which in turn affects positively pupils' academic and social outcomes [7,8]. Grouping practices are commonly used to set-up cooperative learning, but this task is either challenging for teachers to be set up or rarely successful in taking learning to a higher level. Organizing group work considers not only physical space but also pupils' learning and more importantly, the classroom's social pedagogy: children's communication, interaction between peers, the attitudes of individuals, their engagement with tasks, and so on. Furthermore, group work should involve children as co-learners and would include cooperative teamwork [9].

This paper describes "The Game of Pirates" and "The Virtual Maze", two game-based prototypes aiming to support some educational activities for primary schools. The former, would help the encouragement of physical activity whereas the latter, would offer a resource to support children's collaborative work. The next section, reviews previous work in which ubiquitous computing is applied for academic purposes. Subsequently, we present the implementation of The Pirates and Virtual Maze games, and their educational features. Finally, conclusions and future work are presented.

1. Related work

The availability of computer technology in educational contexts opens a window of opportunities to create spaces of interaction between human and computer. There is a significant interest from the scientific community to integrate computing technology to empower educational experiences. In "Savannah" [10], for instance, off the shelf technology, such as wireless routers and mobile devices, was used to create a virtual environment through which children learn about group behavior of lions when hunting for food. Mobile computing was similarly applied in "Ambient Wood" [11], an educational experience that enabled children to collect evidence in a forest ecosystem to complement the theory given in the classroom. Additionally to learning contexts, technology has also been explored to support social integration. In "Pirates" [12], authors create a virtual environment where players explore islands to complete missions and to look for treasure, which allows them to get sturdy ships. Players navigate around the physical space with their PDA, which renders the virtual world of the game. The proximity of a player to an island means the pirate is hunting for gold, and his/her proximity to another player means war! Recently, some authors have been interested in promoting the development of physical activity through the use of existing, sometimes traditional, games. The game "Stop the Bomb" [13], for example, could be considered as a version of the traditional game "Catch Me If You Can."

Computing technology seems to have also found a niche to support children's education on, and practice of, collaborative/cooperative tasks, i.e. teamwork. "Games of Water" [14], for instance, has a vision system installed around a fountain to capture user activity. The more children play around the fountain, the more extensive are the water dancing effects. Kidpad [15] and Klump [16] are two storytelling-like tools that have a shared virtual interface in which children work together colouring, shaping or creating fictional personages. "Virtual Learning Environments" (VLEs), are also applied to create environments that assist people with special needs. The "Collaborative Puzzle Game" [17] seeks to foster collaborative skills in children with Aspergers Syndrome Disorder (ASD). The game play *forces* players to carry out joint actions on digital objects.

To summarize, similarly to previous experiences, our work explores application scenarios for human-computer interactions, and focuses on the implementation of game-like experiences that could support educational activities [18] related to physical activity and peer-to-peer collaboration.

2. The Pirates' Game

The Game of Pirates can be played between individuals or teams. Children compete by physically exercising to obtain the keys that open the treasure chest. The current version of The Game of Pirates considers jumping, pulling a rope and running around as the physical exercises children need to take to a level such that points, and therefore keys, are gained. Figure 1(left), shows the map of the island presented at the start of the game. The map displays information about the island and the different places where keys are hiding. Once the player moves to explore an area, the player is prompted to carry out a specific activity. For example, in Figure 1 (center), she/he is asked to jump vigorously in order to divert the stampede of animals. Whenever the child has completed the indicated level of activity, the player is given one of the keys to the treasure chest (Figure 1, right). The game ends when all keys have been obtained. Players can lose if they run out of time or fail to achieve the required level of activity at any point in the game.

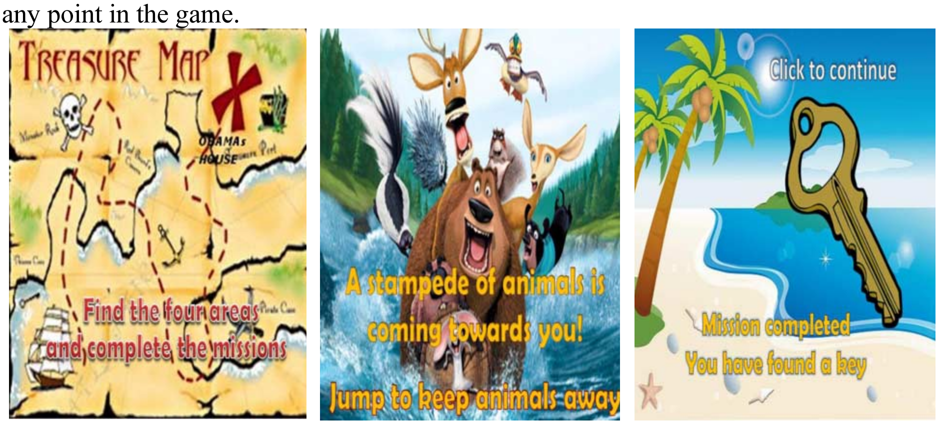


Figure 1. Some of the user's interfaces available along the game play.

Technology supporting The Pirate's Game includes Phidget sensors [19], a web server and a mobile device.

- Sensing technology. Vibration, proximity, light and temperature sensors are currently used. Sensors are attached to a PhidgetInterfaceKit I/O board which sends information back to the server wirelessly.
- Web server. A Tomcat web server registers sensing data into a MySQL database and synchronizes activity information with the mobile client. Whenever the client requests information, the server responds with information that identifies the active sensor, its value, the date and time of its last reading and a flag that indicates if other sensors are active or not. This data is sent in XML format as follows:


```
<?xml version="1.0" encoding="UTF-8">
<sensor_table>
    <Sensor>
        <TimeStamp>2010-09-0708:41:42.0</TimeStamp>
        <Sensor_Id>4</Sensor_Id>
        <Sensor_Name>Weight</Sensor_Name>
        <Sensor_Value>0.980000019073486</Sensor_Value>
        <Sensor_Flag>true</Sensor_Flag>
    </Sensor>
</sensor_table>
```


- Mobile client. An HP iPAQ 6510 Smartphone is the device used by users when playing the game. The game logic on the client side is implemented with the Mscape tool [20], which provides a development environment for experimenting with location-based services.

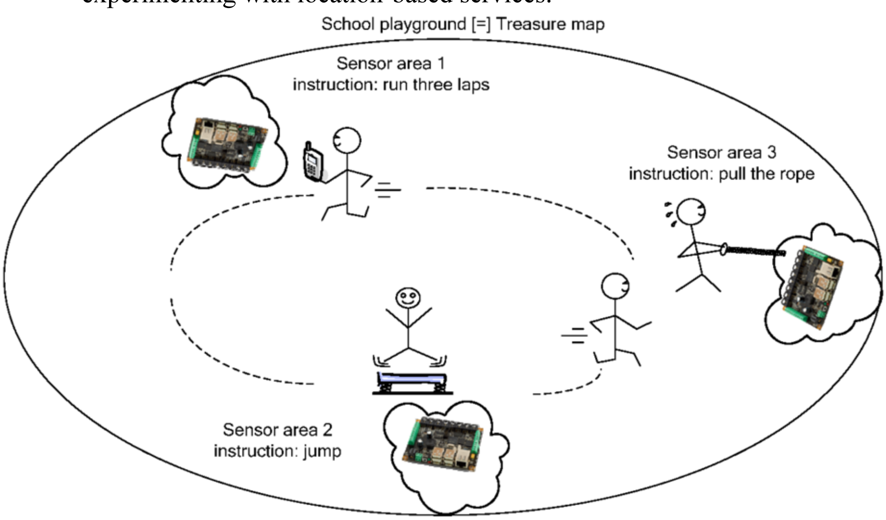


Figure 2. In The Game of Pirates, the player's proximity to an area of sensors starts synchronization between the server and the mobile client, which in turn instructs the user the type of activity to develop. Here the activities are: (1) run three laps, (2) jump, (3) pull the rope.

Figure 2 illustrates a representation of an elementary school playground, the distribution of a set of sensors in space, and some of the instructions the player would receive from the system.

3. The Virtual Maze

The Virtual Maze is a game-based collaborative environment for children of primary school age. The main player uses a mobile device to traverse a virtual labyrinth. A second player, who is in an adjacent room, supports and helps the main player to avoid obstacles and find the exit of the maze.

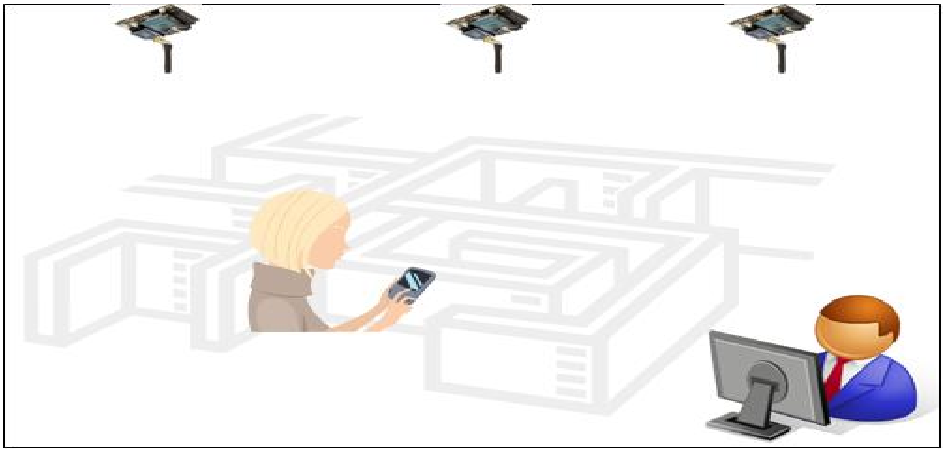


Figure 3. The Virtual Maze game. The Cricket location system is installed on the ceiling to track the user's position. Using his/her mobile device, the player walks through the labyrinth. The collaborator is ready to assist the main player remotely.

As indicated in figure 4, the game encourages social engagements in three levels: navigation, collaboration and interaction.

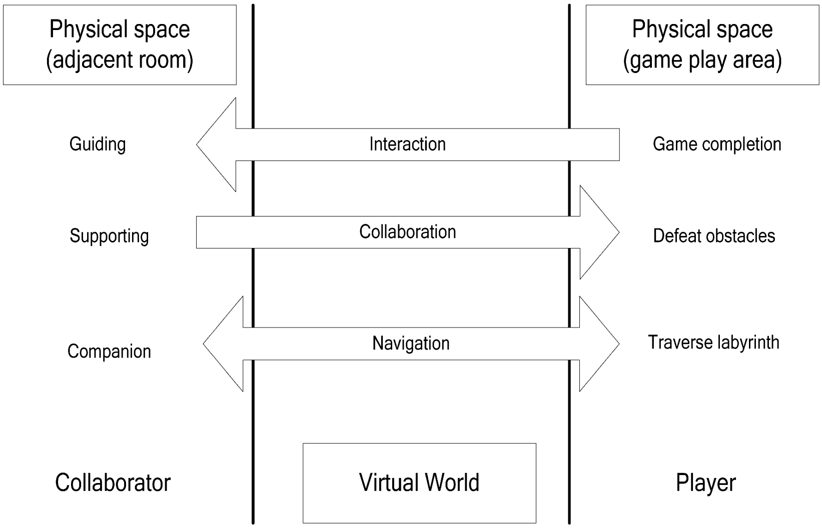


Figure 4. Collaborative approach for The Virtual Maze. Arrows indicate the expected level of engagement between players. In the lower level (navigation) communication between players is fair and in both ways. In the medium level (collaboration) the main player takes into account opinions and suggestions from the collaborator. In the highest level (interaction) the main player demands interventions from the collaborator to complete the game.

- Navigation. The basic level of social engagement is given at the navigation level. The player could request collaboration from some sort of 'companion'. In this case, the collaborator is just an observer for most of the time and only intervenes when the main player request information. For instance, a confident player, experienced with real physical mazes, may choose to go around The Virtual Maze alone and without any help from the collaborator.

- Collaboration. At this level, players share information in order to make joint decisions to progress forwards with the game. For instance, the main player could ask the collaborator what corridor might be the best choice, or which one might contain less problematic obstacles.
- Interaction. We expect there are going to be players who prefer to delegate major responsibilities to the collaborator. That is, the collaborator acts as the *tourist* guide for the main player, who might decide to completely follow instructions given by the collaborator.

In addition, the game has been designed to demand from players, awareness, intrigue, eagerness, uncertainty and need for assistance, so that social engagement can take place. The orchestration and control of the visual content of the game play, we believe, will also help to *force* some degree of collaboration and interaction between players. For instance, the 3D view of the virtual world for the main player, as it happens in a physical maze, is constrained to four walls, whereas the collaborator has a complete view of the labyrinth and its corridors. Therefore, the main player would solicit some assistance and support from the collaborator to walk through the maze. The collaborator can also identify obstacles an instant before the main player, hence prompting communication between players.

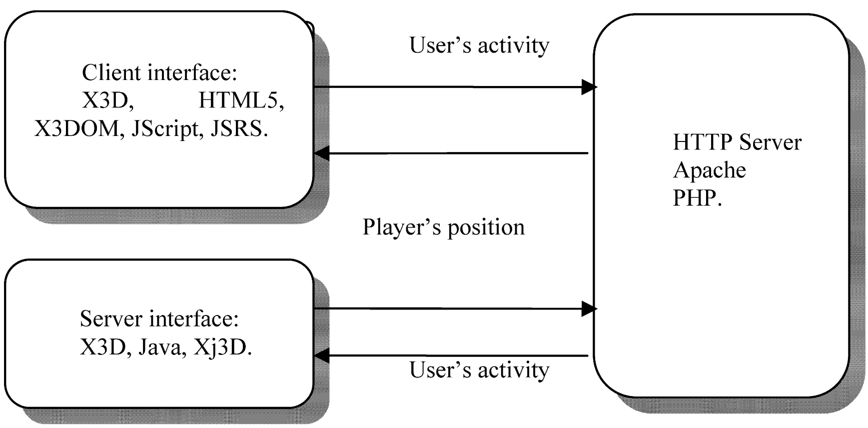


Figure 5. The Virtual Maze's architecture

The technology platform of The Virtual Maze consists of an indoor location system, a web server and a mobile client.

- Location system. The Cricket location system [21], allows us to get 3-axis coordinates from a sensor worn by the player. That information is sent to the web server.

- Server. This socket-based server helps with the processing of the virtual world for the collaborator. X3D, Xj3D and Java3D are the tools used to add functionalities to this user's interface.
- Mobile client. The virtual world of the main player runs on a Motorola XOOM tablet. Some of the web technologies integrated in the client include X3D, HTML5, X3DOM, Jscript and JSRS.

4. Results

Regarding The Game of Pirates, we have developed an initial prototype, and laboratory tests have been conducted to evaluate both the communication from the sensor layer to the user interface and the game play synchronization. These initial tests are promising, encouraging us to continue with the next step of considering making the game a real experience. This includes the design of physical artifacts that will be augmented with the appropriate technology. For example, a specially designed base is required to house the vibration sensors to allow children to carry out the jumping activity. Regarding The Virtual Maze game, we have separately carried out the evaluation of the location system and the virtual world for both the main player and the collaborator. The synchronization of events, both in 3D visualizations, has been manually tested. We have played and evaluated, in this case, that collisions are properly managed by the system; e.g. that the virtual world does not allow users to walk through walls. At this stage, we are merging the location system with the virtual world in order to evaluate the game as a whole.

5. Conclusions and future work

In this article we have presented the application scenario for ubiquitous computing - sensor technology, wireless networks and mobile devices – to create opportunities for human-computer interactions. In particular, given that some of the above technologies are readily available in primary schools, we argued that it is possible to implement augmented spaces that could empower educational activities. The Game of Pirates and The Virtual Maze are the two game-like prototypes through which children could be invited to engage in such activities. The Game of Pirates seeks to offer a resource for the encouragement of physical activity, whereas The Virtual Maze can be used to promote collaborative work.

Future work for The Game of Pirates includes the design of physical artifacts and the integration of technology into them. Once this phase is completed, it shall provide a first experience with the game for elementary school children. User experience with the prototype so obtained will allow us to determine the degree of acceptance by users as well as suggestions for improvements to the game. In addition, we are migrating the client's user interface to an Android-based mobile device in order to add more visual dynamics, hence increasing the interest of children. In terms of The Virtual Maze, we have to complete the integration of the location system within the dynamics of the game play and to test the system with real players; something we have planned for this summer of 2012. Furthermore, we are working with a Chronos EZ-430 development system, which would be integrated as part of the game to collect real data from the user's experience, e.g. hearth rate information. Making video chat available is also part of the proposed future work for The Virtual Maze.

Acknowledgment

We would like to thank the PROMEP for funding the development of The Game of Pirates and The Virtual Maze through the PEMT (Platform for the Experimentation with Mobile Technologies) project, at the Autonomous University of Chihuahua. The authors wish to express their gratitude to reviewers for their invaluable feedback. We would also like to thank Magda Maldonado and Carla Bustillos for their comments.

References

[1] Secretaría de Educación Pública (Department of Public Education), México, Enciclomedia, Visited on June 17th 2010, http://www.enciclomedia.edu.mx.
[2] R.R. Pate, M.G. Davis, T.N. Robinson, Thomas, E.J. Stone, T.L. McKenzie and J.C. Young, . Judith, Promoting physical activity in children and youth, Visited on April 17th 2012, http://circ.ahajournals.org/cgi/reprint/114/11/1214.
[3] J. Byrd, The impact of physical activity and obesity on academic achievement among elementary students, Visited on April 17th 2012, http://cnx.org/content/m14420/latest/.
[4] World Health Organization, Obesity and overweight, Visited on April 17th 2012, http://www.who.int/dietphysicalactivity/publications/facts/obesity/en/.
[5] S. Biddle, N. Cavill, J. Sallis, Young and Active? Young People and Health-Enhancing Physical Activity: Evidence and Implications, Health Education Authority, London, 1998.
[6] N. Jennings-Aburto, F. Nava, A. Bonvecchio, M. Safdie, I. González-Casanova, T. Gust and J. Rivera, Physical activity during the school day in public primary schools in Mexico City, Visited on April 17th 2012, http://academic.research.microsoft.com/Publication/10165154/physical-activity-during-the-school-day-in-public-primary-schools-in-mexico-city.
[7] E. Baines, P. Blatchford and A. Chowne, Improving the Effectiveness of Collaborative Group Work in Primary Schools: Effects on Science Attainment. British Education Research Journal, 33-5 (2007), 663-680.
[8] R. Slavin, E.A. Hurley and A. Chamberlain, Cooperative learning and achievement: theory and research, W.M. Reynolds & G.M. Miller (Eds.), *Handbook of Psychology: Educational Psychology*, 7, 177-198 (2003), Wiley, New York.
[9] P. Blatchford, P. Kutnick, and M. Galton, *Towards a Social Pedagogy of Classroom Group*, International Journal of Educational Research, 39 (2003), 153–172.
[10] S. Benford, D. Rowland, M. Flintham, R. Hull, J. Reid, J. Morrison, K. Facer and B. Clayton, Designing a location-based game simulating lion behaviour, Proc. Advanced Computer Entertainment, 2004.
[11] Y. Rogers, S. Price, G. Fitzpatrick, R. Fleck, E. Harris, H. Smith, C. Randell, H. Muller, C. O'Malley, D. Stanton, M. Thompson, and M. Weal, M, Ambient wood: designing new forms of digital augmentation for learning outdoors, In Procedings of Advanced Computer Entertainment, 3-10 (2004).
[12] S. Björk, J. Falk, R. Hansson and P. Ljungstrand, Pirates! - using the physical world as a game board, In Proceedings of Interact, 2001.
[13] K. Hendrix, G. Yang, D. Van de Mortel, T. Tijs and P. Markopoulos, 2008. Designing a head-up game for children, British Computer Society Conference on Human-Computer Interaction, British Computer Society, LNAI 1793, Springer Verlag, 389-400 (2000).
[14] A. Carreras and N. Parés, Designing an interactive installation for children to experience abstract concepts, new trends on human-computer interaction, Research, Development, New Tools and Methods, Springer, 31-40 (2009).
[15] A. Druin, J. Stewart, B. Bederson and J. Hollan, KidPad: a design collaboration between children, technologists and educators. CHI97 pp. 463-470 (1997).

[16] L.E. Fahlén, C.G. Brown, O. Stahl and C. Carlsson, A space based model for user interaction in shared synthetic environments, InterCHI, 1993.
[17] A. Battocchi, A. Ben-Sasson, G. Esposito, E. Gal, F. Pianesi and D. Tomasini, *Collaborative Puzzle Game: A Tabletop Interface for Fostering Collaborative Skills in Children with Autism Apectrum Disorders*, Journal of Assistive Technologies, 4 (2010), 4-13.
[18] I. Verenikina, P. Harris, P. Lysaght, Child's play: computer games, theories of play and children's development, Young Children and Learning Technologies, Australian Computer Society, 2003.
[19] Phidgets at http://www.phidgets.com/
[20] HP MSCAPE at www.mscapers.com
[21] Cricket Mote at http://www.memsic.com/support/documentation/wireless-sensor-networks/category/7-datasheets.html?download=132%3Acricket-kit.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-267

Adapting e-learning contents based on navigation styles and preferences

Rosa BASAGOITI[a], Asier AZTIRIA[a], Igor ORDOÑEZ[a], Ion PADILLA[a]
[a] *Electronics and Computing Department*
Mondragon University
Mondragon, Spain
rbasagoiti@mondragon.edu

Abstract. Personalized access to e-learning contents is based on suggesting routes to users considering their preferences. Recent studies show that navigation styles can content much of the information related to cognitive learning styles and preferences. These navigation styles can be defined by means of different parameters. In order to develop a system that provides personalized learning contents, it is necessary to automatically extract student's characteristics and preferences. For that, we use data mining techniques that are methods that, basically, use indicators about student's activity to generate models that can direct the content selection process. In this paper, we address the problem of generating student's learning characteristics and preferences, which can make the task of designing adaptation models easier. This work classifies students in each of the next dimensions: (Active/Reflexive, Deductive/Inductive, Visual/Verbal, Sensitive/Intuitive and Local/Global).

Keywords. Navigation styles, e-learning, Data mining

Introduction

Environments that provide intelligent services based on the current situation and preferences of the user are becoming real. So far, most of these environments are physical environments, such as smart homes or intelligent classrooms that considering the principles of Ubiquitous Computing, are able to provide personalized services in an unobtrusive and proactive way. The same principles can be applied to virtual environments where without physical devices, the environment can provide intelligent and personalized services to users.

In the last decade, terms such as e-learning or lifelong learning have emerged introducing new concepts of learning, which demand new contents and tools. Access to e-learning contents creates many challenges, overall when the system has to provide personalized contents to each user. Personalized contents mean contents based on preferences and previous interactions of each user with the system, taking into account that different users have different preferences and habits of navigations.

In order to develop a system that provides personalized learning contents, the following steps are necessary:

1) Design the domain model where a hierarchy of learning goals is set.

2) Define the user model where the current knowledge of the user is calculated respect to goals and where the user's cognitive characteristics and preferences have to be set in order to better fill his needs.
3) Define the media space, where the educational resources description model is written.
4) Define the adaptation model where the concept selection rules are used for concept selection from the domain model.

After designing the Adaptive Educational Hypermedia System, following the above-mentioned steps, the adaptation engine is responsible for interpreting the adaptation rules specified in the Adaptation Model in order to generate personalized learning paths. This process is called adaptive educational hypermedia sequencing.

However, the design of adaptive educational hypermedia systems still requires significant effort from the system's designer. Adaptation rules can help, making their work easier, but there are some problems that have to be addressed since dependencies between educational characteristics of learning resources and users' characteristics are too complex. This complexity introduces several problems on the definition of the required rules [1], namely:

- Inconsistency, when two or more rules are conflicting.
- Confluence, when two or more rules are equivalent.
- Insufficiency, when one or more required rules have not been defined.

There are some approaches that try to avoid these problems [2]. For example courses where all possible paths that match the objective of a student are previously defined. Then the system, adaptively, selects one of the desired paths, at each moment.

The educational hypermedia sequencing process has different abstract layers. The concept identification layer, where some contents are identified as adequate by the knowledge space and defined as goals. The concept selection layer, that has to work over the objects previously identified. These two layers are complex to be defined because, in the first layer, the dependencies between learning concepts are difficult to be defined and, in the second layer, the decision has to be made based on student's characteristics and preferences, when the learning characteristics of the students are not clear yet.

Data mining techniques can be useful in this second step, in order to automatically extract student's characteristics and preferences. Data mining techniques are methods that, basically, use indicators about student's activity to generate models that can direct the content selection process.

In this paper, we address the problem of generating student's learning characteristics and preferences, which can make the task of designing adaptation models easier. This work classifies students in each of the next dimensions: (Active/Reflexive, Deductive/Inductive, Visual/Verbal, Sensitive/Intuitive and Local/Global).

This paper is organized as follows. Section 1 describes different learning styles and related works. Section 2 explains the process of discovering users' navigation styles. Section 3 provides the conclusions of this research.

1. Learning styles and models. Related works

Data mining has been used in web based learning systems in order to improve the learning process [3]. The information generated by web learning systems is so wide that is impossible to analyze this information manually. Some automatic methods can be useful to analyze this data. Data mining techniques can extract new knowledge in order to guide the student during the learning process. This guide can be transparent for the user, personalizing the environment to the student needs, recommending some contents or doing some modifications in the presentation tool. Different approaches can be used in this process in order to integrate data mining tools, for example algorithms can be integrated inside the learning platforms.

All these systems use three main elements to complete the process: a) data collected about the user, b) user model inferred from the collected data and c) adaptation tools that will show the selected elements. Content and link structures can be adapted to better achieve these goals, although the knowledge level and learning styles better define users.

1.1. Learning styles

Identifying learning styles can be fundamental in order to develop personalized learning models. Felder [4] proposes a taxonomy in a scale of 5 dimensions that defines different learning styles (See Table 1).

Table 1. Learning styles

DEFINITIONS	DIMENSIONS		DEFINITIONS
Do	ACTIVE	REFLEXIVE	Think
Learn facts	SENSITIVE	INTUITIVE	Learn concepts
Needs drawings	VISUAL	VERBAL	Needs read or dissertation
Derive facts from facts	INDUCTIVE	DEDUCTIVE	Derive results from principles
Step to step	LOCAL	GLOBAL	Global framework

Previously, Bloom had defined another taxonomy for learning styles [5], which sets a hierarchical classification of learning objects that was later redefined in [6] and can be seen in Table 2.

Table 2. Educational objects hierarchical classification

LEVEL	OBJETIVE	HABILITY
1	Knowledge	terminology, specific facts, ways and means of dealing with specifics, conventions, trends and sequences, classifications and categories, criteria, methodology, abstractions in a field, principles and generalizations, and theoretical structures
2	Comprehension	translation, interpretation, and extrapolation
3	Application	of concepts in the use of abstraction in particular and in concrete situations
4	Analysis	of elements, relationships, and organizational principles
5	Synthesis	of ideas in the production of unique communications and plans
6	Evaluation	leading to judgments about the value of materials and methods for given purposes

1.2. Modeling techniques

Different data mining methods have been used with different objectives in the user modeling task: grouping users by navigation pattern, grouping pages, grouping pages by their content, extracting relationships between pages, finding relevant information in pages, finding learning models, discovering visited page sequences, creating routes for users.

Regarding to the methods used to cover these objectives, visualization techniques are widely used to show information related to the user access to a course, a resource or the participation in a forum [7]. Classification techniques are used when some previously labeled cases belonging to different groups are available. The goal is to classify new cases, for which, the group is unknown. The users can be grouped by navigation patterns and the pages can be grouped by their content. Sometimes, a mixture of different data mining methods has been used but, usually, algorithms that can provide comprehensible output are better, such as Decision Trees (C4.5, C5.0) or Naive Bayes. Some educational data mining algorithms were also used [8] in order to discover groups of students with some similar characteristics with the purpose of determining pedagogical strategies, also for the student performance prediction and final rating. An example of applying modeling techniques can be found in [9].

Association rules are used to obtain rules that associate concepts that are located in different columns of a data base where the user activity has been registered. They have been widely used for extracting customer preferences, relating aspects that frequently occur together. In an e-learning context, it has been used as a way of finding out associations between learning activities. They can be useful to monitor students and instructors' results in conjunction with alarm thresholds. These methods have been used to find relationships between concepts among the groups found by the clustering techniques and also to create automatic recommendations or finding out errors that occur together. They can be useful to optimize an e-learning content, taking into account what content is interesting for students. An example of this use can be found in [10] and [11].

AHA! is an example of an adaptive hypermedia system developed in the Technical University of Eindhoven in Holland [12], where the Apriori algorithm is used. The University of Cordoba has developed a tool that enables to use the resources (course, unit, lesson, exercise, etc.) and arrange the content by levels.

The sequence analysis process allows to analyze sequences of pages seen during a session or different sessions of the same user. It analyzes the order of the pages accessed by the user. These paths can be analyzed alone or aggregated after a clustering process. These results can then be used in order to reorganize the web content, personalize the resource delivery, doing suggestions to students with a similar profile, to evaluate the design of web pages or to identify sequences of interactions that can be indicative of success or problems.

Clustering techniques [13] have also been used in order to discover groups of objects with similar characteristics. The main goal is to discover groups of students with similar behaviors trying to thrust their collaboration and level of activity. Clusters can also be generated to establish different education journeys, fix personalized tutoring hours, etc.

A lot of another techniques are being used, XML and ontologies [14],[15], Neural networks[16], or methods that are feed by the labeling of the resources provided by the users[17], Semantic inference [18], rules [19], or variable definition [20], among others. Another system works providing information to the tutor [20], or selecting and synthesizing specific content documentation [21].

2. Discovering students' navigation styles in Moodle

In order to discover users' preferences when it comes to navigation styles, first of all, it is necessary to collect data. In Adaptive educational hypermedia systems, information referred to user logs, activities they have done, the quizzes they have tried, which ones they passed and level of knowledge, etc. is available. In order to collect data about the user, we used Moodle (Modular Object-Oriented Dynamic Learning Environment), which is a platform to show e-learning contents. Although this platform gives some information about the student interactions, we decided to extend it using log data and adding some extra indicators about user's activity, specifically for each student and course. Some of them are shown in table 3.

These indicators were added by means of a plug-in to the Moodle platform, so that they could be used by teachers and tutors. Table 3 shows the list of indicators that the tutor can use in Moodle in order to monitor each student, each course and the activity of each student in each course.

Table 3. Observatory indicators extracted from moodle.

1. Number of pages seen	4.Use of discussion forum	7. Consumed multimedia resources	10.Comment a entry in the glossary
2. Number of unique pages seen	5.Attempts to do the quiz	8. Number of glossary entries	11.Use of the chat
3. Unique records	6.Sequentiality	9.Actualize an entry in the glossary	12. Use of the Wiki.

2.1. Relating indicators with navigation and learning styles

It is not easy to decide what indicators better characterize students and their learning style. In order to select the most representative ones, a previous interaction of the experts with the system was identified as necessary. Teachers and tutors have information and experience that can be useful in this process. In order to interactively change and evaluate these indicators, a web application has been developed where a tutor can establish relationships between activity indicators, navigation styles and learning styles(obtained by filling questionaries or by some other ways). This has been understood as a validation process of the observatory indicators and a first approach to identify the best indicators in order to discover user learning models. The developed architecture for the whole process can be seen in Figure 1.

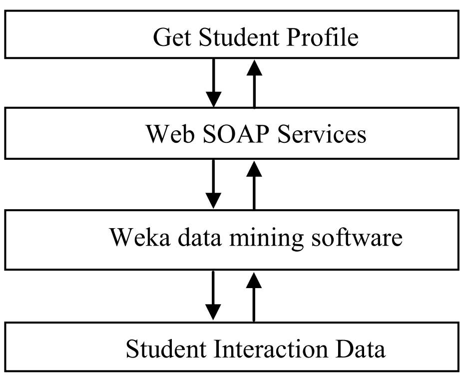


Figure 1: Developed architecture

2.2. Modeling Users

Once indicators have been selected, they are used to model users' learning styles. As mentioned previously we consider different learning styles, also called dimensions (see Table 1). Based on the indicators, the algorithm we have developed classifies students in each one of these dimensions. Figure 2 shows the steps carried out by the algorithm in order to discover the learning style of each student.

Using the dataset and clustering techniques, data across the defined dimensions are divided in different clusters. Each cluster groups students with similar values for such dimensions. Before the clustering process, a prototype selection step (used in data mining to select the cases to build the model) gives the possibility to generate the model using filtered data, discarding noisy data (for example, data from students that begins the course later than the others or that finishes before).

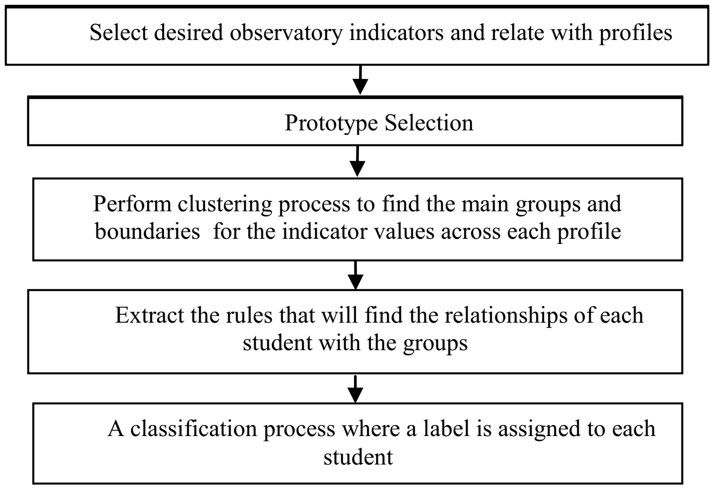


Figure 2: Discovering students' learning styles

Then, once clusters are defined, the algorithm extracts the set of rules that better defines the relationship between the dimensions and the clusters. These rules establish the boundaries for the values of each cluster. Finally, considering each student's indicators and using classification techniques, each student is characterized in each dimension. .

3. Conclusions and Future work

Providing personalized contents based on preferences and navigation styles is a necessary step to provide intelligent services in virtual environments. For that, different techniques can be applied. We have developed a system that models user's navigation styles using clustering and classification techniques. As an initial approach, we have considered the parameter that defines if a user's navigation style is either local or global.

This work is being extended in order to generate more accurate patterns. In that sense, one of the challenges is the validation of the models, due to the fact that it demands the contribution of teacher, tutors, as well as the collaboration of the students.

Analyzing the characteristics of the learning objects that can fit better the preferences of the students for each learning style is a big goal. This way, we can focus not only on the learning content that is shown to the student, but on how it looks the page that he will see.

We also expect to add some extra indicators to our systems, some of them related with engagement, other ones related with emotional status.

Acknowledgment

The authors would like to thank the financial support given by the Spanish Government, Project Reference TSI-020311-2009-6.

References

[1] H. Wu, P. De Bra, (2001). Sufficient Conditions for Well-behaved Adaptive Hypermedia Systems. Lecture Notes in Computer Science, 2198, 148-152.

[2] P. Karampiperis, D. Sampson, Adaptive Learning Resources Sequencing In Educational Hypermedia Systems. Educational Technology & Society, 8 (4), 128-147

[3] S. Preidys, L. Sakalauskas, Analysis Of Students' Study Activities In Virtual Learning Environments Using Data Mining Methods. Technological And Economic Development Of Economy. Baltic Journal on Sustainability.2010.

[4] R. M. Felder, and L. K. Silverman, Learning and Teaching Styles in Engineering Education, Engineering Education, 78(7): 674-681, 1988

[5] Bloom ,Taxonomy of Educational Objectives: The Classification of Educational Goals; pp. 201-207; B. S. (Ed.) Susan Fauer Company, Inc. 1956.

[6] D. R. Krathwohl. A Revision of Bloom's Taxonomy: An Overview. Theory Into Practice, Volume 41, Number 4, Autumn 2002 Copyright C) 2002 College Of Education, The Ohio State University

[7] Nachimas & Hershkovitz , A case study Using visualization for Understanding the Behaviour of the online learner. Proceedings of the International Workshop on Applying Data Mining in e-Learning 2007.

[8] B. Minaei-Bidgoli, D. A. Kashy, G. Kortemeyer, and W. Punch, Predicting Student Performance: an Application of Data Mining Methods with an Educational Web-Based System (LON-CAPA), Frontiers in Education Conference 2003.

[9] S. Baldiris, G. Moreno, R. Fabregat, I. Guarín, R. Llamosa, J. García, Extensiones en SHABOO: sistema hipermedia adaptativo para la enseñanza de la programación orientada a objetos. Enlace Informático 6 .2007.

[10] A. Merceron and K. Yacef, Measuring Correlation of Strong Symmetric Association Rules in Educational Data, chapter in Handbook on Educational Data Mining, C. Romero, S. Ventura, M.Pechenizkiy and R.Baker (Eds). CRC Press,2010.

[11] E. García, C. Romero, S. Ventura y T. Calders. Drawbacks and solutions of applying association rule mining in learning management systems. International Workshop on Applying Data Mining in e-learning (ADML'07). Crete (Greece), 2007.

[12] P. de Bra, AHA! The adaptive hypermedia architecture Conference on Hypertext and Hypermedia archive. Proceedings of the fourteenth ACM conference on Hypertext and hypermedia. 2003.

[13] A. Klasnja-Milicevic, B. Vesin, M. Ivanovic, Z. Budimac: E-learning personalization based on hybrid recommendation strategy and learning style identification. COMPUTERS & EDUCATION . 2011

[14] F. Lazarinis, S. Green, E. Pearson: Focusing on content reusability and interoperability in a personalized hypermedia assessment tool. MULTIMEDIA TOOLS AND APPLICATIONS.

[15] F. Liu , X. Li , G. Gao : Ontology-based personalized recommendation technology research of learning resources. ETP/IITA Conference on System Science and Simulation in Engineering. 2010.

[16] A. Baylari; G. A Montazer: Design a personalized e-learning system based on item response theory and artificial neural network approach. EXPERT SYSTEMS WITH APPLICATIONS 2009.

[17] N. Escudeiro, P. Escudeiro, D. Remenyi: Organizing web Educational Resources According to Specific User Needs. 8th European Conference on e-Learning. 2009.

[18] Y. Blanco-Fernandez, Jose J.Pazos-Arias, Alberto Gil-Solla, Manuel Ramos-Cabrer,

[19] G. Fernandez, A. Sterbini, M. Temperini, Learning objects: A metadata approach. International Conference on Computer as a Tool. 2007.

[20] T. Roselli, V. Rossano, Describing learning scenarios to share teaching experiences. 7th International Conference on Information Technology Based Higher Education and Training. 2006.

[21] S. Abidi, C. Han, S. Abidi, V. Patel, R. Rogers, R. Haux, Patient empowerment via 'pushed' delivery of personalised healthcare educational content over the Internet. 10th World Congress on Medical Informatics (MEDINFO 2001).

[22] P. Clough, N. Ford, M. Stevenson, Personalizing Access To Cultural Heritage Collections Using Pathways. 2011 Workshop on Personalized Access to Cultural Heritage. 2011.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-276

An Entrepreneurship Model for Future Intelligent Educational Environments

Victor CALLAGHAN[a1], Ping ZHENG[b], Hsuan-Yi WU[c]
[a]*School of Computer Science and Electronic Engineering, Essex University*
[b]*Business School, Canterbury Christ Church University*
[c]*Graduate Institute of Business Administration, National Taiwan University*

Abstract. In this paper we introduce a novel entrepreneurial model for funding and managing future intelligent educational environments and other academic business enterprise; the *Faculty-Cooperative*. The goal of this model is to create a business mechanism where academics are both the owners and customers of the IPR they generate, thereby providing synergy to optimize the educational product for the market, provide an embedded sales team and offer a source of investment for the enterprise. Our model is inspired by the cooperative ethos that has historically existed in universities. In addition we draw on parallels with the Western Cooperative movements and Chinese collectives but in a more virtualised form within the university system. To illustrate the *Faculty-Cooperative* approach we examine how it might be applied to the formation of an academic enterprise, FortiTo (a manufacturer of the technology that underpins intelligent educational environments and a producer of learning tools for students engaged in learning these new technologies). Whilst this paper describes the early stages of the development of an entrepreneurial model for academic enterprise, our hope it that this paper will promote discussion and participation in what we hope will be a successful model for funding and managing entrepreneurial academic enterprise.

Keywords. Entrepreneurship, Faculty-Cooperative, Academic Enterprise, Educational Technology, Intelligent Environments, Embedded-Computing.

1. Introduction

1.1. The Academic Market

Perhaps a good way to start a discussion about entrepreneurial business opportunities in academia is to consider the size of the education market we are addressing. The education sector comprises three main areas, school education (K-12), further education (FE), higher education (HE) and in-work training (eg hospitals). Its difficult to get reliable figures but some estimates place the global academic workforce as approaching 100 million with the global education market value being around $2.5 trillion comprising around 97 million students which may triple by 2025 [1]. Some observers have noted this market is larger than the music and car industries. By way of more reliable figures, the UK's British Council has produced a number of reports that quantify the size UK education market as comprising (in 2008/9) approximately

[1] Corresponding Author: Victor Callaghan, School of Computer Science and Electronic Engineering, University of Essex, COLCHESTER, Essex UK, E-mail: vic@essex.ac.uk

500,000 teachers in almost 28,000 schools, over 220,000 teaching staff in FE colleges and 175,000 lecturing staff in 169 HE institutions. The same report suggests that, in 2007, when indirect jobs were taken into account, the total rises to over 668,500 jobs, or 2.6% of the UK total workforce. The higher education market is reported to be worth some $95bn to the UK economy with about $30bn being associated with educational goods and services produced in the UK [2] [3]. By any measure education is a substantial market, with significant funding and a vast pool of actors that our proposed *Faculty Cooperative*" educational enterprise model seeks to work with.

1.2. Academic Collaborative Culture

Universities are, by their nature, scholarly places where teaching staff are motivated by the pursuit and communication of knowledge, above the commercial bounds of the commercial environments they find themselves in. As such there is a tradition of openness and sharing of resources amongst teachers that lives on, despite the advent of commercial market forces that seek to place universities and academics in competition with each other. Thus, university have a natural inclination towards publishing information freely in papers, and sharing technology such as software tools to the benefit of the wider community; the so-called open systems and freeware movement. One example is the GNU project, which seeks to enlist the software community, around open standards, to produce "freeware" software tools such as operating systems and compilers. GNU is especially popular with academics. Open systems can generally be understood as being products based on non-proprietary standards, which are "owned" by the community, thereby promoting interoperability and portability (albeit sometimes managed by a company). Other examples of open systems include Oracle's Java and Open Office. The GNU project points out on their website (http://www.gnu.org/), free (or open) does not mean that the software's users do not have to pay for the software, rather it means they are given a few essential freedoms such as being able to run the program, to study/change the program source code and to redistribute exact/modified copies. Thus, whilst the words "open" or "freeware" may suggest not making a profit, this is not the case but rather it's a freedom of usage, plus harnessing the wider community in defining, building, supporting and funding the product which are also principles that underpin the proposed "*Faculty Cooperative*" educational enterprise model.

1.3. The Cooperative Movement

The vision for people to form cooperative groupings to benefit their wider community can be traced back to the UK in the 18^{th} century when groups, such as the Scottish "Fenwick Weavers Society" (formed in 1769) or the "English Lockhurst Lane Industrial Co-operative Society" (formed in 1832) and now known as the "Heart of England Co-operative Society' became the forerunners of a worldwide movement that saw cooperative groups move from community stores to schools through to business cooperatives. One notable cooperative was the English "Rochdale Society of Equitable Pioneers" (founded in 1844), which established a set of principles that co-operatives still use. These principles include the need to have an open and voluntary membership, the need to avoid unfair discrimination between people, that members should have a sense of Altruism (note that this does not prevent members enjoying financial rewards) and that the enterprise should be funded by the members. [4] There are numerous variations of these principles such as the "Emelianoff's three cooperative business principles" which seek to embody a principle whereby members may receive outputs at-cost" (but to non-members at good profit levels), a "proportionality principle" which seeks to allocate benefits according to stakeholding and a "self-financing principle". Cooperatives remain popular options for organising work with the United Nations estimating that, globally, around 800 million people are members of cooperatives with almost 100 million people being employed by them [5]. Of course, there are numerous potential hybridisations, one of which we describe in this paper which we call the Faculty-Cooperative.

1.4. New Chinese Collectives

As with the western cooperative, another notable movement occurred in China, where land ownership is spilt into two categories: state-owned land and collectively owned land. "A collective" is a basket of land (resources) that is assigned to a community (e.g. a village or town) and distributed to its members for cultivation (benefit) [6]. Collectives were an instrument used in China during the Mao Zedong era as a means to boost agricultural productivity and provide a much-needed measure of food security [7] [8]. As far as production is concerned, the advantages lay in the nature of ownership and control [9]. Under capitalism, the means of production and economic surplus are privately owned, while in the socialist societies, the ownership and economic surplus are legitimately transferred to government, legally – in the name of the people. The distribution of this 'publicly-owned' surplus is subject to claims by all sectors of socialist society and is a deliberate political process [10]. The collectively owned cooperatives are literally owned by the employees, in which the distribution of profit is subject to claims by the collective shareholders [11] [12]. Whilst the reputation of collectives became tarnished in the west, by association with communism and their perceived poor performance, a more dispassionate analysis might reveal that the collective model had some interesting ideas that could be relevant to a modern global and high-tech business. In this paper we examine we argue that the value of collective stakeholding provides a powerful means to motivate and empower faculty to have a hand in investing, directing and benefiting from the fruits of their intellect. While there are various schemes for achieving this, in this paper we are proposing a loosely inspired variation (cherry picking the bits we like and adding aspects we require) to, create a type of academic collective, or as we prefer to phrase it, a "*Faculty-Cooperative*".

In the following sections we will seek to explain the *Faculty-Cooperative* model which is inspired by the about discussion on the academic market, the historic academic culture and the notion of cooperatives and collectives.

2. The Faculty-Cooperative Model

2.1. From Collectives to Cooperatives

Universities might be seen as a form of educational eco-system. Within this eco-system, they might be regarded as a form of government assigned academic collective comprising a group of academics (labelled with a university name, eg Oxford, Cambridge, Essex etc) a resource (buildings, degree conferment rights etc) with the responsibility to use them to the good of the country. However, we ask, is such a collective bounded by the physical limits of a particular university, or is it bounded by a different label related to interest groups or specialities (business studies, computer science). In our "*Faculty Collaborative*" model, we are proposing to introduce one such virtual-collective, based on entrepreneurial academic activities, where academic in differing institutions can collaborate together to advance their entrepreneurial visions.

2.2. From Open Systems, to Open-Innovation, to Open Financing

As was described in section 1.2, academics are, by and large, strong advocates for an open approach to innovation based on well-established principles of openly publishing knowledge and actively seeking to collaborate with fellow researchers. In a recent example, "Living Labs", Universities have extended such open research cooperation to local government and communities engendering cooperation to mutually improve the technology that impacts all our environments [9] [13]. The concept of 'open innovation' gives a strategic emphasis on developing and intensifying collaboration across industry networks and partnerships, opening up their innovation processes in line with the open innovation framework [14]. One important assumption underpinning the concept of 'open innovation' is that an organisation cannot innovate in isolation [15] [16]. Under a turbulent business environment and hyper-competitive market condition, innovation is considered as a major engine to enhance business performance and to strengthen an organisation's competitiveness in the marketplace [17] [18] [19] [20]. In our *Faculty-Cooperative* model we are seeking to build on this principle by devising a model whereby the company structure and investment follows such an open framework by seeking to make the IPR, shareholding (investment) and strategy to be owned by the academic community in as transparent a way as is possible. In the following sections we describe this model from various perspectives, principally the faculty members, the students and the company personnel.

2.3. Perceptions of the Faculty Cooperative Model

2.3.1. A Non-Entrepreneurial Faculty Member Perspective

For a non-entrepreneurial member of University staff, the *Faculty-Cooperative* represents an opportunity for them to become stakeholders in the "tools of their trade". This stakeholding takes the form of being able to contribute to the specification and nature of an educational product and to share in a financial reward from the combined intellect of the academic system that they have committed their life to.

2.3.2. An Entrepreneurial Faculty Member Perspective

For an entrepreneurial member of University staff the *Faculty-Cooperative* provides all the advantages of the non-entrepreneurial member, described in the previous section but provides the academic entrepreneur with a source of finance by offering a large number of low cost shares to the academic community, thereby raising the required capital to fund the company, without seeding control to another single and dominant investor (which is often the case with venture capitalist funding). Furthermore, it offers a pool of tangible and intangible resources to incubate any new ideas in an embryonic state for entrepreneurs aiming to start a new venture with/in the university.

2.3.3. A Non-Entrepreneurial Students Perspective

For a regular student, attending a university, they would be essentially unaware of this organisation but indirectly benefit from better-designed educational tools that arise from within the academic community.

2.3.4. An Entrepreneurial Students Perspective

For an entrepreneurial student, the *Faculty-Cooperative* represents an opportunity for them to apply their newly acquired knowledge, exercise their product innovation and entrepreneurial skills, enrich their CV and become stakeholders in one of the largest and most worthwhile global industries. Apart from that, there is the added bonus of earning some welcome income.

2.3.5. The Customers Perspective

From a customers' prospective (Universities, faculty members, students etc) they receive a better quality product, designed and tested by the leading educational experts. In the same way as there is some enthusiasm for green products that benefit the earths eco-system (the environment debate) then customers (the Universities) can feel good about supporting and improving their own *educational eco-system* via the mutually owned *Faculty-Cooperative*.

2.3.6. The Company Personnel Perspective

For company personnel, the *Faculty-Cooperative* provides a "feel good factor" of being associated with both a worthy cause (the education business, that transforms lives positively) and a secure profitable business (education generates more revenue than the music business) all of which contribute to job satisfaction.

2.3.7. The Business Perspective

With a global workforce of the order 100 million, and market value approaching $2.5 trillion, business prospects for the *Faculty-Cooperative* are good. Postgraduate education (MScs, PhDs) involves students working at the cutting edge of disciplines and companies made up of investors and workers drawn from such auspicious ranks must have some advantage over its competitors. Thus, the combination of a large market and a well qualified set of stakeholders' present a positive business perspective.

3. FortiTo – The *Faculty-Cooperative* Exemplar

In the following section we give an example of company, FortiTo, that we are proposing to operate based on thie *Faculty-Cooperative* model. FortiTo is a spin-off company from the School of Computer Science and Electronic Engineering at the University of Essex.

3.1. FortiTo Market

FortiTo is a company that aims to provide educational technology for the Intelligent Environments and related applications such as embedded-computing, the Internet-of-Things, Ubiquitous and Pervasive Computing etc [21] [22]. All of these applications are based around the use of network connected embedded computers, each which senses and controls (individually or collectively) some part of the built environment. Thus, for example when such systems are placed in a domestic home, and managed by software agents, the "*Smart-Home*" is created. Likewise, if a similar arrangement was used in a classroom (physical or virtual), an "*Intelligent Classroom*" could be created. There are no reliable estimates for the value of this market but a recent report suggested it could reach between 22 billion and 50 billion dollars by 2020 made up of some 16 billion connected devices [21]. These figures are given additional credibility by other findings that show the Chinese market has already reached 30 billion dollars [21].

.

The mBed is based on the ARM Cortex-M3 Core running at 96MHz, with 512KB FLASH, 64KB RAM and various interfaces including Ethernet, USB Device, CAN, SPI, I2C

Figure 1a. mBed

The Raspberry Pi is based on an ARM 1176JZFS, running at 700Mhz, in a Broadcom BCM2835 SoC. The current best specified model has 256Mb RAM, 2 USB port and an Ethernet port.

Figure 1b. Raspberry Pi

The educational technology developed by FortiTo is used to train and teach students the basic skills required for developing products for new high-tech markets such as Intelligent Environments or the Internet-of-Things etc. The most basic component is an embedded computer, which is essentially a small integrated circuit. In order to make it useful for education or industrial prototypes it needs to be added to a carrier; an example is given in figure 1a which shows the popular ARM processor on a baseboard carrier called an mbed that was developed by Philips in partnership with ARM. Although primarily intended for rapid prototyping in industry, its keen pricing and versatility made it a popular choice for Universities.

The main problem with students using an mbed is that it needs a keyboard, display, power supply, I/O and software tools to do anything sensible. Adding such things takes a considerable time to build, considerably longer than most student lab periods allow. Alternatively, academic or technical staff would need to spend time making pre-assembled versions, which would limit the students flexibility and takes up valuable academic time. In some respects, the newly announced Raspberry Pi (see Figure 1b), with its headline grabbing "*$25 computer*" tagline represents such a preassembled system, needing only the addition of a power supply, keyboard and display to become a fully operational computer. While nobody can argue this is amazing value, to make this computer go beyond operating on data and to control real artefacts requires some time consuming electronics design leaving it much in the same problem-space as the mbed for applications involving the Internet-of-Things, Ubiquitous and Pervasive Computing etc The FortiTo company solves these problems by providing a modularised scheme of educational hardware and software technology that offers a family of pluggable hardware boards (*Buzz-Boards*) that can be plugged together to enable students to construct a variety of embedded-computing applications within the timescales available in a typical computing laboratory session. The company assists busy academics by providing all the necessary pedagogical content such as example software and assignment templates that can be customised by the host institution.

3.2. Buzz-Boards

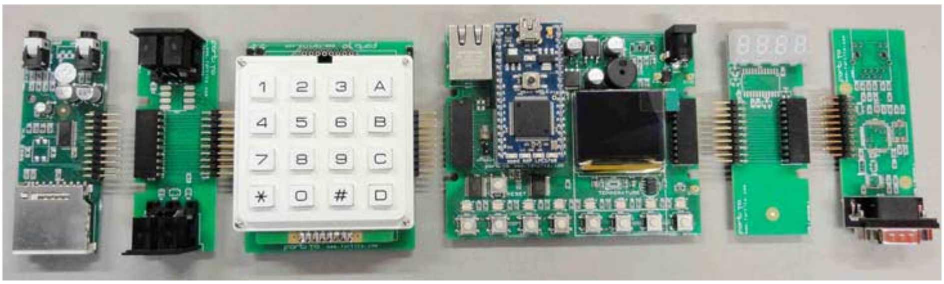


Figure 2. *Buzz Board* Examples (Audio, Midi, KeyPad, Processor, LED & Network Boards)

Figure 3a. An Internet Radio

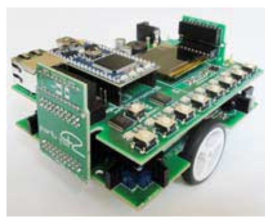


Figure 3b. A mobile Robot

The company's main products are a set of computer boards that can be plugged together to make a variety of applications that are limited only by the students and instructors imaginations. Some examples include *Games* (eg Pacman, Mind Battle etc), *Music* (piano, guitar, tuner, MIDI synthesiser etc), *Media* (video camera, audio etc), *Medical* (fitness, heart rate, body temperature monitor, physiological sensing etc), *Navigation* (inertia 3D sensor, light seeker, GPS etc), *Network Services* (Bluetooth, Wifi, Wired), *Computing Basics* (simple calculator, storage etc), *Weird-Science* (brain wave monitoring, lie detection and emotion sensing systems, random numbers, 'quantum universe splitter' etc), *Mobile Robots* (light seeking, maze escapes, crazy eyes, chatterbox etc), *IT Tools* (multi-meter, oscilloscope, or logic analyser etc), *Product Prototyping* (a bread- & solder- Board for bespoke designs). A few examples of Buzz-Boards are shown in Figure 2.

Buzz-Boards are processor agnostic, in that they work with virtually any popular processor or educational ICT set such as Lego, Arduino and mbed, PIC, ARM, AVR. For programming *Buzz-Boards* work with standard C and C++, the most common embedded-computing languages. The company are also planning a web-based graphical programming environment (*Buzz-Blocks)* for less experienced people, and to allow the products to be used in primary and secondary education. By assembling the *Buzz-Boards* is different ways (ie plugging them together in differing combinations) it is possible to create a variety of hardware application platforms that students can then program to learn differing skills. Two examples are given in figure 3a and 3b.

4. An Implementation Model for the *Faculty-Cooperative*

4.1. The Faculty-Cooperative Principles

As discussed in section 1, much information and analysis exists on the principle underpinning cooperatives, collective and other more modern mutual enterprises such as shareware. From these we have selected the following mix that we feel are appropriate to an academic or faculty cooperative. It would be fair to say we are still in the early stages of developing our *Faculty-Cooperative* model, and one of the aims of this workshop is to raise these issues at the workshop and beyond, so we can refine our principles further; thus the table 1 represents or starting position on this path.

Openness	Support for open implementation standards (eg interfaces)
	Support for open source design standards (eg product specifications)
	Support for open sharing of related work (eg assignments)
Freedom	To use the product for in education without restrictions
	To study and modify the products (eg student project work)
	To profit from the contributors IPR and work (eg faculty or student remuneration)
Collective Stakeholding	A mechanism whereby academics across a number of differing Universities are able to share in the operation of the company.
	A mechanism whereby academics across a number of differing Universities are able to be shareholders (to invest and share in profits)
	A mechanism whereby academics across a number of differing Universities are able to influence the educational product specification
	A mechanism whereby academics across a number of differing Universities involved in the enterprise can receive benefits (eg discounts or direct profit share)

Table 1 – Principles of a *Faculty-Cooperative* company

With the above cooperative philosophy in mind, the questions arises how might this *Faculty-Cooperative* model be implemented? Table 1 sets out the key principles we advocate for the model namely, openness, freedom and collective academic stakeholding.

It should be noted that whilst a collective ethos underpins this model, it recognises that the enterprise is competing in a free market and that the company should operate in the normal way for a commercial company.

4.2. An Example of a Faculty-Cooperative Company Implementation (FortiTo)

Clearly, there are many ways that a *Faculty-Cooperative* model could be translated into an academic company. In the following we briefly discuss how that has been proposed for our exemplary company, FortiTo. Considering '*Openness*', FortiTo is adopting many industry standards such as mbed and RPi processors, I^2C bus technology and C/C++ programming. Considering '*Freedom*' the company makes use of freeware software tools (eg gnu), has opened its interface specifications and computing architecture, so that students and faculty have the important details available for educational assignments and projects. In respect of the '*Collective Stakeholding*', the company is actively seeking to attract membership, gather funding, create product specifications, conduct evaluations and market products in cooperation with as wide a slice of the international educational community as is possible. For example, the company plans to offer members of the educational community 'resource units' (either work packages or financial investment) in return for a shareholding of FortiTo. Finally the company is committed to providing benefits in the form of product discounts and profit share to its members. FortiTo is at a very early stage of its life, having developed a complete product range but is scheduled to commence commercial operations in September 2012, so this paper (written in April 2012) is the first of a series that will follow the application of the *Faculty-Cooperative* principles to an academic start-up and will follow the story of how it fares over a series of studies and papers.

5. Summary

In this paper we have explained how universities have traditionally been based on the ideas of openness, freedom and mutual support that we argued share some of the characteristics of Western Cooperatives and Chinese Collectives. In this paper we have presented an entrepreneurial model for producing future intelligent educational environments and products that seek to embody these values by drawing on the strength of the university system in the form of its faculty and students as a means to innovate, fund and manage academic enterprise. The model is not restricted to future intelligent educational environments but can be used for any academic spin-off company. We have sought to illustrate the implementation of this model based on a company that was formed based on the *Faculty-Cooperative* principles that, although in its infancy, we hope will serve to illustrate the essential principles and form an ongoing case that we will study and report upon at its various stages of growth. Above all, we are presenting this paper so that, through the workshop, we can engage with the wider academic community on exploring the virtues or vices of the *Faculty-Cooperative* model as a means to support academic entrepreneurial enterprise, an approach we hope will prove beneficial to the academic community at large.

Acknowledgements
We are indebted to FortiTo Ltd (www.FortiTo.com), and especially Malcolm Lear and Martin Colley for the information on the *Buzz-Board* product range.

References

[1] http://www.apolloglobal.us/index.php?option=com_content&view=article&id=130&Itemid=106 (Accessed 19th April 2012)

[2] P. Lenton "*Global Value: The Value of Education and Train Exports*", British Council, September 2007"

[3] G. Conlon, A. Litchfield and G. Sadlier "*Estimating the Value to the UK of Education Exports*", Department for Business Innovation and Skills, June 2011

[4] K. Zeuli, R. Cropp "Cooperative Principles and Practices in the 21st Century", Cooperative Extension Publishing, Wisconsin USA, 2004

[5] W. van Diepenbeek "Cooperatives as a Business Organization; Lessons from Cooperative Organization History, *Universiteit Maastricht. Publication, ISBN: 978-90-5681-260-7, 2007*

[6] Z. Zuo, (2001), 'Characteristics and Origin of Pearl River Delta Model', *Journal of Economics & Management Strategy*, 10 (3), pp435-461

[7] P. Zheng '*Understanding entrepreneurial growth and process in emerging business ventures under market socialism in China*', 33rd Institute for Small Business and Entrepreneurship (ISBE) Annual Conference, London, 3rd & 4th November 2010

[8] J. Chen (1998), *Township and Village Enterprises Model Studies*, China Social Science, Beijing.

[9] C. Pierson (1995), *Socialism after communism: the new market socialism*, The Pennsylvania State University Press, University Park Pennsylvania, USA.

[10] H. Davis, R. Scase, (1985), *Western capitalism and state socialism: an introduction*, Basil Blackwell, London.

[11] J. Chen, Y. Liu, X. Liu, L. Ge (2008), *Research On The 30 Years of China's Non-State-Owned Units Reform And Development*, Beijing: Economic &Management Publishing House.

[12] G. Yano, G, M. Shiraishi (2004), 'Efficiency of Chinese Township and Village Enterprises and Property Rights in the 1990s: Case Study of Wuxi', *Comparative Economic Studies,* June, 46 (2), 311-340.

[13] H.Y. Wu "*An Empirical Study of UK Living labs*", International Association for Management of Technology IAMOT 2012 Proceedings

[14] H. Chesbrough (2006), *Open Business Models: How to Thrive in a New Innovation Lanscape*, Boston: Harvard Business School Press.

[15] H. Chesbrough (2003), *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Boston: Harvard Business School Press.

[16] K. Laursen, A. Salter (2006), 'Open for Innovation: The Role of Openness in Explaining Innovative Performance among UK Manufacturing Firms', *Strategic Management Journal*, 27 (2), 131-150.

[17] C. Lechner, M. Dowling (2003), 'Firm networks: external relationships as sources for the growth and competitiveness of entrepreneurial firms', *Entrepreneurship and Regional Development*, 15, 1-26.

[18] C. Lee, K. Lee, J. Pennings (2001), 'Internal Capabilities, External Networks, and Performance: A Study of Technology-based Ventures', *Strategic Management Journal*, 22, 615-640.

[19] D. Lavie (2006), 'The competitive advantage of interconnected firms: an expansion of the resource-based view', *Academy of Management Review*, 31, 638-658.

[20] W.P. Wu (2008), 'Dimensions of social capital and firm competitiveness improvement: the mediating role of information sharing', *Journal of Management Studies*, 45, 122-146.

[21] M. Wang, V. Callaghan, M. Lear, M. Colley "*Teaching Next Generation Computing Skills; The Challenge of Embedded Computing*", Intelligent Campus 2011 (iC'11), Nottingham 26th July 2011

[22] V. Callaghan, "*Buzz-Boarding; Practical Support For Teaching Computing, Based On The Internet-Of-Things*", 1st Annual Conference on the Aiming for Excellence in STEM Learning and Teaching, Imperial College, London & The Royal Geographical Society, 12-13 April 2012

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-286

Intelligent Assistive Interfaces for Editing Mathematics

Dilaksha ATTANAYAKE[a], James DENHOLM-PRICE[b] , Gordon HUNTER[b,1], Eckhard PFLUEGEL[a], Angela WIGMORE[a,c]
[a] *School of Computing & Information Systems*
[b] *School of Mathematics*
Kingston University, KT1 2EE, U.K.
[c] *ESS Ltd., U.K.*

Abstract. In this paper, we discuss the need for improved user interfaces for editing mathematical text, focusing of three types (individuals suffering from various disabilities, people relying heaving on on-line learning systems and ones relying on using portable devices) of people for whom conventional editing approaches are not very satisfactory. After reviewing various recent approaches, we focus on describing the development and evaluation of our own intelligent web-based interfaces, *TalkMaths* and *SWIMS* for editing mathematical text. The former is a speech-based editing interface, the latter a system which assists the user through the predictive and corrective power of statistical language models. It offers options for predicting what will appear next (analogous to predictive text for SMS messages) and identifying likely errors due to simple mistakes on the user's part in order to assist in correcting the errors. Using text-stream input, we investigate the utility of the error identification by studying the proportion of times the correct version of the complete mathematical expression appears within the M most likely alternatives suggested by our system. These systems are currently independent of each other, but we aim to integrate the facilities they provide into a simple intelligent assistive interface.

Keywords. statistical language model; web-based mathematical editors; assistive technology

Introduction

Information and Communication Technology (ICT) has been having a greater and greater impact on education, in the classroom and elsewhere over the last few decades. For example, nowadays, teachers can use smartboards and similar facilities to annotate notes in front of their class, then save the resulting "hybrid" document for their students to download and review in their own time. This greatly reduces the burden on students for taking accurate notes at high speed during classes, which was normal 20 years ago.

However, in some ways, mathematics – which is a core subject of study in the school curriculum in most countries and proficiency in which, at least at an elementary level, is essential for success in a wide range of scientific, technical and commercial fields - is perhaps much less suited to be taught via modern educational ICT than many other disciplines. It is a subject which many students find difficult, partly due to its specialized language and notation. These make working with mathematical equations and formulae a problem for a large proportion of people. This is even more notable when the mathematical expressions to be manipulated are to be included in electronic documents. Typing and editing ordinary text can be both slow and error-prone for non-experts, and this is even more the case for mathematical text, with its non-alphanumeric symbols and typically somewhat complicated two-dimensional layout. Furthermore,

[1] Corresponding Author: G.Hunter@kingston.ac.uk

creating, editing and reading mathematical text (in its conventional form) is particularly difficult for three types of groups of people : individuals suffering from various disabilities [5], people relying heavily on on-line learning systems [24, 25] (particularly distance-learning students), and people relying on using portable devices, such as smartphones and tablet computers, to access learning resources. These factors can severely limit the educational and career opportunities of such groups.

In this paper, we first review some existing recent approaches to addressing these problems. We then focus on our own approaches, namely using spoken input as an option for creating and editing mathematical text in electronic documents, and providing an "intelligent assistant" to aid the user by predicting what may come next and offer a semi-automatic correction facility to rectify mistakes.

Spell checkers, automated correcting facilities and predictive text have been familiar features of word processing and text messaging systems for a number of years. These have aimed to provide "intelligent assistance" to the user in order to make the task of creating and editing ordinary text easier. In this paper, we discuss the development of similar features for an editing system for mathematical text. Although the prototype prediction and correction system described here is a "proof of concept", it is proposed to integrate it with our existing editor, TalkMaths [6], which now supports both spoken and typed input and editing commands, and is available online [6].

1. Problems Accessing and Creating Mathematical Content

As noted above, mathematical text tends to have a complicated layout, with a notation which is not simply related to "regular" natural language. Complicated and/or multi-dimensional mathematical expressions cannot easily be conveyed in narrative terms [26]. Österholm [27] performed a study to find out if reading a mathematics textbook with symbols required different skills from reading one without symbols. He concluded that, whilst the use of non-alphanumeric symbols gave mathematics great strengths, saving time and space, comprehending texts using this specialised language and notation, and also translating written words into algebraic equations, required particular skills which students needed time and considerable effort to acquire [27]. Furthermore, established ways of creating, formatting and editing mathematical text in electronic documents – including LaTeX and MathML, and even the GUI-based equation editors embedded in many modern word processors – are not particularly easy for novices to learn to use. These issues are even more relevant for the three types of people mentioned previously : individuals with many types of disability [5], on-line and distance learners [24, 25], and people relying on portable devices with small screens and keyboards which make accessing and typing mathematical content very difficult indeed. Some previous authors have tried to address these issues by taking novel approaches. We briefly review some of these here.

2. Previous Approaches to Addressing the Problems

2.1. Solutions using Tactile Output

Braille has been a successful tactile medium for nearly 200 years, enabling the blind to read and write. Conventional Braille codes use a two-dimensional representation for each alphanumeric character, but these are arranged in a "linear" format, much like the normal orthography of English, to represent words, phrases and sentences. This is not particularly well-suited to represent mathematics. However, novel extensions and modifications to Braille have allowed blind and other visually-impaired students a much wider access to mathematical resources (e.g. [37]). Nevertheless, all of these coding schemes need to be learned by both the students and teachers, which is a problem for most teachers since they (except specialist teachers of the blind) will probably only teach a rather small number of blind students over many years.

2.2. Solutions using Optical Character Recognition (OCR)

For some groups, typing, but not writing, mathematics is a problem. These would include people with certain types of repetitive strain injuries, people using small mobile devices and, in some cases, on-line distance learners [25]. For such people, an appropriate option might be to write the necessary mathematical expressions using a smart pen or stylus. The characters used could then be identified using an optical character recognition system, and converted into (correctly) typeset mathematical text in electronic form. Some previous authors [38, 39] have developed systems following this approach. However, two remaining issues are how to deal with the possibilities of misidentified symbols (potentially a big problem, since many people have poor "on-screen" handwriting) and mistakes by the user. Previous researchers have used syntactic [40] or statistical [41] approaches in attempts to resolve these issues. The latter approach is to some extent similar to the methodology we use in this paper (see section 4.1 below).

2.3. Solutions using Head Motion or Eye Gaze Direction Monitoring

Although we are not aware of these approaches having been applied to the editing of mathematical text, severely disabled people, including tetraplegics, can interact with computers using systems which monitor motions of their head, eyes, or possibly facial muscles [44]. One such system of particular note is *Dasher*[43], which uses statistical language models (see section 4.1 below) to allocate an appropriate area of a display screen according to the likelihood of the character, word or symbol displayed there being the next item in the sequence being input. *Dasher* can be controlled using a mouse, pointer or any motion or graze tracking system. It should be possible to adapt *Dasher* for use with mathematical editors.

2.4. Solutions using Spoken Input and/or Output

Several groups, including the visually impaired, people with limited use of their hands or arms, and people using portable devices, could benefit by the input and/or output modalities being through speech. There have been a variety of systems attempting to provide synthetic speech descriptions of mathematical text, including *AsTeR* (*Audio Systems for Technical Readings* [31], *MathGenie* [42], *REMathEx* [28], the commercial system *MathPlayer*™ [29], and *AudioMath* [36]. The latter system is open-source, but unfortunately only functions in Portuguese. Previous approaches to allowing spoken input of mathematics include the research prototype systems of Bernareggi & Brigatti [30] (which only works in Italian) and Hanakovič and Nagy [34] (which is restricted to use with the Opera web browser), plus the commercial systems *MathTalk*™ [32] (which is only compatible with certain commercial editiors) and *Math Speak & Write* [33] (which has a rather limited mathematical vocabulary). All of these systems allowing spoken input of mathematics have serious limitations, prompting us to develop our own system, *TalkMaths*.

3. The TalkMaths System

3.1. Overview

The *TalkMaths* system initially started as a desktop application for creating and editing mathematical text, but allowing input by speech alone [4]. With recent developments to the project, the current *TalkMaths* system is now a web-based application [3] and the additional facility for typing input has been added to make the system more useful to a wider audience. The system currently uses a commercial speech recognition system as a front-end. However, since *TalkMaths* works by using a context free grammar to parse a text stream resulting from the speech recognition process, other speech recognition systems could also be employed. Several different editing paradigms (see Figure 1) proposed as a result of earlier work [4, 15] have also been incorporated into this new web-based solution. Three such methods, highlighting all sub-expressions, all individual symbols and all operators, respectively, are illustrated in Figure 1. The appropriate box, corresponding to a particular sub-expression, of the user's choice can then be selected for further editing by specifying the number indexing it.

Figure 1. Different editing paradigms for editing mathematics by speech.

3.2. Speaking Mathematics

Attempts at providing standards for speaking mathematics have been given by various previous authors [16, 14]. These have been specified as formal languages aiming to model, to as great an extent as possible, the natural spoken language constructs that people may use when dictating or teaching mathematics. Our approach, slightly extending that of [14, 4], has been to design our formal language in order to be as close as possible to how mathematically proficient people speak or read mathematical equations and formulae. This should make the formal language relatively easy to learn and use but certain compromises have to be made to avoid potential ambiguities. In particular, for dictating single alphabetical characters (a-z), the names from the NATO pronunciation alphabet [12, 15] must be used. An example of a simple mathematical expression is the equation for velocity under uniform acceleration $v = u + at$ which in our spoken mathematical language would be read as: "*victor equals uniform plus alpha tango*". A more complex example is the formula for the solutions of a general quadratic equation:

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

which would be spoken as "*minus bravo plus or minus square root of bravo squared minus four alpha charlie all over begin two alpha end*". Greek characters, such as α, β , etc. can be inserted using the prefix "greek" before the name of the character. For example, the trigonometric identity : $sin(\alpha + \beta) = sin\ \alpha\ \ cos\ \beta + cos\ \alpha\ sin\ \beta$ would be read as "*sine begin greek alpha plus greek beta end equals sine greek alpha cos greek beta plus cos greek alpha sine greek beta*".

3.3. User Evaluation of TalkMaths Version 1.0

The original version of *TalkMaths* was tested on a group of users, none of whom had any disability. The results of that evaluation, which indicated that non-disabled users, who were more used to using the keyboard and mouse than a speech recognition system, were faster and made fewer errors when creating and editing mathematical text using a conventional editor than when using *TalkMaths* [4, 15]. However, they did find learning to use *TalkMaths* interesting and relatively straightforward. We also had one participant who suffered from Duchenne Muscular Dystrophy and was wheelchair-bound. He was able to use the keyboard and mouse, but found this unpleasant, and he did have some previous experience of using automatic speech recognition systems. He performed much better using *TalkMaths* than did the non-disabled group, and on many tasks was faster using our system than he was with the conventional editor [4]. This illustrates the potential of *TalkMaths* as a useful tool, particularly for people for whom conventional types of interface are not very satisfactory.

4. New Intelligent Features for Prediction and Correction

4.1. Background

A wide variety of existing technological systems employ prediction and/or correction methodologies in an attempt to make the systems more useful and usable. These include automatic (or semi-automatic) correction systems found in word processors and internet search engines (“Showing results for … Search instead for …”) and the prediction systems used in Automatic Speech Recognition (ASR) systems and SMS text message editors on mobile telephones. Although manufacturers of commercial products rarely reveal exactly their secrets, it is understood that correction systems look for “close matches” to what was entered from a database of common words or phrases, whilst prediction systems use statistical models. These models give probabilities of words and word sequences, using information from a large set of previously observed data and evidence from the current situation together with an “inference rule”, such as a Bayesian framework, in order to combine information from more than one source [9].
It has been noted that the majority of human typing and spelling errors are quite minor, often involving just the omission or addition of a single character, typing two characters in the wrong order, or accidentally substituting one character for another (often one adjacent to the correct symbol on the keyboard). The Damerau–Levenshtein distance [1, 2] between two character strings measures how different the strings are by taking account of the minimum numbers of insertions, deletions, substitutions and transpositions of characters required to transform one of the strings into the other. Although one of the original motivations for the development of this metric was to compare the similarity of short pieces of natural language text, it has also been applied in fields such as genetics, for example to study how similar two fragments of DNA are to each other. However, we believe that our present paper is the first application of this metric to descriptions of mathematical expressions.

Statistical language models (SLMs) [18] have been at the core of ASR systems for many years [9, 10] where they use statistics from “past experience” to predict the likelihood of what will be spoken next, and combine this with evidence from the acoustic signal of the speech to decide what words were actually said. More recently, such SLMs have been incorporated into innovative systems for automatic translation between languages, such as Google Translate [11].

The simplest types of SLMs are N-gram models, which use statistics of the occurrences of specific sequences of N consecutive words within a database (or “corpus”) of training material observed in the past. Individual words (N=1) are referred-to as unigrams, pairs of consecutive words (N=2) as bigrams and triplets of consecutive words (N=3) as trigrams (see [19] for more details on use of trigram and similar models). Longer N-grams are not commonly used [9, 21]. N-gram models can be used both for analyzing how likely or common an observed sequence of words is (for example, is a given piece of text more typical of author A or of author B?), or for predicting the most likely candidate words or words to next appear in a sequence. This latter case is used in both ASR systems and in predictive text systems.

4.2. Datasets and Building the SLMs

Some ASR systems require the user to enroll (the process of adapting the system to a specific user, before it can actually be used for speech). In order to do this, a user has to provide samples of his or her speech input, which is then used to train the ASR. However, these samples do not cover all possible speech patterns and hence speech recognition programs tend to find "out of vocabulary" words or improbable word sequences when put to work. We hope our predictive and corrective language models will impose constraints which can be efficiently used in combination with those language and acoustic models built-in to the ASR system to correct these errors, with the aid of minimal intervention by the user, as perhaps the best source of knowledge on what was actually said by the user is the user him/herself.

For the work presented in this paper, we built various trigram-based SLMs for spoken mathematics generated from content extracted from a variety of "tutorial" web sites on elementary and intermediate level mathematics. We obtained approximately 4100 equations from such web sites and converted these to the simplest equivalent "spoken" forms using the formalized language described in Section II B. Table 1 shows the most frequent words found in this corpus of spoken mathematics. The SLMs were built from this data using the CMU Toolkit [7], applying Good-Turing discounting [13] and then evaluated for perplexity, as in our previous work [8]. The complete set of distinct words found in the corpus formed the vocabulary of the system.

Table 1. The most frequent words in our "spoken mathematical expressions" corpus. "x-ray" is the spoken form of the symbol "*x*".

Word	end	begin	x-ray	of	two	power	equals	bracket
Frequency %	7.58	7.57	7.3	7.26	4.82	3.93	3.79	3.61

5. *SWIMS* Prototype System

In this section, we introduce our prototype web-based mathematical document editor, *SWIMS (Speech-based Web Interface for Mathematics using SLMs)*, as a proof of concept. *SWIMS* was developed as a separate module which can be later integrated into the *TalkMaths* system following successful evaluation. The goal of *SWIMS* is to assist the user by predicting and/or correcting his/her input using SLMs prior to parsing, required in order to display the output on the screen using suitable mathematical rendering technology such as *MathML*. For ease of evaluation and for better performance, *SWIMS* has been divided into two units, one to predict the next word(s) in the input and the other to correct user mistakes. The former interface is called "Predictive Mathematics" and the latter "Alternative/Corrective Mathematics". We used *JSON* to store trigram probabilities for our SLM, *JavaScript* for calculating probabilities, and *jQuery* library to communicate between the browser and the *TalkMaths* parsing server. See [8] for a high-level diagram of the system's architecture.

*5.1. Predictive Mathematics Inte*rface

The predictive mathematics interface predicts one or two words ahead of the currently typed or dictated mathematical text. In order to predict one word ahead, the system uses the last two words of the input to match trigram probabilities. In the case that the input is less than two words or there is no matching trigram, then the system will back-off

[17, 20] to bigram probabilities and if this is not successful, unigram probabilities will be used. Two word prediction is a recursive extension of the one word prediction mechanism. Figure 2 shows this applied to the formula for the voltage across a capacitor which is being discharged through a resistor.

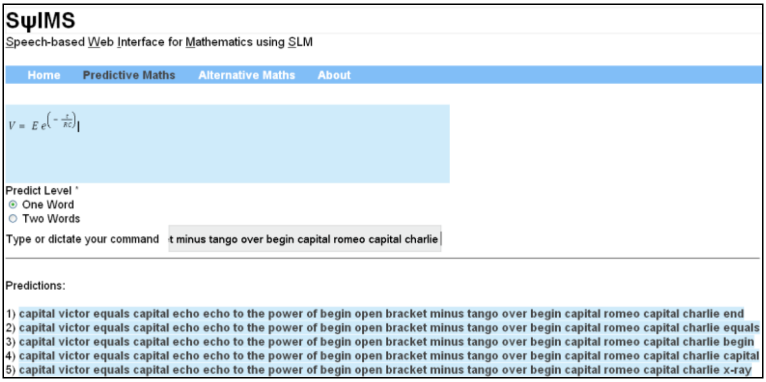


Figure 2. Predictive Mathematics Interface in use. In the top-ranked suggestion, the SLM predicts that "charlie" will be followed by "end".

5.2. Alternative/Corrective Mathematics Interface

To realize correction of errors, we implemented another web interface called Alternative/Corrective Mathematics. Once an "out of vocabulary" (OOV) word is detected, the Damerau – Levenshtein algorithm [1, 2] is used to calculate the Levenshtein distance of the typed word relative to each word in the vocabulary, in order to find suitable candidates for correction of the OOV word in question. Once a list of such candidates has been obtained, SLM probabilities can be used to re-rank the resulting new sequences of words. To illustrate this concept, we designed three variants of correction methods in the Alternative/Corrective Interface of SWIMS. These use Damerau – Levenshtein only, SLM only, and both in combination, respectively. To date, only the first of these has been developed and tested, but we intend to implement and evaluate the other two methods in future work.

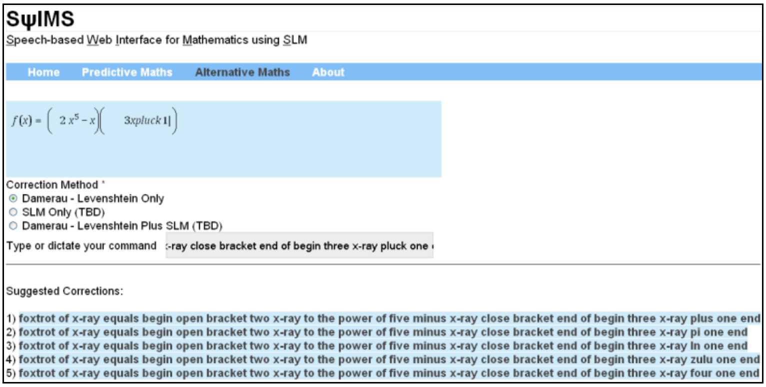


Figure 3. Alternative/Corrective Mathematics Interface in use. In the top-ranked suggestion, the OOV word "pluck" is replaced with "plus".

6. Initial Evaluation

In order to evaluate the predictive power of our statistical language models in the context of our alternative predictive mathematics system (*SWIMS*), we set up three experiments. From our previous studies of perplexities [8], we demonstrated that such

SLMs have the potential to be useful for prediction of mathematical text. The current study empirically evaluates these models when put into practice. For the first two experiments, A1 and A2, we trained a SLM using 90% of our database of spoken mathematical equations (≈3700 expressions). The remaining 10% (≈400) was then used to test the predictions offered by the system, based on the trained model, comparing these with the complete version of each expression. For the third experiment, A3, we varied the size of the training and test data sets in order to monitor the consequential change in the system's prediction performance. In experiment A4, the correction algorithm of the *SWIMS* system has been evaluated by artificially introducing a controlled selection of mistakes into otherwise "correct" expressions.

6.1. Experiments A1 and A2- prediction success rate depends on number of alternatives

Each expression in the test set was run through the interface with the last one (Experiment A1) or two (Experiment A2) word(s) omitted. We then observed the next word(s) predicted by the system, to see if one of highest ranked predictions (by probability) contained the actual missing word(s). The word "end" is normally used as a "context cue" within our specialized language for spoken mathematics, resulting in it being the most common word (see Table 1). Hence, we did not test expressions ending with "end" (A1) or ones which had "end" in the last two words (A2). Table 2 illustrates the percentages of times the correct prediction was included in the list of M "best" suggestions being offered to the user, and how this varied with M. In order to check that the results obtained were consistent, we performed 10 fold cross validation by dividing the complete dataset into 10 "folds", using one fold as the test set whilst the other 9 folds were used to build the SLM in each trial. The results of Experiment A1 are also represented graphically in Figure 4, which shows that the success of the one word ahead prediction increased as the number of suggestions shown to the user was increased, but with diminishing return.

Table 2. Experiment A1: Variation of success rate of one word ahead prediction with number of suggestions offered to the user.

	Size of Training Set	Size of Test Set	Number of Suggestions				
			5	10	15	20	25
Min	3684	407	62%	74%	79%	85%	88%
Mean	3691.8	410.2	63.2%	77.6%	84.4%	88.9%	91.1%
Max	3695	418	66%	81%	87%	91%	94%

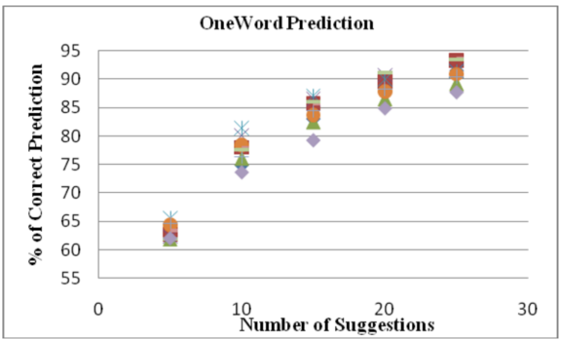


Figure 4: One word ahead prediction success rate increasing with the number of suggestions offered to the user

Experiment A2 evaluated the two words ahead prediction within *SWIMS*, in a similar manner to Experiment A1. The results are summarized in Table 3, and graphically in Figure 5. The trend is similar to that for A1, but the success rates in A2 are lower for a given number of suggestions.

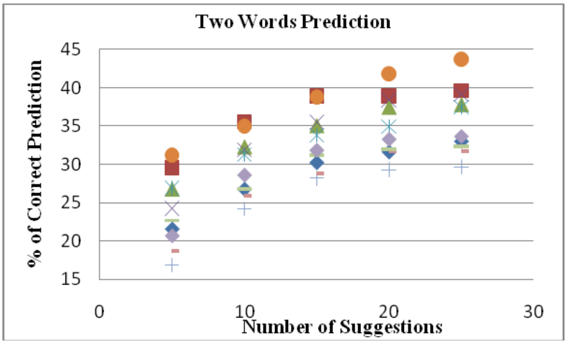


Figure 5: Two word ahead prediction success rate increasing with the number of suggestions offered to the user

Table 3. Experiment A2: Variation of success rate of two word prediction with the number of suggestions offered to the user.

	Size of Training Set	Size of Test Set	Number of Suggestions				
			5	10	15	20	25
Min	3684	407	17%	24%	28%	29%	30%
Mean	3691.8	410.2	24.3%	30.2%	33.6%	35.2%	36.2%
Max	3695	418	31%	36%	39%	42%	44%

6.2. Experiment A3- prediction success rate depends on size of training dataset

In Experiment A3, we observed how the success rate for one word prediction varied as different sized data sets were used to train the SLM. The results are summarized in Table 4.

Table 4. Experiment A3: Variation of success rate of one word ahead prediction with SLM size (5 suggestions per trial). MTrSS is Mean Training Set Size, MTSS Mean Test Set Size and M% Mean %

MTrSS	3691.8	3281.6	2871.4	2461.2	2051	1640.8	1230.6	820.4	410.2
MTSS	410.2	820.4	1230.6	1640.8	2051	2461.2	2871.4	3281.6	3691.8
M%	63.10%	62.60%	62.30%	62.10%	61.30%	60.30%	59.00%	56.50%	52.70%

6.3. Experiment A4- how successful the correction system is at correcting errors

In order to evaluate the performance of the correction algorithm, we artificially introduced some controlled errors into each of 100 expressions selected from a test expressions, then observed the proportion of these where the "correct" version was found within the 5 top ranked alternatives offered by our correction system. This was carried out for each of introducing R characters per expression, deleting R characters per expression and swapping R pairs for adjacent characters, for each of R = 1, 2, 3. The percentage of expressions which were successfully corrected using this approach for each trial are shown in Table 5. It can be seen that our method is extremely successful in correcting up to 3 insertions or transpositions of characters per expression, and fairly successful in correcting cases where up to three characters have been deleted from an expression. However, investigation of its performance in cases involving more complex or larger number of errors will require further experiments.

Table 5. Experiment A5: Variation of success rate (%) of correction using Damerau-Levenshtein method (5 suggestions offered per trial)

Number of Changes	1	2	3
Deletion of characters	95	92	68
Insertion of characters	100	98	97
Swapping pairs of adjacent characters	100	95	91

7. Discussion and Conclusion

From Experiment A1 and A2, we observe that one word ahead and two word ahead prediction success rates can be improved by increasing the number of alternatives, M, suggested to the user. However, the rate of increase of success rate diminishes as M increases, and it would appear that the maximum possible rates are about 90% for one word prediction, but around just 40% for two word prediction. However, if the user has to read too large a number of suggestions, the cognitive load imposed will very great. Thus, the number of options displayed must be limited. Based on our results, we propose that between 5 to 10 suggestions should be offered for one word prediction, giving success rates from 63 to 80%. Displaying the alternative expressions rendered into standard mathematical notation may help ease the reading burden on the user. However, two word prediction is rather less useful unless a large number of suggestions is presented. Experiment A3 showed that a small increase in success rate for one word ahead prediction could be achieved by increasing the amount of training data used, whilst M remained fixed. This is consistent with other studies of the predictive power of models based on other types of text [4, 22 and 23]. Finally, according to the results we obtained in Experiment A4, the Damerau-Levenshtein based correcting method is highly successful at correcting up to 3 errors at the character level errors in an expression. However, when the number or complexity of such errors is increased, the efficiency of correction declines.

At present, the predictions are limited to the vocabulary of the SLM. This implies that each time a new word (in our case, a spoken name of a mathematical entity) is encountered it will need to be added to the system's vocabulary. This should be relatively straightforward for the Damerau-Levenshtein based method. However, in order to modify the SLM, the corpus of mathematical expressions will have to be extended to reflect the change. Although possible in principle, this is not straightforward as one would have to find a considerable additional amount of data in order to update the model. Online learning within an adaptive system may be the solution to this issue.

Our work to date has indicated that the prediction/correction assistive facilities incorporated into *SWIMS* have potential to help make mathematical editing systems, including *TalkMaths*, more powerful and user-friendly. Improving such systems in this manner should in turn make writing and editing mathematics in electronic documents much easier for three groups – the disabled, on-line (particularly "at a distance") learners and people relying heavily on the use of portable devices – for whom these tasks are currently very difficult or even near impossible.

8. Future Work

Our method should be easily adaptable to use in several other closely related domains. For example, computer algebra systems such as Maple and Mathematica have their own language and syntax for mathematical expressions. It should be fairly straightforward for our predictive and corrective system to be integrated with these systems, assuming that enough data on past usage is available to train the models.

We have noted the issue of the high cognitive load on the user when having to read a large number of possible alternatives offered by the system. We aim to provide previews each of these, rendered into standard mathematical format, which should be make the user's task of identifying the correct version easier.

The corrective method we have used is currently not based on probabilities given by SLMs. Our next goal is to make use of SLMs to further enhance the current

Damerau-Levenshtein based correcting method, re-ordering candidate expressions by "likelihood". We also consider possibilities of incorporating a user's own "common usage" statistics, or specialized topic dependent models, in addition to our baseline SLM, so that an adaptive approach can be implemented into our system, fine-tuning it to be appropriate to the current need. Previous authors have tried to improve on N-gram models by combining these with "cache" or "word trigger" models [18, 45] in an attempt to incorporate longer-range statistical relationships between words. These approaches may prove fruitful for our system. We also aim to investigate ways we can adapt the Damerau-Levenshtein algorithm to correct speech recognition errors by applying a phonetically-weighted version to the phonetic transcription of the recognized words. For example, we would expect a higher rate of confusion between two similar phonemes, such as voiceless fricatives /s/ ("s") and /ʃ/ ("sh"), or between voiced plosives /d/ and /g/, than between two phonemes of different types [35].

References

[1] F. Damerau, A technique for computer detection and correction of spelling errors. *Communications of the ACM*, 1964, Vol. 7, Issue 3, pp. 659-664.

[2] V. Levenshtein, Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics-Doklady*, 1966, Vol. 10, Issue 8, pp. 707-710.

[3] D.R. Attanayake, E. Pfluegel, J. C. W. Denholm-Price, G. J. A. Hunter, Architectures for Speech-Based Web Applications, *Proc. 4th International Conference on Semantic E-business and Enterprise Computing (SEEC2011)*, July 20-22, U.K. 2011

[4] A. Wigmore, Speech-Based Creation and Editing of Mathematical Content. Ph.D. Thesis, Kingston University, U.K., 2011

[5] A. I. Karshmer, Access to mathematics and science, *Proc. 11th International Conference on Computers Helping People with Special Needs (ICCHP)*, 2008, Vol. 5105, pp. 873–874.

[6] TalkMaths Project, http://www.TalkMaths.org

[7] P. Clarkson, R. Rosenfeld, Statistical Language Modeling using the CMU-Cambridge Toolkit , *Proc. Eurospeech* 1997, Vol. 5, 2707-2710. Toolkit available on-line http://www.speech.cs.cmu.edu/SLM_info.html

[8] D.R. Attanayake, G. J. A. Hunter, J. C. W. Denholm-Price, E. Pfluegel, Interactive error correction using statistical language models in a client-server interface for editing mathematical text, *Designing Inclusive Systems – Designing Inclusion for Real-World Applications*, Ed. P Langdon et al. Springer Verlag, London, Chapter 13, pp. 125-132.

[9] S. Young, Large Vocabulary Speech Recognition : A review, *IEEE Signal Processing Magazine*, 1996, Vol. 13, Issue. 5, pp. 1-4.

[10] S. Young, Talking to Machines – Statistically Speeching, *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, Denver, Colorado, U.S.A, 2002

[11] Google, Inside Google Translate, 2011, http://translate.google.com/about/index.html

[12] G. Law, Phonetic alphabets (alpha bravo charlie delta), 2007 http://www.faqs.org/faqs/radio/phonetic-alph/full/

[13] I.J. Good, The population frequencies of species and the estimation of population parameters, 1953, *Biometrika*, Vol. 40, (3–4), pp. 237–264

[14] R. Fateman. How can we speak math? www.cs.berkeley.edu/~fateman/papers/speakmath.pdf, 2012.

[15] A. Wigmore, G. Hunter, E. Pfluegel, J. Denholm-Price, M. Colbert TalkMaths Better ! Evaluating and improving an intelligent interface for creating and editing mathematical text. *Proceedings of 6th International Conference on Intelligent Environments*, July 2010, Kuala Lumpur, Malaysia.

[16] L. A. Chang, Handbook for spoken mathematics (Larry's speakeasy). Lawrence Livermore National Laboratory, University of California, USA, 1983.

[17] S. Katz, Estimation of probabilities from sparse data for the language model component of a speech recognizer, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1987, Vol. 35, pp. 400-401

[18] R. Rosenfeld, Two decades of statistical language modeling: where do we go from here?, *Proc. of the IEEE*, 2000, Vol. 88 (8) (August), pp. 1270 – 1278

[19] L.R. Bahl, F. Jelinek & R.L.Mercer, A maximum likelihood approach to continuous speech recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 1983.

[20] R. Kneser and H. Ney, Improved backing-off for m-gram language modeling, *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP),* vol. 1, pp. 181 - 184 , 1995

[21] F. Jelinek, Up from trigrams ! - the struggle for improved language models, *Proceedings of EuroSpeech*, 1991

[22] R.K. Moore, There's No Data Like More Data - but when will enough be enough ? Proceedings of the Institute of Acoustics, 2001, Vol. 23 (3), pp. 19-26

[23] G. Hunter, M. Huckvale, Is it Appropriate to Model Dialogue in the Same Way as Text? A Comparative Study Using the British National Corpus, *Proc. 2006 European Modelling Symposium*, London, U.K, pp. 199-203

[24] T. Sancho-Vinuesa et al, Automatic verbalization of mathematical formulae for web-based learning resources in an on-line environment*, INTED2009 Proceedings*, 2009, Valencia, Spain, pp. 4312–4321

[25] J. Cuartero-Olivera et al, Reading and writing mathematical notation in e-learning environments, 2012, eLC Research Paper Series, Universitat Oberta de Catalunya, Spain, http://elcrps.uoc.edu/ojs/index.php/elcrps/about

[26] Y. Solomon & J. O'Neill, Mathematics and Narrative, *Language and Education*, 1998, Vol. 12 (3), pp. 210-221

[27] M. Österholm, Characterizing Reading Comprehension of Mathematical Texts*, Educational Studies in Mathematics*, 2006, Vol. 63 (3), pp. 325-346

[28] P. Gaura, REMathEx - Reader and Editor of the Mathematical Expressions for Blind Students, *Proceedings of 8th ICCHP 2002*, Springer-Verlag, London, U.K, pp. 486-493.

[29] N. Soiffer, MathPlayer: web-based math accessibility. In Proc. of the 7th int. ACM SIGACCESS conf. on Computers and Accessibility (Assets '05), 2005, ACM, New York, NY, USA, pp. 204-205

[30] C. Bernareggi & V. Brigatti, Writing mathematics by speech: A case study for visually impaired, *Proceedings of 11th ICCHP 2008*, pp. 879–882

[31] T. V. Raman, Audio system for technical readings, 1998, Springer Verlag, Berlin

[32] Metroplex Voice Computing, Inc. mathtalk.com, http://www.mathtalk.com/

[33] C. Guy, M. Jurka, S. Stanek, and R. Fateman, Math speak & write, a computer program to read and hear mathematical input, 2004, Technical report, University of California, Berkeley, Electrical Engineering and Computer Sciences Department

[34] T. Hanakovic & M. Nagy, Speech recognition helps visually impaired people writing mathematical formulas, Proceedings of 10th ICCHP 2006, pp. 1231–1234.

[35] P. Ladefoged & K. Johnson*, A Course in Phonetics* (Sixth Edition), Cengage Learning, 2010

[36] H. Ferreira, D. Freitas, Enhancing the Accessibility of Mathematics for Blind People: The AudioMath Project, 2004, K. Miesenberger et al (eds.) *Proc. ICCHP 2004*, Springer LNCS, vol. 3118, pp. 678–685.

[37] Math in Braille Project, 2011, http://www.mathinbraille.at/en

[38] M. Suzuki, F. Tamari, R. Fukuda, S. Uchida, & T. Kanahori, Infty - an integrated OCR system for mathematical documents. In C. Vanoirbeek, C. Roisin, E. Munson (eds.) *Proceedings of ACM Symposium on Document Engineering*, 2003, pp. 95–104, http://www.inftyproject.org/en/index.html

[39] S. M. Watt and X. Xie, Recognition for Large Sets of Handwritten Mathematical Symbols , *Proc. IEEE Int. Conf. on Document Analysis and Recognition, (ICDAR 2005)*, Korea, IEEE Press, pp. 740-744

[40] A. Fujiyoshi, M. Suzuki, S. Uchida, Syntactic Detection and Correction of Misrecognitions in Mathematical OCR, *Proceedings of The 10th International Conference on Document Analysis and Recognition, ICDAR 2009*, Barcelona, Spain, pp.1360-1364

[41] E. Smirnova & S.M. Watt, Context-Sensitive Mathematical Character Recognition, *Proc. IAPR Int. Conf. on Frontiers in Handwriting Recognition, (ICFHR 2008),* August 19-21 2008, Montreal, Canada,

[42] N. Jacobs, "MathGenie : User's Guide and Teacher's Manual", 2006, http://logicalsoft.net/MathGenie.pdf

[43] K. Vertanen & D.J.C Mackay, Speech Dasher: Fast writing using speech and gaze. *Proceedings of CHI* 2010, U.S.A., pp. 595-598.

[44] P. Zielinski, Opengazer: opensource gaze tracker for ordinary webcams, 2007, http://www.inference.phy.cam.ac.uk/opengazer/

[45] R. Kuhn & R. D. Mori, A cache-based natural language model for speech reproduction, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 1990, Vol.12(6), pp. 570–583

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-298

The InterReality Portal: A Mixed Reality Co-creative Intelligent Learning Environment

Anasol PEÑA-RIOS [a,1], Vic CALLAGHAN [a,b], Michael GARDNER [a] and Mohammed J. ALHADDAD [b]
[a] *Department of Computer Science, University of Essex, UK.*
[b] *Faculty of Computing and Information Technology, King Abdulaziz University, KSA.*

Abstract. Technology has become an inseparable part of everyday life, as we use it at home, at work, in school etc. For education, technology is frequently used as a tool to support classroom-based learning. However, an emerging challenge is the application of technology, beyond the traditional classroom, to a diverse mix of pedagogical techniques. Traditional education is based on the process of transmitting knowledge; however, it is necessary for education to encourage the formation and strengthening of skills such as creativity, innovation, analysis, communication and other capabilities that are important aspects of professional life. This paper extends our previous work towards developing an *InterReality Portal* learning environment. The *InterReality Portal* applies Problem-based Learning (PBL) pedagogy and co-creative learning to the realization of mixed-reality laboratory activities for learning computing. To reach this goal we propose a series of structured lab activities created using IMS Learning Design, with a combination of *Cross-Reality (xReality)* and *Virtual* objects, to produce Internet-of-Things-based computer projects using collaborative interaction between geographically dispersed students.

Keywords. Mixed reality, intelligent learning, learning design, cloud learning, co-creative learning, constructionism, interreality portal, end-user programming, xReality objects.

Introduction

Learning is one of the topics in which progress in technology has enabled major innovative advances in practice (e.g. from web-based education to ubiquitous learning or intelligent tutor agents). However, we need to ask whether advances in technology, that enable learning outside the traditional classroom, are really bringing significant educational gains? Papert's constructionism theory established that the acquisition of long-life learning was a consequence of performing active tasks that construct meaningful tangible objects in the real world and relating them to personal experiences and ideas [1]. New trends on education technology take the view that learning should not be focus on only memorizing information but should also develop problem-solving skills. Problem-based Learning (PBL) is a constructionist student-centred pedagogy where students work on co-creative problem solving, where learning occurs as a side-

[1] Corresponding Author: Anasol Peña-Ríos, School of Computer Science & Electronic Engineering, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK; E-mail: acpena@essex.ac.uk

effect of problem solving [2]. Are these approaches being taken into consideration by the new technology-based learning environments?

In a previous work [3] we presented a Mixed Reality Intelligent Learning (MR-iLearning) Model as a context for an online learning environment. The MR-iLearning Model was based on:

a) *Intelligent Learning* (iLearning), an innovative paradigm of learning that is able to offer ubiquitous personalized content via a context-aware environment [3].
b) A *cloud learning infrastructure* [4], which provides strategic benefits such as storage capacity, resource sharing and adaptation, mobility and accessibility, plus the possibility to keep a unified track of learning progress, thereby providing attributes such as synchronous sharing, asynchronous storage and high-availability [5].
c) The *Instructional Management Systems (IMS), Global Learning Consortium Learning Design specification* for the creation of learning activities. This enables learning objects to be designed to achieve particular goals, regardless of the pedagogical methods utilised [6].

In this paper we explore the application of these concepts to a holistic learning environment. We begin by introducing an innovative online learning environment, the *InterReality Portal* (our work-in-progress test bed), its conceptual architecture and implementation. This is followed by a description of the operation and design of our exemplary virtual lab (based on embedded-computing) in the *InterReality Portal* using *Cross-Reality (xReality)* and *Virtual* objects. After this we discuss the use of end-user programming for customization issues. Finally, we provide conclusions and identify challenges to be addressed in our future research.

1. The InterReality Portal

Grounded on the learning-by-doing vision of constructionism, our research is inspired by an innovative learning environment, derived from a science-fiction prototype, described in "*Tales from a Pod*" [7]. In this story we described how advanced technology might create a futuristic educational environment where learning can be provided within an immersive mixed-reality teaching environment (*pods*). Students can experience personalised learning, along with collaborative learning, through a network of interconnected *pods* distributed geographically to create large-scale education environments. In a similar way, the *InterReality Portal* is able to deliver personalised online content enhanced with co-creative mixed-reality activities that support a PBL pedagogy and co-creative learning [8] in an immersive learning environment.

We define an *InterReality Portal* as a collection of interrelated devices (both real and abstract) comprising a 3D virtual environment, physical objects and software agents that allow people to complete activities at any point of Milgram's Virtuality Continuum [9] [3]. For the application of PBL pedagogy, we propose the use of mixed reality lab activities, that enhance effective problem solving and collaboration skills by enabling students to work in groups and utilise real and virtual objects to co-creatively construct computing systems. The interaction between the elements can be defined as

Cross-Reality (*xReality*) [10]. During this interaction, networked sensors obtain real-world information and pass it to the 3D virtual environment, where the data is processed to trigger events previously programmed by the teacher or student. In a similar way, interaction performed by virtual objects can be reflected into the physical world through a diverse set of physical artefacts.

1.1. Conceptual model

Our *InterReality Portal* has four layers. The first layer is the *Client layer* or real world, where the user interacts directly with the *xReality* objects. The second layer, the *Data Acquisition layer*, is responsible for obtaining real-world information produced by interactions between the user and the *xReality* object or between the user and the learning environment. The main actor at this level is the *Context-awereness* agent, which identifies the object being used in a learning task with the aid of QR codes, cameras and network eventing data. This information is sent to the *Mixed Reality (MR) Agent*, in a subsequent *Event Processing layer*. The MR agent obtains, from the *Resources Repository,* a set of rules and actions (behaviours) available for the object. Finally the MR agent instantiates a virtual representation of the *xReality* object with its properties and rules in the *Virtualization layer*.

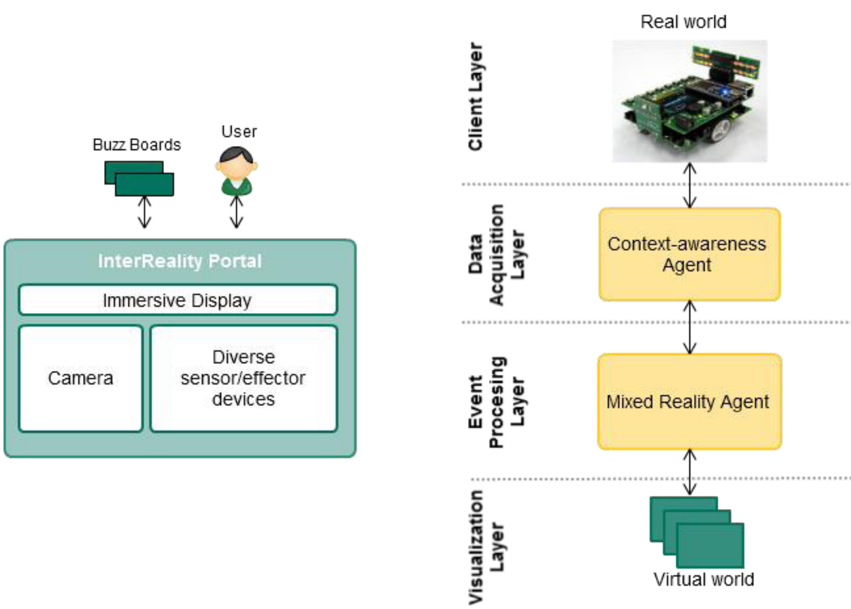


Figure 1. InterReality Portal Conceptual Model.

1.2. Implementation

For the implementation, we are utilising a semi-spherical sectioned screen, the ImmersaStation, manufactured by Immersive Display's Ltd. [11] (Fig. 2). One characteristic of this device is that it allows the completion of mixed-reality activities in a less intrusive way by offering a traditional work-desk that allows a free-range of head movement without the need for any other body instrumentation (e.g. special glasses).

Our prototype includes a camera that, in addition to enabling videoconferencing, is used to read QR codes that allow the automatic identification of actors and objects.

The *InterReality Portal* uses two different types of objects: *xReality* objects and Virtual objects (Fig. 3). *xReality* objects refers to objects that have a physical (or optionally, virtual) form and contain rules, that determine the behaviour and interaction with other objects or the environment (e.g. all FortiTo buzz-boards can work with any other Buzz-Boards and must behave according to the laws of physics, etc.). *Virtual* objects can only exist inside the 3D virtual environment although they can have rules and behaviours associated with them (e.g. software created for the virtual lab activity).

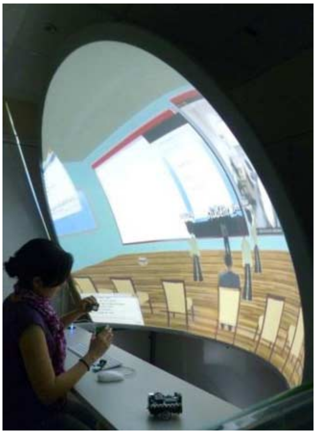


Figure 2. InterReality Portal.

Finally mixed-reality lab activity, based on the above, can be created using combinations of *xReality* objects and *Virtual* objects to create *Internet-of-Things* applications. The *Internet-of-Things* definition establishes a union between the virtual world of 'information' with the real world of 'things' allowing numerous interesting applications to be constructed [12]. The learning goal of our activity is to create a computer science project that includes hardware and software modules, emphasising computing fundamentals.

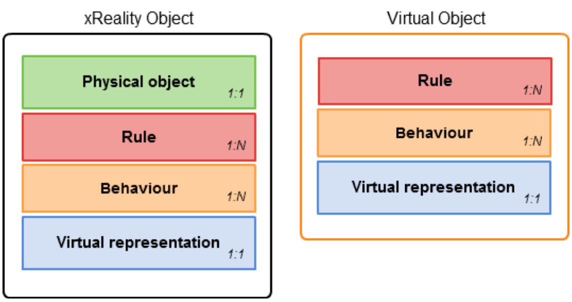


Figure 3. Types of objects.

1.3. 3D Virtual Environment

Virtual worlds are used in distance-learning because they enhance social interaction between remote people giving them a greater sense of presence and engagement within the class. The University of Essex MiRTLE project (Fig. 4) is a virtual learning environment based on teacher/student interaction that links a physical classroom to a virtual classroom, providing an instructive educational setting [13] [14].

The application of this technology, based on Papert's Instructionism, is an example of the use of technology for traditional classroom-based education, where the knowledge is transmitted from teacher to student based on isomorphic concepts [1] [16].

MiRTLE is based on Open Wonderland, a java-based open source toolkit for creating collaborative 3D virtual worlds [15]. Some of the benefits of using this toolkit are the integration with other technologies, such as immersive audio, sharing live desktop applications or documents and the extensibility afforded by the tools that enable the creation of new worlds and features.

By taking MiRTLE as our 3D Virtual World platform for the InterReality Portal, we will experiment with the transformation from an instructionist pedagogy into a constructionist model for co-creative learning.

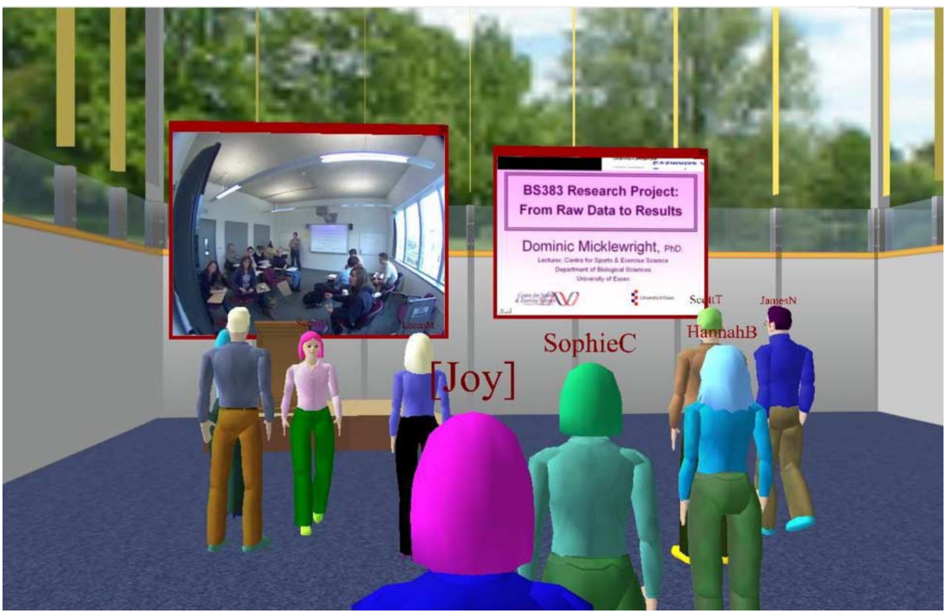


Figure 4. MiRTLE.

1.4. XReality Objects

For the construction of mixed reality lab activities we are utilising Fortito's Buzz-Board Educational Toolkit (Fig. 5) [17] [18]. This educational toolkit comprise over 30 pluggable hardware boards that can be interconnected that, together with software modules, can create a variety of *Internet-of-Things* applications [19] such as mobile robots, mp3 players, heart monitors, etc. The modularised set of objects (hardware & software) allow the students to create projects using the deconstructed appliance model. In this model a number of elementary services can be combined to create higher order functions [20]. The hardware boards communicate via network events. The software modules care developed by students using the C & C++ language.

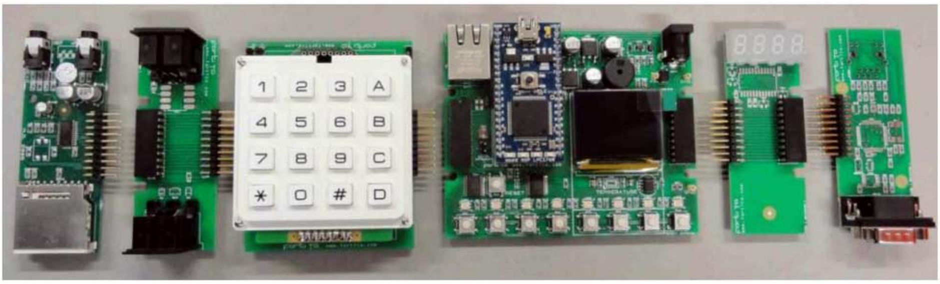


Figure 5. FortiTo Educational Toolkit.

1.5. Virtuality Continuum Learning Activities

In order to create meaningful activities, that enhance and promote learning, it is necessary to establish appropriate learning goals and consider the design of such activities. For our research we have defined a classification for learning activities according to characteristics such as temporality (time learning/teaching) and context (type of action and scope of action) [3]. This classification enables the use of the MR-iLearning Model in different learning environments. However, in our system, the *InterReality Portal* activities are performed within the *Virtuality Learning Continuum*, where the scope (physical-virtual) of the action performed, is relevant to the completion of the task.

According to IMS Learning Design, the structured sequences of learning activities are known as Units of Learning (UoL) and can be preceded by zero or more conditions before starting or completing the tasks [6]. The learner is the person who performs this sequence of actions in order to fulfil one or more inter-related learning objective. In these activities the teacher becomes more of a facilitator than a lecturer, and the learning environment allows a teacher to create learning activities regardless their expertise on computers.

Figures 6 & 7 shows the interaction between system components in a collaborative *Virtuality Learning Continuum* activity. The first diagram relates to the beginning of the learning session. In this the *Context-Awareness Agent* (CAa) identifies the object being used in the learning session with the aid of QR codes, camera and network eventing data. This information is sent to the *Mixed Reality Agent* (MRa), which obtains, from the *Resources Repository (*via the *Content Manager),* a set of rules and actions (behaviours) available for the object. This information is sent to the *UoL Manager*, which constructs the sequence of activities in the UoL. In support of these activities, the *Mixed-Reality Agent* instantiates a virtual representation of the *Buzz-*

Boards and other objects in the *3D Virtual Environment*. Finally the *UoL Manager* starts with the execution of the activities. At this point the *Context-Awareness Agent* and the *Mixed-Reality Agent* will handle a single "*Dual-Reality*" state in which the objects in the virtual world are a copy of the real world and if there is any change performed in either the virtual or the real world, the *Context-Awareness Agent* will perceive it and send the updates to the *Mixed Reality Agent*.

For the co-creative virtual lab, additional learners perform the same steps, and then through the *3D Virtual Environment* they establish communication with other remote learners. As long as the session continues, changes to any of the Inter-Reality Portal objects will result in the *Mixed-Reality Agent* handling the following situations:

a) A change in any *Virtual object* in a given InterReality Portal results in identical changes to all subscribing InterReality portals.

b) A change in a *xReality object* of a given InterReality Portal results in changes in the representation of the real device on all subscribing InterReality portals.

While this synchronization processes occurs between the *Context-Awareness Agent* and the *Mixed Reality Agent,* our connected InterReality Portals extend the single "*Dual-Reality*" concept to a multiple *"Dual-Reality"* states in which all the virtual world views should be symphonises and synchronised with each other.

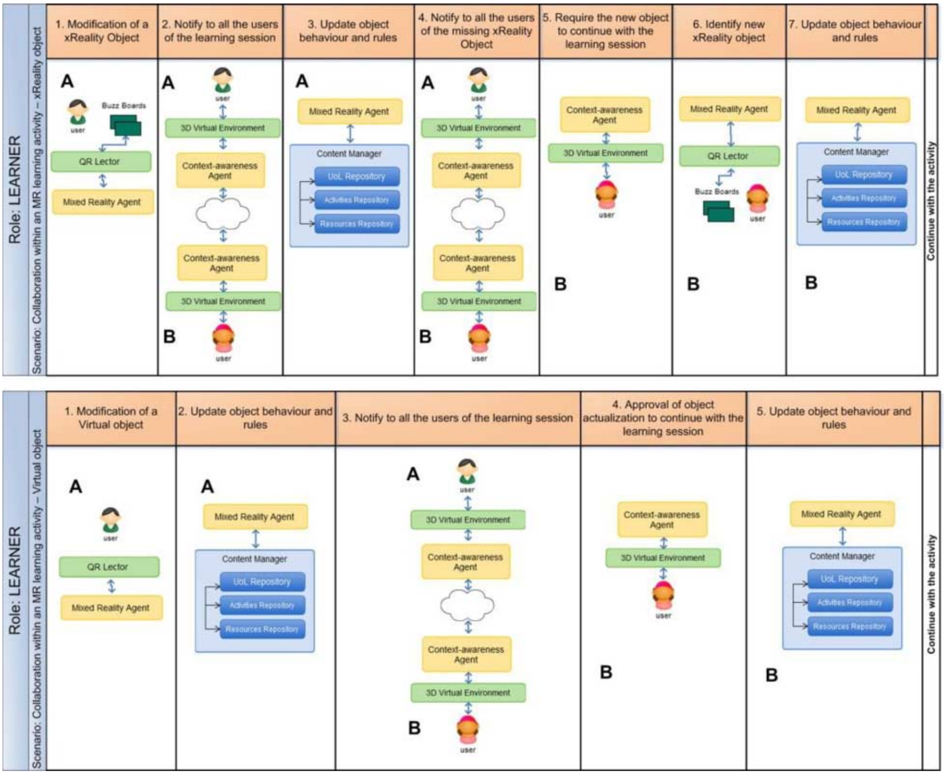


Figure 6. *xReality* and *Virtual Reality* objects interaction in a Collaborative MR learning activity.

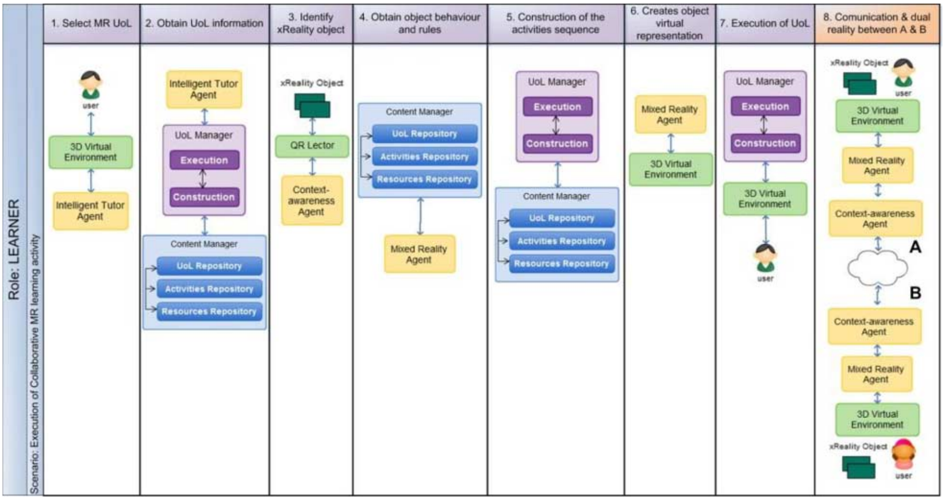


Figure 7. Execution of a Collaborative MR learning activity.

1.6. End-user programming

xReality and *Virtual objects* can be considered as a collection of "atomic" services (software) and devices (hardware) which, could be configured into novel combinations, forming personalised *Internet-of-Things* appliances [20]. To create these mash-ups and participate in the co-creative process of learning it is necessary that the *InterReality Portal* provides a satisfactory end-user programming environment. End-user programming refers to the use of techniques that enable non-technical people to create a set of coordinated actions that the learning environment can perform to mirror some specific designed operations [21].

End-user programming for the *InterReality Portal* can be examined from two different angles:

a) From the learner's view, during the learning session when they interact with the environment and between them; they are programming a series of actions to be executed by the objects.

b) From the instructor's view, when they create a UoL they are establishing a sequence of activities to be performed by the learners during the educational session.

Recently, different approaches to encourage and empower users to create "programs" have been developed. One approach is to use iconic objects (e.g. a graphical icon, a physical artefact) that represent programming constructs and which people "assemble" (e.g. a jigsaw puzzle), to create algorithms without logic programming knowledge [22] [23]. These, combined with *xReality* or *Virtual* object's rules & behaviours and user' profiles, are the mechanisms used in our *InterReality Portal*.

Summary and future work

In this paper we have described a holistic immersive intelligent learning environment that offers a co-creative constructionalist PBS learning to teams of geographically dispersed students. In addition, we have defined and explored how *xReality* and *Virtual* objects can be combined with end-user programming to create an educational mixed-reality object that can be constructed and shared between learners during a *Virtuality Learning Continuum* session. By aggregating such objects students may co-creatively construct a variety of interesting lab assignments. These *Virtuality Continuum* activities and learning sessions are regarded as belonging to the pedagogical schools of "experientialism" and "enactivism" due to the emphasise on interaction with tangible objects, rather than conceptual abstractions [24].

We have discussed how the benefits to differing educational stakeholders. For example, the learner can benefit from a more natural acquisition of long-life knowledge, constructed by the correlation between concepts and real tasks. The instructor can benefit from the standardised construction of learning activities using IMS Learning Design, allowing him or her to share and re-use learning material. An additional benefit may occur as this learning system can also serve as a rapid prototyping system for entrepreneurs or companies developing Internet-of-Things products.

Finally, by way of providing a signpost to our future work, Figure 8 summarises the implementation phases of our research. The first stage involves the implementation and construction of a fully immersive *InterReality Portal*, able to recognise and work with *xReality* and *Virtual* objects. The second stage concerns the use of the *InterReality Portal* to implement a Learning Design example based around the use of *Virtuality Learning Continuum* UoLs, enabling us to evaluate the first scenario of our pedagogical model; the single immersive mixed reality learning environment.

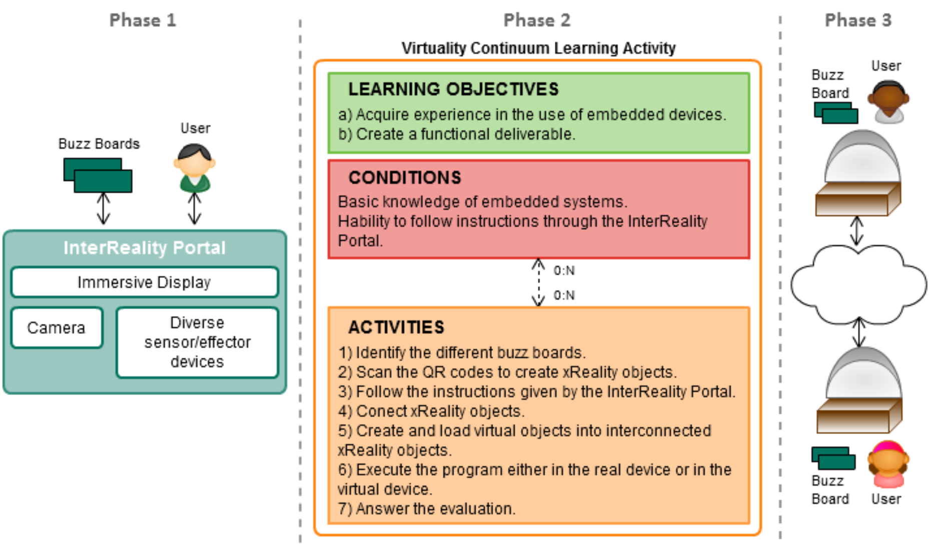


Figure 8. Implementation phases.

The final stage of the implementation consists of the construction of a second *InterReality Portal* to perform the co-creative virtual lab between two people in separate locations. By doing this we will expand usage into a wider geographical and pedagogical context thereby allowing *InterReality Portal* educational resources and immersive environment sharing, based on cloud computing technology, plus a deeper exploration of co-creative PBL pedagogies.

We are currently in the process of integrating the *InterReality Portal* with the learning design concept, and therefore much research will be done over the coming days to answer various research questions ranging from technical issues such as "*how to map the virtual and physical spaces together, in a single and multiple Dual-Reality state*" to educational concerns, such as "*examining whether the interconnectivity between physical and virtual spaces, provides significant advances to collaborative and creative learning*". We look forward to presenting significant progress on these and other issues at forthcoming workshops and conferences.

Acknowledgments

We are pleased to acknowledge King Abdulaziz University, Saudi Arabia for its generous funding of this research project, including the provision of a PhD scholarship to the lead author. In addition, we wish to thank Immersive Displays UK Ltd. and Fortito Ltd. for their support. Finally we are pleased to acknowledge Dr. Jeannette Chin (Essex University) for her important advice with the end-user programming aspects, Prof. Dr. Minjuan Wang (San-Diego State University) for her invaluable guidance on educational technology issues, Malcolm Lear (Essex University) for technical support relating to Buzz-Board technology and Dr. James Dooley (Essex University) for his assistance with the technical aspects of the immersive reality desk.

References

[1] S. Papert & I. Harel, Constructionism, Ablex Publishing Corporation, 1991.

[2] K. -. K. Ngeow & S. Yoon, «Learning to learn: preparing teachers and students for problem-based learning. ERIC Digest.,» 2001. [Online]. Available: http://www.ericdigests.org/2002-2/problem.htm. [Last access: 03-2012].

[3] A. Peña-Ríos, V. Callaghan, M. Gardner & M. J. Alhaddad, «Towards the Next Generation of Learning Environments:An InterReality Learning Portal and Model,» on *8th International Conference on Intelligent Environments 2012*, Guanajuato, Mexico, 2012.

[4] S. Kim, S.-M. Song & Y.-I. Yoon, «Smart Learning Services Based on Smart Cloud Computing,» *Sensors,* vol. 11, nº 8, pp. 7835-7850, 2011.

[5] N. Sultan, «Cloud computing for education: A new dawn?,» *International Journal of InformationManagement,* vol. 30, nº 2, pp. 109-116, 2010.

[6] IMS Global Learning Consortium, «Learning Design Specification,» IMS Global Learning Consortium, 2003. [Online]. Available: http://www.imsglobal.org/learningdesign/. [Last access: 12-2011].

[7] V. Callaghan, «Tales From a Pod,» on *Creative Science Workshop*, Nothingham, 2010.

[8] P. Dillenbourg, «What do you mean by 'collaborative learning'?,» on *Collaborative Learning: Cognitive and Computational Approaches. Advances in Learning and Instruction Series.*, Elsevier Science, Inc., 1999, pp. 1-15.

[9] P. Milgram & F. Kishino, «A taxonomy of virtual reality displays,» *IEICE TRANSACTIONS on*

Information and Systems, Vols. 1 of 2 E77-D, nº 12, pp. 1321-1329, 1994.

[10] J. A. Paradiso & J. A. Landay, «Cross- Reality Environments,» *IEEE Pervasive Computing,* vol. 8, nº 3, pp. 14-15, 2009.

[11] Immersive Display Group, «ImmersaStation,» 2011. [Online]. Available: http://www.immersivedisplay.co.uk/immersastation.php. [Last access: 10-2011].

[12] D. Uckelmann, M. Harrison & F. Michahelles, «An Architectural Approach Towards the Future Internet of Things,» on *Architecting the Internet of Things*, Springer Berlin / Heidelberg, 2011, pp. 1-25.

[13] V. Callaghan, M. Gardner, B. Horan & J. L. Scott, «A Mixed Reality Teaching and Learning Environment,» on *Proceedings of the 1st International Conference on Hybrid Learning and Education*, Hong Kong, China, 2008.

[14] M. Gardner & M.-L. O'Driscoll, «MiRTLE (Mixed-Reality Teaching and Learning Environment): from prototype to production and implementation,» on *AcrossSpaces11 Workshop in conjunction with the EC-TEL 2011*, 2011.

[15] Open Wonderland Foundation, «Open Wonderland,» Open Wonderland Foundation, [Online]. Available: http://openwonderland.org/. [Last access: 04-2012].

[16] M. Driscoll & K. Rowley, «Semiotics: Toward learning-centered instructional design,» on *Instructional development paradigms*, Englewood Cliffs NJ: Educational Technology Publications, 1997, pp. 311-334.

[17] Fortito - Educational Technology, «Buzz Products,» 2011. [Online]. Available: http://www.fortito.com/buzz/. [Last access: 10-2011].

[18] M. Wang, V. Callaghan, M. Lear & M. Colley, «Teaching Next Generation Computing Skills: The Challenge of Embedded Computing,» on *Intelligent Campus 2011 (iC'11)*, Nothingham, 2011.

[19] V. Callaghan, «Buzz-Boarding ; practical support for teaching computing based on the internet-of-things,» on *The Higher Education Academy - STEM*, London, 2012.

[20] J. Chin, V. Callaghan & G. Clarke, «Soft-appliances : A vision for user created networked appliances in digital homes,» *Journal of Ambient Intelligence and Smart Environments,* vol. 1, nº 1, pp. 69-75, 2009.

[21] D. C. Smith, A. Cypher & L. Tesler, «Novice Programming Comes of Age,» *Communications of the ACM,* vol. 43, nº 3, pp. 75-81, 2000.

[22] J. Chin, V. Callaghan & G. Clarke, «A programming-by-example approach to customising digital homes,» on *4th International Conference on Intelligent Environments*, Seattle, 2008.

[23] MIT Media Lab, «Scratch,» MIT Media Lab, [En línea]. Available: http://scratch.mit.edu/. [Last access: 04-2012].

[24] Q. Li, B. Clark & I. Winchester, «Instructional Design and Technology Grounded in Enactivism: A Paradigm Shift?,» *British Journal of Educational Technology,* vol. 41, nº 3, pp. 403-419, 2010.

1st International Workshop on Intelligent Multimodal Interfaces Applied in Skills Transfer, Healthcare and Rehabilitation (IMIASH'12)

Oscar Sandoval (Instituto Tecnológico de Orizaba, México)

Alejandro Rodriguez-González (Universidad Carlos III, Spain)

Paolo Tripicchio (Scuola Superiore Sant'Anna, Italy)

Otniel Portillo-Rodríguez (Universidad Autónoma del Estado de México, México)

This page intentionally left blank

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

Introduction to the Proceedings of IMIASH'12

This section of the workshop proceedings presents contributions from the first International Workshop on Intelligent Multimodal Interfaces applied in skills transfer, healthcare and rehabilitation. The aim of the workshop was to bring together researches, developers and practitioners involved in the research area of machine learning, bio-signals processing, electronics, robotics, mechatronics, virtual & augmented reality, medicine and rehabilitation.

Nowadays, multimodal interfaces have changed and revolutionized the way that people communicate, interact and learn. Undoubtedly, we are living in an age where these technological improvements have increased the human potential at levels that have not been seen before. Day by day, thanks to the science and technology, we are able to see how the fiction is converted into reality and how the multimodal interfaces are naturally integrated in the current life. This strong integration between humans and technology represents an incredible opportunity to investigate the human behavior through this interaction and analyze all the potentialities and advantages when the multimodal interfaces are used in the field of skills transfer, healthcare and rehabilitation.

The use of new technologies has dramatically transformed the conventional methodologies in which humans learn. In one hand Skills Transfer is focused in the human behavior of learning process in which motor skills activities are transmitted from intelligent multimodal interfaces to human beings. The research investigates how these learning processes could be applied using intelligent multimodal systems and virtual environments to transfer knowledge associated to human motor skills. Skills transfer is a trend in the scientific and technological development fields, which focuses on developing systems that can help the users (human beings) to learn different kinds of tasks. These systems are designed to deal with the acquisition, interpretation, storing and transfer of human skills by means of multimodal interfaces, Robotics, Haptics and Virtual Environments (VE) technologies. It also adopts cognitive sciences and interaction design methodologies in order to obtain a digital representation of the skills and to develop techniques for its capturing and rendering.

In other hand the multimodal interfaces in healthcare have been changing and improving the quality of life of the people, principally of handicap people. Therefore, this track is focused in the design of intelligent multimodal systems that improves and accelerates the rehabilitation process of the patients through the use of robotic and electronic devices capable to help the users to carry out different routines in an accurate way.

The workshop organizers deeply thank the conference organizers and the workshop Program Committee for the help on organizing this event.

Oscar Sandoval, Paolo Tripicchio, Otniel Portillo and Alejandro Rodríguez
Workshop Organizers IMIASH'12

J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-312

Knee Rehabilitation with fuzzy Velocity and force control and compensation of weight

Juan Alberto PRADO-PRICHARDO [a,1], Armando SEGOVIA-DE LOS RÍOS [a,b], Otniel PORTILLO-RODRIGUEZ [c]
[a] *Instituto Tecnológico de Toluca, Toluca, Estado de México, México*
[b] *Instituto Nacional de Investigaciones Nucleares*
[c] *Universidad Autónoma del estado de México.*

Abstract. This work reflects the intention of reduce import costs of rehabilitation equipments in México. This rehabilitation requires the integration of several disciplines, not just engineering also medicine. The mechanical design is the main and the most important factor for safety, but also the design considers the detection of force in flexion and extension to increase patient safety which is their main concern. This will redundant with the electronic system, the fuzzy control is implemented for your interaction with nature in biomechanical systems.

Keywords. Knee Rehabilitation, Robotics Rehabilitation

Introduction

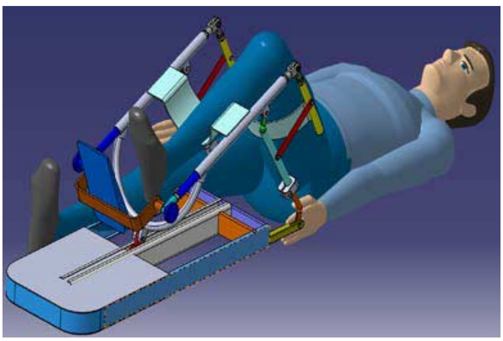
The main causes that can cause a disability of this type are due to infections, accidents, various diseases, lack of cerebral oxygenation or occupational hazard, among others. However there are different ways to classify disabilities, one is according to the degree of disability and the other for its duration [1].Rehabilitation of the human body is one of the fundamental challenges of the century XXI medicine's research. The integration of mechatronics and the development of intelligent mechanisms envisage results that could improve the life quality of the people that have reduced mobility in their extremities. The state-of-the-art in computing, sensing, and control has intensely advanced, but the design and implementation of exoskeletons and active orthoses still requires great efforts, and this requires compliance many principles as the rehabilitation should be intensive, task-based, should require active participation, [2]. One example of the apparatus that can be achieved is the robotic task-practice system that helps patients to recover the reach and grasp skills [3]. Restoring the walking capacity by rehabilitation of the lower limb, a person is able to be more productive along with the psychological advantages that this fact involves. Passive orthoses have been used to support gait in cases of spinal cord injured where muscle fatigue and the joint torques are factors that hinder the total rehabilitation of individuals

[1] Corresponding author.

[4]. In some cases, the spinal cord injured can be restored by a continuous repetition of the natural movements in the affected extremities. This effect is accomplished as a reconnection of the motor neuron generated by the reinforcement learning. The repeatability and the accuracy are the most important factors to be considered when designing active orthosis.

1. Mechanical design

The robotic system is designed to help in the rehabilitation of the knee. The knee requires rehabilitation after a chirurgical intervention, after stroke or any other problem. To simulate the natural movement of the knee, the active orthoses has to move the leg, in the sagittal plane, performing flexion and extension. Figure 1 shows an sketch and a robotic system´s mechanical design. The proposed mechanical design is modeled as a rod-crank-slider mechanism. Transmission of power is done through a worm screw coupled to a DC motor of 440 rpm. We want a light robotic system, for this reason the materials chosen for its development are aluminum and plastic. The reduction in weight allows the portability of the robotic system, ideal for medical purposes in hospitals and at patient's home. The design considers that the equipment is adjustable to different sizes of patients, which in the Mexican population, have the following characteristics: weight of 80 kg (maximum of 100Kg) and height 1.30m to 1.80 m.



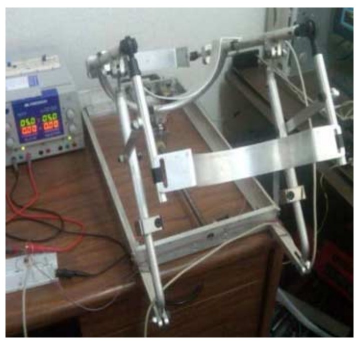


Figure 1: sketch and a Mechanical design.

To obtain the kinematic model, we write the equation for the vector loop position presented in figure 2. The kinematic model expressed in Eq. (1), represents the changes in the position of the four elements:

$$ae^{j\theta_2} - be^{j\theta_3} - ce^{j\theta_4} - de^{j\theta_1} = 0 \quad (1)$$

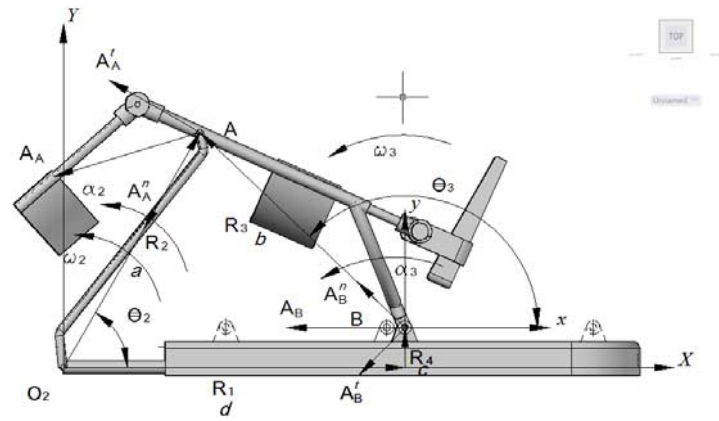


Figure 2. Cinematic model.

Using the Euler form and taking θ_1=0° and θ_4=90° the equation (4) relates the angles θ_2 and θ_3 with length d. In this we consider *a, b* and *c* constant. Length *a* is a specific measure for each patient, d is measured for the encoder incremental.

As this is an active orthoses, care must be taken in the detection of force in flexion and extention, as the robotic system and biomecanic are interacting at the same time can conflict, so you have to sense the reaction force leg and gradually stop the movement if limit is reached, according to the anti-collision algorithm to avoid the most important: prevention of damage or pain to the patient [4], as seen in the figure 2, only whit the difference of control is on the force, this algorithm reduce gradually the voltage when the force is incremented, this looks more graphic in on the fuzzy surface of control figure 8.

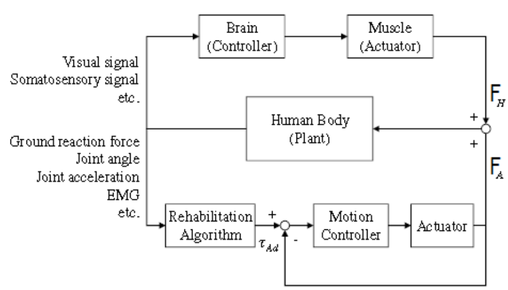


Figure 3. Control of force in orthoses active.

Two strain gauges are mounted in the movable toehold (see figure 4). The strain gauges are implemented due to their low cost, compared with the torque sensor. The sensor force implementation in flexion or extension of the biomechanical coupling has the advantage of limiting the force applied to the leg in order to avoid causing any harm or pain to the patient who has spasticity or other problems.

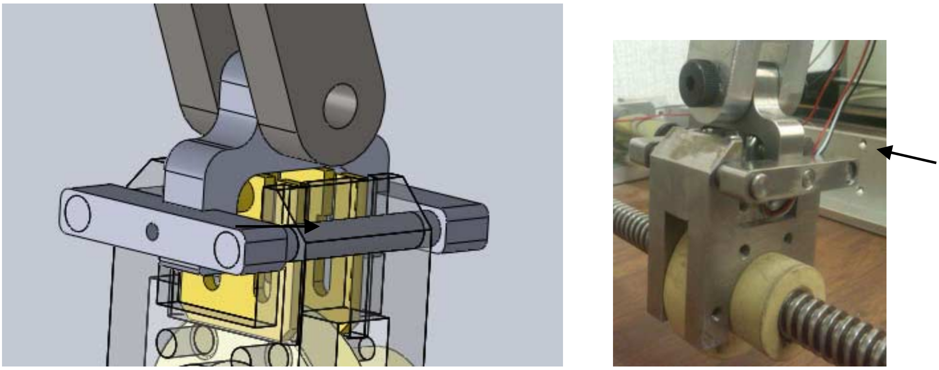


Figure 4. Two Strain Gauges for flexion and extension

However, the gauges sensed force must be compensated due to the weight produced by the leg, and should be calculated in the toehold, but this depends on the inclination's angle on the point support position.

Once compensated the leg weight we can get the effort made by the biomechanical coupling, which is we need to control, (See the vector dynamic model in figure 5).

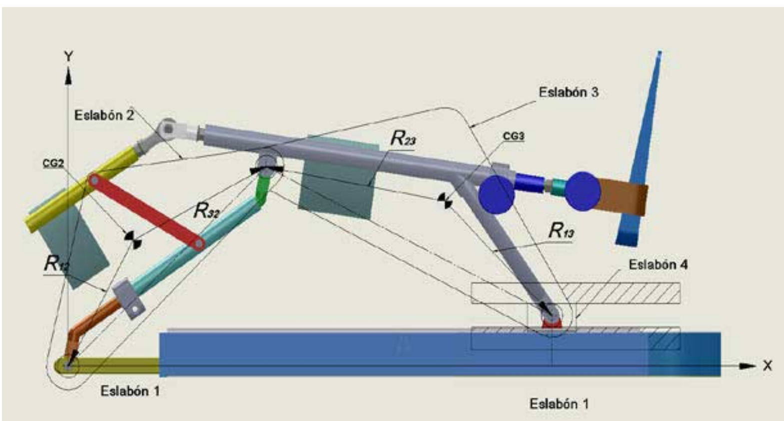


Figure 5. Dynamic Model equivalent.

In figure 4 we are the dynamic model, with the mass centre of body an mechanism considering the thigh and calf are represented as straight cylinders and foot as a right pyramid, we can determine the forces applied to the straing gauge, based on the analysis of free body diagram of link 4, and we obtain:

$$\sum F_x = ma_x \qquad F_{14x} - F_{43x} + F_{Px} = m_4 a_{G4x} \quad (2)$$

$$\sum F_y = ma_y \qquad F_{14y} - F_{43y} + F_{Py} = m_4 a_{G4y} \quad (3)$$

$$\left(R_{14x}F_{14y} - R_{14y}F_{14x}\right) - \left(R_{34x}F_{43y} - R_{34y}F_{43x}\right) + \left(R_{Px}F_{Py} - R_{Py}F_{Px}\right) = I_{G4}\alpha_4 \quad (4)$$

In Equation (2), (3) and (4) we obtain the forces and torques, but we want to analyze is (2), the ΣFx is the force that must be compensated to know the difference of strength between the biomechanical system and device.

2. Electronic control

The hardware used is the Quanser Q4 with the software Matlab and Simulink and xPC Target. The analog input of strain gauges is amplified and offset by an Amplifier Operational, using a single source in order to avoid increase the bipolar sources costs, therefore the voltage obtained from Wheatstone's bridge and the amplifier is 3.3V and based on the force applied to extension or flexion this voltage rises or falls, thereby it can determine if the applied force is an extension or a flexion, this signal is input to the analog channel zero of the card Q4, the other parameter introduced is an incremental encoder on channel zero, This control is implemented in simulink whith state flow and the fuzzy control, the output is the analog output zero with UPM-1503 amplifier and then to the DC motor with belt drive.

3. Fuzzy Control

Fuzzy control was chosen because its system nonlinearities and its interaction with nature in biomechanical systems. This controller is a Mamdani type with five Gaussians membership functions (view figure 6)

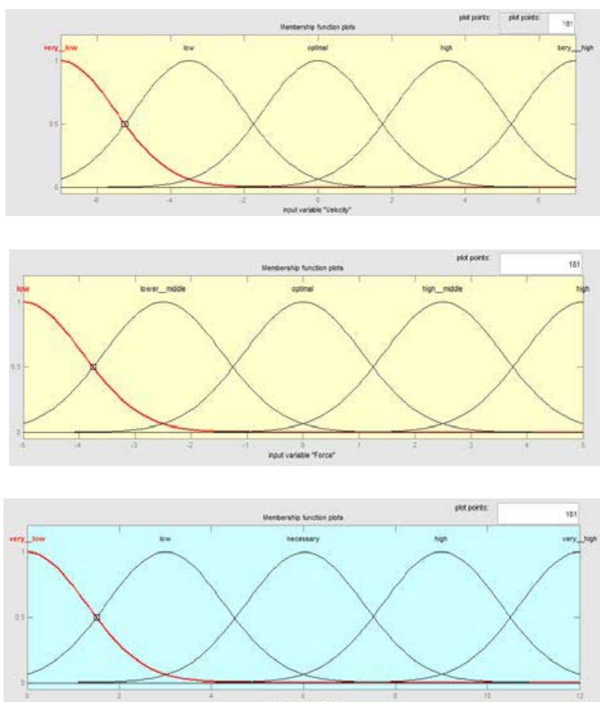


Figure 6. Membership functions Mamdani.

Finally, two input variables: speed and force from an output variable: the voltage applied to the motor, controlled by the rules of inferences, see figure 6. And is designed so that the speed is constant and the force of extension or flexion this bounded to prevent damage to the patient, and you see the surface control in figure 7.

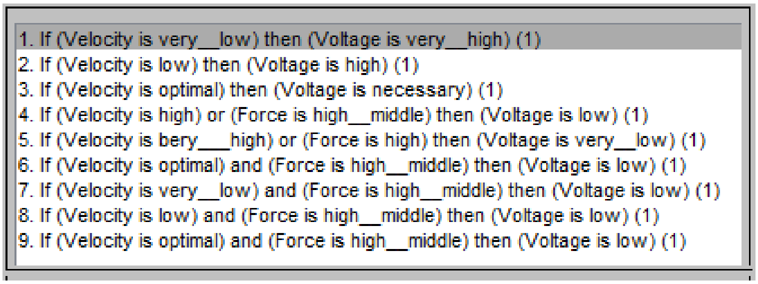

1. If (Velocity is very__low) then (Voltage is very__high) (1)
2. If (Velocity is low) then (Voltage is high) (1)
3. If (Velocity is optimal) then (Voltage is necessary) (1)
4. If (Velocity is high) or (Force is high__middle) then (Voltage is low) (1)
5. If (Velocity is bery___high) or (Force is high) then (Voltage is very__low) (1)
6. If (Velocity is optimal) and (Force is high__middle) then (Voltage is low) (1)
7. If (Velocity is very__low) and (Force is high__middle) then (Voltage is low) (1)
8. If (Velocity is low) and (Force is high__middle) then (Voltage is low) (1)
9. If (Velocity is optimal) and (Force is high__middle) then (Voltage is low) (1)

Figure 7. Inference Rules.

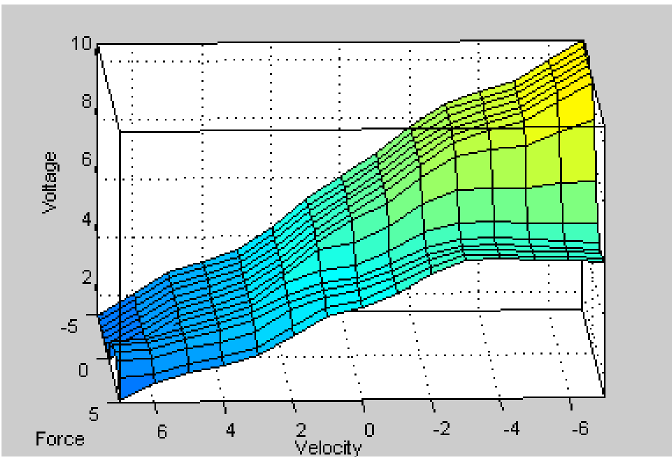


Figure 8. Control Surface.

4. Conclusions

In this paper we proposed a new orthoses multi proposes, trying to maintainer low cost of equipment, This equipment is designed for rehabilitation of many problems, including post operating, after stroke, muscle weakness. But this is not yet tested in patients; this is the next step for probe their efficiency and makes appropriate improvements. This work will be done in "Centro de Rehabilitación y Educación Especial" (CREE Toluca), under supervision of medical doctors and therapists, making a medical protocol for his equipment.

5. Future Work

This equipment will be used to made biofeedback in patients who have neuromotor problems, using EMG signals to regenerate the necessary neural connections and detect the intention of movement.

References

[1] INEGI, *Discapacidad Motriz*, en Las personas con discapacidad en México: una visión censal. Aguascalientes, Ags.: INEGI, 2004, pp. 79-94.

[2] A.M.. Dollar, H. Herr, "Lower Extremity Exoskeletons and Active Orthoses: Challenges and State-of-the-Art," Robotics, IEEE Transactions on , vol.24, no.1, pp.144-158, Feb. 2008

[3] C. Younggeun, J. Gordon, K. Duckho, N. Schweighofer, "An Adaptive Automated Robotic Task-Practice System for Rehabilitation of Arm Functions After Stroke," *Robotics, IEEE Transactions on* , vol.25, no.3, pp.556-568, June 2009

[4] R.J. Farris, H.A. Quintero, T.J. Withrow, M. Goldfarb, "Design of a joint-coupled orthosis for FES-aided gait," Rehabilitation Robotics, 2009. ICORR 2009. IEEE International Conference on, vol., no., pp.246-252, 23-26 June 2009

[5] B. Joonbum, K. Kyoungchul, and T. Masayoshi, "Control Algorithms for Prevention of Impacts in Rehabilitation Systems" 2011 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM2011), July 3-7, 2011

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-319

Ocular and Craniofacial Trauma Treatment Training System: Overview & Eyelid Laceration Module

Mark P. OTTENSMEYER[a,b,1], Bummo AHN[a,b], John CHO MOORE[a], Evgheni MUNTEANU[a], Stefano IORINO[a], Gianluca DE NOVI[a,b], Ryan BARDSLEY[a], Rajesh SHAH[a], Paul NEUMANN[c], Dan LANGSTON[d]
[a] *Simulation Group, Department of Imaging, Massachusetts General Hospital*
[b] *Harvard Medical School*
[c] *NinePoint Medical*
[d] *Dan Langston Studio*

Abstract. Eye trauma in combat-related injuries can be devastating. Body armor improvements increase survivability from many injuries, however the face remains relatively unprotected, especially from IED blast and fragments. The number of physicians with significant eye trauma expertise is limited due to lack of teaching opportunities and evolution of surgical techniques away from microsuturing. To provide a non-animal, non-cadaver training system, a hybrid physical/virtual eye trauma simulation system for surgical technique learning and evaluation is being developed. It includes a silicone/polymer head with replaceable eye trauma modules with realistic anatomy. Position/orientation and grasp force/position of surgical instruments are measured, and real time scoring and feedback methods are under development. These elements are integrated with teaching materials presented on-screen and during training scenarios via augmented reality projection. The initial efforts focus on a lacerated eye lid scenario. Developments towards additional trauma modules with animatronic features are discussed.

Keywords. Eye trauma, surgical gesture measurement, augmented reality

1. Introduction

1.1. Motivation

Craniofacial and ocular injuries have been elements of up to 29% of all battlefield injuries suffered by US, UK and allied forces in recent conflicts in Iraq and Afghanistan. While improvements in armor systems have reduced fatality due to injuries to the torso, the face and head remain relatively vulnerable and casualties present with more injuries to the eyes and face than proportional relative to body surface area. Emphasis on the use of eye protection reduced the prevalence of eye injuries from Operation Desert Storm in 1991 (13%) to Operations Iraqi Freedom and Enduring Freedom in 2001-2005 (6%)[1]. Strict compliance with eye protection usage can reduce eye injuries to 0.5% in some cases [2], however this is typically not

[1] Corresponding Author.

achieved, and rates vary [3]. The eye is sensitive to blast injuries and without proper eye protection, many small fragments that might not penetrate skin, are likely to do damage to the eye and surrounding tissues.

At the same time as soldiers are surviving formerly fatal injuries, but with disfiguring and disabling eye/face wounds, the skills required to address them are not as broadly available as they once were. Microsuturing has been superseded by sutureless methods for certain procedures, so the relevant skills are not as frequently used. Trauma cases are typically referred to trauma centers, so fewer ophthalmologists have exposure to relevant cases, and the uniqueness of each trauma case makes the learning process a lengthy one. In addition, there are financial disincentives to specialize in trauma vs. other ophthalmology specialties, involving reimbursement levels for re-operations for a given case.

Alternative training modes include use of cadaver tissues [4], animal eyes and tissues [5], and simulated anatomy [6]. Cadaver and animal eye tissue typically has different mechanical behavior from live tissue. Further, there is a strong push for military medical training to reduce or eliminate the use of tissue for training purposes. With both tissue training and the use of simulated anatomy, there is a need for senior oversight and guidance, which is an expensive use of a skilled surgeon's time which needs to be balanced against time spent in the OR.

There is a need, therefore, to provide more and better opportunities for self-directed, repeated and repeatable, diverse and relevant synthetic training options. In support of this, we are developing an eye and facial trauma simulation system, integrating not only a physical mannequin, but the enhancements provided by multiple modes of content and guidance presentation, multiple modes of sensing, and the capability for presenting a variety of trauma cases using the same base platform.

1.2. Previous Research

Existing eye surgery simulation systems tend towards opposite ends of the virtual vs. fully-physical spectrum. In the virtual space, there are a variety of simulators focused on non-trauma scenarios, including cataract surgery [7][8]. A recent entrant to the virtual arena is [9]. Primary drawbacks of fully virtual systems is that they either dispense with tactile and force feedback components, in a domain where careful control of applied forces is crucial, or when force feedback is included, it relies on typical commercial haptic interfaces, which are difficult to integrate within the tight confines around the eye. Further, multiple instruments are used during the course of a typical procedure (multiple forceps, needle holders, scissors, accessories such as surgical sponges), which must be "faked" if a common interface handle is used, or detachable connectors designed to allow swapping of instruments.

In the domain of the fully physical are a number of commercial products [6][10]. These have varying levels of accuracy, typically matched closely to the needs for training particular procedures. With careful choice of materials used in fabrication, they can provide accurate tissue behavior, however the automatic performance measurement possibilities of the virtual are unavailable. Performance evaluation depends on the oversight of proctors and senior physicians, and when used in a self-directed fashion, there would be minimal opportunity for immediate feedback or advice on improvement.

One notable effort to combine approaches is that of [11], in which an *ex vivo* human or animal eye was supported in a cup mounted on a force sensor, to evaluate loads applied to the eye. Sensing of instrument motions was not performed.

The general focus of our research group is to combine the most useful features of the physical with the complementary features of the virtual when developing medical training simulators: physical realism, so force feedback is obtained passively from physical objects, combined with tracking and responsiveness capabilities of the virtual – measurement of instrument trajectories, sensing of interventions and automatic provision of real-time feedback, evaluation and guidance.

One example of this work was a collapsed lung response simulator [12][13], which tracked trajectories of "chest tubes" and "chest darts", causing the physical mannequin to respond appropriately to insertion, automatically detecting success or failure and providing immediate feedback to the trainee. Similarly, we developed an early laparoscopy interface which quantitatively measured physical instrument motions while performing basic training exercises, yielding performance scores relative to an aggregate of expert performance results [14] and most recently, a full-body trauma victim mannequin for use in training for severe extremity hemorrhage, which housed an autonomous control unit inside the body of a physical mannequin [15].

1.3. Eye lid anatomy

The system currently focuses on learning techniques relevant to the repair of eye lid lacerations. The eye lid (Figure 1) is made up of the layers of skin, the orbicularis oculi muscle, the orbital septum (not shown), tarsal plate and the palpebral conjunctiva. The muscle layer extends close to the margin of the lid and is a visible dividing line between the inner and outer layers of the lid, known as the grey line. Towards the inner corner of the eye (medial canthus) are the lacrimal ducts, which drain tears away from the eye and into the nasal sinus. Additional features are the eyelashes, growing out in front of the grey line, and the openings of the meibomian glands, behind the grey line.

For the purposes of the initial simulation, we consider lacerations through the margin of the lid, in regions away from the lacrimal ducts, so the users will be primarily be concerned with alignment and approximation of the tarsal plate, skin surface and the margin of the eyelid.

In future versions, we will be including lacrimal ducts, which require more advanced techniques to repair, and the lateral canthal tendons, which are dissected as part of a procedure to relieve pressure from bleeding in the eye socket often due to blunt trauma. Similarly, while the current globe is in a fixed position, future versions will have both passive and active mobility, as described below.

1.4. Eye globe anatomy

Future versions of the simulator will focus on repair of injuries to the anterior anatomy of the eye globe, primarily partial and full-thickness lacerations through the cornea and sclera (white). The major structures of the anterior hemisphere of the eye (Figure 1) are the clear cornea, the anterior chamber, filled with watery aqueous humor, the iris, which contracts and dilates under the control of the ciliary muscle, the lens and capsule, the sclera (white), which is the structural wall of the eye, the choroid layer inside the sclera, and taking up most of the volume of the eye, the jelly-like vitreous humor.

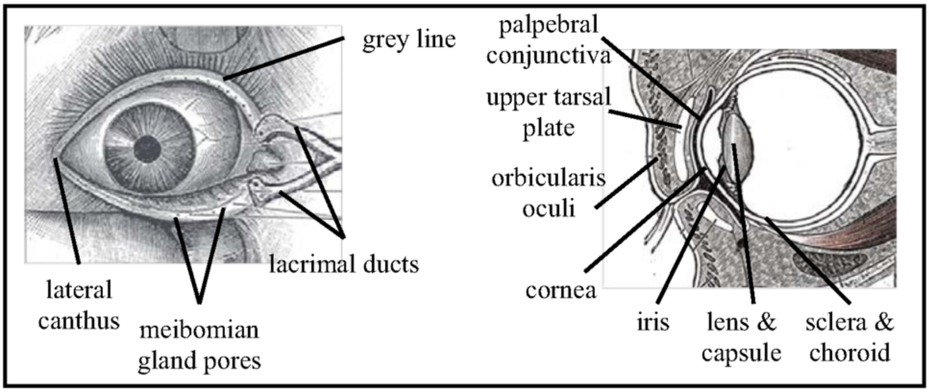


Figure 1: (left) External eye showing lacrimal ducts. (right) Cross section of the eye lid and globe. Adapted from http://en.wikipedia.org/wiki/File:Gray888.png, http://en.wikipedia.org/wiki/File:Gray896.png, http://en.wikipedia.org/wiki/File:Gray1205.png

2. System Overview

Out system is a stand-alone eye/face trauma simulator with augmented reality teaching/evaluation content. It consists of a mannequin head/neck with a replaceable eyelid and eye globe module. Titanium ophthalmic surgical instruments are tracked in six degrees of freedom (DOF) and grasping action and grasping force are measured. In the current version, the AR (augmented reality) hardware is comprised of a Microsoft Kinect and a pico-projector. A lighting system, stand, tray, electronics for the surgical instruments, and a host PC computer complete the system (Figure 2).

The physical components of the system are integrated as part of a learning system, which will keep track of user identity and performance, present teaching content relevant to particular training scenarios, monitor performance of surgical exercises during the particular scenario, and present feedback during and after the scenario, including performance scores.

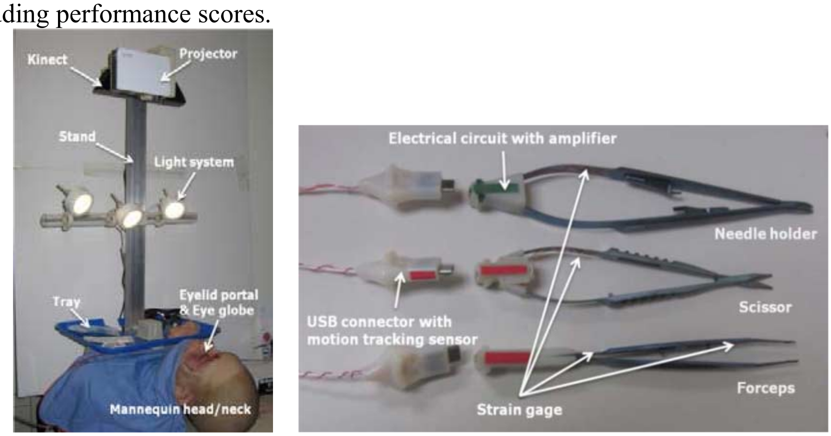


Figure 2: Stand-alone ocular and craniofacial trauma treatment training system: (left) Mannequin head/neck with eyelid portal and eye globe, and hardware architecture for augmented reality (AR) (right) Surgical instruments, hot-swappable connectors and signal conditioning circuitry.

2.1. Anatomical Modeling and Fabrication

At present, the major anatomical components of the system are the mannequin head/neck, multiple copies of a module including the tissue of the eyelids, brow, and upper cheek, and the eye globe itself. We will briefly describe the fabrication techniques for each of these.

2.1.1. Head form – skin and core

With the participation of a volunteer, adult male, we created a life cast (Figure 3) using standard techniques and materials (skin safe silicone, plaster wrap) to create the external impression. This mold was removed from the model and used to cast a clay first copy of the model's head. This clay positive was then textured and retouched by hand and molded again to create the final mold for the head. A plastic positive was cast from the mold and laser scanned, so that a graphical/CAD model with identical dimensions would be available for the simulation. A second positive, made of molding clay, was shaved down to approximate the shape of the underlying skull based on standard thicknesses of soft tissue over the skull. In the eye region, a cavity was created, offsetting the surface, so that replaceable modules fitting within this space could be subsequently fabricated. Finally, the external impression and the internal core were reassembled, with an additional insert occupying the space of the eye portal. A silicone skin, with the correct external and internal contours and eye portal space was cast. The CAD model was used to create a rapid prototyped core for use in mounting the replaceable components and the initial mounting of the tracking system field emitter.

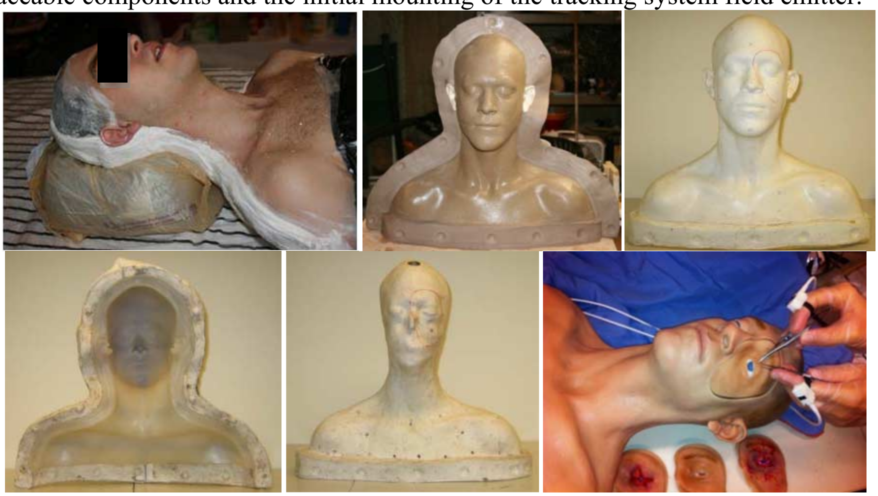


Figure 3: Molding sequence for prototype head-form. Life casting. Clay duplicate. Plastic positive with 3D scanning fiducial marks. Silicone production mold. Plaster core with 3D scanning fiducial marks, air holes for skin casting. Final molded silicone skin, with pigmentation and replaceable trauma modules.

2.1.2. Ocular adnexa

The opening in the skin layer surrounds the trauma module inserts, which include the eyelids and surrounding tissues, both cast in silicone. The eyelids are complex structures that include multiple tissue layers and a variety of complex external anatomical features, however, most significant for our first trauma scenario are the skin and a stiffer, underlying structure called the tarsal plate (Figure 4).

The tarsal plate was designed by first laser scanning the positive plastic casting of the eye portal. The scanned 3D model was then imported into SolidWorks CAD software and the surface model of the eyelid was thickened to create the mold for the tarsal plate. The mold was then printed on an Objet Eden 260V 3D printer (Objet Ltd., USA), capable of printing smooth and clear surfaced parts. During fabrication, a piece of nylon mesh is sandwiched between two tarsal plate mold parts and silicone is injected into the mold. The mesh helps to limit the tendency for punctures and cuts in the silicone to propagate. The part is removed from the mold and placed in the portal mold where silicone is again injected to create the eyelid with tarsal plate.

Figure 4: Molding of eye portal - production mold, early single material eyelid models; eye lids with "tarsal plates" (back and front after suturing during data collection scenario).

2.1.3. Eye globe

The current version of the globe (Figure 5) is similarly fabricated in various grades of silicone. A nylon mesh is wrapped around a two-piece inner mold core. An external core is aligned and attached to the inner, and white pigmented silicone is injected, forming the sclera (white of the eye). A stem at the back of the eye allows for removal of the inner core. Following removal, a molded, clear silicone lens, pigmented iris and clear cornea are bonded in place in the anterior segment of the eye globe, and a sealing plug to fill the stem at the posterior pole of the globe. The plug includes an access hole, filled with silicone adhesive, which is self-sealing after puncture. A syringe is used to evacuate air from the globe and replace it with a material representing the vitreous humor, a jelly-like material that fills the human eye behind the lens. We have employed a number of different materials for this purpose – polyvinyl acetate/borax "slime", egg white and Astroglide lubricant, all of which exhibit a "stringiness" when an instrument contacts the material and is withdrawn, similar to the vitreous humor.

Figure 5: Eye globe fabrication sequence and modeling, example of "stringiness" of materials similar to vitreous humor

2.2. Surgical Instruments, Tracking System and Sensors

2.2.1. Surgical Instruments

In the repair of lid (and later, globe) lacerations, a variety of surgical instruments are necessary. For our simulator, we provide the user with Castroviejo suturing forceps, a Castroviejo curved needle holder and curved Westcott tenotomy scissors (Figure 2 (right)). Additional instruments are available for future procedures, including finer tipped tools for corneo-scleral injuries. During scenarios, information on the instruments' spatial positions, open/close position and tip force, and surgeon's hand motion is required to grade the surgical techniques of users. To detect the instrument motion, we use miniature, magnetic position tracking sensors (Model 180 6DOF sensor, Ascension Technology, USA) and a tracking system (3D Guidance trakSTAR, Ascension Technology, USA). To be compatible with the magnetic tracking system, all of the instruments are fabricated from titanium. The instruments are enhanced with strain gages (CEA-06-062UW-350, Micro-Measurements, USA), which are placed at appropriate sites based on finite element analysis (SolidWorks 2012), to detect the open/close position of the instrument jaws and to measure the force applied between the jaws of the forceps to the tissue (Figure 6). To convert the gage output signals into useful data for our system, we developed miniature circuit boards (actual size 10mm x 16mm) with a microcontroller (Microchip PIC18F2450) and electrical components that are attached to each instrument. This is achieved using housings fabricated using the 3D printer. The microcontroller supports USB communication and has a 10-bit analog to digital converter with 10 different inputs. The board currently communicates with the PC using the USB bus. For each instrument, there are three analog output signals: one is the tool identification signal; the other two are the opening and force states of the tool. The tracking sensors are expensive and are tethered to their interface electronics, so to minimize cost and the number of cables that the surgeon must contend with, we developed interchangeable connectors between the instruments and DAQ/interface board, which also hold and align the tracking sensor with the instruments. Figure 7 shows the pad and component layout of the DAQ/interface board. To the left are three micro-USB connectors that mate with the instrument mini boards. Top and bottom are headers which allow connection to additional boards. To the right are (top to bottom), power supply, USB connection, and a chip programming interface connector. This board can also be used for future motor/pump control and additional sensor DAQ.

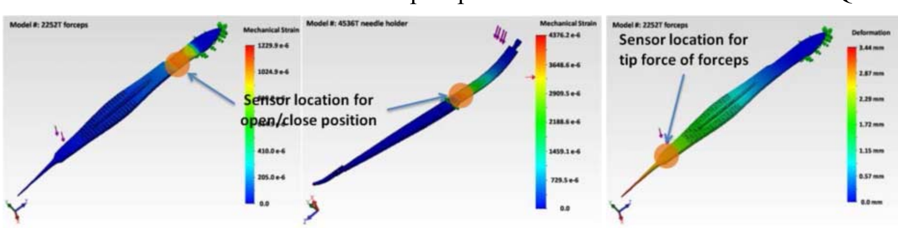


Figure 6: Finite element analysis results for half-model of each instrument: (left) the maximum strain sites at the rear of the instruments are the best place to detect the open/close position of instruments. (right) To measure the tip force we attached the strain gage to the flat inner surface as close as possible to the tip.

2.2.2. AR hardware

To detect hand motions around and display graphical overlays onto the mannequin head, we are using the Microsoft Kinect motion tracking system and a "pico" projector

(Vivitek Qumi). We also developed a portable framing system to ensure precise and reliable registration between the display and tracking systems. The projector and Kinect can be adjusted and locked into the best position. The head of the mannequin can extend over the edge of the table for access similar to a patient's head extending past the end of an operating table. The mannequin is rigidly fixed to the stand, avoiding vibration. A tray provides a site for a single-handed instrument exchange fixture. The lighting system was implemented on the simulator stand to approximate the intensity of surgical lamps and to minimize shadows. All of the parts can be disassembled for ease of transportation; off-site data collection tests are planned.

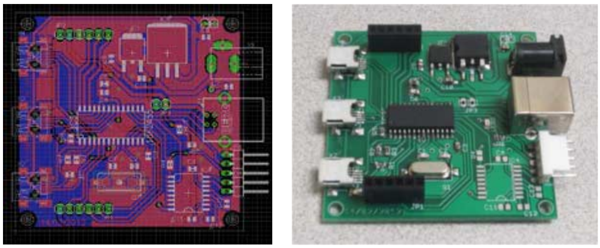


Figure 7: DAQ/interface board showing top-side component positions and board layout. To the left are three micro-USB connectors that mate with the instrument mini-boards. Top and bottom are headers which connect to additional boards. To the right are (top to bottom), power supply, USB connection, and a chip programming interface connector (for installing the firmware on the microcontroller).

2.2.3. Modular eye motion and pupil contraction systems

Future trauma scenarios will require additional functionality in the replaceable lid/globe modules, so we are developing eye motion and pupil contraction components. These motions are intended for first-responder scenarios, in which they can provide clues during diagnosis of the casualty in an emergency incident. As space is limited within a simulator's head, and noise from some types of motors may be distracting to the user, we are considering the use of shape memory alloy (SMA) wire actuation.

In one prototype, the eye motion component consists of two SMA linear actuators (MigaOne-15, MIGA Motor Company), a driver and motion limit detection circuit, a pivoting eye motion plate, ball-and-socket links, connectors, and supporters as shown in Figure 8 (left). The circular plate rotates in two degrees of freedom (pitch and yaw).

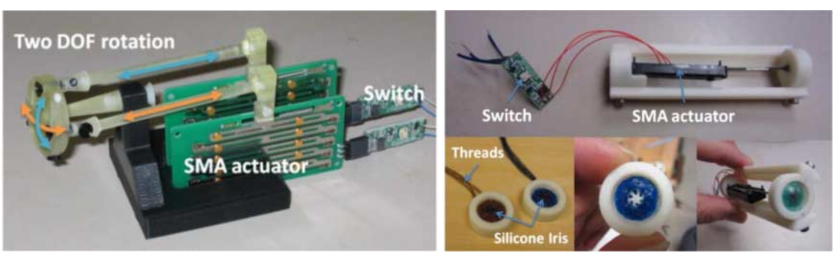


Figure 8: (left) 2DOF rotation eye motion and (right) pupil contraction components using SMA actuators.

In order to develop the pupil contraction component, we observed the range of motion of pupil contraction and dilation, and designed a contracting iris mechanism. Figure 8 (right) shows a silicone iris, supporting parts, fibers linking the iris to the actuator, and an SMA linear actuator (NanoMuscle NM706-Super, MIGA Motor Company). We sutured the inner margin of the silicone iris using thread, which passes through an array of holes. When the SMA actuator contracts, the fibers pull the inner iris margin towards the center of the pupil. When the actuator is shut off, the elastic iris returns to its normally dilated position.

2.3. Content and learning architecture

Our initial developments towards a content presentation and learning architecture is geared towards junior surgeons with limited prior oculoplastics and globe surgery experience, and more senior physicians who may not have frequent trauma patient exposure. We aim to tailor the system to military surgeons, who would benefit from pre-deployment ocular and periocular surgical planning and technique refresher courses.

The planned architecture will allow access to background content and teaching materials on multiple platforms. Smart-phones, personal computers and similar devices would be appropriate at this stage. The user's interaction will begin with a typical login, in which the system recognizes the identity and level of expertise of the user. This will be followed by content access, logged by the system to allow bypass of such background when access to the simulator is available. In the current context, this would include relevant anatomy, suturing technique, operative strategy and surgical tips.

As mentioned, our first scenario is marginal lid laceration, although the current module design also supports non-marginal injury presentation. Injuries are generated by manually cutting through the skin and tarsal plate structures of the module in the location desired by the trainer/course proctor. At the outset of a simulator session, patient history, lab tests, pre-op imaging work and similar information will be presented to the user. Once these items have been reviewed, the simulator is initialized, the user takes hold of selected instruments, and the trauma module is unveiled.

The user, with instrument and hand-tracking enabled, will perform an examination of the injury, verify the absence of foreign bodies and additional injuries (in the initial scenario), and begin to conduct the repair. Tracking and sensor data are recorded to disk for post-processing at present, however we are developing online gesture recognition techniques to allow intra-operative feedback, such as suggestions when a non-standard technique appears to be used.

During the scenario, the user will be able to use two or three instruments simultaneously (surgeons often palm one instrument and pick up an additional one in the same hand before returning it to the tray – cutting of excess suture is one typical example, in which forceps or a needle holder are not put down first), and the tracking system will be able to identify which instrument is in use at a given time. For procedures requiring four or more instruments, the user will hot-swap the connector on the end of an instrument no longer in use, to the newly necessary device. The connector/DAQ system automatically detects its identity.

Following scenario completion, performance statistics will be presented to the user. At present, we can calculate task completion time and path length, as well as timing and sequence of instrument usage. We are developing techniques to identify distinct surgical gestures (e.g. knot tying, suture cutting, tissue grasping), so that finer grained, more useful feedback can be determined and presented to the user.

2.4. Data Acquisition and Event-based Gesture Segmentation Software

At present, software developments have primarily focused on data acquisition and instrument interface tools. The application has the flexibility to use multiple user interface input devices (e.g. mouse, SpaceNavigator, trackball). It records position tracking streams from up to eight different magnetic position tracking sensors, from up to two strain gages on each instrument, and the identity of each instrument. In addition, Kinect outputs, representing hand motions (vs. instrument motions) are being integrated into the system; hand tracking can be used to detect actions like picking up a specific type of suture. All of these data are recorded as separate streams and saved in XML format. Simultaneously, a 3D rendering of the instruments and head form is displayed to the user (Figure 9). The system is easily expandable to track more instruments, the AR loupes when available, and head motion relative to the base once a flexible neck is implemented. The 3D rendering is important as it enables navigation around the scene, useful during the post processing playback as we manually observe task performance and work towards an automated gesture recognition algorithm.

Thresholding functions are built into the DAQ software, allowing preliminary segmentation using features such as closing and opening of scissors (indicating cutting), and the simultaneous closure of forceps jaws and the increase of measured force, suggesting that tissue has been grasped (vs. closing in free-space). Similarly, the entry into or exit from pre-determined regions of interest (RoI's) are recognized as meaningful state transitions. We refer to this as "event-based surgical gesture recognition," and propose that the subsequent recognition of sequences of events will be useful for identifying differences in performance between users. This process will be described in detail in a separate paper.

Figure 9: (left) Developer's interface showing data streams and graphical display of instrument positions. (right) Data stream examples showing discrete events to be segmented using software under development.

3. Data Acquisition Trials

Three of our subject matter expert oculoplastics surgeons have performed marginal laceration repair on our modules, in each case repairing cuts in nearly identical locations and of identical geometry (perpendicular, full-thickness upper lid laceration at lid mid-line, with extension of the laceration parallel to lid margin). They were asked to perform the repair, aligning and connecting the tarsal plate and overlying tissues, while making only commentary that would typically be employed in the OR (rather than in a teaching case, in which additional discussion would be necessary). Following the repair exercise, they performed a series of 15 partial tasks, including a variety of knot tying and suturing techniques, to help us collect data in support of detecting the discrete events within the extended full scenario testing.

In all of these tasks, instrument tracking was recorded, as was video-taping of the operative field, and an external observer's synchronized, manual event detection (time-stamped key presses). Analysis of these data is underway; the analysis techniques are described in an accompanying paper [16]. The results will be used to automate gesture recognition during external user data collection trials in the next few months (these studies will be covered under appropriate human subjects research approvals from our IRB and that of the funding sponsor). Those tests will include both junior and senior surgeons, and we expect to find distinguishing characteristics between them that will allow us to provide meaningful and timely scoring and recommendation to future users.

4. Results

At present, we have a functional system that is in use by our subject matter experts for initial testing purposes. It is capable of:

- Real-time on-screen display of surgical motions with animated instruments
- Display of all trajectory, force, grasp position values in real time
- Detection of thresholds to identify grasp/release

We have identified, through the experts' usage experience, a series of improvements to the eye lid structure (addition of tarsal plate to original, single layer lid, corrections to the material properties of the tarsal plate and skin silicone selection). Similarly, the instrument connectors and cable management have improved, with additional sensor attachment stability, strengthened mating between the connector housing and the instruments, and modifications to the Velcro attachments we use to attach sensor wires to the user's wrists (and out of the way of the instruments).

Most importantly, initial data collections are allowing us to proceed towards the development of gesture segmentation algorithms and scoring methods.

Informal feedback from our surgeon experts suggest that we are close to acceptably realistic tissue stiffness, both around the lid region and around the rest of the head form. Similarly, they approve of the globe geometry, layer thicknesses, cornea and lens clarity, and properties of the vitreous humor.

5. Future work

We are progressing towards a complete system on a number of different fronts.

The eyelid anatomy continues to evolve as we implement additional realism. Meibomian gland pores, adjustments to layer thickness, and the addition of the lacrimal ducts and canthal ligaments are under development, as are methods for including conjunctival and muscle layers.

Revisions to the signal acquisition and conditioning circuitry are under way – the current "mini-boards" exhibit some heating in the circuit that is noticeable to the user and may be affecting the gain and/or offset of the force and grasp position signals.

The pupil contraction and ocular motion components are being integrated into a modified version of the existing globe anatomy, and additional sensors to detect the patency of repairs to corneo-scleral lacerations will be added to the system. These sensors will also be useful to aid in the detection of canthotomy/cantholysis treatment in response to simulated retrobulbar hemorrhage.

Beyond improvements to the current system, the head form will be reconfigured to accept multiple eye portals (left and right, and different trauma states), as well as facial hemorrhage and fracture modules, and airway components, which will support the detection of artificial airway devices and surgical cricothyroidotomy.

6. Acknowledgements

We would like to express our grateful appreciation of our Subject Matter Experts: Dr. Robert Mazzoli, Dr. Suzanne Freitag, M.D., Dr. Daniel Lefebvre, Dr. Justin Kanoff, Dr. Maria Troulis, and Dr. Steven L. Dawson.

This material is based upon work supported by the U.S. Army Medical Research Acquisition Activity under Contract No. W81XWH-11-C-0095. The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Army Medical Research Acquisition Activity.

7. References

[1] J. Breeze, K. McVeigh, J.J. Lee, A.M. Monaghan, A.J. Gibbons, Management of maxillofacial wounds sustained by British service personnel in Afghanistan. *Int. J. Oral Maxillofac. Surg. 2011*;40:483–486.

[2] J. Breeze J, Horsfall I, Hepper A, J. Clasper, Face, neck, and eye protection: adapting body armour to counter the changing patterns of injuries on the battlefield. **Br J Oral Maxillofac Surg. 2011** Dec;49(8):602-6. Epub 2010 Oct 30.

[3] R.A. Mazzoli, Don't Shoot Your Eye Out. *Naval Operational Medical Lessons Learned Newsletter*, Sept. 2008.

[4] EM Helveston, Coffey AC. Surgical practice kit. Ophthalmic suture simulator. *Arch Ophthalmol. 1986* Jun;104(6):923-4.

[5] R. Leuschke, A. Bhandari, B. Sires, B. Hannaford, Low cost eye surgery simulator with skill assessment component, *Stud Health Technol Inform*. 2007; 125:286-91.

[6] A. Yeh, B.T. Chan-Kai, A.K. Lauer, Basic training module for vitreoretinal surgery and the Casey Eye Institute Vitrectomy Indices Tool for Skills Assessment. *Clin Ophthalmol. 2011*;5:1249-56. Epub 2011 Sep 1.

[7] D.J. Solverson, R.A. Mazzoli, W.R. Raymond, M.L. Nelson, E.A. Hansen, M.F. Torres, Bhandari A, Hartranft CD. Virtual reality simulation in acquiring and differentiating basic ophthalmic microsurgical skills. *Simul Healthc*. 2009 Summer;4(2):98-103.

[8] E. Dumortier, J. Dequidt, S. Cotin, C. Duriez, J. Allard, J. Rouland, Computer-Based Training System for Cataract Surgery. *Proc. Third International ICST Conference on Electronic Healthcare for the 21st Century*, 2010.

[9] Millersville Review Staff. Saving Soldiers' Sight. *Review: Millersville University Magazine*. 2011:126(1):2-7

[10] D.D. Carda, D. Van, T.W. Johan, 2012 Model human eye. United States. Eye Care and Cure Pte. Ltd (Singapore, SG). US Patent No.: 8137111,

[11] R. Leuschke, A. Bhandari, B. Sires, B. Hannaford, Low cost eye surgery simulator with skill assessment component. *Stud Health Technol Inform*. 2007;125:286-91.

[12] S.L. Dawson, M.P. Ottensmeyer, S.M. Cotin, P.F. Neumann, Medical procedure training system. Application #: 10/488,415. Publication #: US 2004/0234933 A1. Filing date: Mar 2, 2004.

[13] E.M. Ritter, M.W. Bowyer, Simulation for trauma and combat casualty care. *Minimally Invasive Therapy*. 2005; 14:4-5; 224–234.

[14] N. Stylopoulos, S. Cotin, S.K. Maithel, M. Ottensmeyer, P.G. Jackson, R.S. Bardsley, P.F. Neumann, D.W. Rattner, S.L. Dawson, Computer-enhanced laparoscopic training system (CELTS): bridging the gap. *Surg Endosc*. 2004 May;18(5):782-9.

[15] R.S. Bardsley, M.P. Ottensmeyer, J.T. Samosky, S.L. Dawson. Methods and Apparatus for Autonomous Casualty Simulation. Application #: 12/067,260. Publication # US 2008/0227073 A1. Filing Date: Sep 28, 2006.

[16] G. De Novi, R. Bardsley, R. Shah, M.P. Ottensmeyer, J.C. Moore, B. Ahn. Event-driven Surgical Gesture Segmentation and Task Recognition for Ocular Trauma Simulation. *Proc. IMIASH 2012 - 1st International Workshop on Intelligent Multimodal Interfaces Applied in Skills Transfer, Healthcare and Rehabilitation, Guanajuato, Mexico.* Accepted for publication.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-331

Capturing the rower performance on the SPRINT platform

Alessandro FILIPPESCHI[a], Paolo TRIPICCHIO[a], Massimo SATLER[a] and Emanuele RUFFALDI[a]
[a]*PERCRO laboratory- Scuola Superiore Sant'Anna*

Abstract. Capturing athletes performances with the purpose of skills training in the specific field of rowing sport is here presented, in particular The SPRINT multimodal system is introduced. This system is comprised of a mechanical reproduction of a rowing boat and of a virtual reality system with augmented feedback suited for novice and expert training. This paper details the implementation of an embedded acquisition system capable of measuring all the biomechanical data necessary for the rowing physical simulation increasing the performance with respect to the previous system.

Keywords. Multimodal system, skills training, embedded electronics

Introduction

In the last decades the number of simulators has exponentially increased. They are being used in many different fields such as industry, surgery, army, art and many others. Among them, sport is one of the last where simulators have taken place, from one hand because coaches and athletes are often skeptic about the simulator's capabilities of transferring skills; on the other hand, because the advantages brought by simulators are not as evident as in other application (e.g. flight simulators), where safety and economic issues encourage their development. When a simulator is developed for training purposes, like in sport applications, the designers has to decide which features of the real task must be replicated by the simulator to optimize training effects, under the constraints of the available resources(e.g. money, space, computational resources, etc.). Therefore, often, simulators of a specific application evolve during time, starting from a very basic device, which keeps only the main features of the real situation, and adding step by step new features as the knowledge about the real task and the available technology improve.

Training in sport is a complex argument which comprises many aspects of human behavior[1]. In particular rowing training deals both with the physiological and psychological behavior of the athlete[2]. Rowing coaches are required to address all these issues in order to make the trainee able to successfully cope with the race. Therefore, they need to gather information from the rowers performance in order to give them suitable advices. SPRINT is a multimodal rowing training system that collects in real-time these information, allowing both the coach and the athletes to monitor the ongoing performance.

After showing the rowing bases along with what is currently used for simulating outdoor rowing and for capturing rowers' performances, the paper will briefly describe the SPRINT system whereas it will focus on the embedded capturing components.

1. Professional Rowing

1.1. Rowing Bases

Rowing is an outdoor sport in which rowers propel a boat through the water. Rowers seat on a sliding seat and are fixed to the hull by means of foot-stretchers. Each rower transmits forces to the boat by means of one or two oars. Rowing with two oars per rower is called sculling, whereas rowing with one oar is called sweep rowing. The rower repeats cyclically the same sequence of movements, which is called *stroke* which, in turn, can be segmented in many phases. The most common segmentation in four phases (catch, drive, finish and recovery) is described as a first approach:

1. *Catch.* As the stroke begins, the rower is coiled forward on the sliding seat, with knees bent, arms outstretched. At the catch, the rower raises the hands to place the oar blade vertically into the water.
2. *Drive.* At the beginning of the drive, the rower keeps the coiled pose of the upper body and the legs do all the work. Then the back starts uncoiling propelling in turn the boat. In the last part of the drive arms begin their work drawing the oar blades through the water, while the back stops rotating backward when an angle of approximately 40° with respect to a vertical line has been reached.
3. *Finish.* In the end of the drive, the rower move his hands quickly towards his body, which by this time holds steady in the layback position. During the finish the oar handle is lowered, drawing the oar blade out of the water. At the same time, the rower feathers the blade that turns from vertical to horizontal.
4. *Recovery.* The oar remains out of the water as the rower begins recovery by moving his hands away from the body and past his knees. The body follows the hands and the sliding seat moves forward, with help from the feet and hips, until the knees are fully bent; the rower has already squared the blades and he is ready to raise his hands for the next catch.

The rowing stroke typically lasts between 1.2 and 4 second. In this time lapse rowers have to move practically all their limbs in a quick and coordinate manner, often when they are already fatigued. This means that some movements last only few tenths of milliseconds, therefore an accurate and timing reliable system for capturing rowers' performances is required. See [3] for further information

1.2. Rowing Simulators

Currently, the most diffused rowing simulators (Concept2 Ergometers, Morrisville, VT, USA) have a very simple kinematics: they consist of a sliding seat and a foot-stretcher similar to the boats' ones. Instead of the oars, whose motion is spherical, they provide users with a handle to be moved back and forth on a vertical plane. The handle is connected to a device that provides resistant force by means of a chain and a cable. The device that provides resistance is a fan mounted on a flywheel, during the drive phase most of the athlete's work is stored as kinetic energy in the flywheel, whereas this energy is wasted by the fan during the whole stroke. These simulators diffused because they are affordable for all rowing clubs and they allow for a quantitative evaluation of rowers. Indeed they are used by national rowing associations for ranking rowers and for selecting the country teams. These simulators give information about ongoing performance in terms of force profile, power output, strokes frequency and estimated boat speed. Since all these information are based only on fan's angular speed, boat speed estimation is rough and any information can be provided about technique correctness. Variant of this kind of simulator are the Waterrower (Warren, RI, USA) and Rowperfect[4], that aimed at improving force rendering by using the water instead of the air and by making the device slide under the rower as the hull does in outdoor rowing. Kinematic improvements are implemented in more recent simulators, such as Oartec Rower (Sidney, Australia) and Biorower (Wien, Austria), whereas the most complex simulators were developed for research purposes in the last years (e.g. ETH M3 rowing simulator [5]). SPRINT is comparable to the last as hardware complexity but it was developed for training purposes as it is presented in [6] and briefly summarized in section 2.

1.3. Capturing rowing performance

Rowers require many skills to win elite competitions. The most important skills are high muscular power, optimal aerobic and anaerobic capacity, efficient gesture, ability to perform the correct sequence of movements, ability to manage own energy stock when rowing at high pace and under pressure, ability to coordinate with teammates. Coaches and athletes have always sought for assessing performance in order to find the way to improve it. Most of the effort have been done for the physiological skills: from seventies the strongest rowing federations and clubs started measuring oxygen consumption, lactate and heart frequency to assess rowers' aerobic skills, their capabilities of carrying out effort in presence of lactic acid and establish an easy way for monitoring athletes' status during demanding training tasks and races. Little attention have been paid to technical and coordination skills, that are qualitatively assessed by inspection during the performance, or by means of videotapes analysis after the training session. In both cases there are not direct measures of the rowers' performance and therefore an immediate feedback based on quantitative information. Recent devices embedded on the boat allows for an accurate assessment of outdoor rowing, they are not diffused and they are mostly used for research purposes.

SPRINT aims at training professional rowers under both physiological and technical points of view, therefore it is required to be equipped with a capturing system able to collect in a fast and reliable way both physiological and kinetic information about ongoing performance. Moreover, since training is strongly enhanced when concurrent feedback are available, capturing has to be fast enough to allow data to be captured, processed and fed back suitably for training. The following sections will briefly present SPRINT and the sensing components embedded in it.

2. The SPRINT system

The SPRINT system is composed of a mechanical platform, a sensing system, a software system and devices for providing users with feedbacks. The mechanical platform allows to row as in outdoor rowing with almost the same kinematics and an accurate force rendering. The force rendering is given by an Energy Dissipating Device EDD composed of a flywheel and a fan mounted on. The sensing system is composed of a set of encoders and strain gage-based force sensors that are embedded in SPRINT whose acquisition is the focus of this paper. Moreover, external devices such as VICON for motion capture and cosmed k4 for oxygen consumption measurement can be integrated. The software system is composed of many parts ranging from the acquisition and processing of the signals, to the modeling of the rowing task (shown in [7]) and the management of the feedback to be sent to the users. Finally, many feedback devices are available: vibrating motor equipped belts allow for providing vibrotactile feedback [8]; speaker are available for audio feedback whereas visual cues may be provided by means of an LCD screen (see Figure 1) or by putting the whole system in a cave [9]

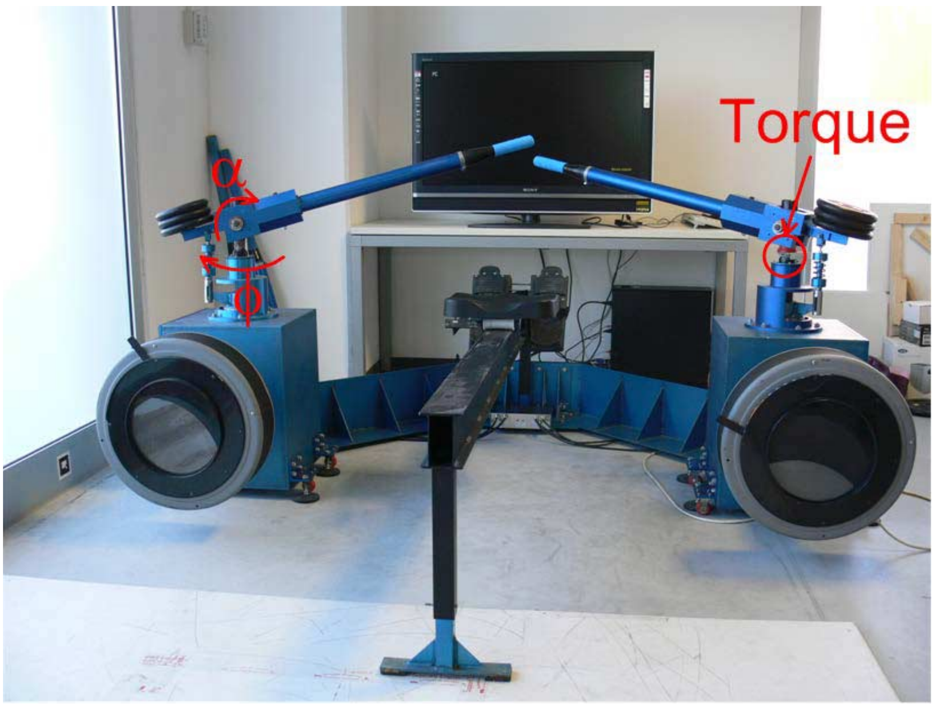


Figure 1 The SPRINT system in a configuration with an LCD display

3. SPRINT sensing system

The electronics embedded in the system before the enhancement described in the following, is shown in[10]. It was composed of two Microchip PIC16F887 Microcontrollers per oar, they were communicating each other via SPI in a master slave configuration. The master then sent signals via USB to the PC. The main issues with that device were the low sampling frequency (60 Hz) that could be obtained to have signal with sufficient resolution, and the synchronization of signal from the two oars.

The TMS320C2000 Microcontroller of Texas Instruments was chosen to substitute the several microcontrollers that were part of the rowing system. This device outperforms the previous as calculation speed, number of I/O ports, and, for analog signals, ADC resolution. Therefore, a reduction of latencies of the several sensorial components present on the platform is expected as well as an improvement the sample frequency in order to get a better resolution of the system signals.

Figure 2 The embedded electronic board and its connection with the sensor on the platform

3.1. Variables

The variables to be captured for assessing rower's performance are the oar rotations and the exerted forces. Oar rotations are described by the angles α and φ, that determine respectively oar handle vertical and horizontal displacements. These angles are measured out by encoders Hengstler RI58-0, which has 5.000 steps per revolution. There are two ways for estimating force on the handle, both requires to measure or calculate the torque on the shaft that bears the oar. A torque sensor composed of a full Wheatstone bridge was mounted on the shaft to have a direct measure of this torque. Because of the frailty of such a sensor, the same torque is estimated by means of a model of the physics of the EDD, which needs as input of the EDD's angular speed. Therefore a further encoder was mounted on the EDD shaft to measure out its speed.

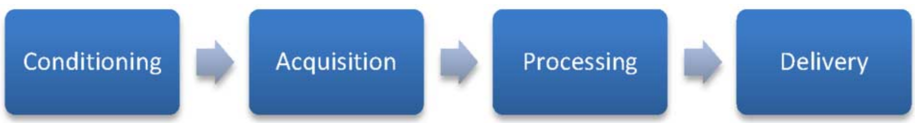
In the end a set of three encoders and a torque sensor outputs have to be captured and processed for each oar (Figure 1).Figure 2shows the placement of the device in the SPRINT system, highlighting the six signals that are captured.

3.2. Signal path

The choice of the computing element was very important in order to ensure that the electronics was low- power consumption, compact and robust, but at the same time powerful enough to handle the computing load for the signal acquisition and processing. The microprocessor has to cope with several requirements: 2 ADC/12bit for force sensor; 6 encoder inputs and 1 serial port for the communication with the host PC. In addition, 32-bit architecture and hardware floating point unit were useful to perform acquisition, processing and communication at a very high frequency.

Only two micro controller classes were found to comply with almost all the requirements: the generation of MPC5X5X from Freescale, and the control DSP F28335 from Texas Instrument. The latter was chosen for the side advantages offered by the Delfino™ control stick which implements in a single low cost module additional features such as: JTAG interface port for debug and programming, the integration with Matlab/Simulink development environment, on-board low-pass filters for ADC inputs, USB-to-Serial channels for programming and debug, and finally high operational frequency. In particular, the evaluation board integrates a TMS320F28335 running at 150MHz with a 32bits floating point unit, a fast and accurate PWM control, an integrated encoder acquisition unit and 68/512Kbyte RAM/Flash memory. Detailed specifications can be found online at Texas Instruments website.

Such a DPS allows to acquire a measure more accurately than with the previous hardware. The signal measure is composed by the following four stages:



3.2.1. Signal Conditioning

The encoder signals are TTL level (0 to 5V) with an high logical level of 5V, while the DPS works with a maximum high logical levels of 3.3V. For this reason, the logical levels must be conditioned before delivering to the DPS. The cheapest solution consists in a voltage divider even if it brings power dissipation. The final design adopts a voltage conversion stage composed by the Texas SN74CB3T3245 8-Bit Fet Bus Switch.

Torque sensors are based on strain gauges. A full Wheatstone bridge configuration was chosen because of self compensation for bending and temperature deformation. Moreover it provides a 4 times amplification of the signal. Since the force sensor signals are always comprised between 0 and 3,3V; the force signals can be connected directly to the DPS inputs avoiding the voltage conversion phase.

3.2.2. Signal Acquisition

The encoder acquisition have been done by interrupt. The TMS320C2000 Microcontroller Target of Texas Instruments has two enhancedQuadrature Encoder Pulse (eQEP) modules, while we need to acquire 6 encoder units. The eQEP modules were used to get the position count of the EDD, while the reaming encoders are read by external interrupt which drive the suitable algorithm for 4X encoders reading. According to data captured on the platform in an all-out exercise by expert rowers the maximum speed that the EDD could reach is 160 rad/s, that, according to the encoder resolution, gives a sampling frequency f_c = 260kHz. Arbitrarily the frequency work of the microcontroller and the eQEP modules was set as 4kHz.

Concerning the force sensor, since the strain gauges provide analog signals, the ADC module has been used to quantify the amount of force. The cell gains and the computing unit have been calibrated to achieve 0.05 Newton accuracy on a full scale of 800 Newton.
The sample frequency of the lecture in the ADC is 4kHz.

3.2.3. Signal processing

When an external interruption is posted, the microcontroller does a specific routine for a determinate encoder in order to increase or decrease the count of the position. To make the calculation of the speed in the EDD the registered count of the encoders was taken at fixed sample time every 1ms and multiplied by a factor in order to get the speed in radians per seconds.

The speed encoders have 2.500 counts per revolution, but since the 4x resolution technique is used, the total count per revolution increases to 10.000 counts, which deliver an angle accuracy of 0,036º per count. Now, assuming that the EDD is moving at 1 rev/s, in 1 second the encoder will count 10.000 in one direction. So, if the system takes a sample every 1ms means that if the encoder counts 10 the real speed will be 1 rev/s.

Finally, the speed factor needed in order to obtain the speed is

$$k_s = \frac{\pi \; eQEPcount}{10} \tag{1}$$

where *eQEPcount* is reset every interrupt.

3.2.4. Signal delivery

This is the final phase of the processes in the microcontroller. The communication between the microcontroller and the PC is made by an emulated serial connection. The specification of the serial connection and the total number of data bytes to be transmitted make the transmission frequency to be upper limited to 400Hz. Therefore the SCI module has been configured to work with a sample time of 0.0025 s. (400 Hz).

3.3. Results

In the following, some examples of data acquired on the SPRINT platform before and after the implementation of the new capturing system are shown. Figure 3 shows oar angles, EDD speed and captured forces before and after the implementation of the new electronics.

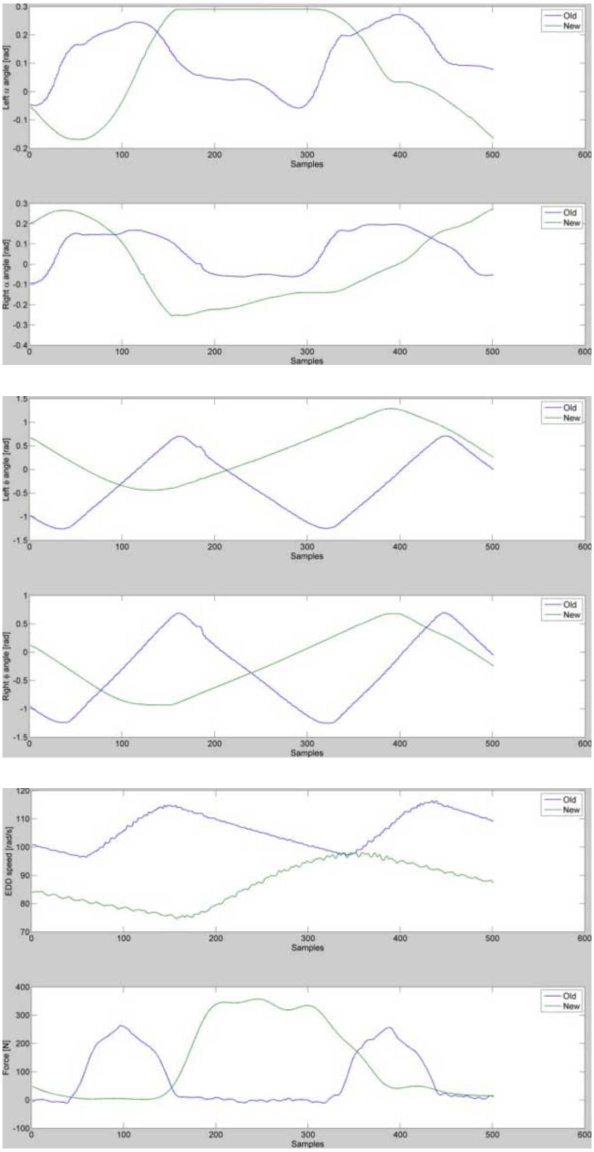


Figure 3 Oar angles, EDD speed and captured forces comparison between previous and current electronic acquisition system., time series were captured at the same stroke rate

It is possible to note that new data are sampled at roughly twice frequency with respect to old ones without loss of resolution. The improvement of the capturing is instead also in terms of resolution and communication smoothness, as shown in Figure 4.

These improvements allowed for a better simulation of the task and a better trainee's evaluation. Thanks to the higher frequency and resolution it has been possible to run the simulation of rowing and the performance evaluation at 125 Hz instead of the previous 60 Hz. It allowed for training timing of the drive phase, in which trainees were taught the correct body limbs motion onsets timing in order to optimize performance [11]. Moreover, it allowed to provide trainees with feedback at a higher frequency. Although it was not necessary for the visual feedback, for which 60Hz was enough, it has been crucial for training scenarios involving vibrotactile feedback, that is often provided for refining performance[11],[12] and hence requires to be provided with minimal latency.

Improvements in resolution (in addition to frequency one), allowed to reduce latency and improve the estimation of the signal derivatives. Since oar angles and fan speed are the fundamental information for the simulation of the rowing physics, they are required to be derived two times. Therefore, with previous electronics a strong filtering of the signals was necessary, thus increasing latency in force and boat motion estimation.

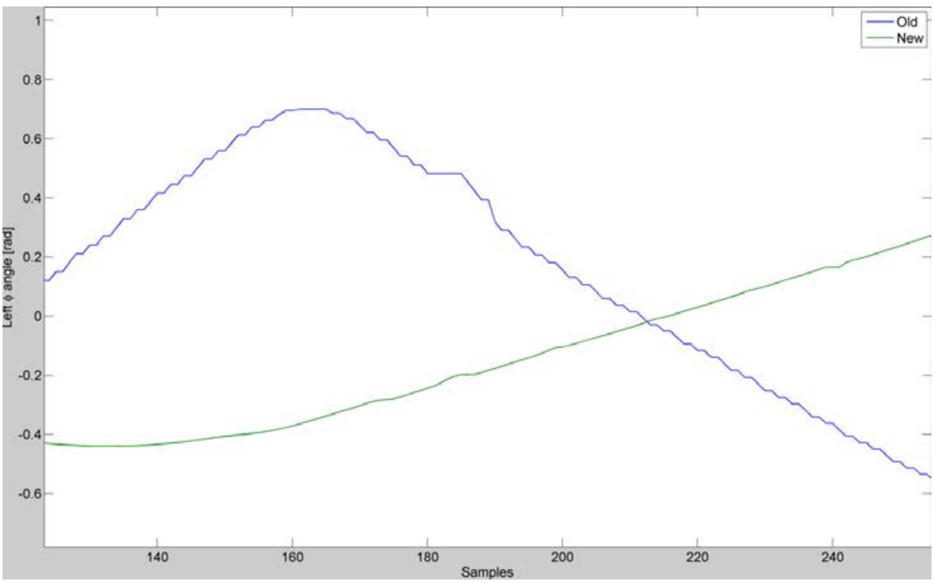


Figure 4 Closer look at resolution improvement with the new embedded electronics

4. Conclusion

This paper showed the electronics improvements of the SPRINT system. These sampling frequency and resolution improvements were and are crucial for the biomechanical analysis of the rowing gesture. They indeed allows for a fast and accurate real-time evaluation of the gesture that is the basis for providing trainees with proper feedbacks. The experiment reported in [11] shows that the obtained improvements effectively contributed to move from overall technique training to training for technique optimization.

References

[1] G. Wulf, G., C.H. Shea. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. Psychonomic Bulletin & Review, 9(2), 185-211.

[2] V. Nolte and et al, Rowing Faster, V. Nolte, Ed. Human Kinetics Publishers, 2005.

[3] FISA rowing handbook. FISA, 2006. [Online]. Available: http://www.worldrowing.com/fisa/resources/

[4] B. Elliot, A. Lyttle, and O. Birkett, "The Rowperfect ergometer: a training aid for on-water single scull rowing," Sports Biomechanics, vol. 1, no. 2, pp. 123–134, 2007.

[5] J. von Zitzewitz, P. Wolf, V. Novakovic, M. Wellner, G. Rauter, A. Brunschweiler, and R. Riener, "Real-time rowing simulator with multimodal feedback," Sports Technology, vol. 1, 2008.

[6] E. Ruffaldi, A. Filippeschi, M. Verlet, C. Hoffmann, and B. Bardy, Skill Training in Multimodal Virtual Environments. Taylor, 2012, no. ISBN 978-143-987-895-8, ch. Design and Evaluation of a multimodal VR Platform for Rowing training.

[7] A. Filippeschi, E. Ruffaldi, A. Frisoli, C. A. Avizzano, M. Varlet, L. Marin, J. Lagarde, B. Bardy, and M. Bergamasco, "Dynamic models of team rowing for a virtual environment rowing training system," The International Journal of Virtual Reality, vol. 4, no. 8, pp. 19–26, 2009

[8] E. Ruffaldi, A. Filippeschi, A. Frisoli, O. Sandoval-Gonzalez, C.A. Avizzano, M. Bergamasco. (2009). Vibrotactile Perception Assessment for a Rowing Training System. Proceedings of the third IEEE Joint conference on Haptics, World Haptics. doi: http://doi.ieeecomputersociety.org/10.1109/WHC.2009.4810849.

[9] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti, "Surround-screen projection-based virtual reality: the design and implementation of the CAVE," in SIGGRAPH '93 Proceedings of the 20th annual conference on Computer graphics and interactive techniques, 1993, pp. 135–142.

[10] E. Ruffaldi, O. Gonzales, A. Filippeschi, P. Tripicchio, A. Frisoli, C. A. Avizzano, and M. Bergamasco, "Integration of Multimodal Technologies for a Rowing Platform," in Proc. of IEEE International Conference on Mechatronics, 2009.

[11] A. Filippeschi, E. Ruffaldi, M. Korman. Preliminary evaluation of timing training accelerator for the SPRINT rowing system. In 2011 SKILLS conference proceedings, 2011, Montpellier. doi 10.1051/bioconf/20110100025

[12] E.Ruffaldi, A. Filippeschi, O. Sandoval-Gonzalez, D. Gopher. Visuo-Vibrotactile Trajectory Training in Rowing Experiment. In 2009 SKILLS Conference Proceedings, 2009, Bilbao.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-341

Event-driven Surgical Gesture Segmentation and Task Recognition for Ocular Trauma Simulation

Gianluca DE NOVI[a,b,1], Ryan BARDSLEY[a], Rajesh SHAH[a],
Mark P. OTTENSMEYER[a,b], John Cho MOORE[a], Bummo AHN[a,b]
[a] *Simulation Group, Dept. of Imaging, Massachusetts General Hospital*
[b] *Harvard Medical School*

Abstract. This paper describes a novel approach for segmenting surgical gestures in an eye trauma treatment simulator, based on the analysis of system and environmental events, generated in order to track user performance. Subdivision of surgical procedures into tasks & sub-tasks allows for application of more accurate metrics (optimized per task), which may more appropriately evaluate user performance. This approach may also be applied in other surgical contexts. For the specific case of eye trauma treatment this system can generate different classes of events that, when combined, may drive a state machine with the capability of subdividing the macro operation(s). The results of this process may be utilized to develop informative feedback to the user in the form of performance metrics and training guidance.

Keywords. Surgical gesture recognition, surgical simulation, surgical training, evaluation metrics.

Introduction

One of the most important aspects of surgical training tools is providing accurate and appropriate evaluation of user performance. This evaluation is a non-trivial process with many facets and approaches [1][2][3][4]. In the context of surgical gesture[2], performance is commonly evaluated through the analysis of movements and trajectories obtained from tracking the path of the surgical instruments, and their associated geometrical (spatial), and temporal features (e.g. smoothness, speed, and compactness) [4]. One method to improve this evaluation process is to reduce the overall surgical procedure into a set of tasks and to apply specific, task-oriented, evaluation metrics to each task [1]. Another important rationale for conducting this reductive process is that it permits selective exclusion of unrelated or extraneous actions (e.g. environment interactions) from potentially computation-intensive analyses. One of the issues related with the task evaluation approach is that the entire procedure must be subdivided into

[1] Corresponding author.

[2] A surgical gesture is the execution of a specific action, performed by moving and/or acting on the surgical tools.

tasks, subtasks and gestures automatically – computationally analyzing the movement data to identify each gesture, subtask, or task (gesture-subtask-task segmentation) [6][9]. A significant body of work has been devoted to this last topic. For example, research has been conducted based on evaluating the human body movements to analyze surgical gestures using human anatomy-driven hidden Markov models (HMM) [7]. While the most common approaches are movement-based, gesture segmentation does not necessarily need to be driven by analysis of the surgeon's (or surgical instruments) movements. The movement signal *is* necessary in order to identify the level of expertise of the surgeon. Efficient movements made by an experienced surgeon are typically more accurate and localized within the surgical field than dispersive and less compact movements made by novices [9].

This paper describes a new approach for surgical gesture segmentation and task recognition that can be performed by analyzing the *events* generated by the system (surgical instruments and environment) instead of the surgeon's (or surgical instruments) *movements*. The key motivation for focusing attention on the events generated by the system is that the events may contribute significant intermediate representations of information. This serves to discretize the movement recognition and analysis process, thereby reducing the complexity of segmentation. We define *events* as actions that can change the state of a state machine, e.g. opening and closing of a specific instrument, grasping with forceps, or entering/exiting from a specific region-of-interest. Assuming a well-structured system, it is possible to track the appropriate number and class of events. From this point, it may be possible to identify specific patterns of events/states that clearly correlate with specific gestures, subtasks and tasks of a surgical procedure.

The focus of this paper is the application of this event-driven approach in an ocular trauma simulator. However, this event-driven approach may be extended to a wider range of surgical procedures.

1. Background

In 2003, our group developed a laparoscopic skills trainer based on the SAGES Fundamentals of Laparoscopic Surgery (FLS) task curriculum. Using a modified commercial tracking system, the Computer-Enhanced Laparoscopic Skills Training Simulator (CELTS) allowed trainees to use actual laparoscopic instruments to perform FLS tasks on physical "task boards" including a silicone suture pad and peg-boards. The CELTS system resulted in the formation of several fundamental design concepts that we have leveraged in our current ocular and craniofacial trauma simulator. CELTS implements a full-color, real-time video display, rather than virtual rendering, to provide the surgical field of view. In combination with natural gravitation effects and standard laparoscopic instruments, CELTS provided trainees with a realistic approximation of procedural elements. The specialized tracking system permitted recording of translation and rotation with additional software-based task metrics. Using this system, our group demonstrated that appropriate performance metrics could be defined and a standardized scoring system could be designed [9].

The metrics generated by CELTS involved post-processing of the entire motion trajectory of the surgeon's gestures, which illustrated distinct differences in performance between experts and novices. However, finer-grained details used to provide real-time feedback and automated generation of suggestions for improvement, could not be extracted using those metrics alone. To address this issue, we propose a method of event-based task segmentation that will enable immediate evaluation and improved feedback.

2. Methods

The methodology utilized for surgical gesture segmentation and task recognition is driven by specific patterns of events generated by the system. By this process, motion path analysis[3] is deferred to a conditional, later tier of analysis. The analysis of a list of discrete events is considerably less computationally expensive, than the continuous analysis on a moving window of data sampled at high rates from one or more surgical instruments paths. This may unambiguously indicate user actions and in cases where it does not, additional path analysis can be conducted on a focused subset of the streamed data. The same approach may be applied both to the subtask and task levels of analysis. As in other methodologies, the first step to perform is to determine the structure of a data frame of measured parameters at each time point. In our case, the parameters measured for each surgical instrument were: time stamp; instrument identity; raw sensor position and orientation; jaw closure; and jaw grasping force. Second order data extends the data frame by deriving instrument tip position, tip velocity, and tip acceleration.

Based on the data types available, we define sets of thresholds. Events are defined as threshold crossings. Observing the data frames over time, the system creates a list of time-stamped events, such as the moment when a spatial region of interest (ROI) is entered or an instrument's jaws are fully closing. This list of events would typically be orders of magnitude smaller than the original data set.

From a detailed task analysis of relevant surgical scenarios, we define *gestures* as meaningful patterns of *events*. The system creates recursively higher order lists of gestures, subtasks, tasks, and ultimately, procedures, corresponding with the different levels of the tasks analysis hierarchy (Table 1,Figure 1).

Identifying actions at each level of the taxonomy provides the opportunity to provide the best evaluation metrics at the most appropriate moment during the scenario. Depending on the training goals, this may be evaluation of decision processes, confirmation of correct sequencing of sub-tasks or gestures, or, when required, detailed analysis of a subset of the motion data within a gesture. At the same time, it allows exclusion from the evaluation process of extraneous motions (e.g. interacting with other colleagues or with the environment).

Table 1. System Taxonomy

Classification	Definition	Examples
First order data	Raw sensor data	• Instruments positions/orientations • Instruments opening/closing • Instruments force
Second order data	Data derived from raw sensor data	• Instruments tip positions • Instrument velocity

[3] Recognition of specific paths, curves and other features by specific features extracted from the path followed by the surgical tool.

Event	A crossing of a defined threshold for a given stream of sensor data or second order data	• Tip of needle holder enters a ROI. • Forceps open • Scissor cut
Gesture	A pattern of events	• Needle holder tip is closed. • Forceps grasp with medium tension • Needle holder pitches 95 degrees about its axis. • Cut w/ scissors
Subtask	A pattern of gestures	• Pass Needle • Two-throw knot tie • One-throw knot tie
Task	A set of subtasks	• 2-1-1 knot placement • Interrupted suture • Running suture • Exploration w/ forceps
Procedure	A set of tasks	• Full thickness eyelid laceration repair

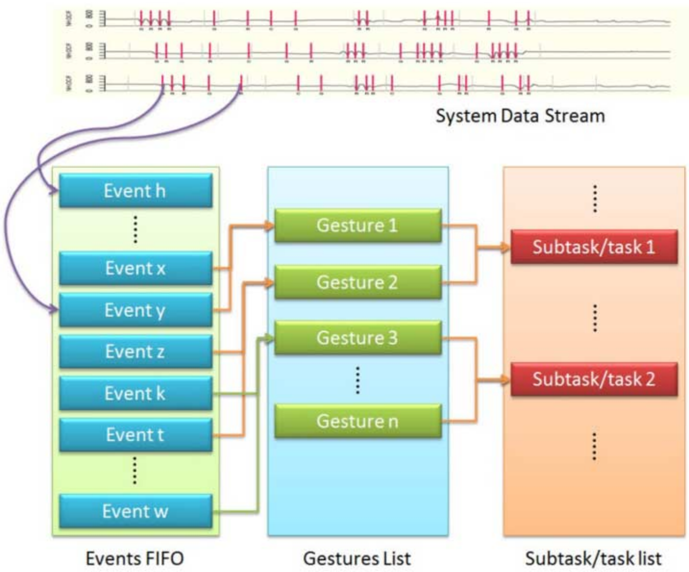


Figure 1. Organization structure of taxonomy.

3. Experimental Setup

The current approach was evaluated using the ocular trauma surgical simulator system that we are developing, described briefly below (Figure 2). The system is described in more detail in an accompanying paper [10].

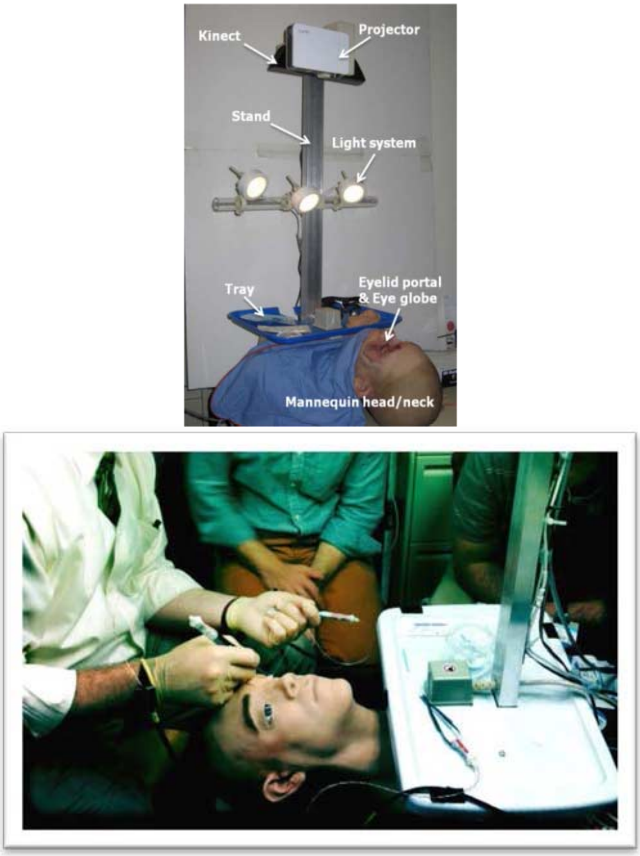


Figure 2. Simulation platform: mannequin head, instrument tray, tracking system, lighting system, overhead projector and Kinect tracker.

The platform is based on a silicone head attached to a vertical stand supporting a 3D camera (Microsoft Kinect^{TM}) used to track the interaction between the user and the simulation content, a set of lights, and a video projector. The mannequin head includes a set of interchangeable portals designed to simulate a range of trauma procedures (Figure 3, left). We are currently focusing on events related to repair of eyelid lacerations.

Figure 3. (left) Eye trauma portals and top view of instruments with position tracking sensors attached (right).

The system provides the user with a set of instruments necessary to perform relevant surgical tasks. Each instrument is connected to a magnetic tracking system (Figure 3, right), which measures instrument positions and orientations in six degrees of freedom

(DOF), and to a sensor circuit able to measure the opening/closing dynamics at the instrument tip and the force applied to the tip of the instrument. The sensors used on the surgical instruments are a set of strain gauges connected with a signal conditioning circuit board.

A series of opthalmic surgical instruments were connected to the magnetic tracking system for this evaluation (Figure 4)

Figure 4. Surgical instruments set: 1-510 Westcott scissors curved tip, 2255T Castroviejo forceps, 4536T Castroviejo needle holder.

The system is controlled by data acquisition software (Figure 5, left), which provides capabilities to record data provided from the instruments/sensors and other environmental parameters. The Graphical User Interface (GUI) also provides real-time 3-dimensional (3D) displays of the mannequin and instrument positions, as well as configurable multi-track views of the recorded data.

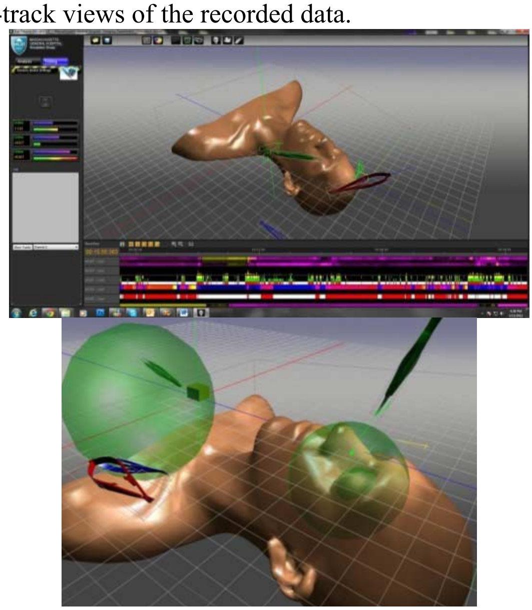


Figure 5. (left) GUI of data acquisition and analysis software. (right) Regions of Interest (ROIs) and instrument interactions with ROIs.

A significant feature provided by the software tool is the management of multiple spherical regions-of-interest (ROI), allowing the creation of maps and generating events for each crossing of an ROI boundary (Figure 5, right).

The software records the signals coming from the system (instrument position/orientation/state,) and displays this information in a timeline. To support the experimental design below, the interface also provides the capability to insert time-stamped markers, using an external interface device (mouse, keyboard, etc). The listing of these keystrokes is shown in Table 2.

Table 2. Direct Observation Markers

0. Null key
1. Needle point pierces tissue
2. Needle eye retrieved from tissue
3. Knot throws initiated
4. Knot pulled taught
5. Scissors retrieved from / returned to tray
6. Cut with scissors
7. Non-tracked hand action
8. User error or digression
9. Verbal comment

4. Experimental Design

The current prototype allowed collection of natural human interaction data in a highly-structured fashion to support initial investigation into our theory of event-driven gesture segmentation and subsequent task recognition.

To develop an understanding of how to parse the data collected by the simulator, it was initially necessary to identify what parameters constitute an *event*. Identification of the unique *types* of these events throughout a surgical procedure, currently focusing on suturing of an eyelid laceration, was also required.

4.1. Experimental apparatus

The subject matter experts (SMEs) were provided with one or two pairs of surgical gloves, depending on personal preference. Several varieties of suture and needle combinations were available to provide the SMEs with familiar interventional options. Across our three subject matter experts, preferences were for Ethicon 5-0 Perma-Hand silk with P-3 needles, Ethicon 6-0 Vicryl with P-3 needles, and Ethicon 6-0 Prolene sutures with P-3 needles. Task portals were inserted into the mannequin head, synthetic blood was applied to the wound, and the lacerated eye was covered with a Fox eye shield to hide the simulated injury until the beginning of the scenario. The prototype's LED lighting units were turned on, and were augmented with a Storz fiber optic Xenon surgical lamp, per the SME's request.

4.2. Protocol for Capturing Expert Data During a Simulated Eyelid Laceration Repair

In order to keep the SME's performance as natural as possible, the protocol was designed to minimize simulator-specific motions. Precautions were taken to avoid having the SME modify his or her natural gestures to accommodate discrepancies in the simulated procedure. Brief conversations with the engineers, comments on known issues with the prototype, and a general hesitation to break potentially fragile components could have distorted how an SME would use the simulator. Surgeons are often also natural teachers and tend to explain their actions during the data capture process. Therefore, sessions began with strict guidance *not* to offer comments, illustrate thoughts by pointing with instruments, or teach using the simulator. The SMEs were instructed to remove the Fox eye shield, investigate the wound, and treat the injury as they would in real life. To ensure the session was as unbiased as possible, SMEs were asked to complete the full eyelid repair before focusing on repetitive tasks so as to avoid any simulator-specific learned behaviors. In addition to the unstructured surgical performance, the structured exercises recorded across all three experts were as follows:

1. Grasp suture from tray and load the needle in the needle holder.
2. Double-bite through both sides of incision, pull length of suture through
3. Double-bite with 2-1-1 surgeon's knot, excess cut with scissors
4. Single-bite through both sides of incision, pull length of suture through
5. Single-bite with 2-1-1 surgeon's knot, excess cut with scissors.
6. Suture placed, pull length of suture through until ready to tie knot
7. Suture placed, tying of 1-1-1 surgeon's knot.
8. **Suture placed, tying of 2-1-1 surgeon's knot.**
9. Suture placed, tying of 3-1-1 surgeon's knot.
10. Suture placed, tying of 3-1 granny-1 adjustable surgeon's knot.
11. Grasp tissue with forceps, light pressure.
12. Grasp tissue with forceps, normal pressure.
13. Grasp tissue with forceps, heavy pressure.
14. Vertical mattress suture.
15. Running suture.

Three eye portals were fabricated from our multi-layer multi-part molds that shared a similar laceration on the upper eyelid. The laceration was a full-thickness wound and was generally repaired with 6 to 8 interrupted sutures. Figure 6 shows the repairs and the sutures that were placed during the later discrete task sessions. Minimally guided performance by all three SMEs was found to be similar.

The system was configured to collect data from the SMEs on both a complete procedure, as well as a specific set of tasks. As the SMEs were performing both the open-ended eyelid laceration repair and the discrete event tasks, an observer manually recorded the data captured by using the keyboard attached to the simulator. These manual keystrokes facilitated the later synchronization of the multiple sources of data: overhead video, 3D path data, and the initial event recognition output. The manual keys were used when parsing the path data to begin to identify trends, distinct markings, and patterns while helping to ignore erroneous data.

Figure 6. Silicone eye portals illustrating the final suture patterns created by three different ocular surgeons.

5. Preliminary Data, Analysis and Discussion

The analysis below focuses on Exercise #8: the tying of a 2-1-1 surgeon's knot. The raw data collected by the simulator was parsed to begin to extract events that characterized this known task. Each track of data was analyzed using three different methods: visualization of the raw data, characterization of the recorded video, and normalization across all three sessions for each of the three SMEs.

Figure 7 shows the open/close position data collected from each of the three SMEs, overlaid with the corresponding manually identified needle grasp and release gestures. One can identify correspondence between the sensor data and the observations, which supports the proposition that an automated event detection and gesture recognition system can reproduce the manual observations.

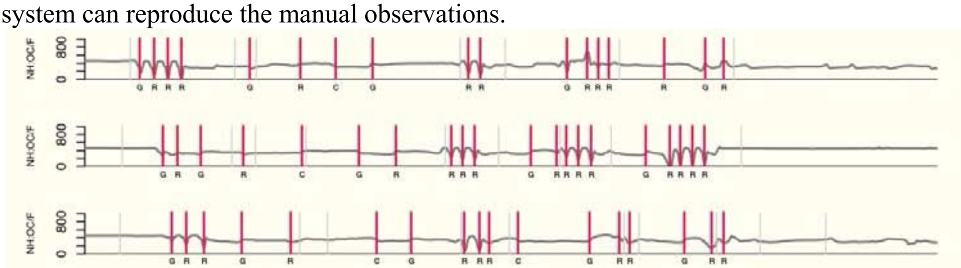


Figure 7. Raw needle holder open/close position data and overlaid observational event markings.

Figure 8 shows a schematic representation of the spatial ROIs around the eye. The radii of these spheres and their centers were established in relation to the real anatomy, and the raw position data processed to identify the boundary crossing events for each instrument. The time evolution of which ROI a given instrument occupies is shown in Figure 9, which presents the motion of scissors, from occupying the instrument tray to cutting a suture and returning. The sequence of events can be interpreted as the gestures: (A) "removal from tray," (B) "entry into working space," (C,C) "cutting of suture ends," (D) "exit from working area" and (E) "return to tray." Over multiple SMEs and multiple repetitions, this sequence of gestures was the same.

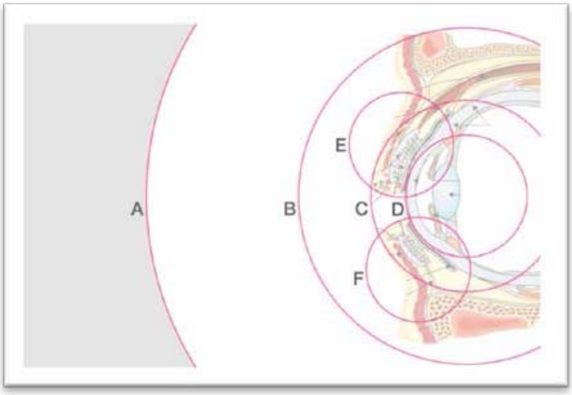


Figure 8. Regions of Interest (ROI) defining areas where the simulator begins to monitor instrument interaction to extract events for gesture segmentation:

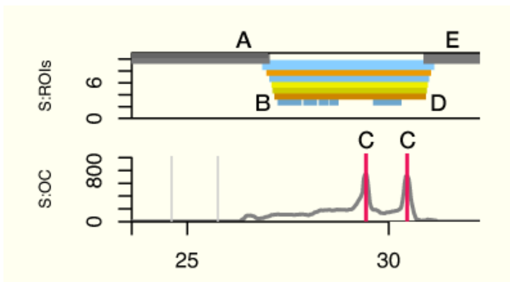


Figure 9. Use of scissors to cut suture ends

Using different combinations of event sequences and their derived gesture sequences, higher levels of the taxonomy can be identified, such as creating a 2-1-1 surgeon's knot. The full sequence is shown in Figure 10.

Both the forceps and the needle holder start within a relatively small ROI near the eye. During the process, the needle holder approaches more closely, entering a smaller ROI. Presence within the smallest ROI and the closure of the needle holder identifies the grasping gesture. The subsequent opening and closings are releases and re-grasping gestures as the needle passes through the tissue. The departure of the instruments from the nested ROIs suggests pulling the needle to draw the suture through the skin, and in combination with orientation thresholds (not shown) the forceps are palmed, pointing away from the suture site, hence apparently moving away, through the sequence of ROIs.

The re-approach and movement into and out of the small ROIs, followed by a needle holder closure (vs. grasp) signal the double wrap of suture around the needle holder and closing of the holder about the short end of the suture to form the knot, which is tightened at 10 seconds with the next passage outwards through the ROIs. The close-up closure and release gestures signal single wraps of suture and short tightening of the next two knot ties. Finally, the needle holder passes out of the eye centered ROIs and is palmed (orientation events not shown) and enters the instrument tray ROI, where the scissors are picked up brought inwards to cut the long end of the suture, to be returned and exchanged with the needle holder again. Together, these gestures constitute the 2-1-1 knot placement subtask.

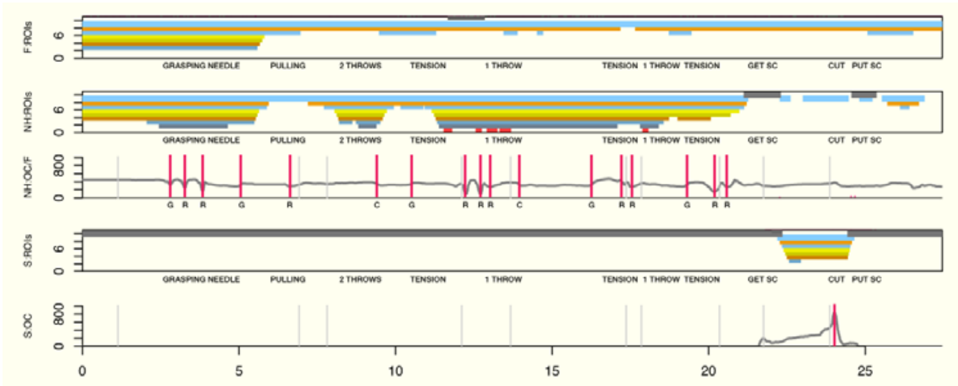


Figure 10. Full sequence of needle insertion and tying of 2-1-1 knot subtask.

This example presents the sequential re-composition of the raw data, events detected and gestures into a subtask. Longer sequences would be further composed into tasks making up a procedure. Similar observations of sets of expert performances will be parsed into the broad taxonomic library that the system under development will use to segment automatically trainee performance data sets, to identify where they perform sequences similarly to the experts, where and how they differ, and ultimately provide feedback in support of the learning process.

6. Conclusion / Ongoing Work

This paper presents the early stage development of a novel approach towards efficient and useful surgical task recognition and performance measurement. It reduces computational load compared with real-time motion trajectory analysis by identifying and operating on small numbers of discrete, low-level events, which are rebuilt into progressively higher level, meaningful taxonomic structures. This tiered system will permit the identification of key elements for evaluation, and trigger feedback appropriate to the level of analysis.

The next phases of the work focus on characterizing the remainder of subtasks that the SMEs performed; extracting additional thresholds to build the event library and the subsequent gesture and sub-task sequence libraries; validation of the analysis algorithms by comparing automated segmentation against manual recognition; and collecting extensive data sets from experts and novices to begin to distinguish how these two groups differ and when those differences rise to the level of triggering corrective feedback. A data collection protocol is currently under IRB review, in which military attending surgeons, fellows and residents will perform laceration repair and elementary exercised so their behavior can be compared. Ultimately, this will form a key element of the ocular trauma surgical simulator and provide new tools for developing other training systems.

Acknowledgements

We would like to express our grateful appreciation of our Subject Matter Experts: Dr. Robert Mazzoli, Dr. Suzanne Freitag, Dr. Daniel Lefebvre, Dr. Justin Kanoff, Dr. Maria Troulis, and Dr. Steven L. Dawson.

This material is based upon work supported by the U.S. Army Medical Research Acquisition Activity under Contract No. W81XWH-11-C-0095. The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office.

Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Army Medical Research Acquisition Activity.

References

[1] I. Hajshirmohammadi, S. Payandeh, Fuzzy Set Theory for performance Evaluation in Surgical Simulator, *Presence 2007*:16(6): 603-622.

[2] L. Moody, C. Baber, T.N. Arvanitis, Objective surgical performance evaluation based on haptic feedback. *Stud Health Technol Inform*. 2002;85:304-10.

[3] C. Sewell, Automatic Performance Evaluation in surgical Simulation, Doctoral thesis, Stanford University, 2007.

[4] C. Feng, H. Haniffa, J. Rozenblit, J. Peng, A. Hamilton, M. Salkini. Surgical Training and Performance Assesment Using a Motion Tracking System. *Proc. 18th European Modeling and Simulation Symposium*, EMSS06-4833796613, 2006.

[5] C.E. Reiley, G.D. Hager, Task versus subtask surgical skill evaluation of robotic minimally invasive surgery, *Med Image Comput Assist Interv. 2009*; 12(Pt 1):435-42.

[6] K. Kahol, N.C. Krishnan, V.N. Balasubramanian, S. Panchanathan, M. Smith, J. Ferrara, Measuring movement expertise in surgical tasks, *Proc. MULTIMEDIA '06, 14th annual ACM international conference on Multimedia*, 719-722, 2006.

[7] N. Padoy, T. Blum, I. Essa, H. Feussner, M.O. Berger, N. Navab, A boosted segmentation method for surgical workflow analysis. *Med Image Comput Comput Assist Interv*. 2007;10(Pt 1):102-9.

[8] S. Speidel, T. Zentek, G. Sudra, T. Gehrig, B.P. Müller-Stich, C. Gutt, R. Dillmann, Recognition of Surgical Skills using Hidden Markov Models, Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling, edited by Michael I. Miga, Kenneth H. Wong, *Proc. of SPIE* Vol. 7261, 726125.

[9] N. Stylopoulos, S. Cotin, S.K. Maithel, M. Ottensmeyer, P.G. Jackson, R.S. Bardsley, P.F. Neuman, D.W. Rattner, S.L. Dawson, Computer-enhanced laparoscopic training system (CELTS): bridging the gap. *Surg Endosc.* 2004 May;18(5):782-9.

[10] M.P. Ottensmeyer, B. Ahn, J.C. Moore, E. Munteanu, S. Iorino, G. De Novi, R. Bardsley, R. Shah, P. Neumann, D. Langston, Ocular and Craniofacial Trauma Treatment Training System: Overview & Eyelid Laceration Module, *Proc. IMIASH 2012 - 1st International Workshop on Intelligent Multimodal Interfaces Applied in Skills Transfer, Healthcare and Rehabilitation*, Guanajuato, Mexico. Accepted for publication.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-353

Development of a natural interaction interface for people with disabilities in a home automation control room

Ruben POSADA-GOMEZ [a], Cristhian Omar RODRIGUEZ-BERNARDO [a], Phares Salathiel LUNA-BRAVO [a], Giner ALOR-HERNANDEZ [a], Albino MARTINEZ-SIBAJA [a], Alejandro RODRIGUEZ-GONZÁLEZ [b]
[a] *División de Estudios de Postgrado e Investigación, Instituto Tecnológico de Orizaba, Avenida Oriente 9 No. 852 Col. Emiliano Zapata, Orizaba, 94320,Veracruz, Mexico*
[b] *Computer Science Department, University Carlos III of Madrid, Av. Universidad, 30, Leganes, 28911,Madrid, Spain*

Abstract. This paper presents an interface for controlling some functions in a room using a natural interaction mechanism. The functions of natural interaction are a group of tools of Human Computer Interaction (HCI) designed to develop gesture control applications, one of the objectives of this work is the development of a domotic application to help people with different disabilities and execute specific tasks in a friendly graphic environment, some of these tasks are turn on/off a lamp, open/close blinds, activate/disable the conditioned air, activate an alarm for asking help, etc.

Keywords. Physical disabilities, Gesture recognition, human interface computer natural interaction interface.

Introduction

Home automation is a field of application development that allows the creation of interfaces using control devices and HCI devices for help people to control some daily activities. Nowadays, automation has contributed significantly to the development of applications to make possible the operation of some basic activities performed at home, but this kind of systems has not a great impact at systems for helping poples with some disabilities. Since every individual needs to have a correct operation in their place of residence, it is pretended to seek the comfort and wellbeing of people with special needs through the relationship of a HCI device and a PC. This paper proposes the development of an interface that allows a person with disabilities to interact more naturally with a home automation environment. Figure 1 shows the proposed control of this paper.

In the room, there are integrated home automation controls such as:

- State of the blinds in a room (either open or closed).
- Lighting of the room.
- Making a telephone call through the interface.

- An alarm to call a nurse if the patient feels bad.
- An access to a camera outside the room or house.

Figure 1. Proposed controls for room automation interface interaction natural.

The human-computer interaction (HCI) is the study, the planning and the design of the interaction between people (users) and computers. It is often considered as the intersection of computer sciences, behavioral sciences, design and several other fields of study. The interaction between users and computers is done at the user interface that includes both the software and hardware [1].

Nowdays there are several HCI devices mainly for entertainment poupuses, among them it highlights the Kinect Microsoft sensor that combines an RGB camera, depth sensor and an array of microphones. The RGB camera is used for adquire the three basic colors at scene, the depth sensor can view a 3D room and microphone array detects isolated voices and ambient noise, all embedded in a single device [2].

Kinect has the ability to obtain a series of models which are based on the detection of the user's body, however large amounts of training data are necessary to fit models, especially in the case of large spaces [3].

There are a few studies for designing user interfaces for the interaction of people with disabilities [3]. Navaratnam *et al* propose an interface for people with mobility impairments in the lower limbs that allows the user to perceive, understand, navigate and interact with an intuitive and natural system [4].

From the standpoint of computer vision, studies are performed applying knowledge and techniques of HCI such as accessibility and usability to help people with cerebral palsy and other disabilities [5].

1. Natural Interaction

Natural Interaction (NI) refers to a concept in person-device interaction based on human sense, especially focused on hearing vision [6]. Daily life examples of using NI are:

- Speech recognition commands, where devices receive instructions from voice commands.
- where the predefined gestures are recognized and interpreted to activate and/or control devices. For example, this control allows users to electronically manage tasks in a room with their own hands.
- Monitoring of body movement, where the whole body movement tracks, analysis and interpretation for the purposes of a specific task

Figure 2 shows a three-layer concept OpenNI.

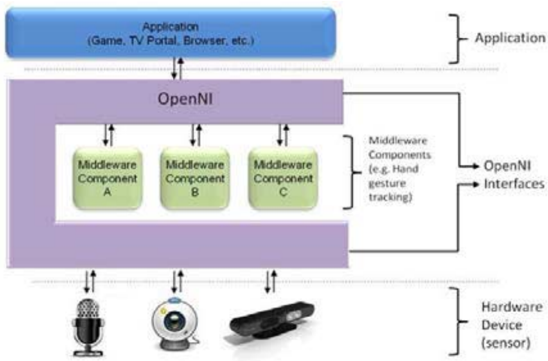


Figure 2. Concept OpenNI.

The modules that are currently supported by OpenNI are:

- 3D sensor
- RGB Camera
- IR Camera
- Audio device(a microphone or an array of microphones).

Since it is possible to program the characteristics of each one of these sensor modules, it is possible to define the gestures for the human interaction with computer. This characteristic is usefull to develop various personalized applications for home automation but is also possible to develop aplication for rehabilitation [5]. The ability of the sensor for the gesture recognition is will be explained in the next section.

2. Gesture recognition.

The sensor can identify different gestures which are processed and understood by the computer, once the gesture is recognized the system executes an asociated action, the gestures used by the system are listed below:

- PUSH. It recognizes the hand movement as a movement of thrust, which is a continuous movement toward the sensor and vice versa.
- STEADY. It recognizes when the hand has been stable for some time, the hand does not move (or barely moves), when the increase of motion is almost 0. STEADY is especially useful among other controls, to be sure to start the next control with the hand completely still.
- WAVE. It recognizes hand motion as a wave motion. A WAVE is a series of changes of address within a timeout. By default four changes of direction are required to identify a WAVE.
- CIRCLE. It recognizes movement of the hand as a circular motion. The Circle detector needs a starting point in any direction to start generating the output. The direction of the clockwise direction is considered as the "positive" and to the left is considered as the "negative".

To develop this application, a push command is used as well as the control of the point, which recognizes the movement of the hand as a point of 2D slider. It allows the recording of events when the hand is on a different element.

3. Graphical interface development.

The graphical interface was developed in C # because of the availability of libraries .NET OpenNI, which allows manipulating gestures control through events that are predefined by the libraries. Test were performed with real images and pose information from motion collection systems. The Figure 3 shows the Graphical User Interface (GUI) developed.

Through Kinect sensor, information is obtained from the hand gestures that are processed by the PC and then sent to a model through a card that works with a PIC18F4550 microcontroller.

Figure 3. Interface Developed.

The development of control by kinect sensor requires 3 main steps:

1. The declaration of variables and configuration
2. The gesture control and
3. The update thread.

3.1. Declaration of variables and configuration.

Figure 4. shows the process of device configuration and state variables must control. The reference ManagedNite.Dll must be declared. This is a namespace that contains the events of gesture recognition.

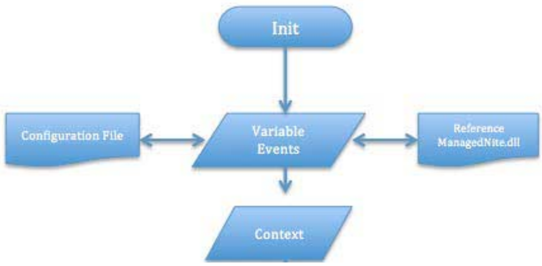


Figure 4. Declaration of variables and configuration.

It also adds the configuration .XML which contains the settings for the different nodes that exist within the kinect sensor such as licensing, screen resolution, hand detection, users, and depth, among others. Then, variables of the events you want to control are declared creating a method to perform specific functions.

Finally, the sensor is initialized and runs parallel update the thread sensor, if a problem occurs while initializing, the context will stop the process and display the error, otherwise the device is placed on hold for a sign of focus.

3.2. Gesure Control.

Figure 5. shows the control of gestures from the gesture of focus to the events of gestures, this is divided into 3 parts:

- Gesture of focus. The kinect has the ability to initialize the context with a bow or a hand movement to activate the sensor. The libraries are designed to recognize different gestures of focus in this application. To do so, you use the Wave and RaiseHand.

- Add.Listener. In this process indicates that event is controlled with this method since gestures depend on Add.Listener.
- Events of gestures. This block defines the events that you want to control. Once declared, the above variables create different methods or events which carry out different activities. These events depend directly on the updating thread. The thread updates the sensor so that they can obtain data and process that information, if the update thread does not start or end events will stop automatically.

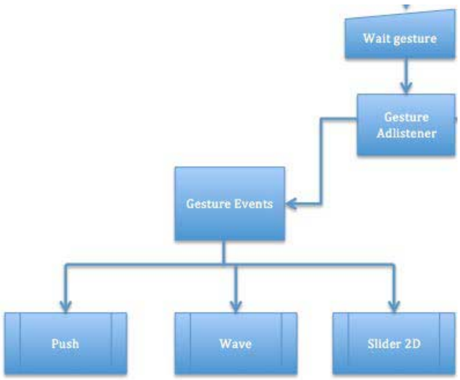


Figure 5. Gesture control

3.3. Updating thread.

Figure 6. shows the last block of the interface, which has the function of keeping the system updated by a thread. This block can also display the values of the different events so that they can perform different actions planned. If the update thread stops, events are automatically stopped because of the lack of information to be processed.

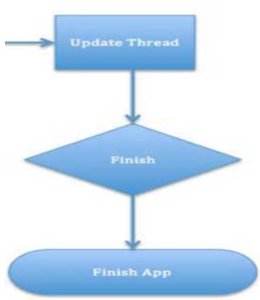


Figure 6. Flowchart update thread.

3.4. Development of control hardware.

In this paper, we developed a gesture recognition interface for an user with motion disabilities that force him to remain at bed in a room. The interface communicates with the corresponding hardware control devices (engine to close the blinds, a relay to light a lamp, etc..). Universal Serial Bus (USB) communication was used to conect the computer with electronic concrol circuits. The USB port has become a popular choice, because of their bandwidth of the connection (480 Mb / s for high speed from the launch of version 2.0 of the USB specification).

From the user standpoint, the benefits of the USB are ease of use, rapid and reliable data transfer, flexibility, low cost, and conservation power [7]. For the development of the USB interface a PIC18F4550 was used. Figure 7. Shows the PIC18F4550 circuit diagram connection. We used the USB communication Bulk mode as this is the preferred mode for data acquisition devices because of their transmision speed [8].

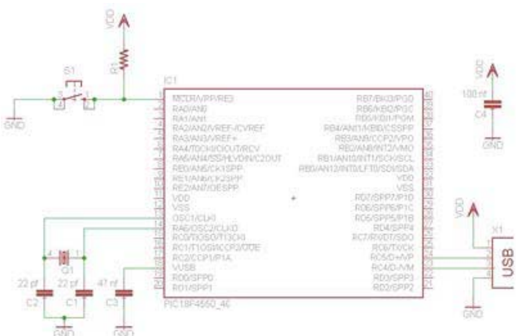


Figure 7. Connection Diagram PIC18F4550 microcontroller.

4. Results.

With the USB communication, it was possible to perform easily all actions prpeviusly mentioned. The control of motors was implemented for open and close the blinds, the room lights and conditioned air was achieved with the use of relays, and the telephone operation was implemented directly with the computer.

User interaction with the graphical interface was excellent and there were no problems to become familiar with the application, the developed system shows the feasibility of improve the actual home automation controls to include an interface with more natural interaction for people with disabilities. Table 1 shows the gestures recognized by the interface and the actions taken with the hardware.

Test of developped system shows good performances of gesture recognition with different users as well as different lighting conditions for controlling the lights room, blinds, conditioned air and telephone operation. It is however necessary to calibrate the gestures recognition system for patients with strong movility disabilities.

Table 1: Results.

Image gestual	**Action Taken**
	Wave salute to activate the application.
	Selecting turn on the weather.
	Lighting the lamp (detector Push) that simulates a click.
	Selecting emergency assistance.

5. Conclusion.

The interaction of HCI devices and home automation is a very useful tool for applications for people with phisical disabilities. This paper demonstrates that the use of gesture recognition system and natural interfaces represents a great help for people with disabilities because the system is not dependant of hardware controls as joystic or keyboards to perform some basic tasks.

Future works will be done to develop the validation of ergonomic techniques of rehabilitation to reduce the actions of therapists. This prototype will be installed on centers working with people with special abilities to assess their performance and response to patients with severe paralysis. A very important conclusion is that it has a great impact on users (especially children) demonstrating its potential not only for controlling home functions but also as a rehabilitation tool.

Based interfaces HCI technologies and home automation devices are a good example of design of accessible technology, so it is possible to extend its use and reduce investment costs in order to maximize the number of potential customers.

References

[1] C. Mauri, T. Ganollers and M. Garcia, Computer vision interaction for people with severe movement restrictions. *Human TechnologyJournal*, Vol 2(I), pp. 38-53, April, 2006.

[2] K. Sung. Recent Videogame Console Technologies, Entertainment Computing, *IEEE Computer Society*, pp- 91-94, Feb. 2011.

[3] R. W. Roe, Towards an inclusive future, *COST219 Project publications*, 2007.

[4] R. Navaratnam, A. Fitzgibbon, R Cipolla, The Joint Manifold Model for Semi-supervised Multi-valued Regression*, IEEE Intl Conf on Computer Vision*, 2007

[5] R. Posada-Gómez, O. O. Sandoval-Gonzalez, E. Sánchez-Muñiz, Development of a Motion Analysis System and Human-Machine Interaction through Digital Image Processing and Virtual Reality*, 21st Internacional Conference on Electronics Communications and Computers (CONIELECOMP 2011)*, Cholula, Puebla, 5p. 28 Febrero – 2 Marzo 2011.

[6] U. Breier, Natural Interaction and transformation of the living room experience, . *v-net.tv.* Jan 2011.

[7] J. Axelson, USB Complete, *Lakeview. Research LLC,* Third Edition, 2005.

[8] R. Posada-Gomez, J. J. Enriquez-Rodríguez, G. Alor-Hernandez, A. Martínez-Sibaja. USB bulk transfers between a PC and a PIC Microcontroller For embedded application*. CERMA '08 Proceedings of the 2008 Electronics, Robotics and Automotive Mechanics Conference,* 2008.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-362

Visuo-Vibrotactile Stimuli Applied for Skills Transfer and Rehabilitation

Ignacio HERRERA-AGUILAR[a], Osvaldo SANDOVAL-GONZALEZ[a], Blanca GONZALEZ-SANCHEZ, Juan JACINTO-VILLEGAS, Alejandro Rodríguez-González[c], Giner ALOR-HERNANDEZ[a] and Otniel PORTILLO-RODRIGUEZ[b]

[a] *Instituto Tecnológico de Orizaba, DEPI, Orizaba, Veracruz México*
[b] *Universidad Autónoma del Estado de México, Facultad de Ingeniería, Toluca, México*
[c] *Computer Science Department, University Carlos III of Madrid, Spain*

Abstract. This paper presents the static and dynamic analysis of visuo-vibrotactile feedback used like skills accelerator for motor skills training. Through this chapter, two experiments are presented in order to evaluate in different fields the use of this skills accelerator. In the first part it was performed a static psychophysics analysis of vibrotactile feedback to understand in details the accuracy in perception under determined conditions like intensity and frequency level. In the second part is presented a dynamic analysis of vibrotactile feedback where the use of diverse stimuli (visual and vibrotactile) are evaluated to understand how these kinds of stimuli can affects the performances of one motor activity under determined conditions. The results indicate improvements of the user's movements using in vibration in certain parameters in order to be useful like a skills accelerator.

Keywords. vibrotactile stimuli, skills transfer, rehabilitation

Introduction

Physical rehabilitation of patients with some form of paralysis and disability is hard work. The paralysis or disability can be caused by various causes, such as stroke, in Mexico there are almost 2.5 millions of people with motor disabilities, these people need rehabilitation therapies and they require specialized equipment to ensure their rehabilitation [1]. The mechanoreceptors in the skins make humans feel different tactile sensations when touching objects. When an object is touched, these sensations are combinations of tactile primitives like the normal indentation, lateral skin stretch, relative tangential motion and vibration. Different researches have explored the tactile sensation as a modality to present information for orientation and navigation in Virtual Environments [2] [3] . Lieberman and Breazeal carried out an experiment in real time with a vibrotactile feedback in a Virtual Environment to compensate the movements and accelerate the human motion learning [4]. In the same line of research Bloomfield performed a Virtual Training via Vibrotactile Arrays [3]. Both experiments show a significant improvement in the human cognition and perception. The vibrotactile feedback in human motion has been widely investigated with different perspectives in the application. Van Erp [6] explores the possibilities of tactile displays in sports applications, and reports an experiment that shows that a tactile feedback system improves rowing efficiency compared to traditional feedback systems. Earlier papers have shown that localized vibrations provide intuitive cues for orientation and navigation. The application of vibration in physical therapies has been widely used. On example of these researches was done by Lindeman [7] that presents a work in the context of a physical therapy application designed to provide more autonomy for patients when performing rehabilitative exercises. This assistive technology has the potential to reduce injuries during therapy due to improper patient joint movement, and decrease the workload of physical therapists, thereby reducing healthcare costs. All these researches have obtained interesting results using vibrotactile devices in human motion skills. However, there is a big area of research in the psychophysics field. Because, it is important to understand the behavior of the human perception under determined conditions like type of movement, velocity of movement, type of vibration, location of the vibrotactile devices and the correlation among different senses like audio and vision.

1. Description of the system

The vibrotactile control system was specially designed to control the vibration in intensity (vibrational acceleration) and pulses of frequency (train of pulses) of four motors. The system presented in Fig 1 consists of one Microchip microcontroller (PIC18F4431) capable to control four PWMs (Pulse Width Modulator) in hardware level with a 12-bit resolution. The power control section uses an IC ULN2803 (Darlington transistors array) which modules the total amount of energy through the PWMs that are applied to the vibration motors. There are two types of communications available: Bluetooth and USB protocols. These protocols transmit the information from the computer to the microcontroller. The vibration motors correspond to the model 12mm Shaftless Vibration Motor of the Precision Microdrives Company [8]. The motor has a nominal speed of 9000rpm generating a vibration of .9G with a current consumption of 120mA as it is presented in Fig 2. The strategy consists in the generation of four PWM signals of 4KHz frequency for each one of the four motors. For the control of vibration in intensity, the microcontroller receives from the computer the parameters of each one of the duty-cycles (0-4096). Once the microcontroller has acquired the values, modulates the duty-cycle of each PWM and send this information to the Electronic Power Circuit ULN2803. This device is a Darlington transistors array that is a compound structure consisting of two bipolar transistors connected in such a way that the current amplified by the first transistor is amplified further by the second one. This configuration gives a much higher current gain (written β, hfe, or hFE) than each transistor taken separately.

Fig 1. Vibrotactile System

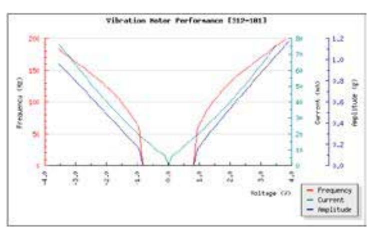


Fig 2. Vibration motor

2. Experiments

The aim of these pyschophysics experiments is to analyze the human perception under determined vibrotactile stimuli. These analyses are an opportunity to know in detail different information about the interaction between the human perception using visuo-vibrotactile stimuli and how could be used like skill accelerator in the learning process of one motor skills activity and rehabilitation. A couple of set experiments (static and dynamic) were carried out in order to understand the human behavior under determined stimuli. On one hand the static analysis is based on the psychophysics experiments of sensitivity of vibration in different modalities. In the other hand the dynamic analysis using visuo-vibrotactile feedback is evaluated during the performance of motor-skills activities.

2.1. Static Analysis of the Vibration's Intensity and Frequency using vibrotactile stimuli

The aim of this set of experiments was to know in detail the human perception using two types of vibration stimuli: intensity and frequency. On one hand, the intensity of the vibration is manipulated through the modulation of the energy delivered to the actuators from 0 to 3.8V in order to vary the rotational speed of the device from 0 to 9000rpm and the r.m.s. acceleration from 0 to 0.6382 mm/sec^2 and constant high frequency of 4KHz. In the other hand, the second vibration consist of a stimuli generated like a train of pulses of low frequency from 0 to 50Hz and acceleration of 0.5mm/sec^2. The experiments were divided in 2 sections. On one hand, the analysis of the inflexion points of sensitivity to determine the linear and non linear zones of the vibrotactile perception is analyzed. In the other hand a psychophysics analysis of vibrotactile sensitivity in intensity and frequency is carried out to determine the ranges of sensitivity according to the zones of contact, the surface area of contact and type of vibrotactile stimuli. In Table 1 is presented thee experiment that were carried out by 6 groups of 15 people were formed and divided depending of the contact area and contact place of the vibrotactile devices. During the experiment, all the people carried out 4 experiments (0.25m/sec^2, 0.5m/sec^2, 10Hz and 20Hz) and 30 trials of each one. The system in each one of the 30 trials renders two stimuli: one represents the pattern stimuli and another one is a stimulus generated by the Weibull function which varies the value of the stimuli according the previous responses of the user. Both stimuli are rendered in a random order in each trial, in the same way all the users carried out the 4 sections of the experiments in random order to avoid any kind of learning or adaptation of the stimuli.

Table 1. Experiments performed by 6 groups in order to identify the sensitivity in the vibration using the intensity and the frequency and the size of the surface area

Group	People	Contact Area	Place	0.25 m/sec^2	0.25 m/sec^2	10 Hz	20 Hz
1	15	226mm^2	Finger	30 trials	30 trials	30 trials	30 trials
2	15	452mm^2	Finger	30 trials	30 trials	30 trials	30 trials
3	15	226mm^2	Forearm	30 trials	30 trials	30 trials	30 trials
4	15	452mm^2	Forearm	30 trials	30 trials	30 trials	30 trials
5	15	226mm^2	Leg	30 trials	30 trials	30 trials	30 trials
6	15	452mm^2	Leg	30 trials	30 trials	30 trials	30 trials

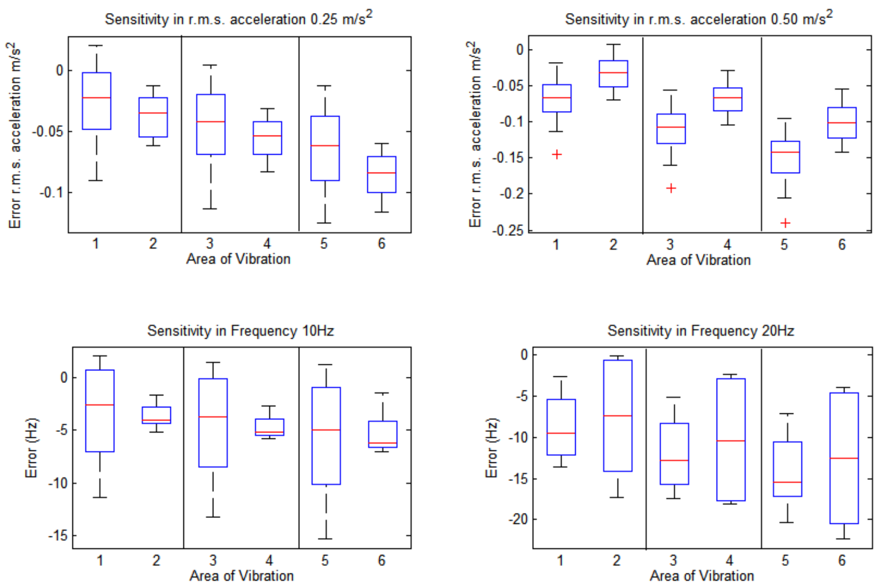


Fig 3. There are 4 plots which represent the vibration in Intensity ($0.25m/s^2$ and $0.5m/s^2$) and Frequency 10Hz and 20 Hz in a contact area of $226mm^2$ and $452mm^2$. Contact Place: (1) Finger $226mm^2$, (2) Finger $452mm^2$, (3) Forearm $226mm^2$, (4) Forearm $452mm^2$, (5) Leg $226mm^2$, (6) Leg $452mm^2$.

The results presented in Fig 3 generated by the perception of the user about vibration in intensity (acceleration) and frequency (train of pulses) shows that the improvement in the perception when the area is stimulated at $452mm^2$ is using an acceleration of $0.50m/s^2$ and frequency of 20Hz. Meanwhile in low intensity of $0.25m/s^2$ (acceleration) and frequency of 10Hz the results were better using a contact area of $226mm^2$. As it was expected the best results in the four tests were obtained by the finger and the worst were obtained by the upper-leg. The important point to comment is that perception of all these zones is located in an acceptable range of sensitivity, and that is important because these areas can be used for training purposes through vibrotactile stimulation, knowing in advance that the users will feel the variation in the simulation during the training. The second important topic to review is the perception according to the type of stimuli. In one hand, the vibration in intensity produced by the r.m.s acceleration (0.25 m/sec^2 and $0.5m/sec^2$) shows that it is better to work with moderate intensity. The users presented better sensitivity using a stimulus of 0.25 m/sec^2 rather than the stimuli of $0.5m/sec^2$. The possible explanation could be that the stimuli saturate the mechanoreceptors at high intensities. It is well known that the human perception presents a non-linear behavior. However, the importance of these experiments is to analyze the zones which present the closest linear response. The error in perception increases with a non-linear behavior when the stimulus is less than 0.25 m/sec^2 or higher than $0.5m/sec^2$.

2.2. Dynamic Analysis of Visuo-Vibrotactile Feedback as an accelerators in Motor Skills Training

Human motion is a fundamental part for the accurate performance in one activity. In the traditional process of learning human beings apply the imitation process to acquire the skills of one activity. Undoubtedly, the visual imitation process has demonstrated a natural instinct action for the acquisition of knowledge. However, although the vision process is one of the most effective and accurate sense to process the information and learn through the imitation, there is one important issue; the vision consumes many resources for the normal brain functionality. It means that the performance in the imitation process depends not only of the type of movement but also in the speed of the type activity. However, the integration of other senses like tactile sense could offer the possibility to augment the perception of the users and help them to perform the activity in an accurate way. The aim of this experiment is to analyse the visual and visuo-vibration feedback in the cognitive process of human being. Vibration is used like a collision feedback to transmit information of a trajectory. In this experiment the goal is to know the correlation between tactile-perception and motor skills performed by user in cognitive process in the brain. During the experiment the user must follow a predefined trajectory plotted in the screen with the movement of his/her hand. The experiment consists of a set of simple linear trajectories. Each trajectory has execution time classified in slow, normal and fast speed (20cm/sec, 40cm/sec and 60 cm/sec). The methodology is based on the idea of collision perception, in other words the user will feel vibration if they are performing the movement in an incorrect position. As is shown in Fig 4 a vibrotactile feedback system will be placed in the wrist of the person. These vibration motors are located in the frontal and back face of the wrist and they will render a proportional vibration depending of the error from 0 to 0.5m/sec^2. During the experiment the participants will carry out seven trials in a random sequence. Each one of these trials corresponds to different kind of visual and visuo-vibrotactile feedback. There is a linear trajectory of 20cm plotted in the screen and the users must follow a pattern sphere in blue color that is moving upwards and downwards with a specific velocity. During each trial of the experiment, the velocity is divided in 3 different speeds (20cm/sec, 40cm/sec and 60 cm/sec).

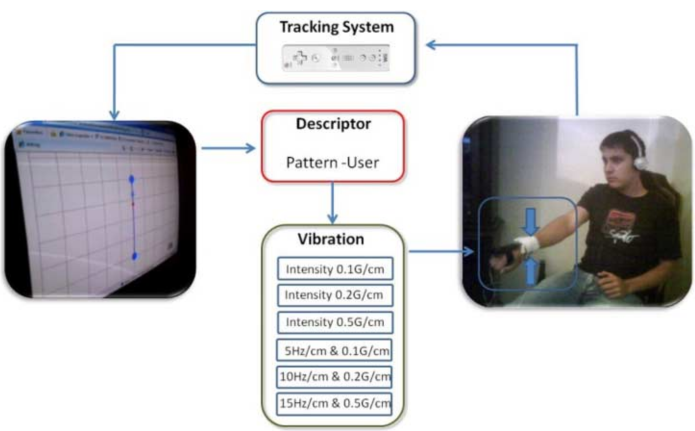


Fig 4. General Structure of the Evaluation

2.2.1. Visual Feedback

This test analyses the motor skills of the user following a simple trajectory by the hand controlled only by vision. The distance of this trajectory is 20cm from top to down. The visual feedback experiment was carried out by 14 people; it was divided in 3 intervals of time (20cm/sec, 40cm/sec and 60 cm/sec). In each interval of time the user must follow the pattern ball in 50 cycles or repetitions. That means that the user carry out 150 repetitions during the complete trial.

2.2.2. Visuo-vibrotactile feedback

This test analyzes the motor skills of the user following a simple trajectory by the arm controlled by visuo-vibrotactile feedback. The distance of this trajectory is 20cm from top to down. This experiment was carried out by 84 people (6 Groups of 14 People); it was divided in 6 tests in 3 intervals of time each. The vibration will act in the following way: If the users perform the movement correctly, the vibration is not rendered. When the actual position of the user is upwards in relation with the pattern position, the vibration motors placed in the frontal face of the wrist will start to vibrate. This vibration will be increased proportionally to the error. This feedback indicates to the user that there is an error in their trajectories and they must move backwards the arm. When the actual position of the user is backwards in relation with the pattern position, the vibration motors placed in the back face of the wrist will start to vibrate. This vibration will be increased proportionally to the error. This feedback indicates to the user that there is an error in their trajectories and they must move back the arm. The 6 tests correspond to the type of vibrotactile stimuli rendered to the users. In Vibration by Intensity there are 3 types of vibration stimuli rendered in intensity form. The stimulus varies proportional to the error in distance generated by the users. These 3 types of vibration are: 0.07 rms m/s^2 per cm, 0.1418 m/s^2 per cm and 0.35 m/s^2 per cm.

In Vibration by Intensity & Frequency there are 3 types of vibration stimuli rendered in intensity & frequency form. In these cases the vibration is changing in intensity and frequency. The user will feel a train of pulses of 5Hz/cm, 10Hz/cm and 15Hz/cm and the intensity of these trains of pulses will vary in : 0.07 rms m/s^2 per cm, 0.1418 m/s^2 per cm and 0.35 m/s^2 per cm depending of the error performed by the user.

3. General Discussion and Analysis of the Results

Fig 5 and Fig 6 shows the results of the visual feedback obtained by the 14 users. It is important to highlight how the error is increased in a proportional way to the velocity in the movement. In a velocity of 60cm/sec the users obtain a huge error of almost 40%. Although the vision is an effective method to follow and control the motor skills activities, it is clear that the efficiency is totally dependent of the velocity and the brain is not capable to process all these visual information and reacts at the appropriate time to correct the trajectories. There is significant difference of 18% of error between the minimum and maximum velocity.

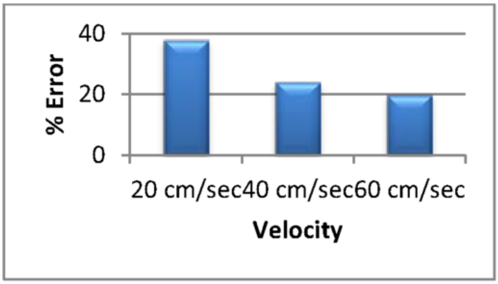


Fig 5. Results of the Visual Feedback in different velocities.

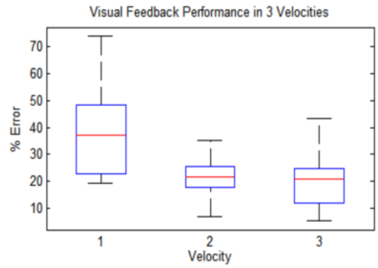
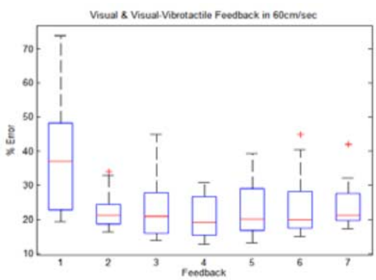
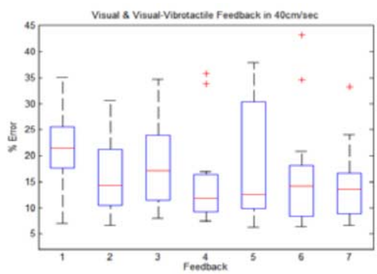


Fig 6. Visual Feedback performance in the 3 velocities. 1) 60cm/sec, 2) 40cm/sec, 3) 20cm/sec

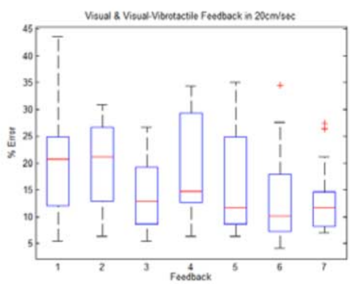
The next analysis shown in Fig 7 (a) (b) (c) presents the results of the seven test using the visual feedback, the three modalities of vibration in intensity and the three modalities of vibration in frequency-intensity in the three different velocities. The maximum errors in the three velocities were obtained by the visual feedback. The best result in 60 cm/sec was obtained by vibration in intensity at 0.35m/s^2 per cm with an error of 21.20%. In 40 cm/sec the vibration in frequency-intensity mode 20Hz/cm & 0.35m/s^2 per cm obtained the best result with an error of 14.73% and finally in the velocity of 20cm/sec the best performance was obtained again by the vibration in frequency-intensity mode (20Hz/cm & 0.35m/s^2) with an error of 13.50%.



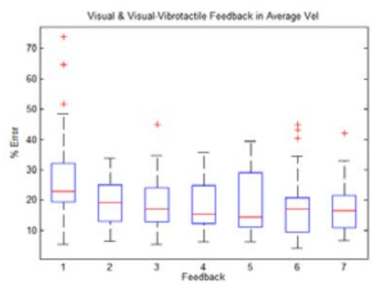
Visual & Visual-Vibrotactile Feedback in 60cm/sec.



Visual & Visual-Vibrotactile Feedback in 40cm/sec.



Visual & Visual-Vibrotactile Feedback in 60cm/sec.



Visual & Visual-Vibrotactile Feedback in Average Velocity.

Fig 7. Visual & Visual-Vibrotactile Feedback in 20cm/sec, 40 cm/sec and 60 cmsec. Feedbacks: 1) Visual, 2) Vis-Vib1, 3) Vis-Vib2, 4) Vis-Vib3, 5) Vis-Vib4, 6) Vis-Vib5, 7) Vis-Vib6.

The interesting point to comment is that although this vibration (vibration in frequency-intensity mode 20Hz/cm & 0.35m/s^2) have obtained the best results in medium and slow velocity, this feedback obtained the second worst performance in high velocity with 23.85%.

The final analysis, shown in Fig 7 (d), presents the results of the seven feedbacks in an average velocity. In other words the average of the 3 velocities was calculated of each one of the 7 feedbacks. On one hand the visual feedbacks obtain the worst performance with 26.43% of error. In other hand, the best results were obtained by vibration in frequency-intensity mode (20Hz/cm & 0.35m/s^2) with 17.36% of error.

4. Conclusions

Through the analysis of the results, it can be concluded that as it was observed during the experiments, the use of visuo-vibrotactile feedback produces in all the cases an improvement in the trajectories performed by the users. These vibrotactile feedbacks in high velocities obtained the best improvement reducing the error almost 20% with respect to the visual feedback. The Visual feedback has a big dependency of the velocity of the performed movement. Higher the velocity higher the error. Faster changes in the relation of vibration/error obtain better results for fast movements.

References

[1] INEGI, Las Personas con Discapacidad en México y sus Características, *INEGI,* pp. págs. 1-2, 2004.

[2] M. Florian and S. Bernt, Sensing and Monitoring Professional Skiers, *PERVASIVE Computing Sports Technologies,* 2005.

[3] Y. T. Kai, M. K. P. Peter, C. Mo, K. H. Sze and Z. Hongfu, Wearable Power Assistive Device for Helping User to Move Their Hand. United States Patent US 2010/0305717 A1, 2 12 2010.

[4] J. L. Breazeal, Development of a wearable Vibrotactile FeedBack Suit for Accelerate Human Motor Learning, *IEEE International Conference on Robotics and Automation*, pp. 4001–4006, 2007.

[5] A. Bloomfield and N. Badler, Virtual Training via vibrotactile arrays, *Teleoperator and Virtual Environments,* vol. 17, 2008.

[6] J. Van Erp and H. Van Veen, Vibrotactile invehicle navigation system, *Transportation Research Part F: Traffic Psychology and Behaviour,* vol. 7, pp. 247-256, 2004.

[7] R. W. Lindeman, Y. Yanigida, K. Hosaka and S. Abe, The tactaPack: A Wireless Sensor/Actuator Package for Physical Therapy Applications, *Proceedings of the IEEE conference on Virtual Reality,* 2006.

[8] P. Microdrives. [Online]. Available: http://www.precisionmicrodrives.com/product_info.php?products_id=93.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-370

Virtual Reality Sports Training

Paolo TRIPICCHIO
Scuola Superiore Sant'Anna, Pisa, Italy

Abstract. The use of virtual reality system in the particular field of sports and skills training is here discussed. The article presents the common architectural parts of such systems showing current examples in the literature where Virtual environments are applied successfully for motor skills learning and technique improvement both for the novices and for the athletes.

Keywords. Virtual Reality, Sports Training, Multimodal System

Introduction

Recently virtual reality applications have been developed with the purpose of sport training[16]. Such multimodal systems are indeed a good help for skills improvements and physical training. In particular they succeed where most of the classical training stops. Such technological training systems have precise records of physiological data and correct performance measurements that are difficult to obtain or estimate without a technological support like in classical fitness or technique training. In addition to such detailed information they can generate appropriate feedback to the user to exhort him/her to improve his/her technique or even to correct in real-time a wrong behavior. Visual stimuli are the preferred and most immediate way to teach a correct movement and display important information to the athletes however, the use of other important modalities can greatly improve the skills learning results. In particular audio feedback has been investigated in [1] where the learning of complex tai-chi movements in the three-dimensional space is proved to be simple to achieve with a localized aural feedback. Haptic feedback add an important sensory cue reproducing the contact sensations when interacting with pure virtual objects. These systems can ease the alteration of test parameters like objects weight or material rigidity. Thanks to this it is possible to easily adapt the sport simulation to the particular athlete and focus the training on a specific aspect that is primary for the athlete improvement.

An interesting aspect not to be undervalued is the realism of such training systems or better the presence and immersion sensations that the system can stimulate on the user. To achieve this goal the visual modality should not be disrupted by introducing other modalities like the haptic one. The problem of encumbrance of mechanical systems and co-presence of humans and devices in the same virtual scene in fact is a central one on the development of immersive virtual systems.

1. Capture of movements

One of the main component of a virtual training application is the tracking system. It is necessary to have a proper capture of human motion to provide correct feedback and display a realistic simulation of the interaction with the virtual world. The technology behind such tracking system can be optical, magnetic, ultrasound or inertial based. Magnetic tracking however has interaction problems with mechanical devices that are not always magnetic shielded so that in some portion of the environment the magnetic field result distorted and the tracking not correct. Inertial sensors require a lot of computation and change of reference systems but are used usually linked to special suits wear by the users. This of course could results in some motion impairment for some specific tasks.

The most used tracking technology results for the abovementioned reasons the optical one. The optical technology field is quite large and the tracking systems available in the market employ a wide number of them. The cheapest technology that is largely used both in the research and entertainment fields is the structured-light 3D scanner technology that consist in projecting a sort of pattern on the environment and reconstructing based on the re-acquired distorted pattern the three-dimensional topology of the scene. A commercial device that utilizes such technology is the Kinect sensor by Microsoft. Other special cameras uses time of flight information to realize a depth image of the scene. Finally classical and more expensive systems uses a set of cameras and special markers to capture the movement of objects in the real world. As example the commercial Vicon system employs spherical retro-reflective markers and infrared light sources to track the motion of the markers with a resolution under the millimeter.

2. Classification and segmentation

Independently of the technology used for the tracking the intended result is to capture the whole body movements or the movements of the specific body part or joints the training should be focused on.

This particular software application usually employ a kinematic model of the human body to extract from the tracked position all the biomechanical data and perform an analysis of the posture of the user. In particular it is also possible to identify the exact movement the user is performing or in general the class of movement its motion belongs to through the use of classifiers and machine learning algorithms. For instance in [1] the authors present a system that perform an online segmentation of the user movements and by means of an artificial neural network is able to recognize the specific gesture he/she intends to do with the purpose of tai-chi learning.

In order to recognize the gesture performed by the user, a state space model approach was selected [2,3]. A common issue with the modeling of gesture with this approach is the characterization of the optimal number of states and the establishment of their boundaries. A dynamic k-means clustering on the training data defines the number of states and their spatial parameters of the gesture without temporal information[4]. This information from the segmented data is then added to the states and finally the spatial information is updated. This produces the state sequence that represents the gesture. The analysis and recognition of this sequence is performed using a simple Finite State Machine (FSM), instead of use complex transitions conditions which depend only on

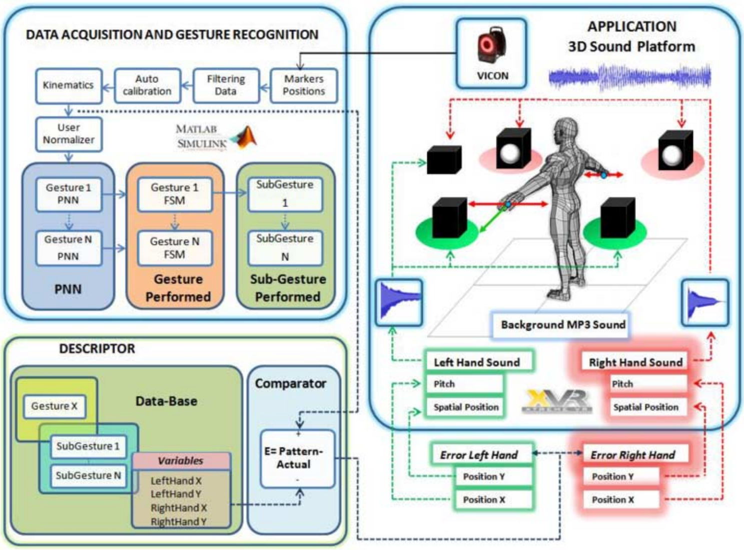


Figure 1. Scheme of the gesture recognition algorithm and output signals on the virtual training simulation.

the correct sequence of states for the gesture to be recognized and eventually on time restrictions.

The novel idea is to use for each gesture a probabilistic neural network(PNN) to evaluate which is the nearest state (centroid in the configuration state) to the current input vector that represents the users body position. The input layer has the same number of neurons as the input vector and the second layer has the same quantity of hidden neurons as states have the gesture. In this architecture, each class node is connected just to one hidden neuron and the number of states (where the gesture is described) defines the quantity of class nodes. Finally, in the last layer, the class (state) with the highest summed activation is computed. A complete reference of the whole system is found in [5]. A number of sixteen variables were used in our configuration space: 2 distances between hands and 2 between elbows, 2 vectors created from the XYZ position from the hands to the chest and 2 vectors created from the XYZ positions from the elbows to the chest.

The comparison and qualification in real-time of the movements performed by the user is computed by the descriptor system. In other words, the descriptor analyzes the differences between the movements executed by the expert and the movement executed by the practictioner, obtaining the values' error and generating the feedback stimuli to correct the movement of the practictioner (equation 1).

$$\Theta_{error} = [P(n+1) - U(n) * Fn] \quad (1)$$

Where Θ_{error} is the difference between the pattern and the user, P is the pattern value, U is the user value, Fn is the normalization factor and n is the actual state.

The system can analyze step-by-step the movement of the user and creates a comparison between the movements performed by the master and user. This algorithm can perform the comparison of 26 variables (Angles (12), positions (12), distance between hands (1) and elbows(1)). For this application we only use 4 measurements which represent the X-Y deviation of each hand with respect to the center of the body. Figure 2 shows the Tai-Chi movements that are recognized by the gesture recognition system and analyzed by the descriptor.

Figure 2. Tai-Chi movements recognized by the system.

Another system [6] focusing on boxing attack /defense training employs a multilayer neural network to discriminate in advance the attack of the user and to teach in third person the corresponding defense strategy to be employed. The system in fact is comprised of two phases , the first is a first person boxing simulation and the second is an offline repetition of what happened from a third point of view to make the boxer understand his/her flaws and improve during the training.

The core of the identification system is based on a Feed-Forward Artificial Neural Network which receives an input vector formed by 4 variables. These variables correspond to the energy levels generated by the following vectors placed in a fixed period of time: Left shoulder to the left hand; Left shoulder to left elbow; Right shoulder to right hand; Right shoulder to right elbow. Where the energy level is computed by: $E = x^2 + y^2 + z^2$.

One important point to take into account is that each body has different dimensions in their joints. Therefore a calibration process must be performed to assure that the values that arrive to the neural network are correctly normalized. This process is done automatically measuring the length of the arm and obtaining a normalization parameter with respect to the pattern dimension.

In each one of the hits the system generates visual stimuli plotting the ideal trajectory that the user must perform. Both in technique and force training the system renders the vibration feedback proportionally to the impact of the hit. At the end of each session, the system replays all the movements performed by the user during the session in third person, giving to the user information about his/her performance.

3. Display technologies

In the realm of the visual feedback the most prominent technologies for sport training are immersive visualization systems like CAVEs [19] that can grant a large virtual workspace and head mounted displays (HMD) , worn on the user head, that can remove

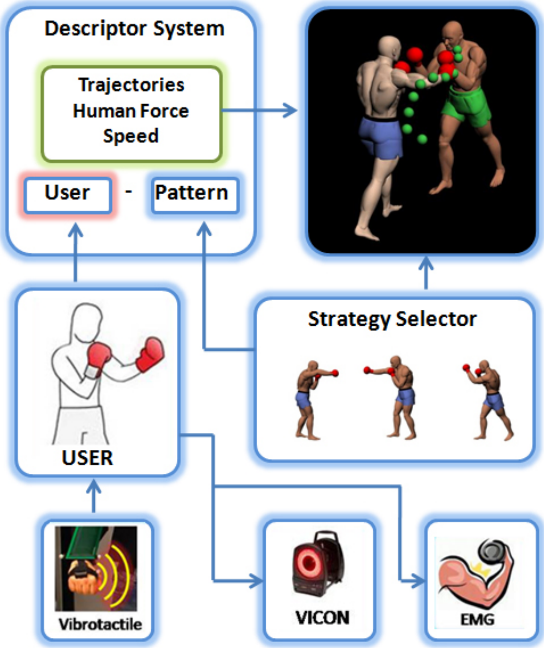


Figure 3. Boxing technique training. The Vicon and EMG are the data acquisition systems. The vibrotactile and Virtual enviroments represents the rendering section. The Descriptor System analyze in real time the movements of the user and the Strategy Selector System choose the ideal movements and trajectories for the training.

the real/virtual barrier immersing the user in a complete virtual world. The first solution allows great resolution of the scene and a free movement of the user inside the virtual environment limited only by the CAVE's walls and eventual device co-existing on the system. The second solution instead frees the optical flow from real devices encumbrance but applies some weight on the user head resulting in an unnatural simulation for prolonged training activities.

Haptic displays play an important role on virtual reality training system giving the user the capability to exert and feel forces for real. The technologies behind such devices are numerous but a main classification can be done based on the perceptual stimulus the haptic interface produces to the user. We have than kinesthetic and tactile displays depending on the appropriate kind of feedback such devices can render. Tactile displays reproduce sensation at the skin level rendering surface information like roughness or softness, material shape etc. Kinesthetic displays instead reproduce forces caused by motion and contact interaction with virtual object.

One of the first application of haptic interface is in pool simulation[20] where the force entities exchanged with the mechanical device are scaled to fit in the usual range of commercial haptic interfaces. However such simulations are intended more as games than as real training platforms. The same talk can be spent for new entertainment applications of the Nintendo's Wii console that had as commercial strength the idea to make people move and stay fit through the use of its innovative controllers playing videogames.

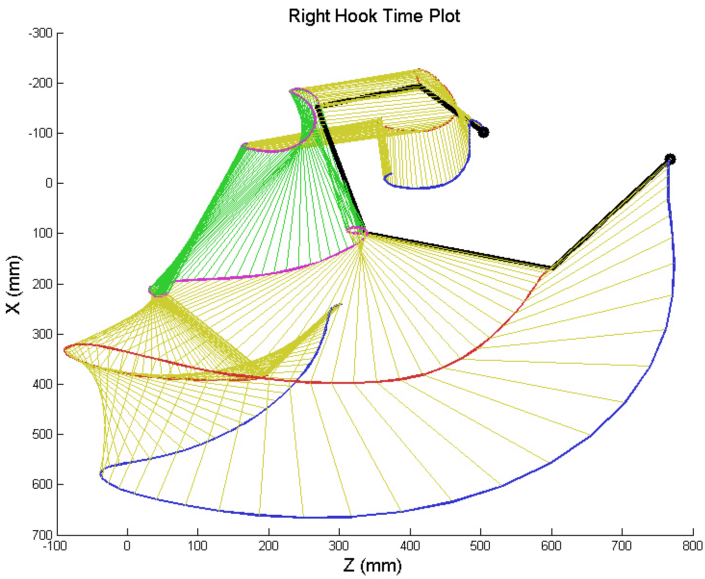


Figure 4. Trajectory plot of the Hook Punch view in the XZ plane.The black larger line shows the configuration of the arms and shoulders of the boxer at hit position. Trajectory of the elbows and hands are reported in red and blue respectively. The shoulders trajectories are displayed in magenta and between them the back segment is shown with a green line. Yellow lines represents the arm configuration on every time instants.

When the force entities in play are big there is however the need to design a specific haptic device with the purpose of reproducing a realistic feedback to the user. An example of this kind of devices can be the one used for car racing simulation or sailing simulation [21]. Here however the device is used to produce the sensation of motion rather than a real contact force.

Lately the employment of vibration motors in many virtual reality application has grown. In literature many gloves for the interaction with virtual 3D worlds are proposed and some of them integrates an exoskeleton to reproduce forces at the fingertips and joints levels while others, more cheaper, make use of pager motors to generate perceivable vibrations on the user hands. This kind of feedback in literature is more intended as a tactile display than a kinesthetic one but it is useful because has a reduced encumbrance and add small perceptual disturbances compared to the use of a classical mechanical haptic interface.

In the real world, when a person touches an object, forces are imposed on the skin. However, this relation between touch and force is not available like a natural link in the Virtual Environments. Haptic interfaces have been applied to solve this problem acting like generator of mechanical impedance. Impedance represents a relationship between forces and displacements in the skin surface and transmits the touching of virtual objects to the human being.

The mechanoreceptors in the skins make humans feel different tactile sensations when touching objects. When object is in touch, these sensations are a combination of tactile primitives like the normal indentation, lateral skin stretch, relative tangential motion and vibration. Different researches have explored the tactile sensation as a modality to present information for orientation and navigation in Virtual Environments. One example is the tactile belts that have been studied by different groups[7,8] as approaches to provide directions in the horizontal plane to blind people.

Lieberman and Breazeal carried out an experiment in real time with a vibrotactile feedback in a Virtual Environment to compensate the movements and accelerate the human motion learning [9]. In the same line of research Bloomfield performed a Virtual Training via vibrotactile Arrays [10]. Both experiments show a significant improvement in the human cognition and perception.

Following this line of research, one approach of skill acquisition using vibrotactile feedback has been studied by Van Erp and others[11] which performed an experiment to study the intrinsic and extrinsic phenomena involved in the cognitive level when a person move his wrist through the combination of certain vibration stimuli, transmitted by five vibrotactile devices located in specific parts of the forearm and hand of the user. The analysis presented interesting results in the field of motion-vibration in order to know the appropriate locations and combination of vibration stimuli.

An interesting example of the application of vibration motors can be found in [12] where thanks to a vibration bracelet users can learn a correct rowing motion to be executed.

The aim of this experiment was to know if a person is capable to correct his/her movements only using the kinesthetic feed-back stimulus. Six participant, five males and one female, aged 25 to 35 were separated in two groups. They are all right handed, one of them was a medium level rower and all of them have at least once experienced haptic devices. The participant sits on a chair between the oarlock and the rail equipped with the vibrotactile device mounted on her/his right arm. The experiment was performed using 4 sensors located in the arm of each person. One in the middle finger to simulate the forward constraint, another one in the elbow to simulate the backward movement. The last two vibration motors are located in the up/down side of the wrist to simulate the upwards and downwards constraints. The objective is to perform a square figure of 20 by 20 degrees in ϕ and α angles (right oar) during 120 sec like a training process. The vibration motors were activated depending of the state of the gesture to simulate physical constraints and give information to the user in order to perform the movement in the best possible way.

The results in figure 5 show how a person is acquiring the correct movement through the time. The black color corresponds to the initial time and the red color corresponds to the final time in the training. A movement almost perfect is done at the end of 120 sec using the vibrotactile feedback.

The Aural feedback is not largely used as training mean but usually just to relax or alienate the user during a training session. In [1] however, the audio is employed as training feedback achieving really good results.

In particular a frequency distortion was generated when an error in the X direction occurred and a position distortion when an error in the Y direction occurred. The distortion was always proportional to the error distance between the correct movement and the performed one.

The results of the experiments conclude that this methodology is an interesting approach for the possible transfer of the knowledge of movements from an intelligent 3D sound platform to the human being using an audio feed-back stimulus (pitch and spatial sound variations) according to the methodologies presented. Moreover, the results sup-

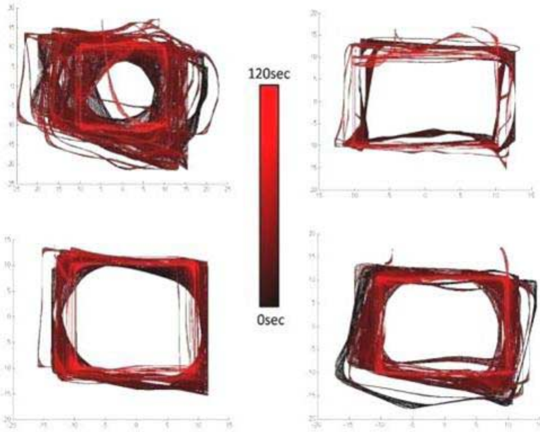


Figure 5. Results of the tactile training for rowing technique learning performed by 4 users during 120 sec.

port the fact that the system not only helps the user to identify an unknown gesture but also trains the people through the repetition process.

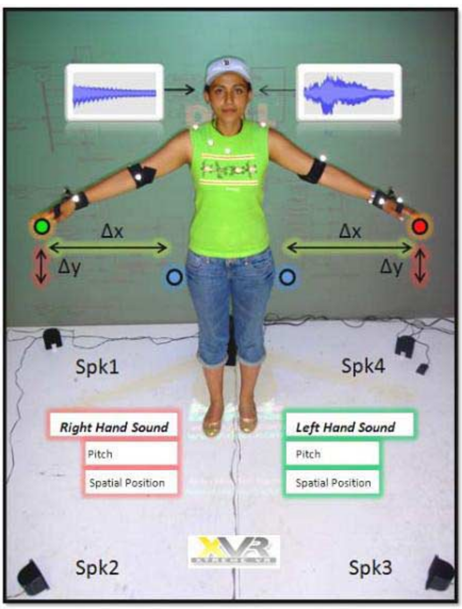


Figure 6. Audio feed-back stimuli when a user is performing a bad movement in the X and Y plane.

4. Physiological data

Lately the introduction of new kind of sensors for the acquisition of human physiological signals directly from the user body has been applied in the field of virtual reality.

The EMG (ElecrtoMioGraphic) signal is a biomedical signal that measures electrical currents generated in muscles during their contraction. These signals are complex to acquire and process, because they are small (0 to 10mV) and highly susceptible to noise. Normally, a sampling frequency from 2 to 4 KHz should be performed in order to obtain good parameters in the behavior and spectrum of the signal.

Therefore, a big drawback is the considerable amount of data that is obtained during the acquisition process of the raw signal using many sensors. The integration of human force based on EMG signals in virtual environments has been normally used in the rehabilitation field to control devices and measure the performance in the movement of the users.

S.Kousidou et al. [13] performed a robotic approach to task based therapy. The work focused on a robotic exoskeleton operating in a 3D volume used in conjunction with a Virtual Environment rehabilitation suite for training patients in relearning daily motor tasks. EMG recordings were used to show the capacity of the system to mediate the level of assistance. Marcello Mulas et al. [14] performed an experiment with a device for the hand rehabilitation. The system was designed for people who have partially lost the ability to control correctly the hand musculature, for example after a stroke or a spinal cord injury. Based on EMG signals, the system can 'understand' the subject evolution to move the hand and thanks to its actuators can help the fingers movement in order to perform the intended task in the virtual environment.

Hideaki Touyama et al.[15] performed a work using the advantages of the EMG signals. It was proposed to use these signals as a promising human interface in immersive multiscreen environments. The user would control virtual objects reflecting the activations of the user's muscles to realize fine operations.

Some new attempts have been made to acquire not only EMG signals but rather ECG (ElectroCardioGraphic) signals directly read form the user body by means of electrodes placed over his/her chest. The ECG device detects and amplifies the tiny electrical changes on the skin that are caused when the heart muscle depolarizes during each heartbeat.

This data can be of great aid to evaluate the physical form of the athletes and to automatically tune a proper training.

Also EEG (ElectroEncephaloGraphic) signals are acquired with electrodes distributed over the scalp of the user to verify at the cognitive level the learning progress of the athletes. However this technology is not yet really diffused for sports training but has a great success on medical applications.

Another interesting information for sports training is where the attention and the focus of the sight is spot. For these reason some virtual reality systems employs special cameras that monitor the activity of the pupil of the user's eye to acquire exactly the direction of his/her gaze.

In [16] in particular also an analysis of the breathing condition of the testing subjects were performed to check if the breathing scheme of each participant was efficient or it was stressed.

5. Hands free interaction paradigm

During human's interaction inside a virtual environment, whenever the virtual representation of the human body comes in contact with a virtual entity, a collision occurs. The collision affects both the virtual and the real worlds. In a standard paradigm the motion of the user causes the collision, affecting the virtual world and at the same time the virtual world produces a feedback on the user. When the interaction is based on haptic interfaces the user is in contact with the interface itself, allowing to take into account the exerted force. In free-space interaction, instead, only the motion is considered. The proposed interaction paradigm uses the force applied by the user in free-space as a new and fundamental input for interacting with virtual environments.

When the user is acting against an object, a force is recorded by EMG sensors placed on the user's body and during contact this force is used for interacting with virtual objects. From a physical simulation point of view this means that the interaction force of the user's avatar with an object is not based on its quantity of motion but instead on exact force exerted by the user. The force transmitted from the user in the real environment to the object in virtual environment has a direction along the velocity of impact and a modulus adapted from the real force. This modulus depends on the specific application, for example in the case of boxing training the force values of the biceps and triceps muscles are averaged. Correspondently, the feedback from the environment to the user is generated by taking into account the exerted force.

The result is a vibrotactile stimulus proportional to the instantaneous applied force and the stiffness of the object.

Integrating four of the aforementioned technologies , in [17] a novel paradigm for hands free virtual reality interaction is presented. In particular a commercial optical system (Vicon) has been used for tracking purpose. Markers where placed over the fist of the user to track position , velocities and accelerations of the hands.

An ad hoc EMG system with sensors has been designed to acquire forces data coming from the biceps and triceps of the users so that it was possible not only to determine the correspondent force applied by the user but also to distinguish the particular punch that the user was performing.

As it is possible to see from the figure 7 the activation of the two muscles are different depending of the specific attack the athletes is going to perform.

Vibration motors have been placed as well over the users fingers to feedback an intensity proportional to the impact force with a virtual boxing bag.

Finally all the presented technology where immersed together with the user inside an L shaped visualization system (a particular CAVE with only two big projection screens).

The proposed system has a wide area of application as target. It's versatility allows the system to be adopted on a large variety of virtual reality application where free interaction is a key factor. A typical situation is one where the user should be able to touch objects in the virtual world and grab as well as move them within the environment to accomplish some specific task. In particular the vibration feedback adds a big perceptual aid to the interaction making possible to understand exactly when a collision occurs[18].

The presented technology is also employable in shape recognition tasks where objects may be occluded from the user sight. All other applications that need an estimation of the users applied force benefits from the accurate recording of the EMG capturing system. In particular when the interaction with the scene is mediated by a virtual tool the capacity of the proposed system to detect the exerted force during manipulation is essential. A wide area of VR applications that can adopt this system is the one of medical practice/training where precise measurements of position, acceleration and forces is fundamental. Another specific area is the one involving manual skills and sport activities.

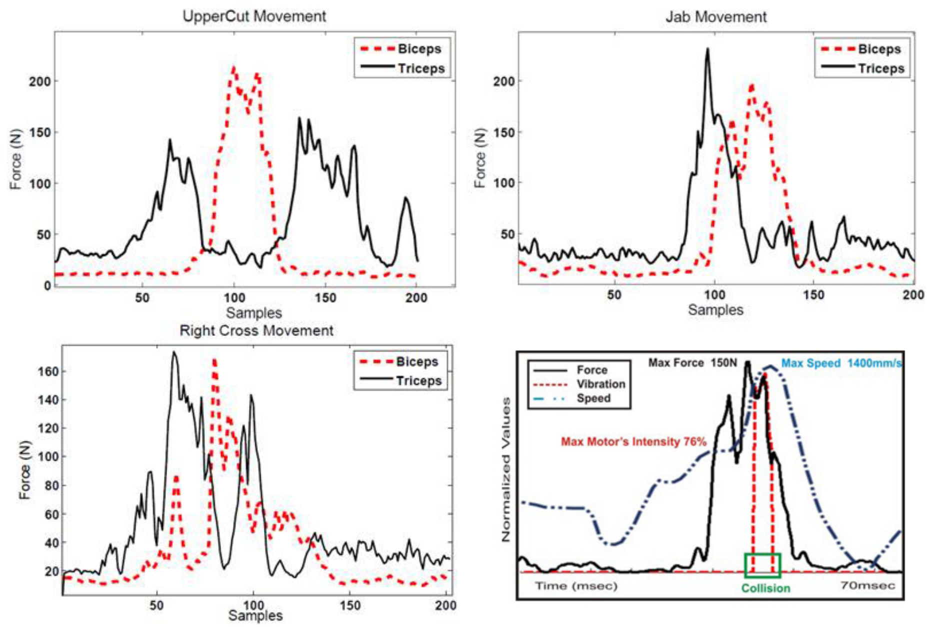


Figure 7. Force profiles captured by the EMG sensors for each fist and correspondence between vibrational stimulus, exerted force and user's movement.

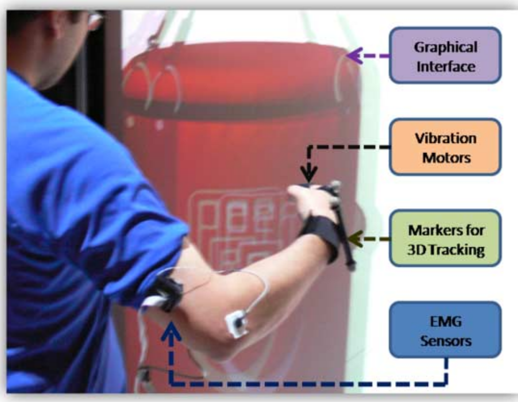


Figure 8. Multimodal rendering components of the free hand interaction system

6. Summary

This article presented the technologies and some results from virtual reality sports training literature. In particular it stressed the importance of motion tracking and analysis that make these kind of systems really useful in human skills training. It also shows how different feedback modalities can contribute to improve the users performances. The result is that virtual reality systems are proved to be valuable means to get competitive level in the sports training and can speed up the learning process and help the athletes identify their own limits. However, these kind of systems are not thought as complete substitutes of classical training but as a complementary training with controllable performances.

References

[1] O. Sandoval-Gonzlez, O. Portillo-Rodriguez, L. Montes-de-Oca, P. Tripicchio, E. Ruffaldi, C. A. Avizzano and M. Bergamasco, *Spatial Sound and Frequency Distortion for Enactive Transfer of Skills in Tai-Chi Movements to Visual Impaired People*, Enactive 2008.

[2] A.F. Bobick and A.D. Wilson. *A state-based approach to the representation and recognition of gesture*. 12, s.l. : Pattern Analysis and Machine Intelligence, IEEE Transactions, Dec. 1997, Vol. 19.

[3] M. Casdagli, S. Eubank, J. D. Farmer, J. Gibson. *State-space reconstruction in the presence of noise*. Physica D: Nonlinear Phenomena, Volume 51, Issue 1-3, p. 52-98, 1991.

[4] A.K. Jain, M.N. Murty and P.J. Flynn. *Data Clustering: A review*. ACM Computing Surveys, 1999, Vol. 31.

[5] O. Portillo-Rodriguez, O. Sandoval-Gonzlez, C.A. Avizzano, E. Ruffaldi, D. Vercelli, M. Bergamasco. *Development of a 3D Real Time Gesture Recognition Methodology for Virtual Environment Control*. ROMAN, 2008.

[6] O. Sandoval-Gonzlez, P. Tripicchio, E. Ruffaldi, A. Filippeschi, C. A. Avizzano and M. Bergamasco. *Skills Accelerators Using Multimodal Systems for Boxing Training*, SKILLS 2009

[7] K. Tsukada and M. Yasumura. *Activebelt: Belt-type wearable tactile display for directional navigation*. In UbiComp 2004, 2004.

[8] S. K. Nagel, C. Carl, T. Kringe, R. Martin and P. Konig. *Beyond sensory substitution learning the sixth sense*. Journal of Neural Engineering, 2(4):R13+, December 2005.

[9] J. Lieberman and C. Breazeal. *Tikl: Development of a wearable vibrotactile feedback suit for improved human motor learning*. Robotics, IEEE Transactions on, 23(5):919926, Oct. 2007.

[10] A. Bloomfield and N. I. Badler. *Virtual training via vibrotactile arrays*. Presence: Teleoper. Virtual Environ.,7(2):103120, 2008.

[11] J. B.F. Van Erp and H. A.H.C. Van Veen. *Vibrotactile in vehicle navigation system*. Transportation Research Part F: Traffic Psychology and Behaviour, 7(4-5):247 256, 2004.

[12] E. Ruffaldi, O. Sandoval-Gonzalez, A. Filippeschi, P. Tripicchio, A. Frisoli, C.A. Avizzano and M. Bergamasco, *Integration of Multimodal Technologies for a Rowing Platform*, ICM 2009.

[13] S. Kousidou, N. Tsagarakis, D.G. Caldwell, and C. Smith. *Assistive exoskeleton for task based physiotherapy in 3-dimensional space*. Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. The First IEEE/RAS-EMBS International Conference on, pages 266271, Feb. 2006.

[14] G. Gini , M. Mulas , M. Folgheraiter. *An emg-controlled exoskeleton for hand rehabilitation. Rehabilitation Robotics*, 2005. ICORR 2005. 9th International Conference on, pages 371374, July 2005.

[15] H. Touyama, K. Hirota, and M. Hirose. *Implementation of electromyogram interface in cabin immersive multiscreen display*. Virtual Reality Conference, 2006, pages 273278, March 2006.

[16] E. Ruffaldi, A. Filippeschi, M. Verlet, C. Hoffmann, and B. Bardy,*Skill Training in Multimodal Virtual Environments*. Taylor, 2012, no. ISBN 978-143-987-895-8, ch. *Design and Evaluation of a multimodal VR Platform for Rowing training*.

[17] P. Tripicchio, O. Sandoval-Gonzalez, A. Filippeschi, E. Ruffaldi, C. A. Avizzano and M. Bergamasco, *Human forces in hands free interaction: a new paradigm for immersive virtual environments*. RO-MAN 09.

[18] A.M. Murray, R.L. Klatzky, and P.K. Khosla. *Psychophysical characterization and testbed validation of a wearable vibrotactile glove for telemanipulation*. Presence: Teleoperators and Virtual Environments, 12(2):156182, 2003.

[19] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon and J. C. Hart. *The CAVE: Audio Visual Experience Automatic Virtual Environment*, Communications of the ACM, vol. 35(6), 1992, pp. 6472.

[20] E. Ruffaldi, A. Frisoli, M. Bergamasco, C. Gottlieb, F. Tecchia. *A haptic toolkit for the development of immersive and web-enabled games*,Proceedings of the ACM symposium on Virtual reality software and technology, VRST 06

[21] C. A. Avizzano, P. Tripicchio, L. Joale and M. Bergamasco. *Design of a Motion Based Sailing Simulator*, RO-MAN 2012

This page intentionally left blank

2nd Int. Workshop on Reliability of Intelligent Environments (WoRIE'12)

Miguel Hornos (Universidad de Grandada, Spain)
Juan Carlos Augusto (University of Ulster, UK)
Pablo A. Haya Coll (Universidad Autónoma de Madrid, Spain)

This page intentionally left blank

J.A. Botía et al. (Eds.)
IOS Press, 2012

Introduction to the Proceedings of WoRIE'12

This section of the Workshops Proceedings presents the selected contributions for the *2nd International Workshop on the Reliability of Intelligent Environments* (WoRIE 2012), to be held within the *8th International Conference on Intelligent Environments* (IE'12) in Guanajuato (México) on 26 of June 2012. The previous edition of this workshop was held in Salamanca on April 2011, within the *2nd International Symposium on Ambient Intelligence* (ISAmI'11). The idea is to consolidate this event with future editions, by creating a strong and solid bridge between both intelligent environments developers and software engineering communities with the aim of collaboratively working on the topics included in the workshop.

It is undeniable the relevance that Intelligent Environments will play within the important changes which are foreseen in the society for a near future. Many of them are due to the ageing of the population of developed countries. Consequently, many governments look ahead and are allocating funds for strategically critical areas in order to identify key issues and find effective solutions. Not surprisingly, many institutions are launching research and development programs focused on health-care, elderly people, quality of life, social inclusion, education, etc. Not only innovation is required for systems supporting such a new assisted, interactive and collaborative world, but also the reliability increment of such systems, especially in those cases in which safety is a critical issue, where a small error in their design can even put human lives at risk. Think for example on the development of an intelligent environment to support independent living for vulnerable (e.g. elderly, disabled or chronic ill) people, which usually includes tasks such as detecting whether the house occupant has fallen or is unwell in some way. Moreover, since these intelligent environments have a multidisciplinary nature, involves very diverse and complex technologies, and even considers distinct requirements for different user groups, their development is a very difficult and error prone challenge. It is therefore necessary to apply formal specification, modelling, verification, simulation or some other software engineering method or technique included in the topics of this workshop to analyze and establish the correctness of such systems, as well as to increase their reliability.

Most of the selected papers include both theoretical contents and empirical or real case studies to support the validity of the proposals, which are addressed to different scopes and applications, such as middleware, privacy management, user interface modelling, simulation and model checking, but all of them with the clear objective of increasing the reliability of the intelligent environments to which they are applied. Moreover, we will have the keynote entitled *Towards Reliability of Intelligent Environments through Testing Services and Applications*, by Juan A. Botía, to whom we thank for accepting our invitation to be our invited keynote speaker of this edition.

We hope the keynote and selected papers will contribute to a successful workshop with the audience, presenters, and researchers participating in mutually beneficial discussion. We sincerely trust that the workshop will be an appropriate forum where the sharing of knowledge and experiences about the reliability of intelligent environments and related systems (such as Pervasive/Ubiquitous Computing systems, Ambient Intelligence systems, Smart Environments, Multi-Agent Systems, etc.) promotes new advances in both research and development. We also expect that readers enjoy the papers included in the proceedings.

We would like to sincerely thank IE'12 Workshops Chairs for their constant help and support, WoRIE'12 Program Committee members for their excellent work and invaluable support during the review process, and most importantly authors of the submitted papers for their very interesting and high quality contributions, which have made possible to successfully organize the present edition of this workshop.

Miguel J. Hornos, Juan Carlos Augusto and Pablo A. Haya
Co-chairs of WoRIE'12

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-387

Towards Reliability of Intelligent Environments through Testing Services and Applications

Juan A. BOTÍA [a,1]
[a] *Departamento de Ingeniería de la Información y las Comunicaciones, Universidad de Murcia, Spain*

Outline of the Keynote

An intelligent environment (IE) is one of the most complex distributed systems existing nowadays. As a distributed system, it is not merely a set of computers working together but it also includes mainly sensors and actuators in the broad sense of the term. If we assume that a system is reliable when the individual pieces' reliability is crucial for the system being reliable as a whole, then a simple fault in a concrete hardware element can make the whole system unavailable. Moreover, the diversity of hardware standards available in the industry worsen the situation. Unfortunately, the problem with software is not simpler. In IEs, it is usually the case that a high number of heterogeneous components, from different vendors, must be integrated together to enable interaction between them. For example, there are a variety of discovery protocols to allow automatic discovery of new components and services. Choosing one single discovery protocol implies the impossibility of using other components not implementing such protocol. And the final result is, usually, an IE that integrates different service discovery protocols using a wrapper based approach for its integration. The same situation may occur when using different kinds of motes for sensors and their corresponding network protocols. It is the same case for software components developed with different APIs. Should the IEs engineer consider all these different aspects when addressing reliability? It is actually realistic to assume that an intelligent environment equipped with components from no unique vendor may be reliable?

But also, reliability has to be considered regarding the intelligent behaviour an IE must show. An IE can be seen as a system devoted to act on behalf of the user or set of users, with the aim of giving some kind of assistance. The specific assistance given will depend on the features of the physical environment the system is deployed at (e.g. the working place, at home, in a small city). But all services have something in common: they behave in an intelligent manner. And such intelligence comes from their ability to manage the user's context [1]. Context is maintained with the aim of knowing the what, where, when and why about the user' state. This knowledge about the user enables

[1]Facultad de Informática, Campus de Espinardo. Murcia, Spain. E-30071. E-mail: juanbot@um.es

proactivity on the system's behaviour. How to measure the effectiveness of an IE to act depending on context? Are there methodologies for the engineering of context-aware systems that can be used to face the problem with guarantee of success?

In the following years, we will face a number of exciting challenges that will be key to the industrial success of IE's. And all of them are related in some way to the reliability concept. One of the most interesting ones is about how to fill the gap between a system developed at the lab and a system deployed in a real scenario working 24 hours a day and 7 days a week? What are the approaches for intensively testing services and applications at the lab, in order to minimise the probability of fault in production mode? All these questions will be hopefully addressed and discussed at the WORIE'12 event to be held at Guanajuato, with the main aim of identifying the most challenging research questions for the next years.

Acknowledgements

This work is partially supported by the Spanish national research project TIN2011-28335-C02-02, Foundations for the Development of AAL Services and Applications, through the Bureau of Innovation and Economy

References

[1] A.K. Dey, Understanding and using context, *Personal and Ubiquitous Computing* **5** (2001), 4-7.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-389

MidBlocks: A Supervising Middleware for Reliable Intelligent Environments

Rafael BAQUERO[a,1], José RODRIGUEZ[a], Sonia MENDOZA[a]
and Dominique DECOUCHANT[b]
[a] *Departamento de Computación, CINVESTAV, México*
[b] *Univarsidad Autónoma Metropolitana, Unidad Cuajimalpa, México*

Abstract. Intelligent Environments (IEs) are expected to have a dramatic impact on our daily lives in the near future. Environments that can assist users in their daily activities, aid the elderly in continuing productive and independent lives, keep people safe and help them make better use of resources can provide great benefits. However, due to the multidisciplinary nature of IEs, the complexity of the technologies involved, and the distinct requirements of different user groups, the development of full scale IEs is a difficult and error prone task. Although a robust design is a fundamental part in the development of an IE, proper supervision of the system for failure detection is important to guarantee the reliability of the systems. In this paper we propose MidBlocks, an event-based component supervising middleware. Through the use of MidBlocks, Intelligent Environments can perform constant supervision of their components and communication links in order to warn users in case of a system malfunction or to perform corrective actions such as dynamic system reconfiguration.

Keywords. Intelligent Environments, Middleware, Distributed Systems, Component Supervision

Introduction

Intelligent Environments are heterogeneous distributed sensor-actuator systems including multimedia presentation services, building automation and control components, intelligent physical objects, wireless sensor network nodes, nomadic personal or shared devices, and many other systems and entities [1]. It is expected that IEs will enrich the lives of humans by adapting the environment to the inhabitant's activities and by providing services such as assistance to users with special needs. In some cases, failure of the IE can have serious consequences to the inhabitants of the environment. Therefore a fundamental issue in the development of any IE is dependability. Dependability is the ability of a system to deliver service that can justifiably be trusted [2]. However complex systems can, and sometimes do, fail. Being able to detect system failures to warn users or to take corrective actions is an important aspect of any IE.

Due to their nature, Intelligent Environments are complex systems that require the integration of multiple components. A frequent solution employed to link these

[1] Corresponding Author: Rafael Baquero S., Depto. de Computación, CINVESTAV, Av. IPN 2508, San Pedro Zacatenco, México D.F., México, 07360. E-mail: rbaquero@computacion.cs.cinvestav.mx

components is the use of a middleware. A middleware is a communications mechanism through which the components of a distributed system can interact. To improve reliability, some systems such as industrial control systems employ the communications mechanism to simultaneously perform component supervision tasks.

In this paper we present MidBlocks, an event-based publish/subscribe middleware which provides a communication mechanism for distributed systems while simultaneously supervising the components of the system.

The rest of this paper is organized as follows. In section 1 we describe the basic operation of event-based publish/subscribe middleware. Next in section 2 we mention other work performed in the area of AmI systems dependability. Afterwards, in section 3, we describe the architecture and operation of MidBlocks. In section 4 we provide a brief description of FunBlocks, a modular, minimalist framework for the integration of AmI systems which uses MidBlocks. Finally in section 5 we describe some future work which we plan on carrying out on MidBlocks.

1. Event-Based Systems

A middleware is a software layer which allows the components of a distributed system to interact [3]. Many different types of middleware have been developed, each based on a different paradigm, and targeted towards the solution of a certain class of problems.

A common class of middleware is the event-based publish/subscribe model. In an event-based middleware entities exchange information through discrete packets called events [4]. An event represents any discrete transition that has occurred and is signaled from one entity to a number of other entities. In the publish/subscribe model producers publish the information they can generate on an event manager, while consumers subscribe to the information they want to receive. The middleware can further be provided with a store and forward entity. When a store and forward entity is used producers send events to the entity that stores them and forwards them to consumers. The use of an event-based, publish/subscribe middleware with a store and forward mechanism allows full space, time, and synchronization decoupling between event producers and consumers. This type of middleware is used in applications that have strong requirements in terms of reliability but do not need a high data throughput [5].

2. Dependability in Middleware and Ambient Intelligence Systems

Dependability in middleware and Ambient Intelligence systems has been a subject of interest in the last few years. In [6] Rice and Beresford explore build dependable pervasive computing applications which utilize a distributed middleware. The authors define a dependable system as one that either performs within specification, or provides suitable feedback to users when faults occur. They use accountability as a mechanism for constructing a dependable system, where accountable applications are capable of describing why a particular action was (or was not) taken.

In [7] Coronato presents Uranus, a middleware aimed at the development of Ambient Assisted Living and vital signs monitoring applications. The objective of Uranus is to provide a middleware infrastructure for the rapid prototyping of Ambient Intelligence applications for healthcare, with a certain degree of dependability. Uranus is capable of providing some self-organizing capabilities through the use of monitors such as a Connection Monitor and a Battery Monitor. Uranus has been used in the development of a Long-Term Monitoring application and a Smart Hospital application.

3. MidBlocks: A Component Supervising Middleware

As mentioned before, a solution to improve the reliability of a system is to perform supervision functions through the communication mechanism used for component interaction. Examples of this type of solution are the pneumatic control systems used in early building automation and the popular 4-20mA current-loop popular in industrial measurement and control systems.

In early building automation systems compressed air was used to transmit information between the components of the system and simultaneously supervise the system. To achieve this, components were linked with hoses that carried compressed air which varied in the range of 3-15psi. 3psi represents a live zero while 15psi represents 100%. Any pressure below 3psi is a dead zero and constitutes an alarm condition [8].

The same principle is used in the 4-20mA current-loop. In this scheme a current flowing through a wire to the sensor or actuator is varied between 4 and 20mA, with 4mA representing a live zero and 20mA representing 100%. If the current drops below 4mA this signals either a device failure or an open link while a current above 20mA signals a device failure or a short circuit [8, 9].

To supervise components in distributed systems, a frequent solution is to use heartbeat messages. In this type of scheme a special kind of message, called a heartbeat message, is periodically sent by the component to other components in the system. The main disadvantage of this type of scheme is that depending on the number of system components, the frequency with which heartbeat messages are sent and the bandwidth of the communications links, the load on the communications infrastructure can be significant.

The event based store and forward middleware is particularly well suited to provide a supervising mechanism without the need of a separate heartbeat message. Events generated by the system components can be considered as a type of heartbeat message. This coupled with the need of an event management entity for the store and forward mechanism, which can simultaneously be used to supervise the interval between events, provides the basis for a complete component supervision solution. This is the principle we use in MidBlocks, an event based, store and forward, component supervising middleware.

3.1. MidBlocks Description

MidBlocks is an event based, store and forward type of middleware in which the store and forward entity additionally performs component supervision tasks. Every component of a MidBlocks based system specifies a maximum time interval between the events that it generates. If a component exceeds this maximum time interval, called the Maximum Event Interval (MEI), MidBlocks will generate a failure event notifying the component's failure to those components that have subscribed to failure events.

A useful feature of MidBlocks is that it does not assume any particular type of communication mechanism or underlying hardware. To achieve this MidBlocks makes use of a layered design in which events are formed independent of any communications mechanism or hardware assumptions and as the events descend through the layers information required by the specific communication mechanism and hardware chosen is added. As a result MidBlocks can be used to provide component supervision in a large array of different systems.

In order to use MidBlocks in a particular system only two elemental requirements have to be satisfied:

- Each component of the system must be uniquely identifiable.
- A communication link must be present, capable of sending and receiving discrete data packets between the components and the store and forward/component supervising module.

3.2. MidBlocks Events

All components in MidBlocks must be assigned a unique identifier. This way event messages can be related to the component from which the message originated. Component identifiers in MidBlocks are 48 bits long analogous to the MAC address used in IEEE 802 networks. These identifiers can either be assigned statically during system installation or dynamically if an appropriate component registration mechanism is provided. The only requirement is that the identifiers are system wide unique.

All MidBlocks events have the same basic structure which is the 48 bit component id, followed by an 8 bit event type id and, depending on the type of event, additional event data (figure 1).

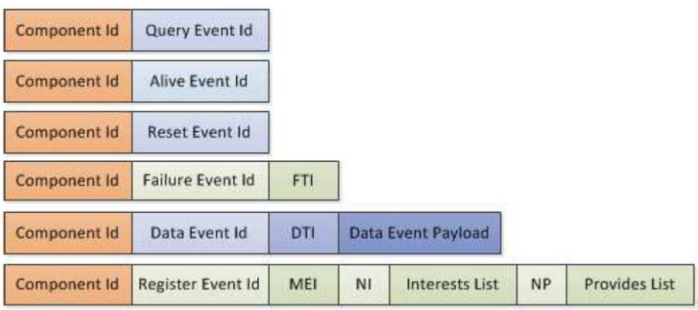


Figure 1. MidBlocks event structure.

MidBlocks uses the following six types of events:

- *Query events.* When a component has exceeded its maximum MEI the Component Supervisor (CS) module sends a query event to the component to determine if it is responsive. If the component does not respond within an MEI then a critical failure event is sent to those components subscribed to failure events.

- *Alive events.* There are two type of Alive events with different event type id that can be issued by components. The first one is issued in response to a query event. While the second one is issued when the component has not generated any events and its MEI is close to expiring.

- *Reset events.* When an event is received by a component that is not registered with the Message Receiver then a Reset event is issued immediately back to the component. Upon reception of a Reset event the component should clear any pending outbound events and perform its registration procedure.
- *Failure events.* Failure events are generated when messages are not received or cannot be delivered to components. Failure events have an 8 bit Failure Type Id (FTI) which indicates the type of failure detected.
- *Data events.* Data events are the normal events exchanged between components and the store and forward entity in MidBlocks. After the event type id, Data events have a 64 bit Data Type Id field. The possible values of the DTI are application specific and are assigned by the system developers using MidBlocks. After the DTI field comes the actual event data. MidBlocks does not assume any particular structure or size for this data.
- *Register events.* Register events are used to register components with the store and forward/component supervision entity. MidBlocks handles two values for the Register Event Id. The first value is used for components that have interest in receiving failure events, while the second value is used for components that do not need to be notified of failure events. Following the event type id Register events have the following fields:
  1. *MEI.* This is a 32 bit field containing the Maximum Event Interval for this particular component expressed in milliseconds. This value is sent using network byte order.
  2. *NI.* The Number of Interests field (NI) is a 64 bit field indicating the number of data event types that this component has interest in. This value is sent using network byte order.
  3. *Interest List field.* This field lists the DTIs of all the data types that this component is interested in receiving.
  4. *NP.* The Number of Provides field (NP) is a 64 bit field indicating the number of data event types that this component provides. This value is sent using network byte order.
  5. *Provides List field.* This field lists the DTIs of all the data types that this component provides.

3.3. MidBlocks Architecture

MidBlocks can be broadly separated into two distinct parts (see figure 2): A component side part (CMPS) and a Store and Forward/Component Supervising entity (SFCS). As with any other type of store and forward event-based middleware the primary task performed by MidBlocks is to deliver events from producers to consumers. However, unlike other types of middleware MidBlocks supervises the components to ensure that at least one event is sent during a components MEI and that events are being accepted by the components. To achieve this MidBlocks requires event message supervising functions which constantly monitor the events being accepted from components by the SFCP and the events delivered to components from the SFCP. In order to perform adequate supervision MidBlocks does not support direct component to component interaction.

In the following sections we will begin by describing the operation of the CMPS part of MidBlocks and afterwards describe the SFCS entity.

3.3.1. MidBlocks CMPS

When a component is first introduced into the system it must perform a registration procedure. The purpose of the registration procedure is to allow the SFCS to configure the queuing and component supervision mechanisms adequately to handle the events of this component and to supervise the component. Registration of the component with the SFCS is performed by the CMPS Component Registration Module (CMPS-CR) upon request of the component's main process (see figure 3). The registration sequence has to be initiated by the main process in order to ensure that the communications hardware and software have been properly initialized and are able to send and receive messages. Components must use one of two Register Event Ids to indicate whether or not they are interested in receiving Failure events.

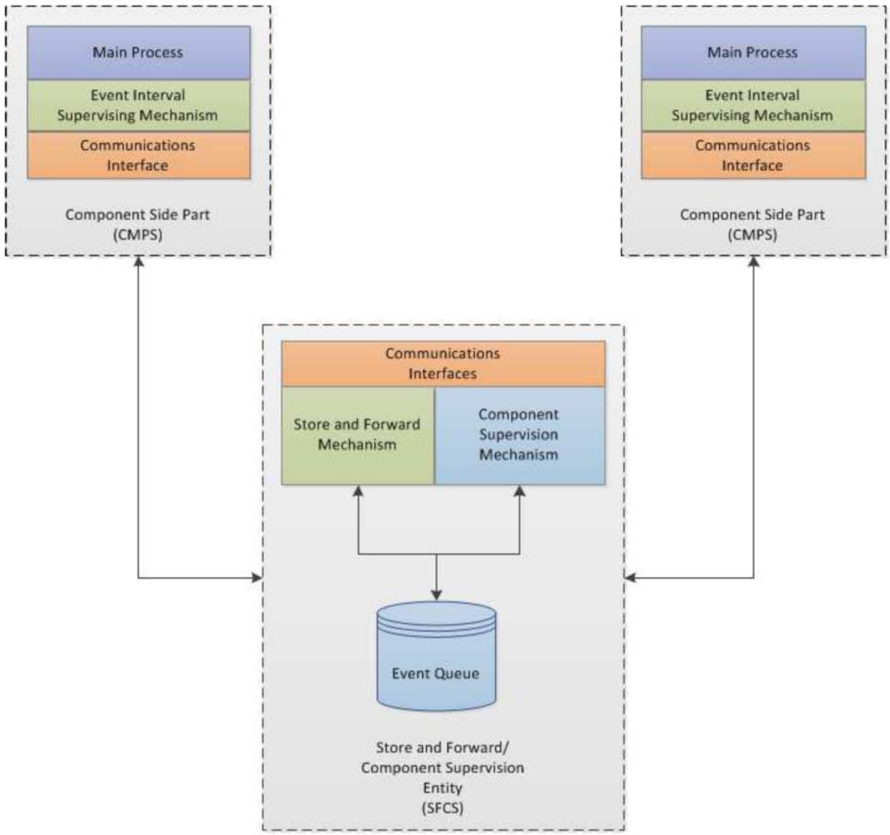


Figure 2. MidBlocks architecture.

Inbound event messages arrive at the CMPS from the SFCS through a communications interface and are received by the CMPS Message Receiver (CMPS-MR). Four types of events can be received by a component: Query, Reset, Data, and Failure events. Failure events are received only by components that have expressed interest in receiving this type of events by using the corresponding Register Event Id during their registration procedure. Depending upon the type of event received by the CMPS-MR the following tasks are performed:

- *Query event.* When a Query event is received, the CMPS-MR immediately notifies the CMPS Message Dispatcher/MEI Supervisor (CMPS-MD) to send an appropriate Alive event. The Query event is then discarded and no further processing of the event occurs.
- *Reset event.* Upon reception of a Reset event the following sequence of actions occurs:
  1. CMPS-MR signals the CMPS-MD to block any further event message sending, with the exception of Register events.
  2. The CMPS-MR then introduces the event into the CMPS Local Event Queue (CMPS-EQ) and will discard any further Reset events until it is signaled by the CMPS-MD that a Register event has been sent.
  3. The CMPS-MD will not accept any events from the component's Main Process until it receives a Register event from the CMPS-CR.
  4. When the component's Main Process extracts the event from the queue it must signal the CMPS-CR to perform the registration procedure to register the component with the SFCS.
- *Data and Failure events.* Data and Failure events are placed directly in the CMPS-EQ for processing by the component's Main Process.

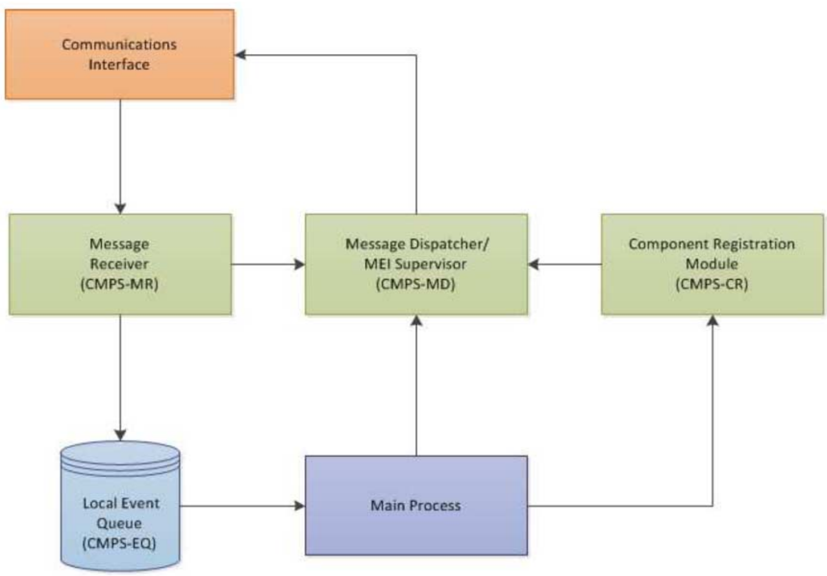


Figure 3. MidBlocks component side (CMPS) block diagram.

Outbound data events are sent through the CMPS-MD and the communications interface. The CMPS-MD monitors the time interval since the last event was sent and, in case it approaches the component's MEI, it issues an Alive event to signal that the component is operating correctly.

3.3.2. MidBlocks SFCS Event Reception

On the SFCS side (see figure 4) event messages are received by the SFCS Message Receiver (SFCS-MR) from the communications interfaces. Three types of messages can be received by the SFCS: Alive, Register and Data events. The first task performed by the SFCS-MR is to verify the event type and the Component Id. If the event is not a Register event and the component's data is not registered in the SFCS Component Register (SFCS-CR) then a Reset event for the component is issued through the same communications interface from which the message arrived. Since Reset events are not entered into the SFCS Event Queue (SFCS-EQ) and are not sent by the SFCS Message Dispatcher (SFCS-MD) the data in the Component Id field of the Reset event is not important.

After verifying that the event is valid one of the following actions is performed depending on the event type:

- *Alive and Register events.* Alive and Register events are directly placed in the message queue without any further processing from the SFCS-MR.
- *Data event.* When the event received is a Data event the SFCS-MR scans the SFCS-CR to determine which components are subscribed to the type of data event received. Next the SFCS-MR attaches a Time-To-Live (TTL) field to the event, and inserts one copy of the event into the SFCS-EQ for each recipient. The inserted events are tagged as New events prior to insertion as explained in the following section.

3.3.3. MidBlocks SFCS Event Processing

Events in the SFCS-EQ can be tagged with four states:

- *New event.*
- *Ready for Delivery.*
- *Unable to Deliver.*
- *Processed.*

Any event in the SFCS-EQ can only be in one of these states at any given time.

The SFCS Component Supervisor (SFCS-CS) constantly scans the SFCS-EQ for events that require processing. Depending on the event type and the state of the event one of the following actions is performed:

- *New event tag - Alive event.* When the SFCS-CS encounters an Alive event tagged as a New event, it resets the timer associated with the producer component and tags the event as Processed. If the Alive event was issued in response to a Query event then the SFCS-CS inserts a Failure-Non critical event into the SFCS-EQ for each Failure event subscriber. Failure events have the same structure as Data events in the SFCS-EQ meaning that the event has a TTL field and a state tag. When the Failure event is created it is tagged as Ready for Delivery.
- *New event tag - Data event.* Data events are handled in much the same way as Alive events. The SFCS-CS resets the timer associated with the producer component and tags the event as Ready for Delivery.
- *New event tag - Register event.* On encountering a Register event tagged as a New event, the SFCS-CS creates and initializes a new timer for the component and inserts the component's data into the SFCS-CR. The Register event is then tagged as Processed.

- *Ready for Delivery*. Events tagged as Ready for Delivery are not handled by the SFCS-CS but are handled by the SFCS-MD. How the SFCS-MD handles Ready for Delivery events is described later on in this section.
- *Unable to Deliver*. This tag indicates that the SFCS-MD was unable to deliver the event after TTL attempts. When encountering a message tagged as Unable to Deliver the SFCS-CS performs the following actions:
  1. The destination component's data is removed from the SFCS-CR.
  2. Any events in the SFCS-EQ destined for the component, including the current event, are tagged as Processed.
  3. Failure events of type Critical are inserted into the SFCS-EQ indicating that the component has failed.
- *Processed*. Processed are handled by the SFCS Garbage Collector (SFCS-GC). The SFCS-GC scans the SFCS-EQ for events tagged as Processed and removes the events from the SFCS-EQ.

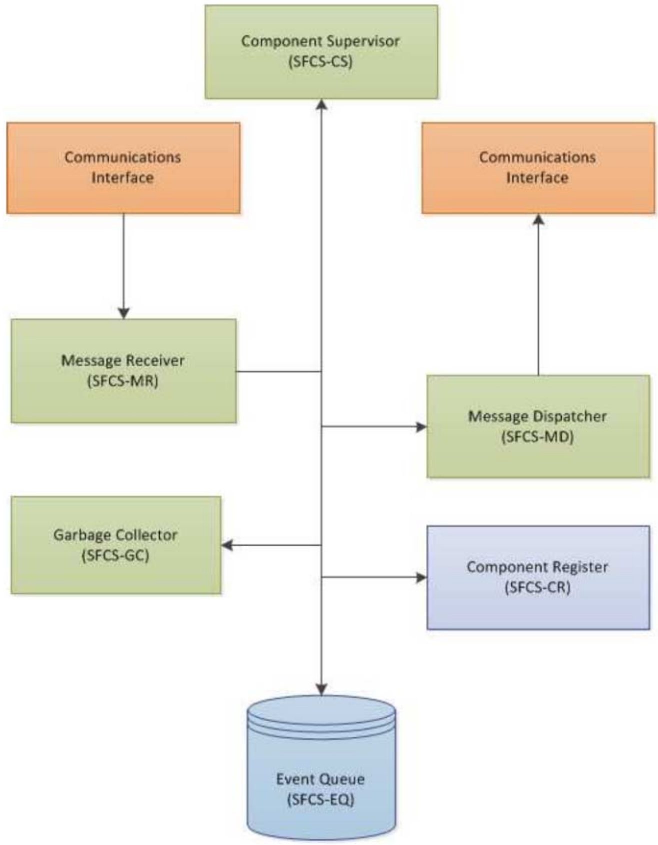


Figure 4. MidBlocks Store and Forward/Component Supervision (SFCS) entity block diagram.

The SFCS-CS also keeps track of the time elapsed since component's last event was received. If a component exceeds its MEI the SFCS-CS inserts a Query event destined for the component into the SFCS-EQ. Additionally the SFCS-CS inserts a non-critical failure event into SFCS-EQ and resets the timer associated with the component.

3.3.4. MidBlocks SFCS Event Delivery

Data events marked as Ready for Delivery are read from the SFCS-EQ by the SFCS-MD on a FIFO basis and delivered to consumers via the appropriate communications interface. If the event is delivered successfully then it is tagged as Processed, otherwise the event's TTL field is decremented by one. If the event's TTL field reaches zero then the event is tagged as Unable to Deliver and no further attempts will be made to deliver the event.

3.4. MidBlocks Communications Interfaces

To allow use with a wide variety of devices and communications technologies, MidBlocks employs a 3 layer design for its communication interfaces as shown in figure 5.

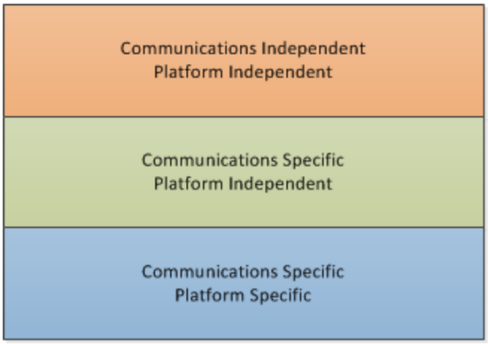


Figure 5. MidBlocks communications interface layers.

In the first layer event messages are formed without using communication specific data such as IP addresses or node ids. Once the core event messages are formed, communication specific parameters such as IP addresses are added to the message. Finally, in the last layer, platform specific routines are invoked to deliver the message.

Currently, the CMPS can only use a single communications interface in each component, however the SFCS can use multiple communications interfaces both for inbound as for outbound messages. This allows the SFCS to function as a gateway which permits components with different types of communication mechanisms to exchange events.

4. Current Projects with MidBlocks

A reduced version of MidBlocks is currently being used in FunBlocks [10-12]. FunBlocks is an event based, minimalist, modular framework for the development of Ambient Intelligence systems. Figure 6 shows a diagram of the basic components of FunBlocks

FunBlocks approaches the development of AmI systems from the point of view of distributed control systems. FunBlocks makes use of the function block abstraction described in the IEC 61499 standard for distributed control systems [13, 14].

Developing an AmI system using the FunBlocks framework consists of developing or reusing function blocks which provide the desired functionality, and afterwards joining these blocks through the services provided by the framework.

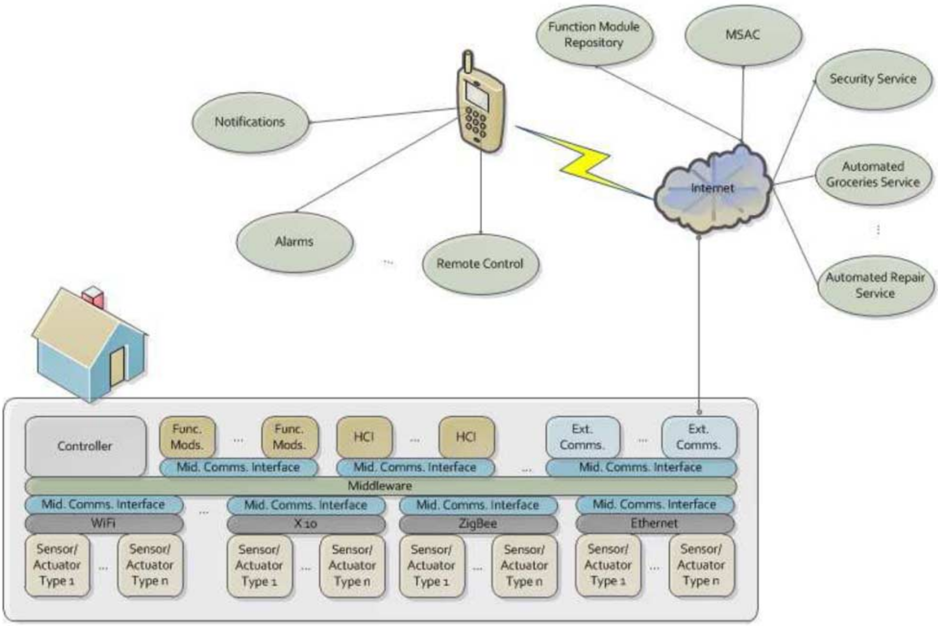


Figure 6. FunBlocks diagram

Thanks to its communications technology independent design, the use of MidBlocks allows FunBlocks to handle different types of sensor and actuator communications protocols such as WiFi, ZigBee and RS-485 buses.

5. Conclusions and Future Work

Although MidBlocks novel design increases the reliability of a system due to its components and communication link supervision functions a weak link in the design is the use of a single Store and Forward/Component Supervising entity. The SFCS represents a single point of failure in the system. To improve the reliability of a MidBlocks based systems the design of MidBlocks must be improved to accommodate redundant or cooperating SFCS.

Asides from the reliability increase to MidBlocks based IEs, the use of multiple SFCS brings added flexibility not currently available in MidBlocks. The use of multiple SFCS would allow, for example, the development of IEs with self-configuring and self-healing capabilities.

References

[1] L. Roalter, A. Möller, S. Diewald, and M. Kranz, Developing Intelligent Environments: A Development Tool Chain for Creation, Testing and Simulation of Smart and Intelligent Environments, *Seventh International Conference on Intelligent Environments*, Nottingham, United Kingdom, 2011, 214-221.

[2] A. Avižienis, J.-C. Laprie, and B. Randell, Fundamental Concepts of Dependability, Newcastle University CS-TR-739, 2001. Available: http://www.cs.ncl.ac.uk/research/trs/papers/739.pdf (May 5, 2012).

[3] P. A. Bernstein, Middleware. A Model for Distributed System Services, *Communications of the ACM* **39 (**1996), 86-98.

[4] S. Tarkoma and K. Raatikainen, State of the Art Review of Distributed Event Systems, Helsinki University Computer Science Department, 2006. Available: http://www.minema.di.fc.ul.pt/reports/MinemaEventsReport_final.pdf (May 5, 2012).

[5] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec, The Many Faces of Publish/Subscribe, *ACM Computing Surveys* **35** (2003), 114-131.

[6] A. C. Rice and A. R. Beresford, Dependability and Accountability for Context-Aware Middleware Systems, *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops*, 2006, 378-382.

[7] A. Coronato, Uranus: A Middleware Architecture for Dependable AAL and Vital Signs Monitoring Applications, *Sensors* **12** (2012), 3145-3161.

[8] Building Automation Products Inc., Understanding 4-20 mA Current Loops [Application Note]. Available: http://www.bapihvac.com/CatalogPDFs/I_App_Notes/Understanding_Current_Loops.pdf (Jan 14, 2012).

[9] National Semiconductor, Application Note 300. Simple Circuit Detects Loss of 4-20 mA Signal [Application Note 300]. Available: http://www.ti.com/lit/an/snoa605a/snoa605a.pdf (Jan 14, 2012).

[10] R. Baquero, J. G. Rodríguez, S. Mendoza, and D. Decouchant, Towards a Uniform Sensor Handling Scheme for Ambient Intelligence Systems, *8th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE 2011)*, Mérida Yucatán, México, 2011, 952-957.

[11] R. Baquero, J. G. Rodríguez, S. Mendoza, and D. Decouchant, Towards a Modular Scheme for the Integration of Ambient Intelligence Systems, *5th International Symposium on Ubiquitous Computing and Ambient Intelligence (UCAmI 2011)*, Riviera Maya, Mexico, 2011.

[12] R. Baquero, J. G. Rodríguez, S. Mendoza, and D. Decouchant, FunBlocks. A Modular Framework for AmI System Development, *Sensors,* 2012 (Submitted).

[13] R. Lewis, *Modelling Control Systems Using IEC 61499*, The Institution of Engineering Technology, England, 2008.

[14] V. Viatkin, *IEC 61499 Function Blocks for Embedded and Distributed Control Systems Design*, International Society of Automation, Research Triangle Park, North Carolina , U.S.A., 2007.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-401

Consistency in Context-Aware Behavior: a Model Checking Approach

Davy PREUVENEERS [a,1], Yolande BERBERS [a]
[a] *IBBT-DistriNet, KU Leuven, 3001 Leuven, Belgium*

Abstract. Context-aware systems have been in the research epicenter for more than a decade. By sensing, reasoning and acting on various kinds of relevant situational information, context-aware systems can make well-informed decisions and manifest autonomous behavior through pre-defined and user customizable rules that define what to do in particular circumstances. However, there is still a significant gap between deploying such context-aware systems in lab environments under ideal circumstances and deploying them in the real world with non-specialist end-users. These end-users may experience a lack of understanding or frustration when a system does not behave as expected, especially in the presence of unforeseen human interventions, exceptional circumstances and unexpected events. This paper investigates challenges and opportunities of using model checking as an approach for improving the reliability and consistency of smart environment applications whose behavior is driven by context-aware adaptation rules, and reports on initial results with the SPIN model checker.

Keywords. context, event-condition-action rules, model checking

Introduction

In the ubiquitous computing paradigm, devices and applications are able to interact with one another and often have an awareness of the situation of their users to create a smart environment that is proactive and supportive. Context-aware systems [1] are able to adapt their behavior to the current context without explicit user intervention by taking environmental context into account. As Dey [2] stated, context is any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and the application themselves.

However, the vision of ubiquitous computing as described in Mark Weiser's seminal article has still not become reality. The complexity of integrating ubiquitous computing in our daily lives and making our environment intelligent, proactive and, moreover intuitive to use, has proven to be very high. As these ubicomp features interact with each other in sophisticated ways [3], so does the complexity that will adversely impact the reliability [4] of the smart environment. Recent work [5] confirms that little is reported in the literature on methodologies to increase the reliability of software for the devel-

[1]Corresponding Author: Davy Preuveneers, Dept. Computer Science, Celestijnenlaan 200A, B-3001 Heverlee, Belgium; E-mail: davy.preuveneers@cs.kuleuven.be.

opment of intelligent environments. Assistive technologies are often too naively or too optimistically developed assuming that users always understand what these systems are doing or that systems know what users want.

In this paper, we will address the challenge of robustness and mitigating the increased potential for failure in a smart environment that evolves over time (i.e. in a dynamic context). We will particularly focus on smart context-aware systems whose dynamic behavior is driven by context-aware decision and adaptation rules (following the typical *event-condition-action* paradigm). The complexity of ensuring robustness is caused by the fact that ensuring consistency among these rules is not straightforward and that many features may interact with one another [6]. Robust intelligent systems [7] ensure acceptable performance, not only under ordinary conditions but also under unusual conditions that stress the developers' assumptions of the system. However, it is difficult for a developer *at design time* to discover and eliminate all errors when the application will be active outside the intended limited context. As a result, fragile context-aware applications are subject to subtle or more severe errors that only make their presence known in unusual circumstances. This problem is aggravated when users can customize *at runtime* the contextual triggers (event-condition-action rules) that modify the behavior of a ubicomp system, causing inconsistencies among these rules (e.g. turning on the heating and the air conditioning at the same time).

In section 1, we present a use case scenario in the smart home environment from which we distill some context-aware adaptation rules. These rules will be used in section 2 to discuss more in detail the kind of inconsistencies that may emerge from these rules. In section 3 we will elaborate on our model checking approach and present some qualitative results with the SPIN model checker in section 4. We conclude and propose further work in section 5.

1. Motivating Scenario: a Smart Home Use Case

Jack and Jill have installed a smart home automation system. It connects all the devices and appliances in their home so they can communicate with each other and also with their owners. Their ZigBee-based smart home solution provides applications related to lighting, home security, home theater and entertainment, and thermostat regulation. It is equipped with sensors to measure the temperature at different places in the house, motion sensors to detect someone's presence, ambient light sensors, as well as other sensors for outdoor use. The home automation system also comes with switches and dimmers to control the curtains and window blinds, the lighting, the heating and the air conditioning in each room. A control panel provides easy remote access to each device.

Jack and Jill set up the system to dynamically adapt to the current context rather than merely relying on pre-programmed timers. Sensor values will drive the behavior of the home automation system. Initially, they would like to configure the basic lightning, heating and air conditioning. They want the home automation system to behave according to the following preferences:

- **Heating:** The heating should only be turned on if it is less than 19 °C inside the house. Sensors in the house will monitor the environment and the thermostat will aim for a temperature of 21 °C.

$$\textbf{if}\ (temperature_{in} < 19^{\circ}C)\ \textbf{then}\ heating = on \qquad (1)$$

$$\textbf{if } (temperature_{in} > 21°C) \textbf{ then } heating = off \tag{2}$$

- **Lighting:** When it is getting dark inside, the lights should be turned on only in those rooms where people are present. The lights should be turned off automatically if no presence is detected for more than 5 minutes.

$$\textbf{if } (brightness_{in} = low) \textbf{ and } (motion = true) \textbf{ then } lights = on \tag{3}$$

$$\textbf{if } (motion = false)_{\Delta t = 5min} \textbf{ then } lights = off \tag{4}$$

- **Air conditioning:** If it is getting more than 27 °C in the house, then turn on the air conditioning until the temperature drops to 24 °C. Motion sensors will scan the room for occupancy and when no movement is detected it will send a signal to the air conditioning unit to switch it off.

$$\textbf{if } (temperature_{in} > 27°C) \textbf{ then } airconditioning = on \tag{5}$$

$$\textbf{if } (temperature_{in} < 24°C) \textbf{ then } airconditioning = off \tag{6}$$

$$\textbf{if } (motion = false) \textbf{ then } airconditioning = off \tag{7}$$

Later on, Jack and Jill also add outdoor sensors to their smart home automation system and they want to update the previous behavior and objectives with the following automation preferences:

- **Lightning:** If it is daylight outside, open the curtains or window blinds and also turn off the lights.

$$\textbf{if } (brightness_{out} = high) \textbf{ then } curtains = open \textbf{ and } lights = off \tag{8}$$

- **Air condition:** If it is more than 30 °C outside, close the curtains and window blinds to reduce the temperature increase inside the house.

$$\textbf{if } (temperature_{out} > 30°C) \textbf{ then } curtains = close \tag{9}$$

The above rules are a typical illustration of how a layman end-user (i.e. not a domain expert) would specify the preferred behavior of the smart system in terms of the storyline described above. Although these rules seem obvious at first sight, the careful reader will have already noticed that dependencies between these high-level automation objectives can be a recipe for trouble (e.g. possible contradictions in rules (5) and (7)) with the intended smart behavior escalating quickly into chaos.

2. Classification of Inconsistencies

Errors and inconsistencies cannot be avoided for systems with humans-in-the-loop, because people are not always consistent, dependable, responsible, trustworthy, etc. They are not designed to perform, and they definitely not always perform as designed (as envisioned by the developer of the system). The following sources for inconsistencies can be explained in terms of three causes:

1. **System factors:** devices behaving erratically
2. **Human factors:** people behaving erratically
3. **Unknown contexts:** unstated operational/environmental conditions

In this section, we will provide an overview of inconsistencies that can emerge due to dependencies or conflicts among the above rules.

2.1. Non-Deterministic System Behavior

Deterministic systems are characterized by the fact that all accepted behavior is adequately expressed for all possible circumstances. In non-deterministic systems, especially those with humans-in-the-loop, the current state of a system is not always predictable. This causes uncertainty about the behavior of a system, and a lack of understanding of why a system does not behave as expected. Without rule (2), the state of the heating would not be complete because the rules do not specify the contextual constraints for each state of the system. However, even with rule (2) the state of the system remains undetermined for a temperature of 20 °C. This leads to the following requirement:

> *The contextual constraints for each system state should be specified non-ambiguously so that the system's behavior can be explained with respect to the current context.*

2.2. Erratic System Behavior

In order to mitigate the increased potential for failure in a smart environment, systems should be prepared for failure rather than trying to prevent every failure (i.e. design for failure). Smart systems that can tolerate partial failures, deal with contextual inconsistencies (e.g. an erroneous sensor reading) or improper human handling, and minimize the impact of a failure through graceful degradation will improve the robustness and the quality of experience for the end-users. In rules (1) and (2) we assume that if the temperature is below 19 °C, the heating will be turned on. If there is a failure in the actuator, the temperature will not increase. Such failure can be detected with a rule that will monitor the execution or progress of other rules.

$$\textbf{if } (temperature_{in} < 19°C) \textbf{ then } (temperature_{in,t=0..10min} > 19°C) \tag{10}$$

As we turn on the heating, the effect of this actuation is not immediate. With this rule, we express that the actuation will lead to a new sensed contextual reality somewhere in the next 10 minutes. If this does not happen, the system can detect its own failure. For the sake of legibility, the above simplified notation means that somewhere in the next 10 minutes the temperature will be above $19°C$ (i.e. existential quantification), and not that this statement is true for the whole period (i.e. universal quantification). Similar to pre-conditions and post-conditions in object-oriented programming, assertions can help monitor and verify the expected behavior of the system. This leads to the following requirement:

> *To detect erratic system behavior early, every rule that modifies the system or the environment should be accompanied with an assertion that verifies the expected state of the system or the environment after the actuation.*

This assertion can be implemented as a rule. In the above example, a temperature sensor is used to verify the state of the environment during the next 10 minutes. However, note that the temperature sensor used to implement the assertion can also produce erroneous values, i.e. the verification can fail. This means that we may be able to detect erratic system behavior, but we may not always accurately identify the cause of the error.

2.3. Conflicting Context Conditions and Actions

When multiple rules affect the same system component, inconsistencies in the contextual triggers may cause a non-deterministic component state. These inconsistencies are often due to unbounded and/or overlapping contextual circumstances. For example, with rules (8) and (9) the state of the curtains is uncertain during a hot summer day because both rules would be triggered. Rules (5) also contradicts with (7) when $motion = false$ and $temperature_{in} > 27°C$. This leads to the following requirement:

To avoid conflicting operations or effects, each context should give rise to only one state for each component in the system.

We can relax this requirement if we are able to prioritize the execution of the rules, i.e. know the order in which the rules should fire. This example investigated contextual conflicts for an individual system component. Next, we will show how inconsistencies can emerge due to separate system components affecting the environment at the same time in opposite ways.

2.4. Hidden Dependencies and Invalid Global States

Consistent behavior of the individual components does not guarantee reliable functioning of the home automation system as a whole. For example, the temperature in a room can be affected by the heating, the air conditioning and the state of the curtains. Obviously, turning on the heating and the air conditioning at the same time is not a good idea. However, these implicit dependencies are nowhere modeled in the above rules. These hidden dependencies should be made as explicit as possible. At the least, invalid global states of the system or the environment need to be explicitly modeled – independent of the context-driven adaptation rules – to guarantee that certain inconsistent behavior does not occur or is detected early.

Invalid global states should be explicitly modeled.

These kind of assertions are similar to invariants in object-oriented programming that guarantee that certain predicates must always be true before and after a sequence of operations.

2.5. Non-Terminating Loops

A difficult problem in rule based systems is the presence of loops. The consequence statements of one rule can cause another rule to fire, possibly causing non-terminating loops. Intelligent environments are typically characterized by non-terminating processes, but some of these execution loops can cause unwanted behavior. For example, we want the temperature to gradually increase, so that there is no initial overshoot causing the air conditioning to become active, and vice versa. The problem with detecting these inconsistencies is the temporal aspect of the temperature variations in a loop. If such a cycle would take hours to complete (e.g. heating turned on in the morning and air conditioning turned on in the afternoon), there is not a problem. However, if this cycle would be much shorter, there is something wrong. Therefore, we should make sure that:

Context-aware adaptation rules should have safeguards in place to avoid non-terminating loops of unwanted behavior.

As loops cannot always be detected through static analysis of the rules, safeguards can for example express how often a rule can be executed in a given time period.

In the above subsections, we identified several requirements for rule-based dynamic and adaptive smart environment applications that would help to make these systems more reliable. In the following section, we will investigate and discuss which kind of inconsistencies can be detected at design time of the smart system through simulation and verification with the SPIN model checker.

3. A Model Checking Approach with SPIN

Model checking in software engineering [8] has been around for more than a decade to formally verify models of distributed software systems. The motivation for using model checking to verify context-driven applications is straightforward. There are often too many possible configurations and contextual situations that an application can be in, that it is impossible to verify all these combinations with a set of test cases. In this section, we will discuss the feasibility of using the SPIN model checker, and highlight the major benefits and weaknesses.

3.1. Promela

The SPIN (Simple Promela Interpreter) model checker relies on a model of the system. This model is specified in the Promela language (PROcess MEta-LAnguage). The objective is to model rules as concurrent processes and to model the behavior of the system as a finite state system.

The following example is a simple Promela model that represents a subset of the functionality of the home automation system. We use the Linear Temporal Logic (LTL) formula `[]( !airco || !heating )` to simulate and verify that the airco and the heating are never turned on at the same time.

```
mtype = { airco_on, airco_off, heating_on, heating_off }
mtype = { activated }

chan system_actions = [0] of { mtype };
chan sensor_reaction = [0] of { mtype, mtype };
```

```
bool airco = false;
bool heating = false;
int temperature = 18;

ltl p1 { []( !airco || !heating ) }

active [1] proctype Airco()
{
end:    do
        :: (temperature < 24) -> system_actions!airco_off; airco=false;
        :: (temperature > 27) -> system_actions!airco_on; airco=true;
        od
}

active [1] proctype Heating()
{
end:    do
        :: (temperature < 19) -> system_actions!heating_on; heating=true;
        :: (temperature > 21) -> system_actions!heating_off; heating=false;
        od
}

active [1] proctype Environment()
{
end:    do
        ::system_actions?heating_on -> sensor_reaction!activated,heating_on;
        ::system_actions?heating_off -> sensor_reaction!activated,heating_off;
        ::system_actions?airco_on -> sensor_reaction!activated,airco_on;
        ::system_actions?airco_off -> sensor_reaction!activated,airco_off;
        od
}

active [1] proctype Sensor()
{
end:    do
        ::sensor_reaction?activated,heating_on ->
            if
            :: (temperature < 5) -> temperature = temperature + 7;
            :: (temperature < 10) -> temperature = temperature + 5;
            :: (temperature < 15) -> temperature = temperature + 3;
            :: else temperature = temperature + 1;
            fi
        ::sensor_reaction?activated,airco_on ->
            if
            :: (temperature > 50) -> temperature = temperature - 12;
            :: (temperature > 35) -> temperature = temperature - 10;
            :: (temperature > 30) -> temperature = temperature - 8;
            :: else temperature = temperature - 1;
            fi
        ::sensor_reaction?_,_ -> temperature = temperature + 2;
        ::sensor_reaction?_,_ -> temperature = temperature - 2;
        od
}
```

Note that it is not our objective to create Promela models by hand, but rather to au-

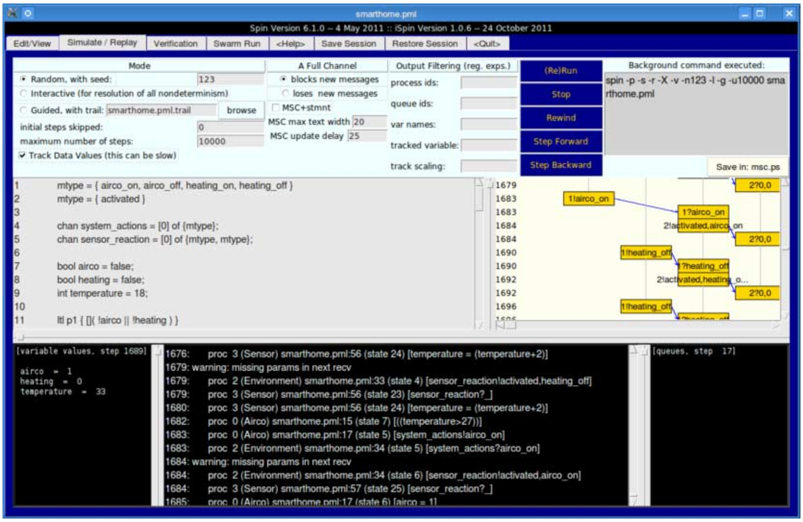


Figure 1. Simulation with the SPIN model checker

tomatically generate them from a simple rule set description that specifies the desired autonomous behavior and adaptation in the smart environment. The reason for this approach is twofold:

- The rule set is easier to specify and maintain, especially for users unfamiliar with the peculiarities and syntax of Promela.
- The same rule set can be used for transformation to other notations as input for other consistency checkers.

For example, we are exploring similar transformation techniques to OWL2RL [9] to model inconsistency rules and analyze them based on its underlying propositional logic. How the transformation from the rule set description to Promela is carried out is beyond the scope of this paper.

3.2. Simulation and Verification

In Figure 1, we show a simulation of this model with SPIN. Verifying this model results in the following success message:

```
spin -a  smarthome.pml
ltl p1: [] ((! (airco)) || (! (heating)))
gcc -DMEMLIM=1024 -O2 -DXUSAFE -DNOCLAIM -w -o pan pan.c
./pan -m10000  -a
Pid: 25147
error: max search depth too small

(Spin Version 6.1.0 -- 4 May 2011)
+ Partial Order Reduction

...
```

```
pan: elapsed time 0.62 seconds
No errors found -- did you verify all claims?
```

The aforementioned Promela does not represent the full behavior as reflected by the rules in the previous section. Some of the missing features (such as the lights and the motion) can easily be included in a similar fashion. However, it seemed not straightforward to model all features of the smart system behavior.

As a detailed discussion of the full Promela model is outside the scope of this paper, we will highlight in the following subsections which concerns were hard to model and which aspects we were not able to represent.

4. Qualitative Evaluation

The above Promela model is only a brief example of what can be modeled with Promela and analyzed with SPIN. After a preliminary assessment, it is clear that it definitely does not address all the requirements that we outlined before. We can make the following observations:

1. The first requirement argues that it should be possible to explain the state of a system based on the current context, as well as the rules that gave rise to the current state. This concern is often referred to as the intelligibility [10] of a system, a user centered property leading to a better understanding and stronger feelings of trust.
2. The second requirement identifies inconsistencies with respect to expected behavior. Obviously, these errors cannot be addressed purely with design time verification tools. Runtime support is necessary to ascertain whether the intended outcome of a rule has been achieved or not.
3. The third requirement states that all components should work together to achieve the same global new context. Leelaprute et al. [3] have how SPIN can be used to detect such feature interactions. We could have specified in our example:

   ```
   ltl p0 { !<> []( airco || heating ) }
   ```

4. The forth requirement stating which global states should never happen, can be implemented easily with Linear Temporal Logic formulas using negation and the ‘always’ operator represented as [] P, stating that P is always true in all system executions. In our example, we specified:

   ```
   ltl p1 { []( !airco || !heating ) }
   ```

   Note that this LTL formula is more strict than the previous one. With the above `p0` formula, we are sure there can never be a situation where both the airco and heating are on indefinitely, whereas with the `p1` formula, we are sure that the airco and heating are never on at the same time.
5. Non-terminating loops cannot (to the best of our knowledge) be expressed easily in Promela. Many works that use SPIN to detect feature interaction assume concurrency while representing their correctness claims as LTL formulas. To address this requirement, we need expressiveness (and most likely runtime support) to specify undesired interleaved processes (e.g. the airco and heating taking turns every 5 minutes).

In the following subsections, we will highlight the major benefits and weaknesses of SPIN and Promela with respect to the requirements.

4.1. Benefits

Counter example: One of the main advantages of SPIN is the ability to produce a counter example with a trace of events that lead to a situation where the claims to be verified are not confirmed. However, we found cases where the produced counter example was not realistic (e.g. the sun shining at night) and we had to finetune the model to exclude these false positives.

Non-deterministic system behavior: Following a particular modeling approach (using channels and the Promela `timeout` concept, it is possible to use a simulation approach where at some point the simulation blocks due to a particular context for which no further autonomous behavior is specified (e.g. a temperature of 20 °C with rules (1) and (2)). However, it is not guaranteed that through simulation you will find all these incomplete context specifications.

Conflicting conditions/actions: It is possible to identify different contexts that give rise to conflicting actions (cfr. rules (8) and (9) for an illustration). The technique to follow is based on the `atomic` construct and the implementation of a concept similar to mutexes. If a rule modifies the state of a component, it increases and decreases the mutex for this component in an `atomic` construct. If each context triggers one rule only, the mutex for each component will never be higher than 1.

4.2. Weaknesses

Explicit notion of time: While the Linear Temporal Logic allows for expressing qualitative properties with respect to time, it is not possible to quantitatively express time. We may guarantee that eventually it will get warmer if the heating is turned on. However, it would be much better if we could also specify that this should happen in the next 10 minutes. See rule (10) for an example of this concern. The inability to express time creates a similar concern regarding the non-terminating loops. It is impossible to specify the minimal length of acceptable adaptation cycle (e.g. heating turned on in the morning and air conditioning in the afternoon).

Modeling external influences: It is not straightforward to model a smart environment accurately. The measured temperature in a smart environment is not only the result of state of the air conditioning, the heating, the curtains, etc. but also influenced by external factors such as the temperature outside. In the previous Promela model, we pursued a simplistic approach by having the temperature randomly go up or down with a degree (exploiting the non-deterministic selection in the Promela `do`-loop statement). However, this is not a realistic representation of the real world, which makes the outcome of the simulation subject to interpretation.

Hidden dependencies: It is very easy to overlook dependencies among context variables and the impact of one activity on multiple contextual variables. For example, closing the curtains will not only keep the heat outside on a sunny day, but will also affect the brightness inside the house. If the behavior is not accurately described, it is unlikely we will be able to capture all inconsistencies in the context-aware rules through verification.

State space explosion: This is a common problem with model checkers. In the aforementioned example, the possible context combinations were rather small due to the limited amount of contextual state variables. As a result, we did not experience any scalability problems or memory constraints that would force us to simplify the abstractions in our model. However, for more complex situations with multiple context variables and context-aware rules and assertions, it is fairly likely that the exponential state space explosion will become a critical concern to verify the consistency of the context-aware rules.

No systematic extraction techniques: The current approach assumes the existence of a Promela that resembles the behavior of the context-aware systems. Classical approaches start off with an initial design of the system which is then manually handcrafted into an abstract Promela verification model. If the verification succeeds, the system is implemented. Unfortunately, this process does not guarantee that the implementation is consistent with the Promela model. Modern approaches start off with the implementation from which the verification model is automatically abstracted. Unfortunately, to the best of our knowledge, such systematic techniques have not been proposed for context-aware systems.

5. Conclusion

In this paper, we investigated an approach to expose context dependent faults in a smart environment system that is driven by rules following the event-condition-action paradigm. We started off with some general observations of how inconsistencies in context-aware rules can lead to erratic system behavior, either through faults in these rules or due to interferences between these rules. We listed some requirements that would help to mitigate the risk of failures, and we experimented with the SPIN model checker to investigate its feasibility as a model checker to debug these inconsistencies. We illustrated the process with a simple smart home automation scenario as a typical example of a smart environment. Experiments with students have shown that formal verification of a context-aware system by means of model checking is not easy, even after an initial training with SPIN on some less complex examples. Many consider it more complex than implementing the system itself, in some cases even too complex. This experience is well captured in one of the quotes of Brian Kernighan, well-known for his co-authorship of the first book on the C programming language:

> *"Everyone knows that debugging is twice as hard as writing a program in the first place. So if you are as clever as you can be when you write it, how will you ever debug it?"* - Brian Kernighan

In the preliminary experiments we carried out with SPIN, we found the lack of an explicit representation of time to be counter-intuitive. Being able to verify that a certain property will hold some time in the future does not allow us to really express what people expect from a system.

Another aspect we are further exploring is the automatic translation of a given rule set and a list of global assertions into a compatible Promela model, and how verification failures can be traced back to the original rule set specification. For this purpose, we need to explicitly model how state changes in a single component of the smart system can interfere with the state of another component (e.g. turning on both the heating and the air conditioning at the same). In a similar way, given that the effects of a state change (e.g. the heater being turned on) are not immediate, we are exploring models to realistically represent how an actuator change influences what a sensor will measure over time. The purpose of such models would be to capture the expectations of the system. Any significant deviations from such common or expected behavior could be identified as abnormal behavior more easily, which would hopefully lead to an increased reliability and robustness against unforeseen human interventions, exceptional circumstances and unexpected events in smart environments.

In some experiments, the simulator of SPIN provided us with useful feedback with contextual situations where erratic behavior would occur. However, we also identified that some of the traces in SPIN were not realistic. In one of the traces, SPIN produced a counter example that claimed an error would occur if the sun is shining at night. Obviously, this error occurred due to hidden dependencies in the context variables. We are currently investigating approaches to capture these facts into the Promela model.

Acknowledgments

This research is partially funded by the Interuniversity Attraction Poles Programme Belgian State, Belgian Science Policy, and by the Research Fund KU Leuven.

References

[1] M. Baldauf, S. Dustdar, and F. Rosenberg. A survey on context-aware systems. *Int. J. Ad Hoc Ubiquitous Comput.* **2**:4 (2007), 263–277.

[2] G.D. Abowd, A.K. Dey, P.J. Brown, N. Davies, M. Smith, and P. Steggles. Towards a better understanding of context and context-awareness. In *HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, 304–307, London, UK, Springer-Verlag, 1999.

[3] P. Leelaprute, T. Matsuo, T. Tsuchiya, and T. Kikuno. Detecting Feature Interactions in Home Appliance Networks. In *Proceedings of the 2008 Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, 895–903, IEEE Computer Society,2008.

[4] R. Cook. How Complex Systems Fail. Technical report, Cognitive technologies Laboratory, University of Chicago, 2000.

[5] J.C. Augusto and P.J. McCullagh. Safety Considerations in the Development of Intelligent Environments. In P. Novais, D. Preuveneers, and J.M. Corchado, editors, *ISAmI*, *Advances in Intelligent and Soft Computing* **92** (2011), Springer, 197–204.

[6] H. Igaki and M. Nakamura. Modeling and Detecting Feature Interactions among Integrated Services of Home Network Systems. *IEICE Transactions* **93-D**:4 (2010), 822–833.

[7] A. Schuster. *Robust Intelligent Systems*. Springer Publishing Company, Incorporated, 1 edition, 2008.

[8] G.J. Holzmann. The Model Checker SPIN. *IEEE Trans. Softw. Eng.* **23**:5 (1997), 279–295.

[9] D. Reynolds. OWL 2 RL in RIF. http://www.w3.org/2005/rules/wiki/OWLRL, 2009.

[10] B.Y. Lim, A.K. Dey, and D. Avrahami. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, 2119–2128, New York, NY, USA, ACM, 2009.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-413

Using Simulation and Verification to Inform the Development of Intelligent Environments

Juan Carlos AUGUSTO[a,1] and Miguel J. HORNOS[b]
[a]*School of Computing and Mathematics, University of Ulster, UK*
[b]*Software Engineering Department, University of Granada, Spain*

Abstract. This paper advocates for a more widespread adoption of Software Engineering techniques aimed at informing the design and development stages of Intelligent Environment systems. Reliability is a fundamental principle of all software systems, and even more of Intelligent Environments, which are meant to support people in real life activities. We present a methodological guide which provides strategies and suggestions on how to model, simulate and verify these types of systems. One of our aims is to encourage their use by colleagues working in this area and gather their feedback in this experience.

Keywords. Intelligent Environments, Simulation, Verification, Spin

Introduction

Intelligent Environments (IE) [1] are very complex systems which are made up of a blend of components coming from different areas, such as networks of sensors and actuators, communication protocols, ubiquitous/pervasive computing, human-computer interaction and artificial intelligence, among others. The software that controls these systems is difficult to develop and error prone, since it is usually made up of diverse distributed processes which run concurrently. Given their mix of hardware and software components and their potentially rich interaction with humans, there are many opportunities for failure. These systems can be deployed in a variety of domains, for example: in home environments, to monitor and assist people; in educational organizations, to support teaching and learning activities; in smart cars, to make cars more efficient and make driving safer; in agriculture, to make it more productive; and in line production, to manufacture more efficiently.

IE are specific examples of reactive systems [2], which are those that react to any stimulus (or event) that occurs in the environment, maintaining a continuous interaction with it. Informal reasoning is not appropriate for the development of such systems due to the interleaving of the many concurrent events that may occur. It is therefore necessary to apply techniques and methods that allow us to design and get more reliable IE [3]. Software testing [4], which is by far the most used technique to improve software quality, is not an adequate technique for exhaustively checking the correctness of complex systems in general, and of IE in particular, since it consists on executing

[1] Corresponding Author: jc.augusto@ulster.ac.uk

only a sampling of all possible runs of the system to be checked. Thus, the application of formal methods is an important complement to analyze and establish the correctness of such systems, especially in applications in which safety is a critical issue, where a small error in their design can even put human lives at risk.

We will focus then on the possibilities offered to IE by a verification technique called model checking [5], which consists on verifying whether any execution of the transition system that represents the system to be checked violates (i.e. does not meet) a certain property, specified by a temporal logic formula. The Software Engineering (SE) community often uses this methods, techniques and tools for verifying systems and software at the earliest stages of their development. However, amongst IE developers these formal methods and tools are rarely used, though there are some exceptions which have started to appear recently [6-12]. Most of these works are isolated efforts which are based on very specific and specialized methods and tools, reason why they are rarely known in practice by the IE community. As a result, the corresponding proposals are mostly adopted only by their developers.

The present work aims to take a first step to build a bridge of understanding and collaboration between both SE and IE communities, by encouraging the use of methods and tools which are well-known by software engineers among researchers and developers of IE, as well as their convenient adaptation to this type of systems, to make the resulting IE systems more reliable. To do this, we provide a methodology with examples based on our experience in particular projects of this area which try to clarify how these methods and tools can be applied to similar projects or systems.

1. Simulation and Verification

There are an interesting number of methods and tools in Software Engineering which can support software engineers to develop their systems. We explained in the previous section our interest on those which can provide simulation and verification capabilities. From all those options we have chosen the Spin model checker [13] because it is efficient, user-friendly, free and of widespread use.

Spin is focused on the concept of models, users have to create them using its *modeling* language, called Promela (acronym for PROcess MEta-LAnguage), which is intended to easily make good abstractions of the system to be modelled, putting emphasis on the description of processes and their interactions through message channels and variables.

Once a model has been created, the *simulator* included in Spin can be used to get some insights on the scenarios described in the model. Simulation can only help to debug the models, but no amount of simulation runs can prove that a model satisfy a given property. To achieve this we have to perform a formal *verification* (by applying model checking), which consists on ascertaining the correctness of a model (with respect to a certain property) by exhaustively checking whether all the executions implied by the model satisfy such property usually expressed as a LTL (Linear Temporal Logic) [14] formula. Core concepts in LTL are the use of the following temporal operators to express how states in the system relate to each other: `<>p`: Eventually (in the future) `p` holds, i.e. `p` is fulfilled at some future state (from the current state forward); `[]p`: Always (in the future) `p` holds, i.e. from the current state forward `p` is true; and `pUq`: `p` is true (from the current state) until `q` is true.

They will be used in the next section to represent situations of interest we want to check on the scenario we consider to explain our proposal.

Given space constraints here we cannot explain SPIN and its use in detail so we refer the reader to our guide on how to install Spin as well as an explanation of its main functionalities using its graphical user interface in [15] and we focus on explaining how SPIN can be used to support the development of IE.

2. Methodology and sample models

Promela models follow the traditional programming template, where each conceptual computing unit consists of a declaration section followed by a section where the behaviour of the system to external inputs is described. This section proposes a methodology to model sensor-based systems that is explained by examples. To give examples on how to use Promela to model intelligent environments we will focus on smart home applications and we will unfold models of increasing complexity. These models however will share the following common structure:

```
...Definition of Global data structures (e.g., sensors)...
/* ------------------------------------------------------*/
active [...] proctype HumanActivity ()
{...(local data structures)...
(behaviour modelling)
}
/* ------------------------------------------------------*/
active [...] proctype SensorActuator ()
{...(local data structures)...
(behaviour modelling)
}
/* ------------------------------------------------------*/
active [1] proctype Central_Unit ()
{...(local data structures)...
(behaviour modelling)
}
```

That is, we will assume we want to model mainly a home equipped with sensors and actuators, which is inhabited and will receive visitors. The effect that the actions of occupants and visitors (i.e., human activity) will have on the sensors and actuators will be collected by a central processing unit. Each of these can be represented by a *proctype* (i.e. the declaration of a process) in Promela. These can be represented in many different ways and degrees of detail or abstraction, we do not want to be prescriptive and each developing group will find the level of usage that is useful for their goals. We only provide examples which we hope will motivate colleagues to use Spin and other tools to reason about their development and in doing so they will develop more carefully engineered, and hence safer, systems.

Although this first model is very simple in terms of the level of detail and insight that can offer on what happens inside the environment being modelled, it is useful to explain some basic principles before we add more information (and complexity). There are several design choices which can be different; for example, instead of using specific channels to model the communication between each pair of processes (proctypes) that have interaction, a common single channel could have been set up with a more complex message structure complementing the message contents with both sender and receiver IDs. Other models can be created to allow more than one occupant and/or visitor, identifying them, distinguish them amongst several possible actions, consider a wide range of specific sensors, etc.

Observe that the generic model previously shown has a process named `SensorActuator`. This is because both sensors and actuators can be treated in that process. Alternatively, we can define two processes, one for sensors and another for actuators. Some of the subsequent models only have sensors, because that is what we want to model. Further below models will also have a process for actuators.

Throughout the rest of the article we will illustrate the usefulness and relevance of tools like Spin to the development of IE by showing how core concepts of a typical application can be addressed. To this end we will make use of one of the most popular application domains within IE, smart homes for Ambient Assisted Living.

2.1. Creating a first basic model

The process of using SPIN involves the investigation of ideas through the use of a number of models which represent partial views of the system. Usually, for any interesting system, the complexity is such that we cannot model the system in full detail, so we can either aspire to model the entire system at a very generic level or specific parts of the system in more detail. In any case modelling as it is used in SPIN has to be differentiated from programming. In modelling we do not aim at creating a fully working implementation, but rather to put some ideas and strategies to the test. Let us assume we want to develop a smart home, a first model can look like Model 1.

```
/* Model 1: basic model to illustrate strategy on representing some smart
home concepts and interaction.  */
        mtype = {occ_act, vis_act};   /* occupants/visitors action */
        mtype = {activated};    /* sensor reaction */
        chan activities = [0] of {mtype};
        chan sensor_reaction = [0] of {mtype};
/* -----------------------------------------------------*/
active [1] proctype Occupant ()
{
end:    do
        ::activities!occ_act
        od
}
/* -----------------------------------------------------*/
active [1] proctype Visitor ()
{
end:    do
        ::activities!vis_act
        od
}
/* -----------------------------------------------------*/
active [1] proctype Sensor ()
{
end:    do
        :: activities?_ ->sensor_reaction!activated
        od
}
/* -----------------------------------------------------*/
active [1] proctypeCentral_Unit ()
{
end:    do
        ::sensor_reaction?activated
        od
}
```

Figure 1 illustrates the use of the simulation facility in Spin, which shows in the left hand side the Promela Model 1, on the right the Message Sequence Chart (MSC) with the messages sent and received by each of the processes through channels, and at the bottom the sequence of executed lines in the Promela model plus their effect on data (left, in this case empty as we do not use variables) and channels (right).

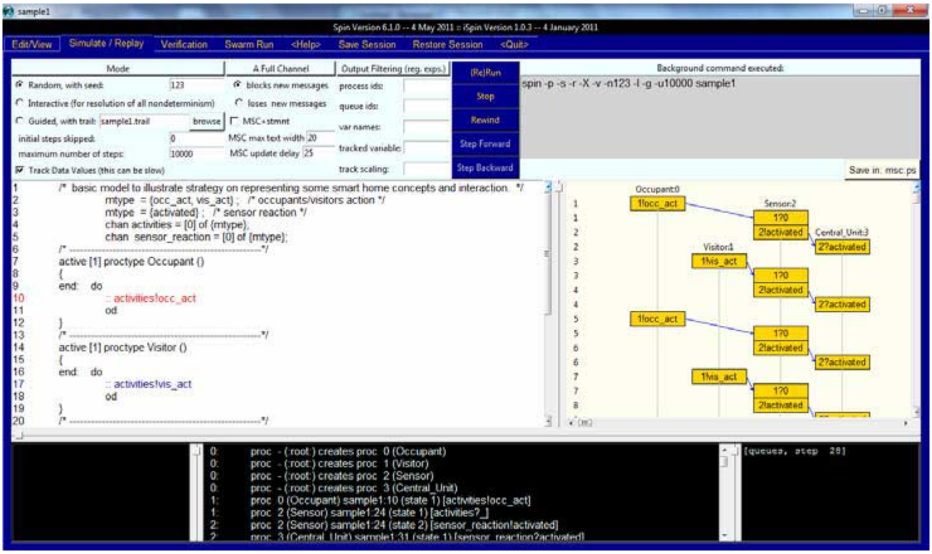


Figure 1. A simulation run of the Model 1.

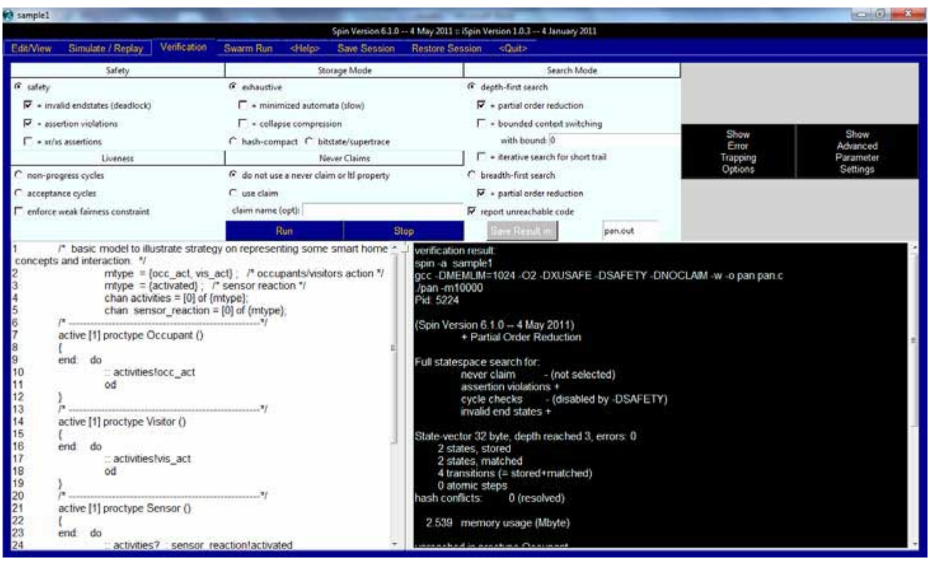


Figure 2. Safety verification with no errors found for Model 1.

The model is very simple and the MSC on the right hand side shows the messages going from both the occupant and visitor processes to the sensing device (these denoting that they have performed some actions which have been perceived by a sensor) and this causing a message sent to the central unit (representing that the sensor activation is being sent to a program or control box which will process it). Notice, we used `activities?_` as part of the contents of `proctype Sensor`, since at this stage we only want to detect activity regardless of whether it is `occ_act` or `vis_act`. A simple safety verification shows that there are a priori no errors of this category, as shown at the black right-bottom panel in Figure 2.

2.2. Investigating the behavioural properties of the model

If we want to conduct a more thorough investigation of the behaviour represented by this model, we can use properties expressed in Linear Temporal Logic (LTL). That can be achieved in iSpin by embedding an inline property specification which we can then refer to in the verification console. Say we want to check that *whenever each time a human performs a detectable action, then eventually that will be notified to the central processing unit, where it will be processed*. One way to achieve this (not necessarily the only one or the best one) is to modify the model as shown in Model 2.

```
/*  Model 2: basic model to illustrate strategy on representing some smart
home concepts and interaction.  */
        mtype = {occ_act, vis_act}; /* occupants/visitors action */
        mtype = {activated};    /* sensor reaction */
        chan activities = [0] of {mtype};
        chan sensor_reaction = [0] of {mtype};
        bool act, received;
ltl p1 { [](act -> <>received) }
/* whenever a human performs a detectable action, then eventually that will
be notified to the central processing unit */
/* -----------------------------------------------------*/
active [1] proctype Occupant ()
{
end:    do
        ::activities!occ_act; atomic{act=true; act=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctype Visitor ()
{
end:    do
        ::activities!vis_act; atomic{act=true; act=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctype Sensor ()
{
end:    do
        :: activities?_; sensor_reaction!activated
        od
}
/* -----------------------------------------------------*/
active [1] proctype Central_Unit ()
{
end:    do
        ::sensor_reaction?activated;
        atomic{received=true; received=false}
        od
}
```

In this previous model we introduced a way to detect when we have reached the two meaningful conditions, for this case: "a human performs a detectable action" and "notified to the central unit". Human actions can be either coming from the house occupant or from a visitor. So we can use a Boolean variable to detect when an activity has been performed and other Boolean variable to detect when the central unit is notified that a sensor has been activated. Model 2 has new code with respect to Model 1: the declaration of the auxiliary variables, the formal statement of the behavioural property to be checked, and the two insertions in the code of the proctypes to detect when the system has reached the meaningful states we are interested to correlate. Figure 3 shows the successful check of the property we described above. In the centre of the image, it can be seen we have asked iSpin to check the consistency of property named as `p1` against our model, and on the right-bottom hand side, the system diagnosis is that there are no errors.

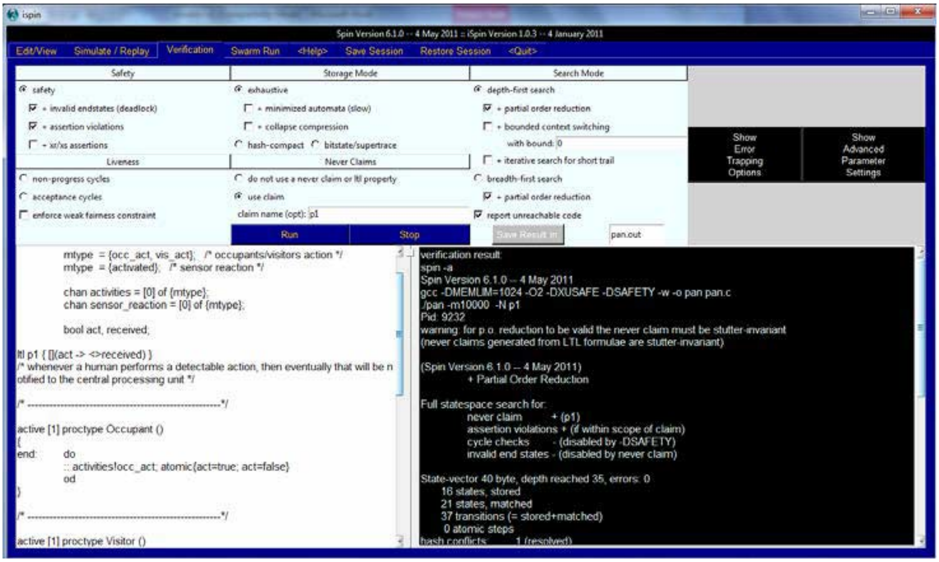


Figure 3. Checking that the model satisfies the property p1.

2.3. Refining the basic model

Of course, the model above does not distinguish which one of the potential users was responsible for the sensor activation. If we want to check for each one separately we need to distinguish them further. Now imagine that for some of those activities we want the system to do something specific, this could be recording the activity in a log or database, sending a notification message to a mobile phone or interacting with the humans in the house (for example to issue a greeting message or a recommendation). Let us call this action from the system `actuation`. We may not want the system to react to every single instance of sensor triggering detected with an actuation so we can insert code which allow Spin to non-deterministically choose when to actuate and when not to. Model 3 incorporates all these changes and also includes a property to attempt to check whether it is true that *whenever a human (occupant or visitor) performs a detectable action, this will result on an actuation.* We will explain more about this property further down.

```
/* Model 3: basic model to illustrate strategy on representing some smart
home concepts and interaction.  */
   mtype = {occ_act, vis_act}; /* occupants/visitors action */
   mtype = {activated, request}; /* measuring sensing and actuation */
   chan activities = [0] of {mtype};
   chan sensor_reaction = [0] of {mtype};
   chan actuation = [0] of {mtype};
   bool oact, vact, received, acted;
ltl p1 { []((oact || vact) -> <>acted) }
```

```
/* whenever a human (occupant or visitor) performs a detectable action, this
will result on an actuation */
/* -----------------------------------------------------*/
active [1] proctype Occupant ()
{
end:    do
        ::activities!occ_act; atomic{oact=true; oact=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctype Visitor ()
{
end:    do
        ::activities!vis_act; atomic{vact=true; vact=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctype Sensor ()
{
end:    do
        :: activities?_; sensor_reaction!activated
        od
}
/* -----------------------------------------------------*/
active [1] proctype Actuator ()
{
end:    do
        ::actuation?request;
        atomic {acted=true; acted=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctypeCentral_Unit ()
{
end:    do
        ::sensor_reaction?activated;
        atomic {received=true; received=false};
        if
        ::actuation!request
        :: skip
        fi
        od
}
```

Figure 4 shows the simulation of this model. The MSC window shows that actuation is not always used, as we expected, since we inserted code to allow Spin to choose over this behaviour. Notice that when we check the property `p1` included in Model 3 (see Figure 5), the result does not reflect the behaviour observed in the simulation. The check performed in Figure says the property is correct for Model 3. However, we have seen in the simulation that not all human activity results in actuation. Where is the problem? The semantic gap is in the interpretation of the LTL formula checked, i.e. `[]((oact||vact)-><>acted)`, which actually is satisfied when `acted` is true in after either `oact` or `vact` have occurred, but without taking into account frequency or correspondence relationships between sensing and actuation events. Thus, as the system executes infinite loops, it sees that there is always an opportunity to have either `oact` or `vact` followed by `acted` (it does not matter how close in time they are, since temporal distance is not measured by the <> operator, which means 'sometimes in the future'). In other words, the above formula *does not necessarily mean* that *each time a human (occupant or visitor) performs a detectable action, this will be strictly followed by an actuation*.

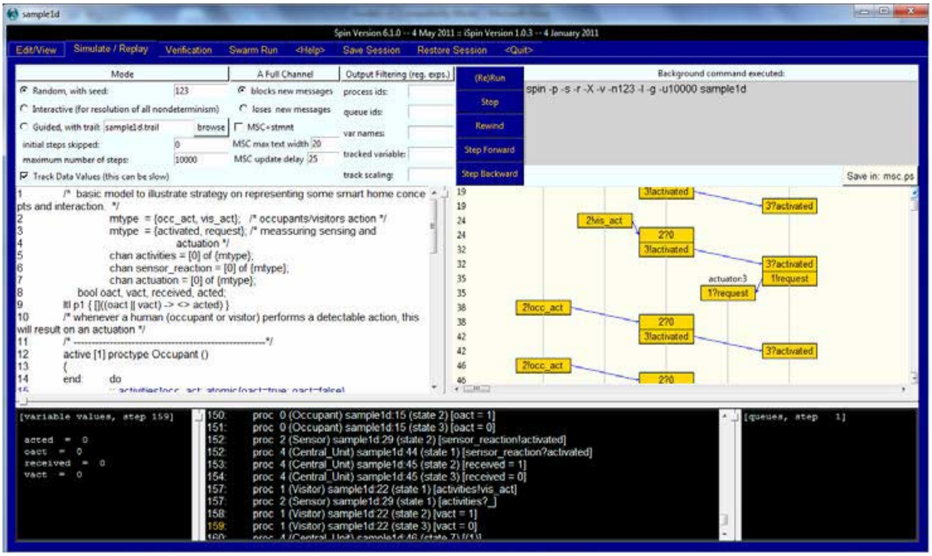


Figure 4.Simulation of the new model.

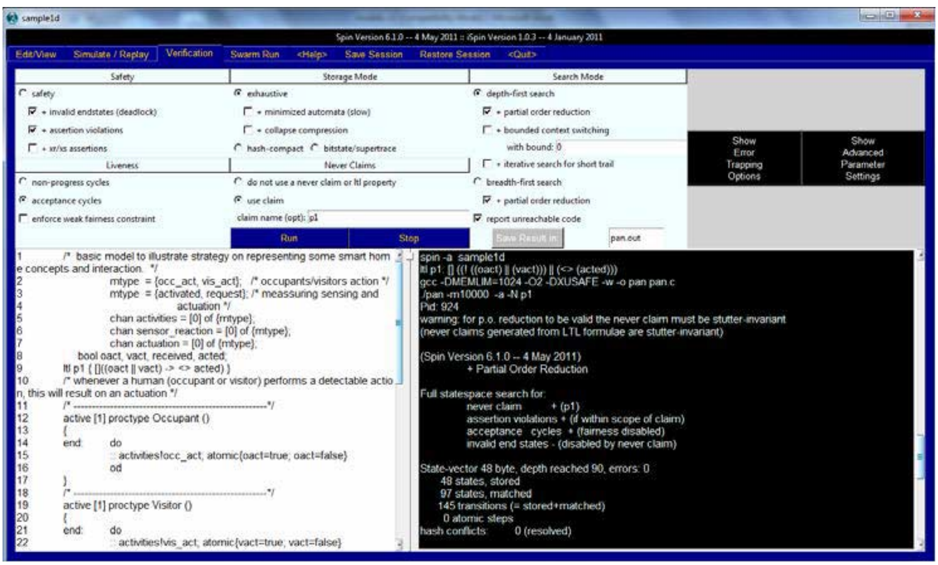


Figure 5. The model satisfies the property p1.

To explain this more clearly, imagine the following two sequences, σ_1 and σ_2, of sensing (s_i) and actuation (a_j) events:

σ_1: …s_1, a_1, s_2, a_2, s_3, a_3, s_4, a_4…

σ_2: …s_1, a_1, s_2, s_3, s_4, a_4…

Should all possible achievable sequences produced by our model were as in σ_1 then we could say of that it is true that "*whenever a human (occupant or visitor) performs a detectable action, this will result on an actuation*". However, the system behaves more like in σ_2. Thus, the formula `[]((oact||vact)-><>acted)` was found valid by iSpin, because for each s_i there is an a_j sometimes in the future (it does not matter how close or far away they are in the sequence).

Going back to our model, if we want to be more precise and be able to track the consequence of specific actions, we need to invest a bit more in the model. Model 4 provides a variant which allows us to be more precise in the follow up of user actions.

```
/* Model 4: basic model to illustrate strategy on representing some smart
home concepts and interaction.  */
        mtype = {occ_act, vis_act}; /* occupants/visitors action */
        mtype = {activated, request, occ, vis, unknown};
                                      /* measuring sensing and actuation */
        chan activities = [0] of {mtype};
        chan sensor_reaction = [0] of {mtype, mtype};
        chan actuation = [0] of {mtype};
        bool oact, vact, received, acted;
        mtype Whom;
        mtype Source=unknown;
ltl p1 { Source==unknown U (Source==vis U Source==unknown) }
/* only visitors are identified */
/* -----------------------------------------------------*/
active [1] proctype Occupant ()
{
end:    do
        ::activities!occ_act; atomic{oact=true; oact=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctype Visitor ()
{
end:    do
        ::activities!vis_act; atomic{vact=true; vact=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctype Sensor ()
{
end:    do
        ::activities?occ_act; sensor_reaction!activated, occ
        ::activities?vis_act; sensor_reaction!activated, vis
        od
}
/* -----------------------------------------------------*/
active [1] proctype Actuator ()
{
end:    do
        ::actuation?Source;
          atomic {printf("Source%d\n",Source);
                  Source=unknown;
                  acted=true;
                  acted=false}
        od
}
/* -----------------------------------------------------*/
active [1] proctypeCentral_Unit ()
{
end:    do
        ::sensor_reaction?activated, Whom;
          atomic {received=true; received=false};
          if
          :: Whom==vis; actuation!vis
          :: Whom==occ; skip
          fi
        od
}
```

The following simulation (displayed in Figure 6) shows how some requests reaching the central unit are not necessarily processed and Spin is non-deterministically interleaving the different processes and before reaching actuator it allows the detection of another action by the visitor. Eventually both are communicated. Then an action from the occupant is detected but not communicated through `actuation`, as we encoded the `Central_Unit` not to do so (using `skip`, i.e. ignore).

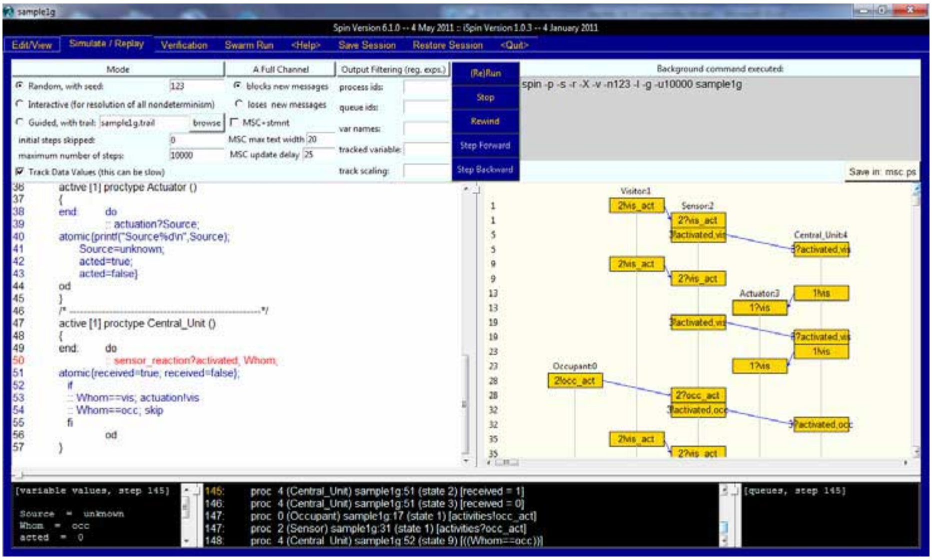


Figure 6. Simulation showing that some requests to the central unit are not processed.

A verification of the LTL formula `[](Source==unknown U (Source==vis U Source==unknown))`, states that only visitors are identified. The idea behind this property is that the detection of unknown and visitor alternates. Model 4 includes a variable declaration section where `Source` is initialized to `unknown`; after this, the model only allows source of sensor triggering to be explicitly identified when it has been labelled as coming from a visitor. Once a visitor triggering has been detected and identified, the system is set to `unknown` again.

If we for example would like to check if the system is actually capable to identify visitors as the source of an event triggering, we can check whether it is always true the formula `[](Source==unknown)` by naming this property `p2` and inserting it in the model, iSpin will show us that this property is not valid for this model. We can see a counterexample by visiting the simulation panel. The counterexample will show that Spin has a way to reach situations where visitors can be detected, hence contradicting the claim.

3. Concluding Remarks

We have presented strategies and suggestions on how to model, simulate and verify different features of Intelligent Environments by using Spin. There are many ways to achieve the same; we do not claim the methodology followed in this paper is the optimal one under all criteria. We also understand the use of Spin has limitations in modelling all the complexity of IE systems, but the “good should not be enemy of the perfect", that is, not because we do not have a perfect tool we should preclude ourselves from enjoying the benefits we can get through tools like Spin.

The most important benefit is on the stimulation of our mind in different directions than those we would have explored without the aid of the tool. To benefit the most out of this exercise it is important to create the habit of reflection associated with the results of a simulation or a verification run. When a simulation or a verification run gives us an unexpected result, it is important we ask ourselves: *Can this happen in the real system if I apply the implementation strategy of this model?* 'Fixing' the model until it behaves as we want is, in certain way, secondary.

Our proposal does not want to be prescriptive but rather inspirational. It aims to build bridges between the IE and SE communities by showing examples to developers in the IE area on how well-known tools from the SE area are very relevant to them. We have introduced schemas to model the classical concepts of IE providing a sequence of models of increasing complexity. We have chosen Spin given it is widely known and highly respected. Many other simulation and verification methods and tools are applicable. We hope this article inspires our IE community to increase the adoption of similar approaches towards developing the systems they deploy.

References

[1] H. Nakashima, H. Aghajan, J.C. Augusto (Eds.), *Handbook on Ambient Intelligence and Smart Environments*, Springer, New York, 2009.

[2] K. Schneider, *Verification of Reactive Systems: Formal Methods and Algorithms*, Springer-Verlag, Berlin, 2004.

[3] J.C. Augusto, P.J. McCullagh, J.-A. Augusto-Walkden, Living without a safety net in an Intelligent Environment, *ICST Trans. Ambient Syst.* **11**, 10–12 (2011), e6, doi:10.4108/trans.amsys.2011.e6.

[4] G.J. Myers, C. Sandler, T. Badgett, *The Art of Software Testing, 3rd edition*, John Wiley & Sons, Hoboken, 2011.

[5] B. Bérard, M. Bidoit, A. Finkel, F. Laroussinie, A. Petit, L. Petrucci, P. Schnoebelen, *Systems and Software Verification: Model-Checking Techniques and Tools*, Springer, Berlin, 2001.

[6] J.C. Augusto, P.J. McCullagh, Ambient Intelligence: Concepts and Applications, *Int. J. Comput. Sci. Inf. Syst.* **4**, 1 (2007), 1–28.

[7] J.C. Augusto, Increasing reliability in the development of intelligent environments, *Proceedings of the Fifth International Conference on Intelligent Environments* (IE'09), IOS Press, pp. 134–141, 2009.

[8] F. Corno, M. Sanaullah, Design Time Methodology for the Formal Verification of Intelligent Domotic Environments, *Proceedings of the Second International Symposium on Ambient Intelligence* (ISAmI'11), Springer-Verlag, pp. 9–16, 2011.

[9] A. Coronato, G. De Pietro, Formal Specification and Verification of Ubiquitous and Pervasive Systems, *ACM Trans. Auton. Adapt. Syst.* **6**, 1 (2011), 9:1–9:6, doi:1921641.1921650.

[10] J.C. Augusto, H. Zheng, M. Mulvenna, H. Wang, W. Carswell, P. Jeffers, Design and Modelling of the Nocturnal AAL Care System, *Proceedings of the Second International Symposium on Ambient Intelligence* (ISAmI'11), Springer-Verlag, pp. 109–116, 2011.

[11] K. Benghazi, M.V. Hurtado, M.J. Hornos, M.L. Rodríguez, C. Rodríguez-Domínguez, A.B. Pelegrina, M.J. Rodríguez-Fórtiz, Enabling correct design and formal analysis of Ambient Assisted Living systems, *J. Syst. Softw.* **85**, 3 (2012), 498–510.

[12] J.C. Augusto, M.J. Hornos, Designing more Reliable MAS-based Ambient Intelligence Systems, in: T. Bosse (Ed.), *Agent-Based Approaches to Ambient Intelligence*, IOS Press, Amsterdam, 2012 (in press).

[13] G.J. Holzmann, *The SPIN Model Checker: Primer and Reference Manual*, Addision-Wesley Professional, Boston, 2003.

[14] A. Pnueli, The temporal logic of programs, *Proceedings of the Eighteenth IEEE Symposium on Foundations of Computer Science* (FOCS'77), IEEE, pp. 46–57, 1977.

[15] M.J. Hornos, J.C. Augusto, Installation process and main functionalities of the Spin model checker, http://hdl.handle.net/10481/19601, 2012.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-425

A Proposal for Evaluating the Privacy Management through "Fair Trade" Metaphor

Abraham ESQUIVEL [a,1], Pablo A. HAYA [b] and Xavier ALAMÁN [c]
[a] *Instituto Tecnológico Superior Zacatecas Norte, México.*
[b] *Instituto de Ingeniería del Conocimiento, Universidad Autónoma de Madrid, Spain.*
[c] *Universidad Autónoma de Madrid, Spain.*

Abstract. This article presents a proposal for evaluating a privacy mechanism in pervasive environments. These environments are capable of sensing personal information anywhere and at anytime. This implies a risk to privacy that might not be assumed if a clear and trustable privacy management model is not provided. We have devised a privacy solution dealing with user's privacy as a tradable good for obtaining environment's services. Thus, users gain access to more valuable services as they share more personal information. This strategy, combined with optimistic access control and transactions registries mechanisms, enhances users' confidence in the system while encouraging them to share their information, with the consequent benefit for the community.

Keywords. Privacy, pervasive computing environments.

Introduction

Pervasive computing is an emerging field of research with everyday growing possibilities over active environments. Initially, most problems found in pervasive environments appeared in technical areas. Several research projects have been oriented on that direction. In particular, great efforts have been made to accomplish a seamless integration between devices operating in such environments [1]. Although technological achievements in this area have been of outstanding importance, social factors will also be key roles in the success or failure of pervasive environments. In this sense, a ma jor step in leveraging such environments comes from overcoming their ethical and psychological issues [2]. In fact, ethical and psychological problems are not strictly from pervasive environments and can be found in many software applications involving social interaction e.g. instant messengers or shared calendars. Clearly, the use of such applications entails new social problems, in particular those concerning user privacy.

Privacy is a dynamic phenomenon; its configuration has as many variations as variations has context i.e. a single change on context can trigger a change on privacy prefer-

[1]Corresponding Author: Abraham Esquivel, Instituto Tecnológico Superior Zacatecas Norte, México; E-mail: abraham.esquivel@gmail.com

ences. In consequence, privacy management should be a continuous negotiation in which the definition of the public and private boundaries will depend directly on user's context. Therefore, privacy management is a constant process of limits regulation. Those limits to accessibility of personal information determine the "sincerity" or "openness", or "distrust" or "closeness" characterizing the user and his/her current context.

But, how can the user establish those limits? A first approximation can be to manually configure privacy, assigning the desired level to each source of information. The key problem of this solution relays on the nature of privacy and pervasive environments: the degree of privacy desired for a source of information depends not only on the user and the source but also on the context. In other words, for every source and every person there can be as many different privacy configurations as different contexts they can be involved in. To configure a priori each possible arising situation for every source of information can be an overwhelming task that justifies the use of automatic management solutions.

On the other hand, the success of such automatic management solutions depends directly on the trust the user place on them. This trust depends directly on the following requirements: a) The model must be simple enough to be understood by a non-technical user, and b) The user must be able to modify the automatic configuration at any moment. In developing our framework we focused particularly on the importance of usability. Especially in this case, confidence and trust are synonyms of usability.

Following this premises, we present a privacy solution dealing with user's privacy as a tradable good for obtaining environment's services. Thus, users gain access to more valuable services as they share more personal information. This strategy enhances users' confidence in the system while encouraging them to share their information.

1. "Fair Trade" Metaphor

The "Fair Trade" Metaphor [3] constitutes the keystone of our design. Privacy is a dynamic phenomenon depending on the user and its context: each individual has as many privacy issues as different situation he/she is involved in. In general, we can say that privacy management is a permanent negotiation where the private and public boundaries are constantly changing in response to the circumstances.

Furthermore, we pose that a particular privacy configuration depends on the user sharing a service, and the service itself. In consequence, depending on the users and services involved, user's privacy needs will change.

One immediate problem arising from those considerations is privacy management. It seems reasonable that users would like to control their privacy boundaries. However, a very dynamic context can frustrate user's expectation due to an overwhelming amount of information. This is particularly true in ubiquitous environments where users must decide privacy configurations for every data gathered or deduced from sensors. We will show that two different kinds of management are required depending on the nature of the user's information.

According to the nature of the user's personal data, we classify information in two categories: **long-term information** and **short-term information**. The former comprises information with a low changing rate -e.g. telephone number, social security number, name, surname or postal direction- The latter, on the other hand, contains information of changing nature -e.g. location, activity or cohabitants. These two types of information have their own strengths and weakness. Thus, since static information is probably valid in the future, a single unwanted access will endanger the information in the long term. Consequently we can categorize this information as especially sensible in a time line. On the contrary, dynamic information may change with time, for what it could seem less sensible, on the other hand, is precisely this kind of information what really describes a person?s way of living -e.g. where you are, doing what and with who. Therefore, even not especially sensible in a time line, this kind of information is more directly related to what we understand by privacy than the previous one. Summarizing, we pose that, in order to design privacy management solutions, we should be extremely aware of the idiosyncrasy of the information for which it is going to be designed.

With this purpose, we define dynamic information as what Dey identify as context [4]. This author split context information in four main variables: identity, location, activity and time. These four concepts will guide the implementation proposed in section 2.

1.1. Long-term Personal Data

As we described before, privacy data could be split on two categories. One kind of information is named long-time personal data. This means that the value of this data remains constant with time.

In our environment model [5] each entity has a set of properties and relations associated. Regarding privacy issues, each property or relation must be included in one of the previous groups (i.e. long-term or short-term information). Depending on the assigned group, each property or relation will require a different privacy management policy. Long-term data are insensible to context changes, meaning that once the value is revealed to an unwanted receiver, the data is compromised forever. In this case, the main concern must be who has rights to access the information while when, where, or how is accessed are not relevant. Accordingly, the privacy mechanism associated to long-term personal data is based on a restrictive approach, in which the user must specify which individuals can access each piece of long-term information.

Thus, each property or relation has an associated privacy level, varying, from closeness to openness, in the following range.

- **Private.** A property or a relationship is tagged as private when the owner is the only one with access rights. Nobody, except the owner, can read the value and, if an entity is specified as private, all its properties and relations will be considered private too.
- **Protected.** The owner can specify who has permission to view the information by defining a list of receivers. Only members of the list can access the information. Any non-member will see the information as private.
- **Public.** There are no access restrictions to the information. Data is open to everyone for consulting. This policy can be applied to the whole entity or to a particular property or relation.

1.2. Short-term Personal Data

Contrary to long term information, properties or relations more likely to change over time are classified as short-term information. Many examples can be found in an active

Table 1. Context variables are filtered using different granularity depending on its privacy level.

	Context variables		
Privacy Levels	Identity	Location	Activity
Restricted	Anonymous	Not available	Not Available
Diffuse	Alias	Building Level	High level activity
Open	Real	Room level/GPS data	Raw data

environment in which sensors are constantly gathering information about the inhabitants and the environment itself. Similarly, information inferred from short-term raw data can be categorized as short-term information too. As a result, there is a huge amount of available information categorized as short-term, from inhabitant spatial-temporal coordinates such as location or activity information, to environment information such as temperature, humidity, lighting levels and appliances? status and so on. As a consequence of our previous research [3] we consider only three types of short-term information: identity, location and activity. Even though identity can be considered ?as it is in fact- as a long-term information, it must be also kept within the short-term ones due to the indexing function it bears. Thus, without identity, nor location neither activity has any sense.

Short-term information is classified according to two discrete axes: privacy level and personal data. As illustrated in Table 1, the accuracy of the information varies with the privacy level. This accuracy is characterized by different degrees of granularity of each variable. Although different scales and values can be considered, our three levels approach is motivated by simplicity of use. We believe that more levels could lead to a non-feasible approach partly due to technical issues, but more importantly, because the model should remain simple enough to be comprehensible by non-technical user. Nevertheless, as we will suggest in future work, we will study if a four level approach is also reasonable.

The three privacy levels are:

- **Restricted.** In this level a non-disclosure of information is achieved. Thereby, user's information is not distributed to the community.
- **Diffuse.** Personal data can be retrieved with some level of distortion. Thus, for each variable, an intermediate value is established hinting at the real one without revealing it. This distorted value can be automatically obtained from the real one or hand-written in the user preferences.
- **Open.** The information is revealed without any modification. Hence, the community receives the information as accurate as possible. Besides privacy configuration, the accuracy of the data will depend on the sensors' quality and deployed infrastructure.

The main challenge relies on how to assign these privacy levels to short-term information. One possible approach would be to apply the same strategy that for long-term information. However, due to the constantly changing nature of context, an approach that requires full-control over privacy will force the definition of particular privacy settings for each situation, meaning that users must anticipate every possible situation, or review their privacy preferences every time the situation changes. While the former is non-feasible the latter can be achieved by following an incremental approach, with the system asking for new privacy settings every time a new situation occurs and then remembering it. In this way, the system acquires, step by step, privacy settings close to the ideal. A real example can be found in firewall applications. On the other hand, the main problem with this kind of solutions is that users might find annoying ?as they truly are in the case of firewalls-the repeated interruptions of the system, even more if we consider that notifications might be sent through ambient interfaces such as speakers. A completely opposite approach would be to consider totally automatic privacy control as an option. This implies that the system automatically infers the privacy configuration for each situation in behaves of the user, whom stays aside. However, we believe that it will be hard for end-users to feel comfortable in delegating control over their privacy to an automatic system, unable to explain its behaviour [6]. Although automatic decision-control techniques can be very competent, user direct control is particularly relevant when the variables to be controlled are personal data.

Summarising, and following P. Maes [7] guide on trust and competence as main issues in developing software applications, we would like to emphasize the importance of choosing a comprehensible and useful solution. Where comprehensible means that the user has enough knowledge of the how, when, and why the system reacts to trust it; and useful means that the system is sufficiently easy to use and powerful to achieve a high degree of competence over the problem. If the solution is not trustable, users will end up setting their privacy so restricted that context-aware services would not have enough information to work properly. If the solution is not competent either the users will have problems with their settings or the services with their sources. This is the main reason to adopt an intermediate solution, simple enough for the common understanding, but rather flexible to fulfill users and services expectations.

The idea behind the "Fair Trade" metaphor is the stimulation of sharing information, hypothesizing that users will accept to harm their privacy if, on return, they receive valuable services. In addition to the services they obtain in exchange of their information, users must be able to track the flow of their disclosed data. Our approach emphasizes the "fair trade" metaphor in the following way: users configure their privacy settings according to the amount of information they would like to know from others, assuming that others would be able to access their own information in the same terms. For example, following the previous approach, if a user defines his privacy level as Diffuse, his personal data will be shown after being filtered but he would only be able to retrieve filtered information from others in the best case, even from those with an Open privacy level. At the end, the amount of shared information grows with information needs, making context-aware services more useful and valuable.

Despite all this, another issue of considerable importance for the success of the metaphor has been the optimistic access control included in the system [8]. This access control mechanism provides free information access, adding a logging mechanism of what data is visualized, who accessed it and when was accessed. As a result, this mechanism acts as a dissuasive measure to prevent abuses of the system combined with the punitive actions that can be taken in case ill-intentioned behaviour patterns are detected.

Summarizing, our privacy management proposal relays on trading quality of service for information. Users give their context information data in exchange of better services. A recording mechanism is provided to mitigate the risk of personal information disclosure. Users can revert their privacy preferences to a more restricted one if they believe somebody is misusing their context information.

2. AmIChat: A Context and Privacy Aware Instant Messaging Application

While developing our prototype, we focused particularly on supporting privacy control following the "Fair Trade" metaphor. In section 1 we classified personal data in two categories according to its durability, we motivated a different privacy management approach for each category and we explained both approaches. Even though AmiChat provides information access and privacy control to both long-term and short-term information this article will only explain in detail the latter, being the one in which the "Fair Trade" metaphor is applied.

In AmiChat short-term information is defined in terms of location and activity and the three privacy levels defined in section 1.2 are incorporated to the IM:

1. **Restricted.** Privacy has the highest level of closeness. Personal information remains secret; as a consequence the user looses IM quality of service. When the user connects to the IM his/her identity will be shown as "anonymous", as they will be all their buddies' identity too, regardless their privacy configuration. Albeit this lack of context information (identity, activity and location) the IM service remains active so the user can send and receive messages even though he does not know to whom or from whom.
2. **Diffuse.** In this privacy level, information is retrieved with certain granularity. User's real identity is replaced with an alias (see Figure 1). As a consequence, the user will only view the alias of his/her buddies with diffuse or open privacy levels. The activity to be shown is selected, as in many commercial IMs, from a list with three different options: "Write me!", "Away" and "Don't disturb". Regarding location, the result value is obtained from environment sensors after being filtered by the privacy middleware. Location information is gathered from two different kind of sensor. On one hand, working spaces such offices and laboratories are outfitted with RFID readers. Users getting in and out of environment should pass their RFID cards. Events on change locations are notified to the IM server. Additionally, a spatial model provides a hierarchy of environments including relationships between them, such inclusion or adjacency. On the other hand, IM monitors the user's personal computer status by means of a resident program that informs if the user is connected or is away. The previous spatial model also represents inclusion relationship between resources and environments. Thus, if a user is using her PC, IM can infer where she is.
3. **Open.** Personal information is completely open, hence the user is able to see others' information with the highest possible resolution: the one defined by the information owner. The identity shown is the real one and the location and activity values are directly inferred from sensors (e.g. "At the computer abraxas, but not working" or "At room lab_b403") Thus, personal information is traded for quality of service (see Figure 1).

Finally, since sensing technologies might not be available in every place, the IM client allows configuring manually an alternative version of the location and activity values. Nevertheless, information obtained directly from sensors will always be preferred to information stated manually, in order to avoid conscious cheating from malicious users.

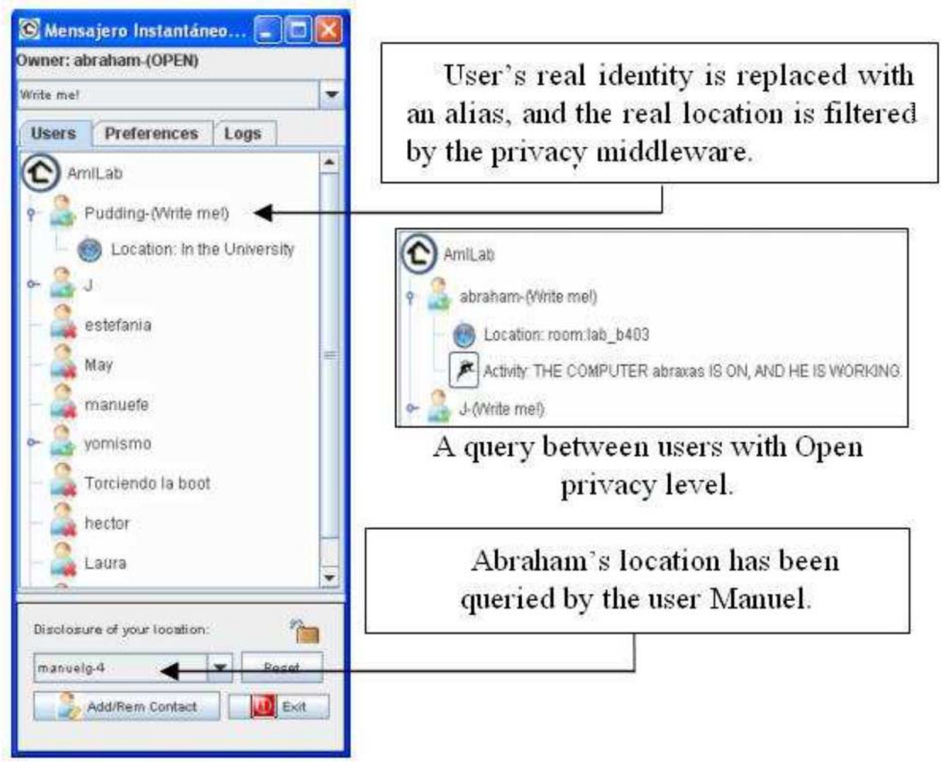


Figure 1. Comparative between Diffuse (left side) and Open (right side) privacy levels.

2.1. Supporting Mechanisms

As we already mentioned, the "Fair Trade" metaphor requires additionally considerations to achieve a satisfactory user experience. In this sense, our prototype uses two optimistic access control premises: free information access (whenever the privacy level allows it), and information access log. These premises, rather than preventing information access as the default policy, they complement the "Fair Trade" metaphor through data sharing: users can see others' information and, more importantly, they can check who is accessing their personal data. Due to the underneath "Fair Trade" policy, if user A access B's information with some granularity x, B will be able to check in the logs A's access with the same granularity x. This information helps B regulating his/her privacy level at any moment if he/she feels uncomfortable. Additionally, A knows in advance that the logs and the "Fair Trade" policy combined will show B the same amount of A's information as he/she is retrieving from B, explicitly entangling personal and alien privacies and deterring abuses such as requesting synchronously and repeatedly another user's location.

Considering these assumptions, we have included several logging mechanisms in the IM. Firstly, the number of location and activity accesses is summarized in the main window. A drop down list shows the requesters and the number of times each one has accessed the information. This list is ordered by the access number, as a measure of privacy risk. Secondly, the IM also keeps detailed record of each access in two logs (one for short-term and one for long-term information) where the user can view when each access has been produced. At that moment, the IM supports exclusion from the contact list as the only punitive measure against users considered to be doing an incorrect use of the application.

3. Context, Purposes and Characteristics of Study

In this work, we propose privacy issues in social and context aware applications. To address this challenge we propose the "Fair Trade" metaphor as the underlying policy of the privacy control mechanism. Categorizing personal information in two different classes: long-term and short-term information, we found the metaphor especially suitable to deal with short-term personal data, what is usually understood as context information. On the other hand, a traditional approach has been proposed to handle long-term information.

To validate our hypothesis we developed a context-aware privacy-sensitive instant messenger with which to evaluate how feasible is the "Fair Trade" metaphor as a privacy control mechanism. The prototype runs on top of the infrastructure deployed in the Universidad Autnoma de Madrid [5] and Instituto Tecnolgico Superior Zacatecas Norte, a working version is been used daily by the researcher staff of our group. We are working on making an evaluation on a broad variety of end-users, in order to determine whether the number of different privacy levels has to be grown or not and if user's black lists should be public.

The goal of our study is evaluate the possibility of encourage users to share their private information promoting the use of contextual services by the community. The instant messenger was developed under "Fair Trade" metaphor, and will be used as a means of communication between habitants of two real active environments, and able to use context information produced by sensors of active environments.

The proposed methodology is as follows:

- A previous interview to the candidates, and pre-questionnaires.
- Installation of Instant Messaging Client.
- Use on a daily basis (two weeks approximately).
- Second interview (feedback effects) and post-questionnaires.
- Recovery of the transaction log file for review.

The interview consists of three sections. The first one contains some questions, which attempt to determine the perception of privacy by the user, classifying different kind of information according to his/her importance. The second one, determine preferred privacy control mechanisms by the user, and finally, the candidate makes an assessment of the mechanisms available for the system in order to alert when someone accesses to his/her private information.

Currently, the call was launched for candidates who wish to participate in the experiment, waiting for results to validate the privacy control through "Fair Trade" metaphor.

4. Conclusions and Future Work

At the time of this writing, we are waiting for candidates who want to help by participating in the experiment. Formal test will also allow a deeper study of the possible punitive measures to apply on abuses —from voting alternatives to access denial— as well as to identify the most sensible variables, from the user perspective —those that users feel as more private- as well as from the abuser one— those being most attacked. Finally, we should mention our current work on integrating two pervasive environments far away from each other, this brings the opportunity of studying privacy in environments in which all communication are digital and, thus, more sensible to our privacy control.

Acknowledgement

This work could not have been made possible without the support of AmILab laboratory of the Universidad Autnoma de Madrid and Instituto Tecnolgico Superior Zacatecas Norte, México.

References

[1] J. Hightower, G. Borriello, Location Systems for Ubiquitous Computing, *IEEE Computer* **34**, 8 (2001), 57–66.
[2] A. Stone, The Dark Side of Pervasive Computing, *IEEE Pervasive Computing* **2**, 1 (2003), 4–8.
[3] A. Esquivel, A, P.A. Haya, M. García-Herranz, X. Alamán, Managing Pervasive Environment Privacy Using the"fair trade" Metaphor, *Cooperative Information Systems (CoopIS), On The Move (OTM) Workshops*, LNCS **4806** (2007), 804–813.
[4] K.A. Dey, Understanding and Using Context, *Personal Ubiquitous Computing* **5**, 1 (2001), 4–7.
[5] P.A. Haya, G. Montoro, X. Alamán, A prototype of a context-based architecture for intelligent home environments, *Cooperative Information Systems (CoopIS), On The Move (OTM)*, LNCS **3290** (2004), 477–491.
[6] D. Norman, How might people interact with agents, *Communications of the ACM* **37**, 7 (1994), 68–71.
[7] P. Maes, Agents that Reduce Work and Information Overload, *Communications of the ACM* **37**, 7 (1994), 31–40.
[8] D. Povey, Optimistic Security: A New Access Control Paradigm, *WNSP: New Security Paradigms Workshop, National Information Systems Security Conference* (1999), 40–45.

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-080-2-434

FollowMe: A Bigraphical Approach

Martin HENSON [a,1], James DOOLEY [a], Luke WHITTINGTON [a] and
Abdullah AL MALAISE AL GHAMDI [b]
[a] *School of Computer Science and Electronic Engineering, University of Essex*
[b] *Faculty of Computing and Information Technology, King Abdulaziz University*

Abstract. In this paper we illustrate the use of modelling techniques using bigraphs to specify and refine elementary aspects of the FollowMe framework. This framework provides the seamless migration of bi-directional user interfaces for users as they navigate between zones within an intelligent environment.

Keywords. bigraph, bigraphical reactive system, refinement, intelligent environment, FollowMe.

Introduction

This paper explores formal methods in the context of intelligent environments. There are many roles such methods might play - for example in providing some evidence of reliability through a verification process. For example, UPPAAL and TCTL [1] have been used in the analysis of AmI applications [2] and SPIN [3] in an AAL application [4]. Here, however, we look at another important possibility: the use of such methods to express and communicate precise designs, and to relate designs at different levels of abstraction, providing perspectives on complex system that abstract away detail.

The framework we use here to examine these ideas is *bigraphs*. The definitive account appears in [5], though the ideas have been around much longer. We begin an examination, to be elaborated in future work, using this framework of *FollowMe*: a proposal for a persistent GUI, that allows for user interfaces to follow a user through the physical world, jumping between local display devices [6]. Bigraphs can be elegantly diagrammed, and we provide a few examples. For reasons of space as much as anything, it is impossible to provide such diagrams throughout this presentation and an algebraic notation is used (see Section 1.8 below). The reader may like to refer to [7,8] for a very gentle introduction and further motivation.

The organisation of the paper is as follows. The the next section we provide a brief overview of bigraph framework, at least at a level of presentation that is adequate for the present work. In the next section we outline the issue that the FollowMe proposal addresses. After that we look at some ways to express some very elementary aspects of this in bigraphical form. Our purpose here is to focus more on the principles and possibilities inherent in deploying the formal method in the context of an application in pervasive computing - so the analysis is highly simplified: we intend to provide a

[1] Corresponding Author: University of Essex, Wivenhoe Park, Colchester, Essex, UK.
E-mail: hensm@essex.ac.uk

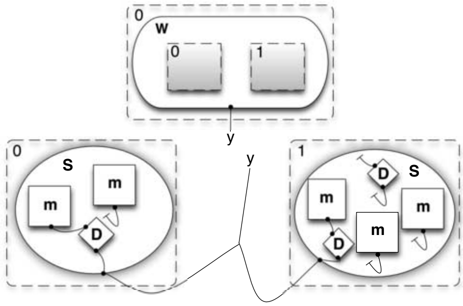


Figure 1. Two Bigraphs

comprehensive analysis in future work. In addition to abstract specifications we look at a the refinement of a specification using a technique developed in [9]. We end some conclusions and related work - especially as that is likely to be crucial to lines for our future work.

1. Bigraphical Reactive Systems

1.1. Bigraphs

A *bigraphical reactive system* (BRS) is a formal model of an interactive system. Underlying this is the notion of a bigraph - comprising two graphs: a *place graph* and a *link graph*. The former captures the spatial organisation of the system, while the latter captures connectivity within the system. Then there is the notion of *reaction rules* that can, in general non-deterministically, rewrite parts of a bigraph and express the potential dynamics of the system. Figure 1 shows how two example bigraphs (one above the other) can be diagrammed. The outer dotted rectangles are called *regions*. The shapes labelled W, S, m and D within these regions are called *nodes*. The labels themselves are called *controls*. The black dots on various nodes are *ports* and the lines connecting ports are *edges*. The inner dotted rectangles (shaded) are *sites*. Finally, the label y on the unconnected ends of two edges are *names*: the name on the upward-facing edge is an *outer name*, while that on the downward-facing edge is an *inner name*.

These diagrammatic representations of bigraphs are very suggestive, and combine the place and link graphs very helpfully. It is important, though, to understand how those two graphs may be untangled from the diagrams. What follows is a relatively informal account; consult [5] for precise details.

1.2. Signatures

The nodes in a bigraph correspond to various places, operations, entities, *etc.* and are classified as such by assigning *controls*. In Figure 1, the controls are W, S, m, and D. Controls are assigned an *arity*, that indicate the number of ports. In Figure 1 the arity of the W, S and m nodes is 1, and for D nodes is 2. A *signature*, κ, gathers all this information together. For our example bigraphs, κ is $\{W:1, S:1, m:1, D:2\}$.

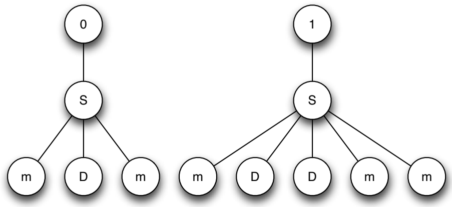


Figure 2. A Place Graph

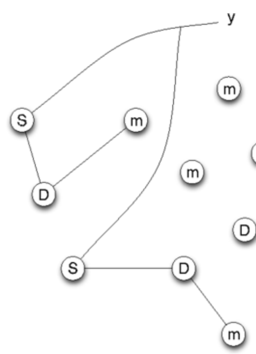


Figure 3. A Link Graph

1.3. Place and Link Graphs

The place graph represents the *nesting* of the regions, nodes and sites in a diagram. For example, the place graph of the lower bigraph of Figure 1 is given in Figure 2. In general, as is the case here, the place graph will be a *forest* of trees - when the bigraph comprises more than one distinct region, each region will be the root of a separate tree.

The link graph represents the connectivity of the nodes in a diagram. For example, the link graph of the lower diagram of Figure 1 is given in Figure 3 As is the case here, a connection may link more than two nodes. So the place graph is, in general, a *hypergraph*.

1.4. Interfaces

A bigraph has two *faces*: an *outer* and an *inner* interface. Each is a pair comprising a face for the place graph and a face for the link graph. A face for a place graph is a natural number. In this framework we treat a natural number as the set of numbers that are strictly smaller. That is, $n = \{0, 1, \cdots, n-1\}$. The inner face of the place graph for the upper bigraph is 2. That is, $\{0, 1\}$ - the bigraph containing two sites. The inner face of the place graph for the lower bigraph is 0. That is, $\{\}$ - the bigraph contains no sites. The outer face of the upper bigraph is 1 and of the lower bigraph is 2, because they comprise one region and two regions, respectively. A face for a link graph is a set of labels - those attached to the unconnected edges. In this framework we write, for example, the set of labels $\{x, y, z\}$ as $\{xyz\}$. The inner face of the link graph for the upper bigraph is $\{y\}$ -

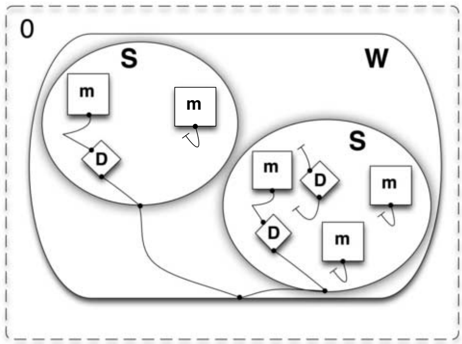


Figure 4. Result of Composition

one inner name. Its outer face is $\{\}$ - no outer names. The inner face of the link graph for the lower bigraph is $\{\}$ - no inner name, while the outer face is $\{y\}$ - one outer name.

The *place graph interface* for the upper bigraph is written $2 \to 1$ - from inner face to outer face. Thus, the place graph interface for the lower bigraph is $0 \to 2$. The *link graph interface* for the upper bigraph is written $\{y\} \to \{\}$ - again from inner face to outer face. Thus, the link graph interface for the lower diagram is $\{\} \to \{y\}$. Given all this, the interface for the upper diagram is $\langle 2, \{y\} \rangle \to \langle 1, \{\} \rangle$ and the interface for the lower diagram is $\langle 0, \{\} \rangle \to \langle 2, \{y\} \rangle$.

1.5. Composition

Note that the outer face of the lower diagram matches the inner face of the upper diagram. This means that the two diagrams may be *composed*. Think of sites receiving regions and inner names receiving outer names. When we compose these two diagrams in this way we obtain Figure 4. The interface of this resulting bigraph is $\langle 0, \{\} \rangle \to \langle 1, \{\} \rangle$. Generally, we will write $\langle k, X \rangle$ (*etc.*) for an interface.

1.6. Reaction Rules

Bigraphs evolve according to *reaction rules*. Figure 5 shows such a rule. The rule involves nodes from Figure 1. We showed various ports in Figure 1 as explicitly disconnected. In Figure 5 we have simply not shown the disconnected ports at all. Here an external member *m* of a space *S* (which may contain other entities - the site) enters the space. The member and the space are contained in the same region, and are therefore proximal before the reaction takes place. Generally there will be a number of reaction rules comprising a bigraphical reactive system. In any given system state (bigraph), several rules may apply, and indeed a given rule may apply in many ways. Thus, evolution of the system under reaction will generally be non-deterministic. A reaction rule, generally, has the form (R, R', η) where R, the *redex* is transformed into R', the *reactum*. η can often be omitted - but is sometimes needed in order to determine how sites in the reactum are related to sites in the redex. We often write reaction rules as $R \to R'$.

Let κ be a signature and B be a set of reaction rules over κ. Then (κ, B) is a *Bigraphical Reactive System* (BRS).

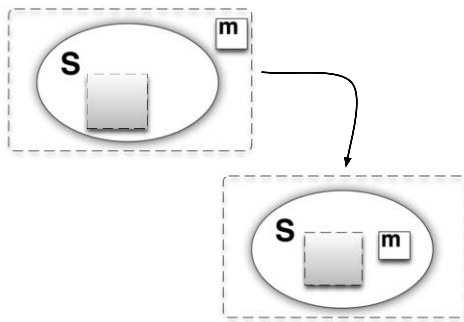


Figure 5. A Reaction Rule

1.7. Bigraphs Formally

We will look at refinement in Section 4, so we need a mathematically precise description of what we have introduced so far.

Definition 1.1 *(See [5] for further details) A bigraph is a quintuple*

$$(V, E, cntrl, prnt, link) : \langle k, X \rangle \rightarrow \langle m, Y \rangle$$

where V is a finite set of nodes, E is a finite set of edges, cntrl is a map from V to the signature κ that assigns controls to nodes, prnt is a (partial) map that takes each site and node (where it has one) to its parent node or region, and link is a map that takes each inner name and port to the edge or outer name to which it is connected.

1.8. Bigraphical Terms

Bigraphical diagrams are extremely helpful and suggestive. However, it is also helpful to have a more concise means to express them. A comprehensive axiomatisation is provided in [10], the language of which is often used, variously extended, for such presentation. For bigraphs A and B, we write $A \parallel B$ for juxtaposition of their regions (distal combination), and $A \mid B$ for juxtaposition within a single region (proximal combination). $U(V)$ represents the containment of V within U. $/x.U$ connects all occurrences of the outer name x in U by an edge (we will generalise this to allow a set of outer names *e.g.* $/X.U$ - writing the set X without brackets and commas). We write $-_i$ for the site labelled i and x/Y for the connection from the inner names in the set Y (which may be empty) to the outer name x (again writing the set Y without brackets and commas). Finally, for a node K with n ports named $\vec{x}$, we write $K_{\vec{x}}(U)$ for *ion* of K (containing U) - that is, K connected to outer names $\vec{x}$.

Given all this, we can write the two bigraphs of Figure 1 as

$$/z.W_z(-_0 \mid -_1) \mid z/y$$

and:

$$/x.S_y(m_x \mid D_{xy} \mid /w.m_w) \mid\mid /x.S_y(m_x \mid D_{xy} \mid /w.m_w \mid /w.m_w \mid /wz.D_{wz})$$

The reaction rule in Figure 5 can be written:

$$/z.S_z(-_0) \mid /y.m_y \;\rightarrow\; /z.S_z(-_0 \mid /y.m_y)$$

2. An Overview of FollowMe

Consider an environment that can be decomposed into zones, where each zone has networked-enabled display devices (and interaction methods such as a touch screen, mouse, keyboard, *etc.*), consider a user, or users, who migrate from zone to zone, and assume that there is a means to detect that migration. The task is to migrate the users' current set of application interfaces from the devices they leave to those at which they arrive. Devices themselves may be mobile and move from zone to zone. This is an example of context-aware computing, of course, and raises additional issues of security (does the user have the right to use the devices), trust (can the user rely on the devices), discovery (what are the devices usable within a zone?), and multiple users (issues of sharing, priority, availability, *etc.*)

This topic, of a persistent GUI in a context-aware scenario, has a history, including *teleporting* using the X-window system to direct client application input/output between different displays [11]. This makes use of an active badge to resolve user location and to prompt display migration, but is not able to dynamically switch the display device used by the application. Another approach is the *Everywhere Displays* project of IBM [12,13] which removes the need for multiple network-enabled display devices and is based on projecting interfaces onto surfaces and using cameras to track the user. There are however, limitations regarding the scale of deployment of such an approach beyond a single zone.

The *FollowMe* approach [6,14] is based on HTML5 technology that permits page-to-server and server-to-page asynchronous event notifications - and by achieving this within a browser, is essentially a device-independent approach.

In this paper we do not by any means reach the level of detail discussed in [6] in terms of the underlying technology, remaining at a much higher level of description and simplifying enormously in order to illustrate some fundamental ideas of modelling in this framework.

3. Basic Modeling

We imagine an *intelligent environment S* with a single port representing a network local to that space. Within the environment there are *zones*. Zones have a single port representing their local network connection. An example might be the *iSpace* at the University of Essex, which is an apartment comprising a living room, kitchen, hall, study, bedroom, and bathroom. Within the environment there are *devices D*, which may be mobile - moving from zone to zone. A device has two ports - the first representing a local (to a zone) network connection and the the second an attachment to a user of the device. Users *U* may also move from zone to zone. Each has a single port representing a potential attachment to a device. This crude picture captures very little - but it is the basic backdrop against which more specific details of the FollowMe approach can be modelled. We have effectively outlined a signature and can describe a simple BRS that captures these high-level considerations. In fact the signature is probably insufficient - as we may wish to avoid zones containing environments, though perhaps not zones containing zones. Such constraints can be enforced, but we will not dwell on the technical details in this paper.

3.1. Mobile Devices

In this section we will assume only one intelligent environment (one S node) at the outermost level of any system state and we will assume no nested zones.

We will look at the issue of a device migrating from one zone to another. Here is our first reaction rule (Rule R1).

$$Z_{x_0}(D_{x_1x_2} \mid -_0) \mid Z_{x_3}(-_1) \rightarrow Z_{x_0}(-_0) \mid Z_{x_0}(D_{x_1x_2} \mid -_1)$$

This is already quite complex. The outer names (the x_i) may connect variously to entities comprising the context into which the redex of the rule matches. For example, consider the following situation.

$$/x_0x_1x_2x_3.S_{x_0}(Z_{x_0}(D_{x_1x_2} \mid -_0) \mid Z_{x_3}(-_1))$$

The local network connection of the leftmost zone into the local network of the space. The device is neither connected to a user nor to the network. The rightmost zone is not connected to the network. However, consider this slightly different situation.

$$/u_0u_1.S_{u_0}(Z_{u_0}(D_{u_0u_1} \mid -_0) \mid Z_{u_0}(-_1))$$

In this case both the zones and the device connect to the network connection of the space. The device does not connect to any user.

To see the reaction rule in action, consider the following bigraph. Once again, note that we create the edges at the outermost level.

$$/v_0v_1v_2.S_{v_0}(Z_{v_0}(U_{v_1} \mid D_{v_0v_1}) \mid Z_{v_2}())$$

When the reaction takes place, the device moves from one zone to the other, stretching (so to speak) the links to the user and the leftmost zone's network (note that the rightmost zone is unconnected).

$$/v_0v_1v_2.S_{v_0}(Z_{v_0}(U_{v_1}) \mid Z_{v_2}(D_{v_0v_1}))$$

Although we shall not provide any, we might develop rules that specify user interaction across the network through a connected device. If we insist that all such interaction requires that the user and the device are in the same zone, and that the device is connected to the local (zone) network. Then the stretched links we have generated in our example are merely *notional* connections (that is, they are not usable). The only reactions we might provide would be to break such notational links. For example, this reaction rule removes the unusable connection between a device in a zone other than its (former) user.

$$/z.Z_{x_0}(-_0 \mid U_z) \mid Z_{x_1}(-_1 \mid /x.D_{xz}) \rightarrow Z_{x_0}(-_0 \mid /z.U_z) \mid Z_{x_1}(-_1 \mid /xz.D_{xz})$$

Note that the closure of x at the device, prevents the device from connecting to any entity in an instance of the site $-_1$ and to anything within an enclosing context. We are insisting that the device, having moved unceremoniously away from its user, may not engage in any network interactions. Thus this reaction cannot take place until the unusable network link has been removed. That would be captured by this reaction rule.

$$/z.Z_z(-_0) \mid Z_u(-_1 \mid D_{zv}) \rightarrow Z_z(-_0) \mid Z_u(-_1 \mid /x.D_{xv})$$

This analysis illustrates well the subtleties of both the example and the framework we are using. The reader may find it useful to draw the “node and link" diagrams that correspond to the bigraphical terms we have presented here.

3.2. Cloud and Spaces

We now want to model a situation in which the users’ data is stored in a cloud that is potentially remote from the environment. This data is served to the device to which a user is connected. We will add, then, an entity C of arity 1, to represent the cloud server. The port is the wide-area network connection to the environment (and probably much else besides). An environment can connect and disconnect to its cloud server. This is captured by the following reaction rules.

$$/x.C_x(-_0) \mid\mid /x.S_x(-_1) \;\rightarrow\; /x.(C_x(-_0) \mid\mid S_x(-_1))$$
$$/x.(C_x(-_0) \mid\mid S_x(-_1)) \;\rightarrow\; /x.C_x(-_0) \mid\mid /x.S_x(-_1)$$

Note that the combination is distal. The two sites allow both the cloud and the environment to contain arbitrary entities. We have specified a disconnection that is not graceful - there is no guarantee that a user’s session will terminate smoothly.

Using the distal combination does not force the cloud server to be remote from the environment. If B is a single building then $B(-_0 \mid -_1)$ can host both the server and the environment locally. On the other hand if *Mexico* and *Portugal* are the countries one imagines, then $Mexico(-_0) \mid\mid Portugal(-_1)$ can separately host the cloud and the environment.

We shall keep the model simple, not identifying *specific* cloud-based data for a particular user. We want to bring over the cloud-based data when a user is connected to a device which is connected to the cloud through its enclosing zone and environment.

$$/x.(C_x(-_0) \mid\mid S_x(Z_x(/y.(U_y \mid D_{xy}) \mid -_1) \;\rightarrow$$
$$/x.(C_x(-_0) \mid\mid S_x(Z_x(/y.(U_y \mid D_{xy}(-_0)) \mid -_1)$$

A user may disconnect from a device.

$$/y.(U_y \mid D_{xy}(-_0)) \;\rightarrow\; /y.U_y \mid /y.D_{xy}$$

That is, when a user disconnects from a device, her session is immediately terminated. This is, perhaps, too abrupt. To improve on this, first let us permit the user to change the local data. Let d be a datum of arity 0.

$$/y.(U_y \mid D_{xy}(-_0)) \;\rightarrow\; /y.(U_y \mid D_{xy}(d \mid -_0))$$

Of course, this is the barest indication only as to how a full account would be constructed - though the reaction here allows an arbitrary number of such data changes.

Given this, we can modify the disconnection reaction above, so that the local data is saved - transferred to the cloud on disconnection.

$$/x.(C_x(-_0) \mid\mid S_x(Z_x(/y.(U_y \mid D_{xy}(-_1)) \mid -_2) \rightarrow$$
$$/x.(C_x(-_1) \mid\mid S_x(Z_x(/y.U_y \mid /y.D_{xy}) \mid -_2)$$

If the user were later to move to another zone and connect to an available device, the previous session would then be reinstated to that new device.

4. Refinement

Consider very simple reactions that would allow a user to connect to, and disconnect from, an available device.

$$/y.U_y \mid /y.D_{xy} \rightarrow /y.(U_y \mid D_{xy})$$
$$/y.(U_y|D_{xy}) \rightarrow /y.U_y \mid /y.D_{xy}$$

We might wish to insist that a user has the necessary credentials to make such a connection. So we might refine the system - requiring that the user has a *code* (a nullary node K) with which to attach to an appropriate device (with the same code) . The code disappears from the device once the user connects. The code should become available again once the user relinquishes control of the device.

$$/y.U_y(K) \mid /y.D_{xy}(K) \rightarrow /y.(U_y(K) \mid D_{xy})$$
$$/y.(U_y(K)|D_{xy}) \rightarrow /y.U_y(K) \mid /y.D_{xy}(K)$$

In the first regime, any user can connect to any available device. In the second, a user may have no code at all - or a code different from that of a given device - so there is a potential reduction of non-determinism in moving from one model to the other.

There should be a *formal* relationship between these models. Is the implementation (that includes the security code) correct? That is, does it conform to the specification (without the security code)?

An answer might be: when any behaviour of the implementation is consistent with a behaviour of the specification. The notion of behaviour is obviously captured by the reaction rules. So imagine, for a BRS A, a set $Tr(A)$ (we make this a little more precise below) of *traces* each of which is a record of the evolution of a bigraph of A under its reaction rules. Then to a first approximation we want $Tr(C) \subseteq Tr(A)$. That is, every behaviour of the implementation BRS, C, is a behaviour of the specification BRS, A. However, this cannot be right, since the implementation comprises bigraphs different from those of the specification (in our example they contain K nodes that do not exist in the specification). So we need some abstraction mapping F from the implementation to the specification that negotiates the differences - and then we want $F(Tr(C)) \subseteq Tr(A)$.

A little work has been undertaken in the area of refinement, notably, in [9], where, among other things, a notion of *safe vertical refinement* is introduced, which builds on the narrative above. We provide a précis of the definitions and results here; the reader is encouraged to consult the source for full details.

Definition 4.1 *(from [9]) A* trace *for a BRS A is a (possibly infinite) sequence of bigraphs of A $\langle a_0, a_1, ...\rangle$ such that $a_i \rightarrow a_{i+1}$ is a reaction of A. We write Tr(A) for all the traces of a BRS A. If F is a functor from BRS A to BRS A' then we extend F to traces of A by applying it point-wise.*

First, note that $Tr(A)$ is *prefix closed*. That is, if t' is a prefix of the sequence (trace) t and $t \in Tr(A)$ then $t' \in Tr(A)$. In particular, the empty trace is in $Tr(A)$. Second, we need the notion of a *functor*. A full description is beyond the scope of this paper - however, essentially, a functor F from A to A' is a pair of maps from interfaces of A to interfaces of A' and a map from bigraphs of A to bigraphs of A' (respecting interfaces) that, in particular, preserves composition (as described in Section 1.5 above).

Definition 4.2 *(from [9])*

$$A \sqsubseteq_F C =_{df} F(Tr(C)) \subseteq Tr(A)$$

That is, abstraction (specification) A refines to concrete (implementation) C when the traces of C (abstracted to the signature of A via F) are all traces of A.

What kinds of functors lead to refinements? Clearly not every functor will do - since $F(t)$ for a trace t must be a trace of the specification BRS, and not just some sequence of bigraphs of that specification. The authors of [9] introduce a notion of a *safe abstraction functor*, which has the following pleasant property: if F is a safe abstraction functor from C to A then C is a refinement of A. One form of safe abstraction functor is a *hiding functor*, and that is what we need here.

Proposition 4.1 *(adapted from [9]) Let C be a BRS over signature Σ and let H be a set of controls of Σ. Let A be a BRS over the signature Σ restricted to controls not in H. Then the* hiding functor *F defined as follows is a safe abstraction functor. F is the identity mapping on interfaces. Then, for any bigraph β, of C: $(V,E,cntrl,prnt,link) : \langle k,X\rangle \rightarrow \langle m,Y\rangle$ we define $F(\beta)$ to be $(V',E,cntrl',prnt',link) : \langle k,X\rangle \rightarrow \langle m,Y\rangle$ where V' is that subset of V whose controls are not in H, $cntrl'$ is cntrl restricted to V', and where:*

$$prnt'(l) =_{df} \begin{cases} prnt(l) & when\ cntrl(prnt(l)) \notin H \\ prnt'(prnt(l)) & otherwise \end{cases}$$

Since a hiding functor is a safe abstraction functor, it leads immediately to a refinement. In our example we simply take H to be $\{K\}$ - the security code.

5. Conclusions, Related and Future Work

The objective of this paper is to introduce the basis notions of bigraphical reactive systems as a basis for modelling systems within an intelligent environment framework, and we have done little more than hint in the general direction of how a sophisticated modelling of the FollowMe framework could be taken.

Previous work in modelling context-aware systems has established that reconfigurations of systems (such as we have illustrated here) are not difficult to model. However, querying the state of a system - establishing in particular that some state of affairs is *not* the case - is hard. Moreover, the straightforward approach we have taken here takes it as given that the user has perfect knowledge of the context - for example is fully aware of what devices are available in the zone they occupy. A more sophisticated approach would allow the user to have a representation of the context. That representation may be incomplete - for example, some devices present in the actual context may not be present in the representation. Such an approach has been termed *Plato-graphic* [15,16]. Such models comprise three components: a *context*, a *proxy*, and an *agent*. The context and agent are *disjoint*, meaning that their signatures have no common entities: the agent interacts with the world (context) only through a representation (proxy).

Safe refinement is not sufficient, since the empty trace (the system that does nothing) is always a safe refinement of any specification. [9] consider *liveness* too - something that needs considerably more attention.

Consideration of such issues is beyond the scope of this paper, but one of the authors (Whittington) is working on a systematic analysis of the FollowMe framework which will involve modelling the user's representation of the actual environment plato-graphically, the refinement of descriptions at various level of detail, and using refinement in order to provide system perspectives - for example from the user's point of view as they navigate the environment. This analysis will also capture the additional issues of security, trust, discovery, and competition between multiple users.

Acknowledgements

The authors wish to thank Hani Hagras, Gaëtan Pruvost, and Christian Wagner, who collaborated on the development of the FollowMe design and implementation. This work has been undertaken as part of the *ScaleUp* project, which is funded by King Abdulaziz University, KSA, and as part of the EU FP7 funded ATRACO project.

References

[1] B. Bérard, M. Bidoit, A. Finkel, F. Laroussinie, A. Petit, L. Petrucci, and Ph. Schnoebelen, *Systems and Software Verification, Model-Checking Techniques and Tools*, Springer, 2001.

[2] J. C. Augusto and P. J. McCullagh, Ambient intelligence: Concepts and applications,*Comput. Sci. Inf. Syst.* **4**, 1 (2007), 1–27.

[3] G. J. Holzmann, *The SPIN Model Checker - primer and reference manual*, Addison-Wesley, 2004.

[4] J. C. Augusto, H. Zheng, M. D. Mulvenna, H. Wang, W. Carswell, and W. P. Jeffers, Design and Modelling of the Nocturnal AAL Care System, in *Ambient Intelligence - Software and Applications - 2nd International Symposium on Ambient Intelligence* 2011, 109–116.

[5] R. Milner, *The Space and Motion of Communicating Agents*, Cambridge University Press, 2009.

[6] J. Dooley, C. Wagner, H. Hagras, and G. Pruvost, FollowMe: The Persistent GUI, in [14], IEEE Computer Society 2011, 123–126.

[7] M. Henson, J. Dooley, L. Whittington, and A. A. G. A. Malaise, Towards Simple and Effective Formal Methods for Intelligent Environments, in [8] *(forthcoming)*, 2012.

[8] J. C. Augusto and V. Callaghan, Eds., *Eighth International Conference on Intelligent Environments (IEâĂŹ12)*, IEEE Computer Society, 2012.

[9] G. Perrone, S. Debois, and T. Hildebrandt, Bigraphical refinement, *Electronic Proceedings in Theoretical Computer Science* **55** (2011), 20–36.

[10] R. Milner, Axioms for bigraphical structure, *Mathematical Structures in Computer Science* **15**, 6 (2005), 1005–1032.

[11] T. Richardson, F. Bennett, G. Mapp, and A. Hopper, A Ubiquitous, Personalized Computing Environment for All: Teleporting in an X Window System Environment, *IEEE Personal Communications* **1**, 3 (1994), 6–12.

[12] G. Pingali, C. Pinhanez, T. Levas, R. Kjeldsen, and M. Podlaseck, User-following displays, in *IEEE International Conference on Multimedia and Expo (ICME)*, 2002.

[13] G. Pingali, C. Pinhanez, A. Levas, R. Kjeldsen, M. Podlaseck, H. Chen, and N. Sukaviriya, Steerable interfaces for pervasive computing spaces, in *Pervasive Computing and Communications (PerCom)* 2003, 315–322.

[14] *Sixth International Symposium on Parallel Computing in Electrical Engineering (PARELEC)*, IEEE Computer Society, 2011.

[15] L. Birkedal, S. Debois, E. Elsborg, T. Hildebrandt, and H. Niss, Bigraphical models of context-aware systems, in [16], 2006.

[16] L. Aceto and A. Ingólfsdóttir, Eds., *Foundations of Software Science and Computation Structures, 9th International Conference (FOSSACS)* LNCS 3921, Springer, 2006.

This page intentionally left blank

1st International Workshop on Improving Industrial Automation Using the Intelligent Environments Paradigm (AIE'12)

Juan Carlos Orozco (ACELAB Industrial Automation, México)
Victor Zamudio (Instituto Tecnológico de León, México)

Workshop Proceedings of the 8th International Conference on Intelligent Environments
J.A. Botía et al. (Eds.)
IOS Press, 2012

Preface to the Workshop on Improving Industrial Automation Using the Intelligent Environments Paradigm IA´12

Intelligent Environments (IE) refer to physical spaces in which IT and other pervasive computing technology are woven and used to achieve specific goals for the user, the environment or both. IE enriches user experience, better manages the environment's resources, and increases user awareness of that environment. Several aspects of Industrial Automation solutions can be improved using the paradigm of Intelligent Environments. For example improved man machine interfaces, training using intelligent e-learning techniques, augmented reality for in field equipment documentation to name a few. Industrial Automation present in most modern factories already include Programmable Logic Controllers (PLCs), Human Machine Interfaces (HMIs), Sensors, Actuators in some cases networked to a centralized Supervisory Control And Data Acquisition System (SCADA). We could use these elements as part of Intelligent Environment solutions to lower the cost of implementation. The objective of this workshop is first to review the current state of the art for *Intelligent Environments* (IE), then present some possible examples of areas where this technology can improve the field of *Industrial Automation* (IA).

To achieve these goals we will have our Keynote speaker Prof. Vic Callaghan, leader of the Intelligent Environments Group (IEG) and director of the interdisciplinary Digital Lifestyles Centre (DLC) at Essex University. Prof. Callaghan will present us an overview and current state of the art of IE. Our second Keynote speaker Dr Keith Prettyjohns, CEO and Chairman of Innovasic, will explain how to manage real-time control over Ethernet in busy industrial Intelligent Environments (IE). Victor Zamudio, will present a short video of an application of Augmented Reality.

In the last part of the workshop Juan Carlos Orozco will present a manufacturing company study case from which we will have an interactive discussion with the workshop participants about how to improve the automation systems of this company using IE technologies. This study case includes some possible uses of IE tools to improve IA, other uses of IE tools and strategies to this area are expected to emerge from the interactive discussion with the multidisciplinary audience. The notes and conclusions from this discussion will be published in the workshop website: http://www.acelab.com/AutomationIE/

Juan Carlos Orozco and Víctor Zamudio
Co-chairs IA'12

This page intentionally left blank

Subject Index

Author Index

This page intentionally left blank

This page intentionally left blank