



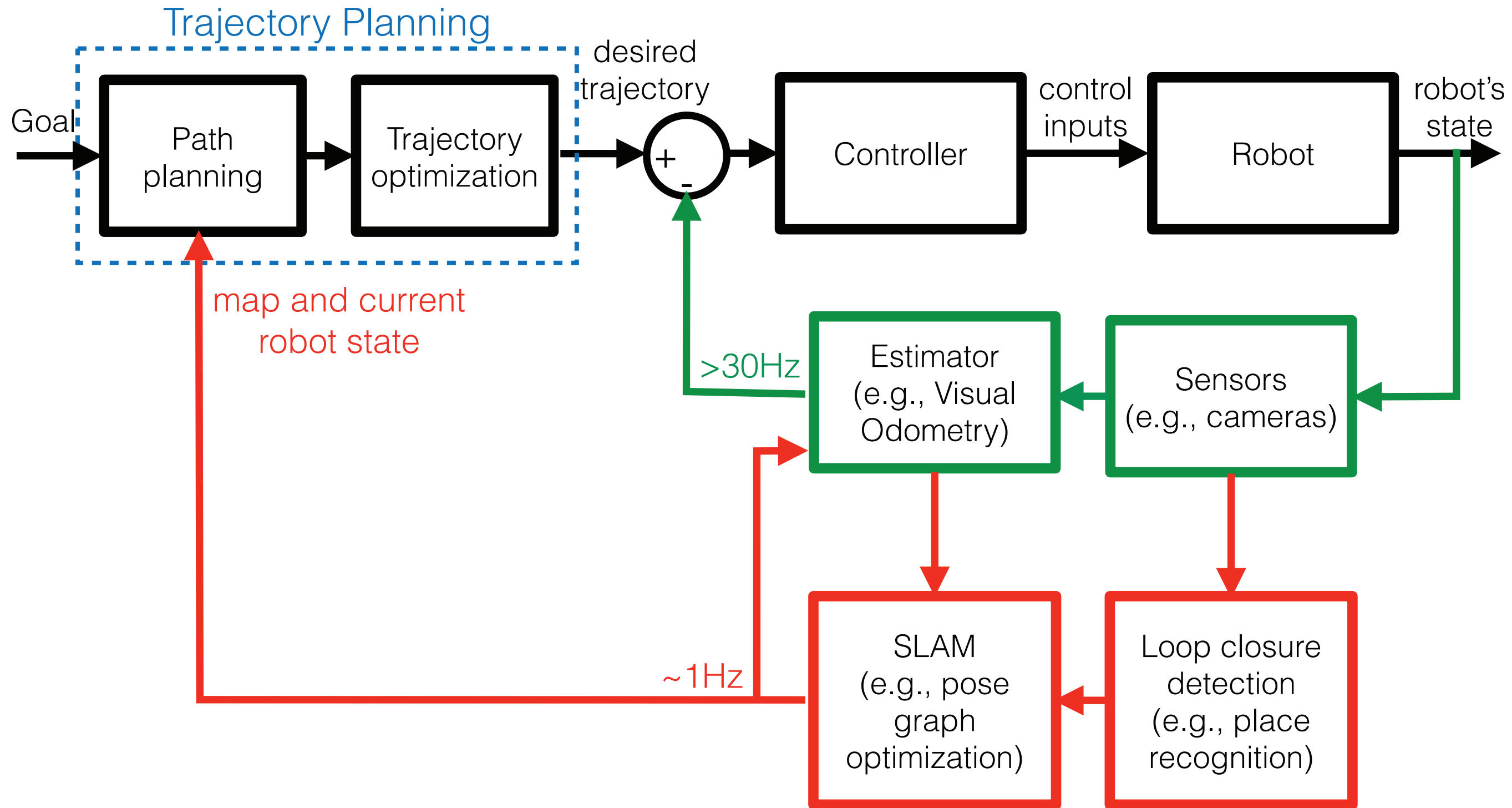
16.485: VNAV - Visual Navigation for Autonomous Vehicles

Luca Carlone

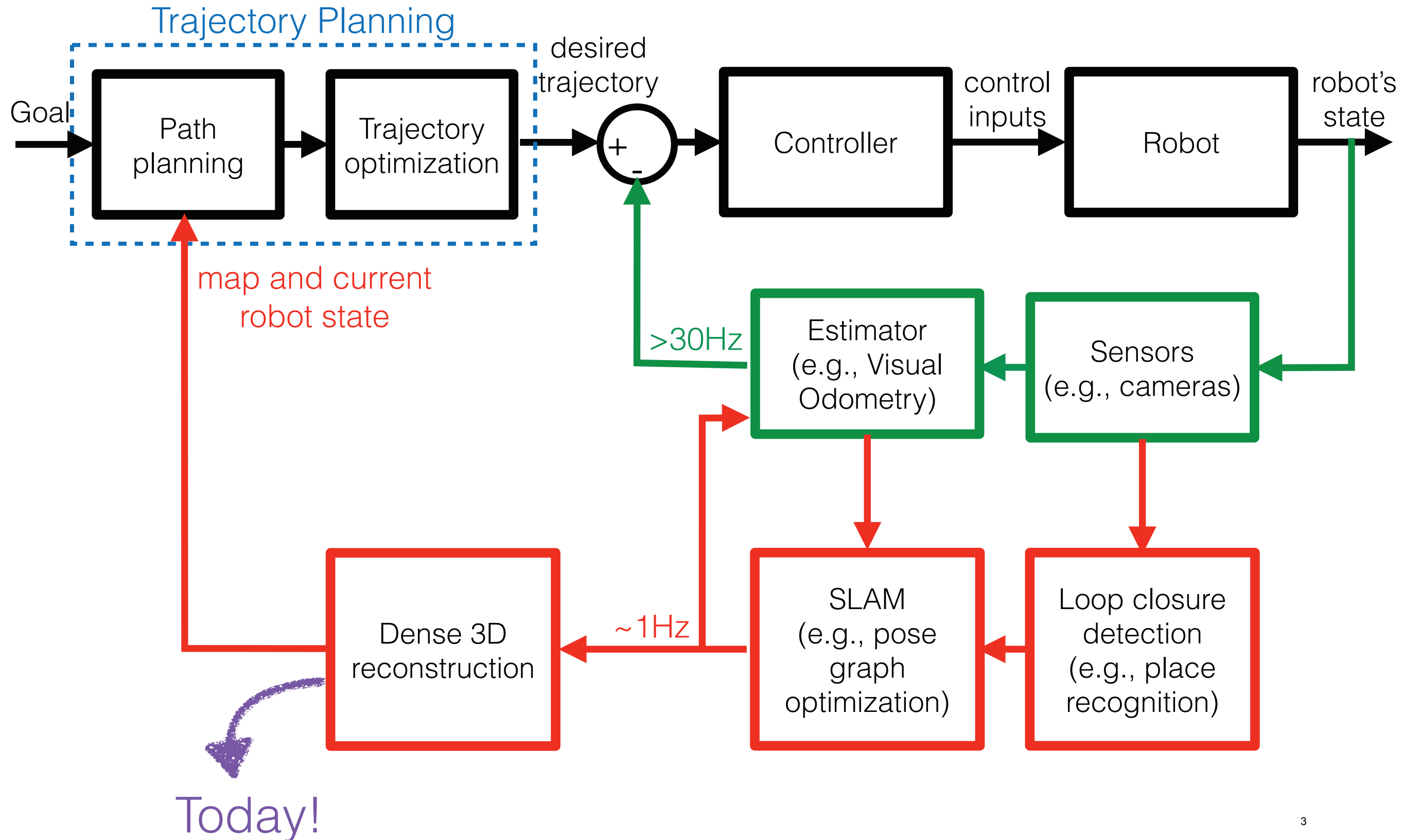
Lecture 25: Advanced topics:
Dense 3D Reconstruction



Big Picture



Big Picture



Today

- Dense Reconstruction
 - 3D representations
 - (Some) Multi-view Stereo
 - Depth fusion
- Final thoughts

Figure 1 in R. A. Newcombe et al., "KinectFusion: Real-time dense surface mapping and tracking," 2011 10th IEEE International Symposium on Mixed and Augmented Reality, Basel, Switzerland, 2011, pp. 127-136, doi: 10.1109/ISMAR.2011.6092378. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

KinectFusion: Real-Time Dense Surface Mapping and Tracking*

Richard A. Newcombe Imperial College London	Shahram Izadi Microsoft Research	Otmar Hilliges Microsoft Research	David Molyneaux Microsoft Research Lancaster University	David Kim Microsoft Research Newcastle University
Andrew J. Davison Imperial College London	Pushmeet Kohli Microsoft Research	Jamie Shotton Microsoft Research	Steve Hodges Microsoft Research	Andrew Fitzgibbon Microsoft Research



2011

Figure 1: Example output from our system, generated in real-time with a handheld Kinect depth camera and no other sensing infrastructure. Normal maps (colour) and Phong-shaded renderings (greyscale) from our dense reconstruction system are shown. On the left for comparison is an example of the live, incomplete, and noisy data from the Kinect sensor (used as input to our system).

Multi-View Stereo: A 2015 Tutorial

Yasutaka Furukawa
Washington University in St. Louis
furukawa@wustl.edu

Carlos Hernández
Google Inc.
carloshernandez@google.com

ElasticFusion: Dense SLAM Without A Pose Graph

Thomas Whelan*, Stefan Leutenegger*, Renato F. Salas-Moreno[†], Ben Glocker[†] and Andrew J. Davison*

*Dyson Robotics Laboratory at Imperial College, Department of Computing, Imperial College London, UK

[†]Department of Computing, Imperial College London, UK

{t.whelan,s.leutenegger,r.salas-moreno10,b.glocker,a.davison}@imperial.ac.uk

2016

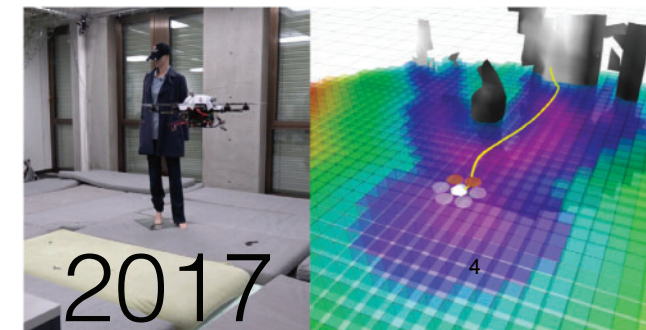
Figure 1 in H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart and J. Nieto, "Voxblox: Incremental 3D Euclidean Signed Distance Fields for on-board MAV planning," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, 2017, pp. 1366-1373, doi: 10.1109/IROS.2017.8202315. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Voxblox: Incremental 3D Euclidean Signed Distance Fields for On-Board MAV Planning

Helen Oleynikova, Zachary Taylor, Marius Fehr, Roland Siegwart, and Juan Nieto
Autonomous Systems Lab, ETH Zürich

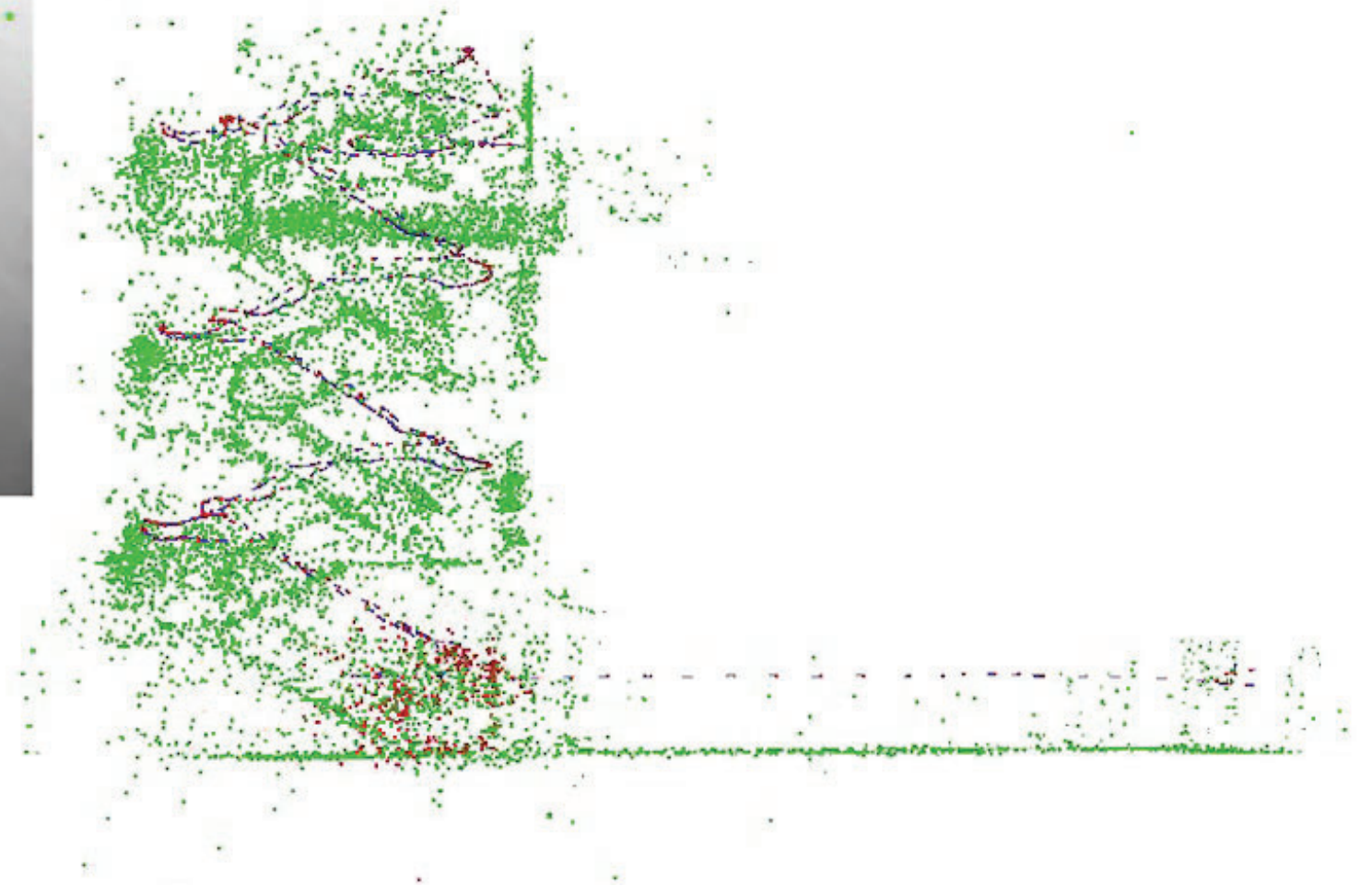
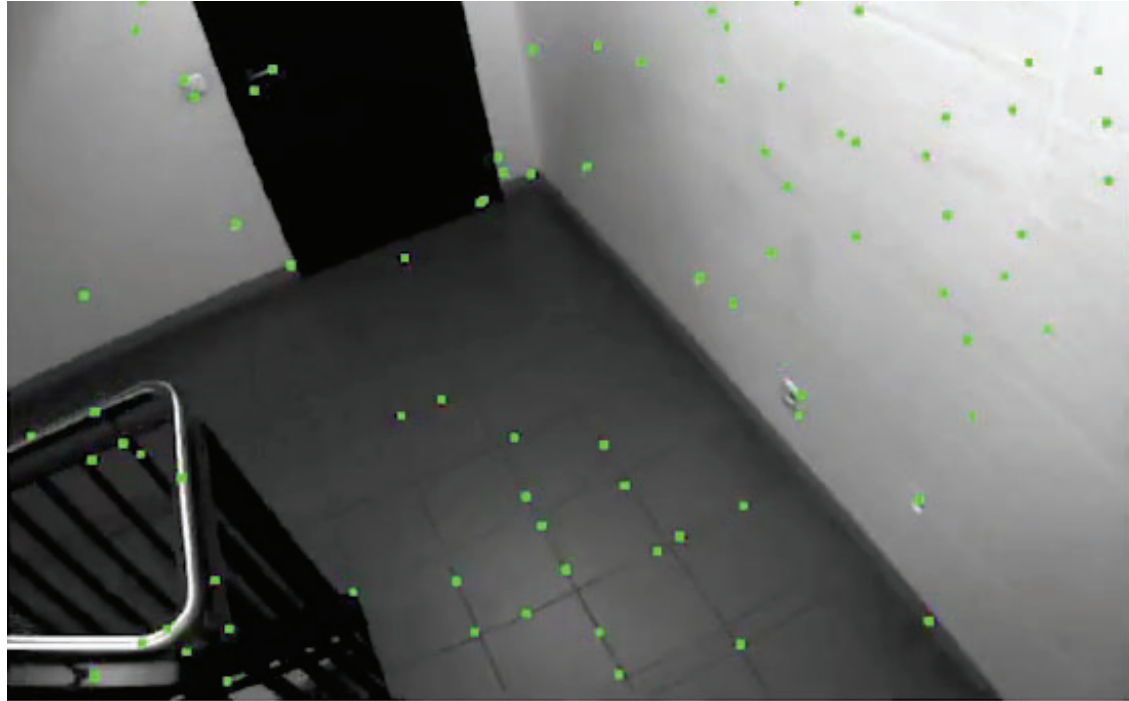
Abstract—Micro Aerial Vehicles (MAVs) that operate in unstructured, unexplored environments require fast and flexible local planning, which can replan when new parts of the map are explored. Trajectory optimization methods fulfill these needs, but require obstacle distance information, which can be given by Euclidean Signed Distance Fields (ESDFs).

We propose a method to incrementally build ESDFs from Truncated Signed Distance Fields (TSDFs), a common implicit surface representation used in computer graphics and vision. TSDFs are fast to build and smooth out sensor noise over many observations, and are designed to produce surface meshes. Meshes allow human operators to get a better assessment of the robot's environment, and set high-level mission goals.



2017

Point Clouds



Point Clouds

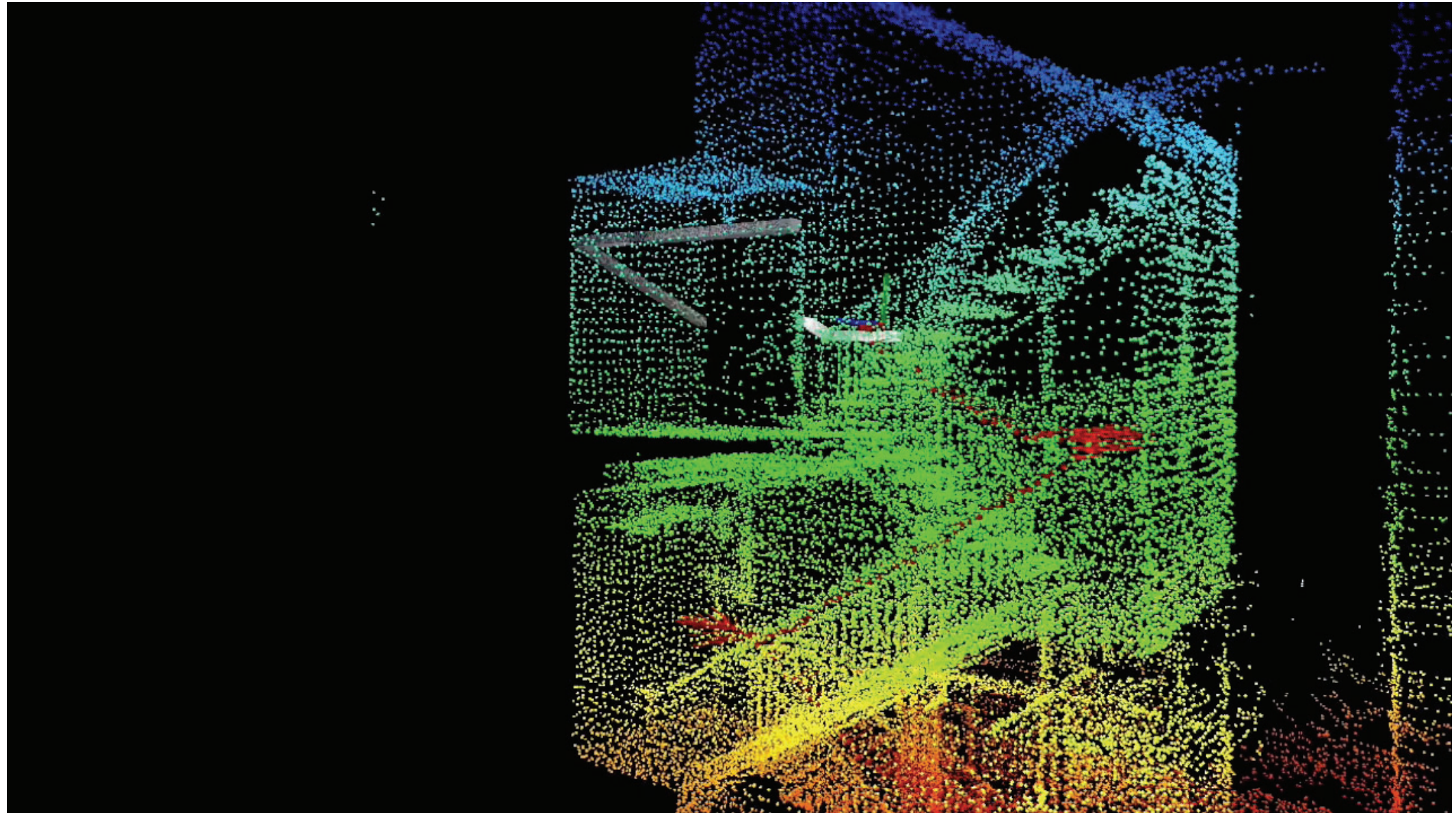


✓/✗
No, if Dense



✓/✗
No, if Sparse

Point Clouds

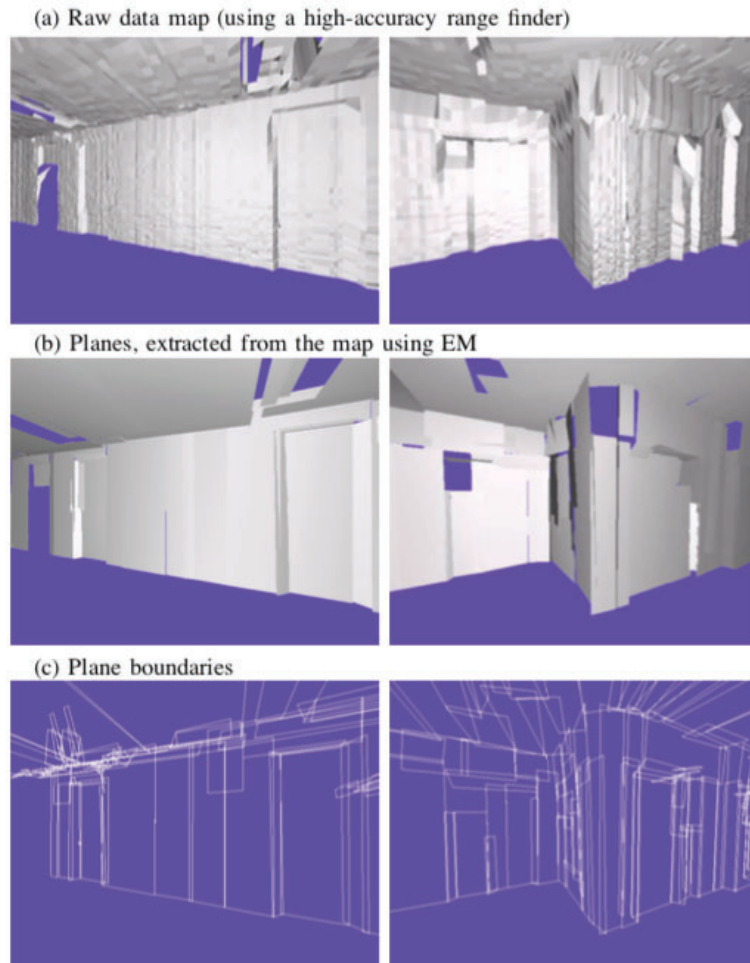


Map representation	3D Topology?	Lightweight?	Filters Noise/Outliers?	Semantics?	Generality
Point Clouds	✗	✓/✗ No, if Dense	✗	✓/✗ No, if Sparse	✓

Geometric Primitives

Point, lines, planes

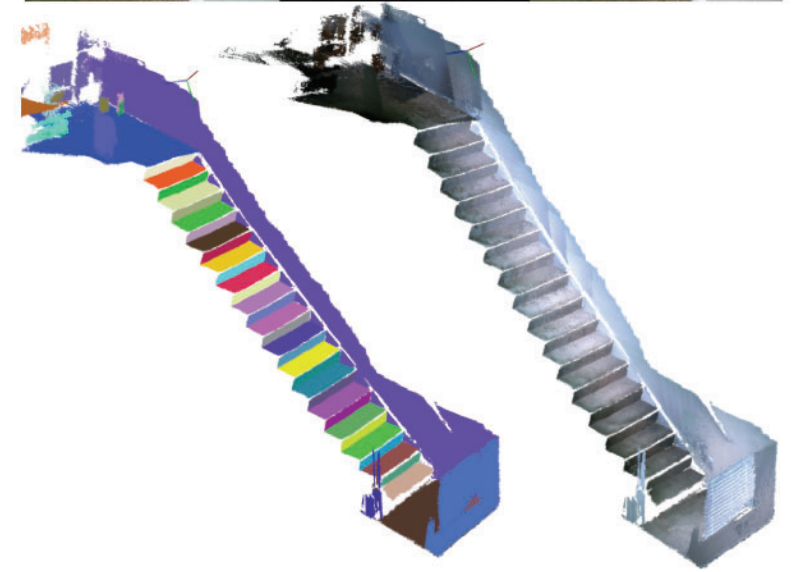
[Thrun et al. 2004]



© Thrun et al. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

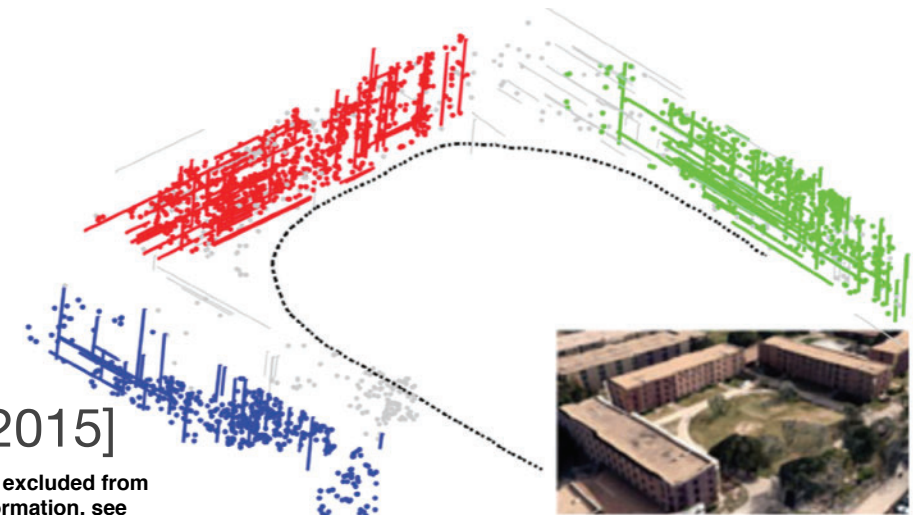
[Kaess 2015]

Figure 1 in Michael Kaess, "Simultaneous Localization and Mapping with Infinite Planes." June 2015 Proceedings - IEEE International Conference on Robotics and Automation 2015:4605-4611. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>



[Lu et al. 2015]

© Lu et al. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>



Map representation	3D Topology?	Lightweight?	Filters Noise/Outliers?	Semantics?	Generality
Point Clouds	✗	✓/✗ No, if Dense	✗	✓/✗ No, if Sparse	✓
Geometric primitives	✗	✓	✓	✓/✗ No, if Sparse	✗ ⁷

Object-based Maps

[Salas-Moreno et al, 2014]

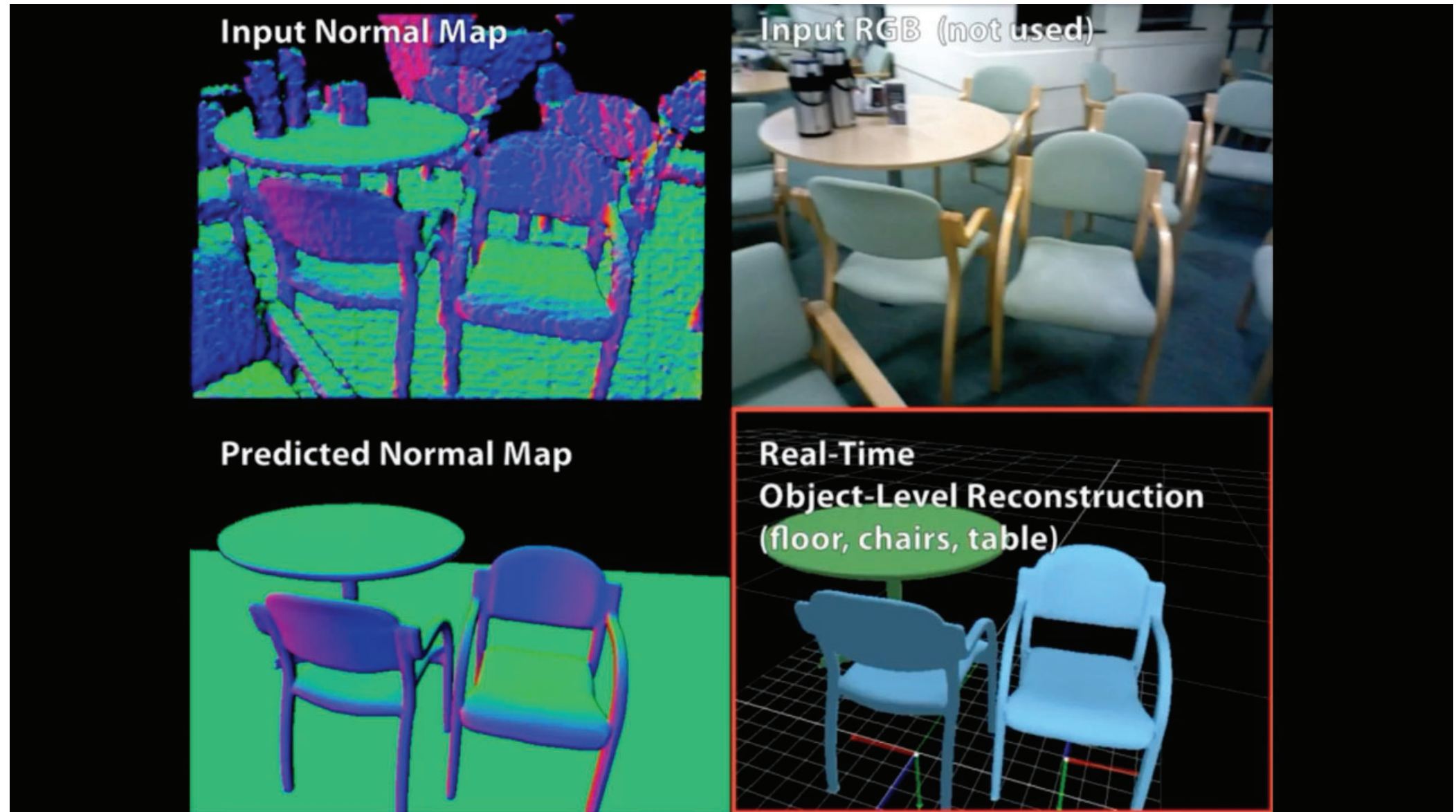


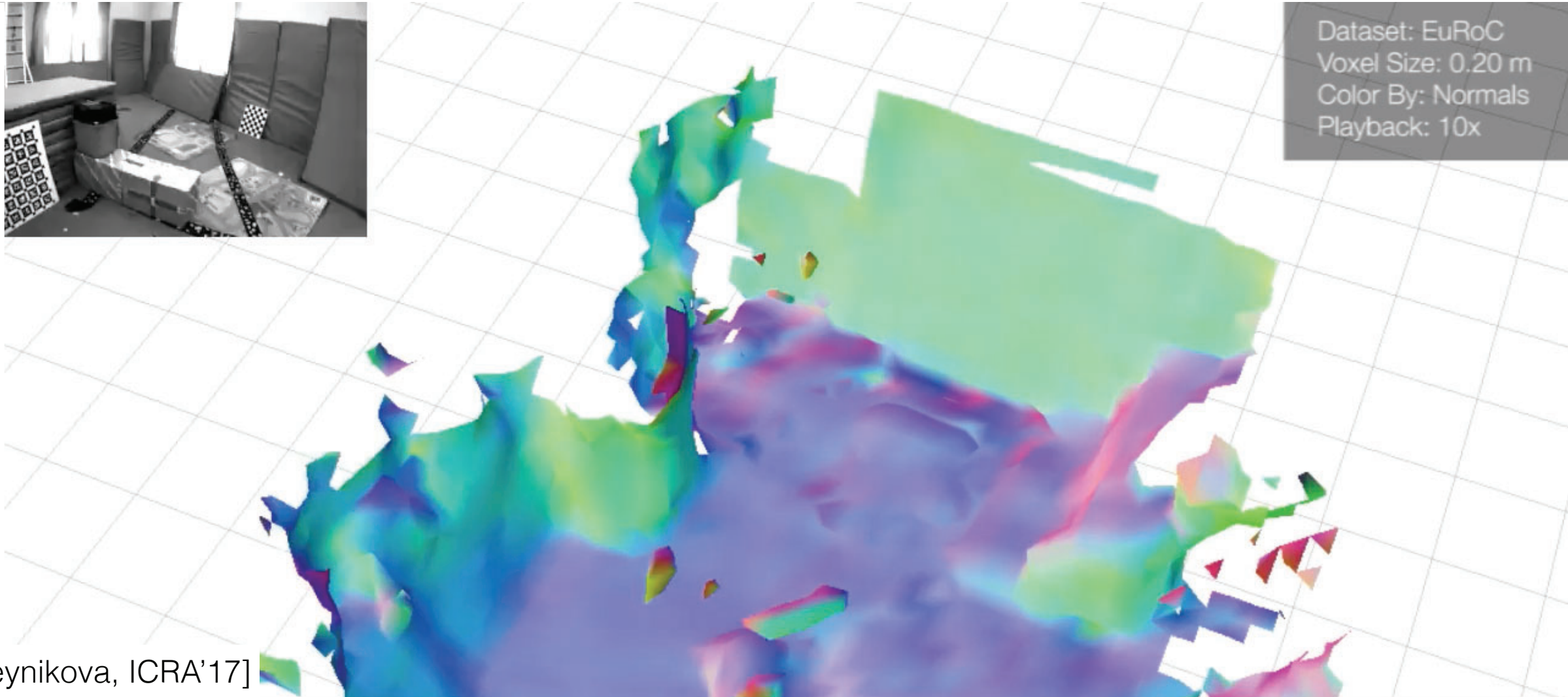
Figure 3 in R. F. Salas-Moreno, R. A. Newcombe, H. Strasdat, P. H. J. Kelly and A. J. Davison, "SLAM++: Simultaneous Localisation and Mapping at the Level of Objects," 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 2013, pp. 1352-1359, doi: 10.1109/CVPR.2013.178 © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Map representation	3D Topology?	Lightweight?	Filters Noise/Outliers?	Semantics?	Generality
Point Clouds	✗	✓/✗ No, if Dense	✗	✓/✗ No, if Sparse	✓
primitives & objects	✗	✓	✓	✓/✗ No, if Sparse	✗ ⁸

Volumetric Methods: Voxels/Octrees



Dataset: EuRoC
 Voxel Size: 0.20 m
 Color By: Normals
 Playback: 10x

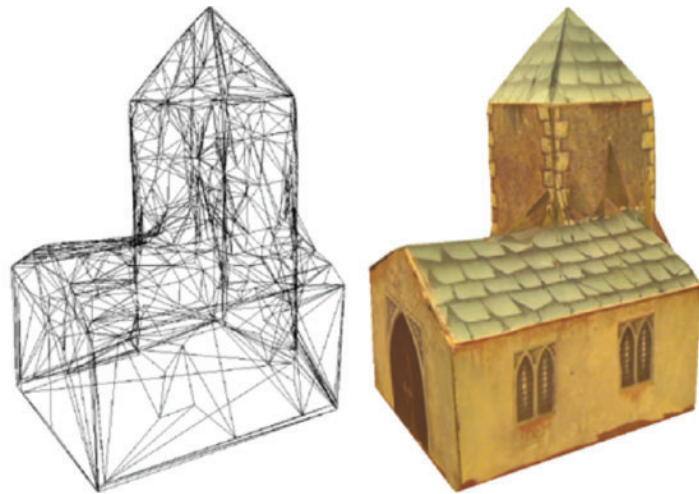


[Oleynikova, ICRA'17]

© Oleynikova et al. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Map representation	3D Topology?	Lightweight?	Filters Noise/Outliers?	Semantics?	Generality
Point Clouds	✗	✓/✗ No, if Dense	✗	✓/✗ No, if Sparse	✓
primitives & objects	✗	✓	✓	✓/✗ No, if Sparse	✗
Voxels	✓	✓/✗ No, if small voxel	✓	✓/✗ No, if large	✓ ⁹

Meshes



Map representation	3D Topology ?	Lightweight?	Filters Noise/Outliers?	Semantics?	Generality
Point Clouds	✗	✓/✗ No, if Dense	✗	✓/✗ No, if Sparse	✓
primitives & objects	✗	✓	✓	✓/✗ No, if Sparse	✗
Voxels	✓	✓/✗ No, if small voxel	✓	✓/✗ No, if large voxel	✓
3D Mesh	✓	✓	✗	✓	✓

Today

- Dense Reconstruction
 - 3D representations
 - (Some) Multi-view Stereo
 - Depth fusion
- Final thoughts

Figure 1 in R. A. Newcombe et al., "KinectFusion: Real-time dense surface mapping and tracking," 2011 10th IEEE International Symposium on Mixed and Augmented Reality, Basel, Switzerland, 2011, pp. 127-136, doi: 10.1109/ISMAR.2011.6092378. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

KinectFusion: Real-Time Dense Surface Mapping and Tracking*

Richard A. Newcombe Imperial College London	Shahram Izadi Microsoft Research	Otmar Hilliges Microsoft Research	David Molyneaux Microsoft Research Lancaster University	David Kim Microsoft Research Newcastle University
Andrew J. Davison Imperial College London	Pushmeet Kohli Microsoft Research	Jamie Shotton Microsoft Research	Steve Hodges Microsoft Research	Andrew Fitzgibbon Microsoft Research



2011

Figure 1: Example output from our system, generated in real-time with a handheld Kinect depth camera and no other sensing infrastructure. Normal maps (colour) and Phong-shaded renderings (greyscale) from our dense reconstruction system are shown. On the left for comparison is an example of the live, incomplete, and noisy data from the Kinect sensor (used as input to our system).

Multi-View Stereo: A 2015 Tutorial

Yasutaka Furukawa
Washington University in St. Louis
furukawa@wustl.edu

Carlos Hernández
Google Inc.
carloshernandez@google.com

ElasticFusion: Dense SLAM Without A Pose Graph

Thomas Whelan*, Stefan Leutenegger*, Renato F. Salas-Moreno[†], Ben Glocker[†] and Andrew J. Davison*

*Dyson Robotics Laboratory at Imperial College, Department of Computing, Imperial College London, UK

[†]Department of Computing, Imperial College London, UK

{t.whelan,s.leutenegger,r.salas-moreno10,b.glocker,a.davison}@imperial.ac.uk

2016

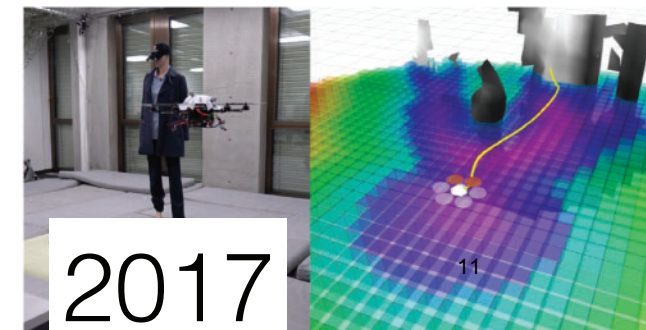
Figure 1 in H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart and J. Nieto, "Voxblox: Incremental 3D Euclidean Signed Distance Fields for on-board MAV planning," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, 2017, pp. 1366-1373, doi: 10.1109/IROS.2017.8202315. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Voxblox: Incremental 3D Euclidean Signed Distance Fields for On-Board MAV Planning

Helen Oleynikova, Zachary Taylor, Marius Fehr, Roland Siegwart, and Juan Nieto
Autonomous Systems Lab, ETH Zürich

Abstract—Micro Aerial Vehicles (MAVs) that operate in unstructured, unexplored environments require fast and flexible local planning, which can replan when new parts of the map are explored. Trajectory optimization methods fulfill these needs, but require obstacle distance information, which can be given by Euclidean Signed Distance Fields (ESDFs).

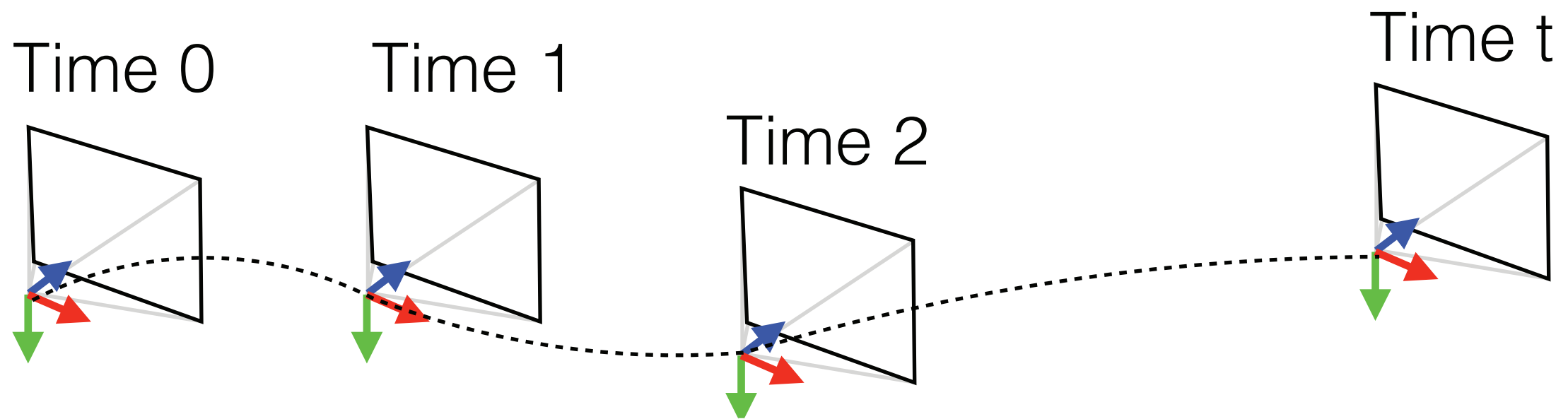
We propose a method to incrementally build ESDFs from Truncated Signed Distance Fields (TSDFs), a common implicit surface representation used in computer graphics and vision. TSDFs are fast to build and smooth out sensor noise over many observations, and are designed to produce surface meshes. Meshes allow human operators to get a better assessment of the robot's environment, and set high-level mission goals.



2017

Multi-view Stereo

From previous lectures: we know how to use SLAM to get a good estimate of the poses of the cameras



© N. Snavely. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>



Stereo



Multi-view stereo

Towards Internet-scale Multi-view Stereo

CVPR 2010

Yasutaka Furukawa¹ Brian Curless²

Steven M. Seitz^{1,2} Richard Szeliski³

Google Inc.¹

University of Washington²

Microsoft Research³

Multi-view Stereo

The Visual Turing Test for Scene Reconstruction Supplementary Video

Qi Shan⁺ Riley Adams⁺ Brian Curless⁺

Yasutaka Furukawa^{*} Steve Seitz^{+*}

⁺University of Washington ^{*}Google

3DV 2013

Multi-view Stereo

Patch-based methods:

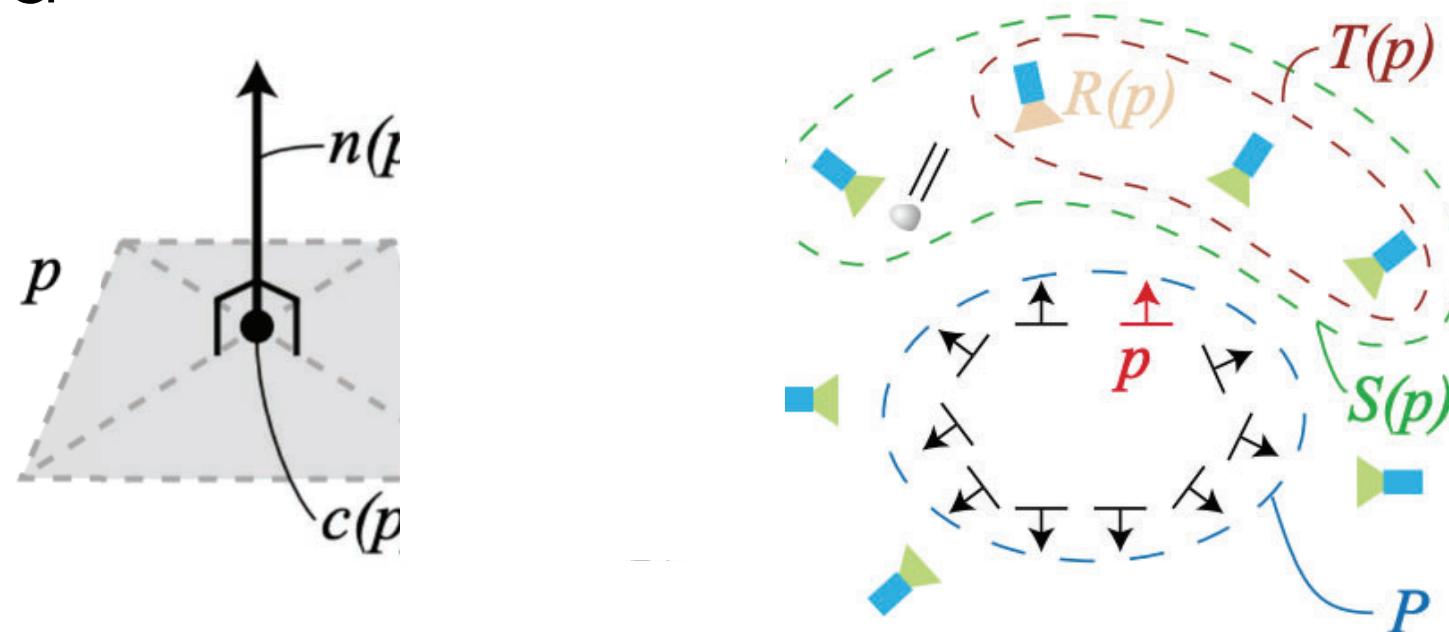


Figure 2. Definition of a patch (left) and of the images associated with it (right). See text for the details.

Estimate normal and center of patch to maximize **photometric consistency**:

$$C_{ij}(p) = \rho(I_i(\Omega(\pi_i(p))), I_j(\Omega(\pi_j(p))))$$

Projection To camera
3D point

Matching Score
Image Intensity
Rectangular Patch

Example of matching score:

$$1 - \sum_{x,y} |W_1(x,y) - W_2(x,y)|^2$$

Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multi-View Stereopsis," 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 2007, pp. 1-8, doi: 10.1109/CVPR.2007.383246. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Multi-view Stereo

Enforcing regularity: Markov Random Fields

Find depth k_p of point “p” such that point is photo-consistent and depth changes smoothly..

$$E(\{k_p\}) = \sum_p \Phi(k_p) + \sum_{(p,q) \in \mathcal{N}} \Psi(k_p, k_q)$$

Unary potentials

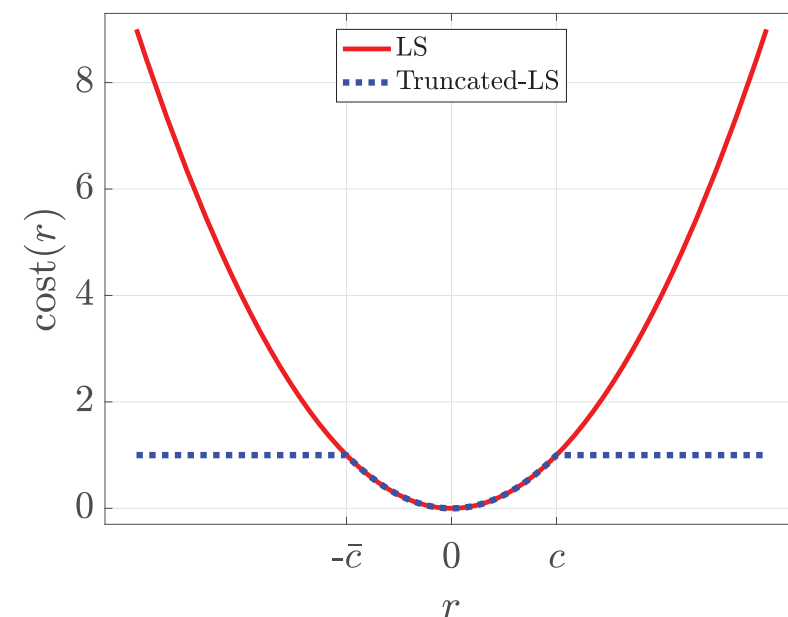
(similar to previous slides)

$$\Phi(k_p = d) = \min(\tau_u, 1 - \mathcal{C}(p, d))$$

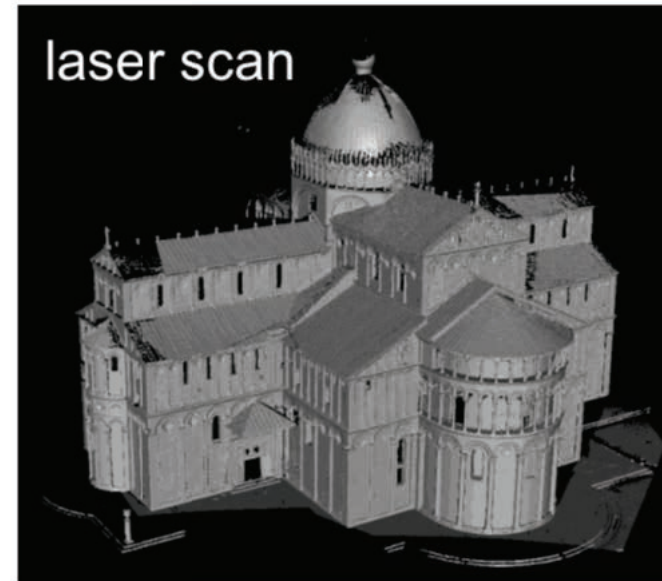
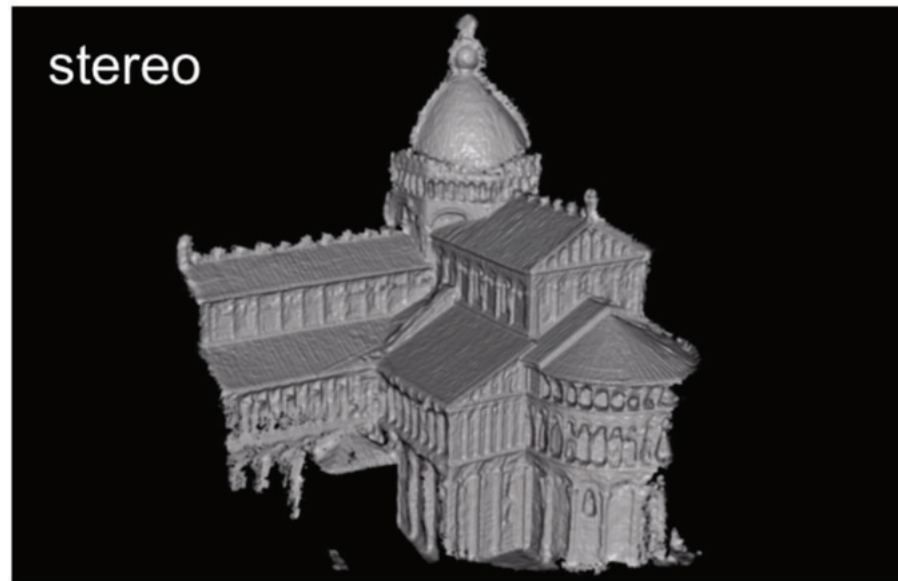
Pairwise potentials

$$\Psi(k_p = d_1, k_q = d_2) = \min(\tau_p, |d_1 - d_2|)$$

Depth is typically discretized before solving..



How Accurate is Multi-view Stereo?

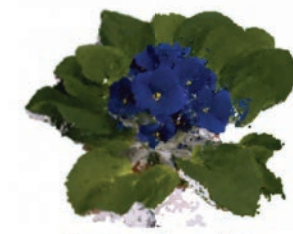


Space Carving Results: African Violet



Input Image (1 of 45)

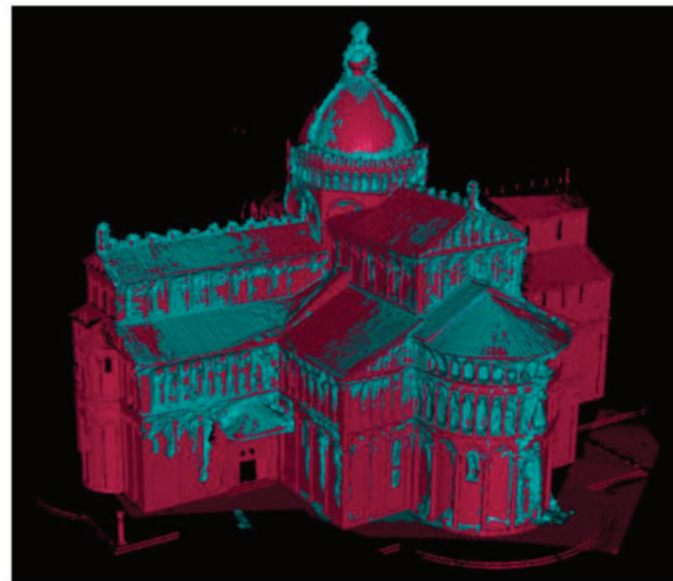
Reconstruction



Reconstruction

Reconstruction

Source: S. Seitz



Comparison: 90% of points within 0.128 m of laser scan (building height 51m)

Space Carving Results: Hand



Input Image (1 of 100)



Views of Reconstruction

M. Goesele, N. Snavely, B. Curless, H. Hoppe, S. Seitz, [Multi-View Stereo for Community Photo Collections](#), ICCV 2007

Figure 7 in M. Goesele, N. Snavely, B. Curless, H. Hoppe and S. M. Seitz, "Multi-View Stereo for Community Photo Collections," 2007 IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007, pp. 1-8, doi: 10.1109/ICCV.2007.4408933. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Many methods: volumetric stereo, space carving, Shape from silhouettes, carved visual hull

Today

- Dense Reconstruction
 - 3D representations
 - (Some) Multi-view Stereo
 - Depth fusion
- Final thoughts

Figure 1 in H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart and J. Nieto, "Voxblox: Incremental 3D Euclidean Signed Distance Fields for on-board MAV planning," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, 2017, pp. 1366-1373, doi: 10.1109/IROS.2017.8202315. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

KinectFusion: Real-Time Dense Surface Mapping and Tracking*

Richard A. Newcombe Imperial College London	Shahram Izadi Microsoft Research	Otmar Hilliges Microsoft Research	David Molyneaux Microsoft Research Lancaster University	David Kim Microsoft Research Newcastle University
Andrew J. Davison Imperial College London	Pushmeet Kohli Microsoft Research	Jamie Shotton Microsoft Research	Steve Hodges Microsoft Research	Andrew Fitzgibbon Microsoft Research



2011

Multi-View Stereo: A 2015 Tutorial

Yasutaka Furukawa
Washington University in St. Louis
furukawa@wustl.edu

Carlos Hernández
Google Inc.
carloshernandez@google.com

ElasticFusion: Dense SLAM Without A Pose Graph

Thomas Whelan*, Stefan Leutenegger*, Renato F. Salas-Moreno[†], Ben Glocker[†] and Andrew J. Davison*

*Dyson Robotics Laboratory at Imperial College, Department of Computing, Imperial College London, UK

[†]Department of Computing, Imperial College London, UK

{t.whelan,s.leutenegger,r.salas-moreno10,b.glocker,a.davison}@imperial.ac.uk

2016

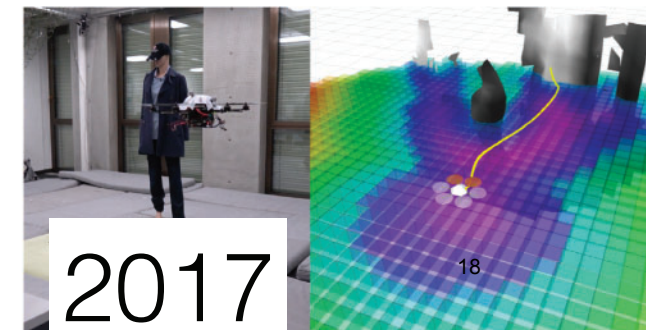
Figure 1 in H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart and J. Nieto, "Voxblox: Incremental 3D Euclidean Signed Distance Fields for on-board MAV planning," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, 2017, pp. 1366-1373, doi: 10.1109/IROS.2017.8202315. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Voxblox: Incremental 3D Euclidean Signed Distance Fields for On-Board MAV Planning

Helen Oleynikova, Zachary Taylor, Marius Fehr, Roland Siegwart, and Juan Nieto
Autonomous Systems Lab, ETH Zürich

Abstract—Micro Aerial Vehicles (MAVs) that operate in unstructured, unexplored environments require fast and flexible local planning, which can replan when new parts of the map are explored. Trajectory optimization methods fulfill these needs, but require obstacle distance information, which can be given by Euclidean Signed Distance Fields (ESDFs).

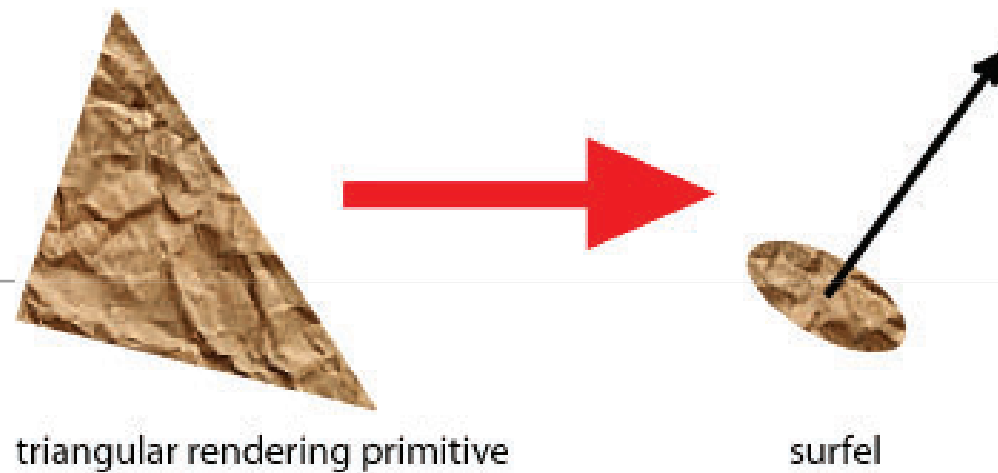
We propose a method to incrementally build ESDFs from Truncated Signed Distance Fields (TSDFs), a common implicit surface representation used in computer graphics and vision. TSDFs are fast to build and smooth out sensor noise over many observations, and are designed to produce surface meshes. Meshes allow human operators to get a better assessment of the robot's environment, and set high-level mission goals.



2017

Figure 1: Example output from our system, generated in real-time with a handheld Kinect depth camera and no other sensing infrastructure. Normal maps (colour) and Phong-shaded renderings (greyscale) from our dense reconstruction system are shown. On the left for comparison is an example of the live, incomplete, and noisy data from the Kinect sensor (used as input to our system).

Surfels



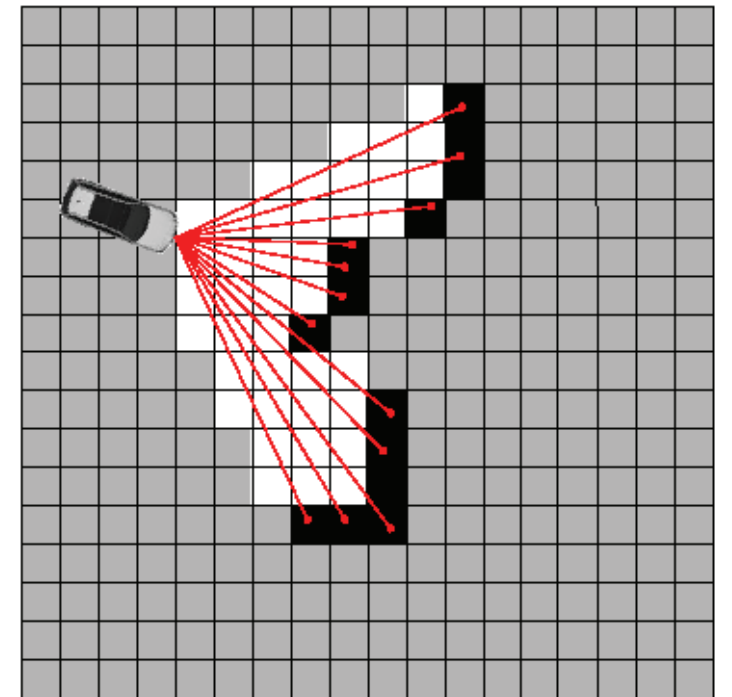
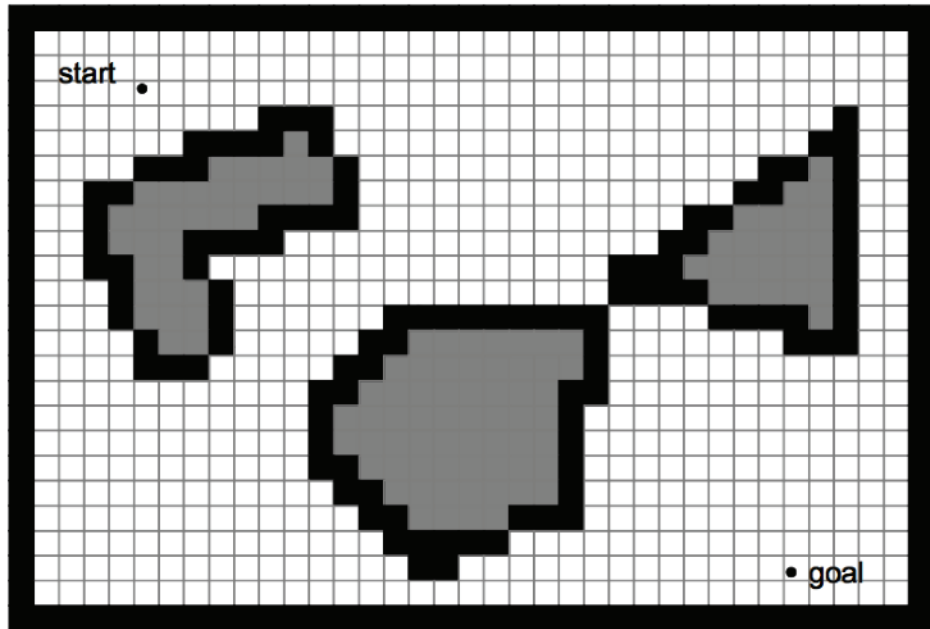
ElasticFusion: Dense SLAM Without A Pose Graph

Thomas Whelan, Stefan Leutenegger, Renato Salas-Moreno, Ben Glocker, Andrew Davison

Imperial College London

note: still based on RGB-D (contrarily to multi-view stereo)

A Gentle Start: 2D Occupancy Grid Maps



- discretize the environment into cells
- Each cell holds real number $[0, 1]$, representing the probability of the cell being occupied

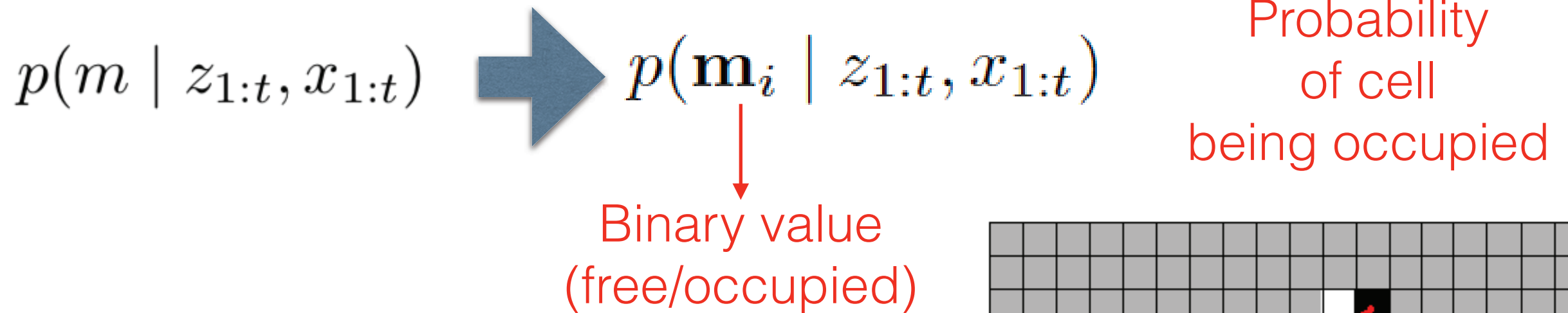
Map posterior

$$p(m \mid z_{1:t}, x_{1:t})$$

Unknown
Map

Known sensor
Depth and robot poses

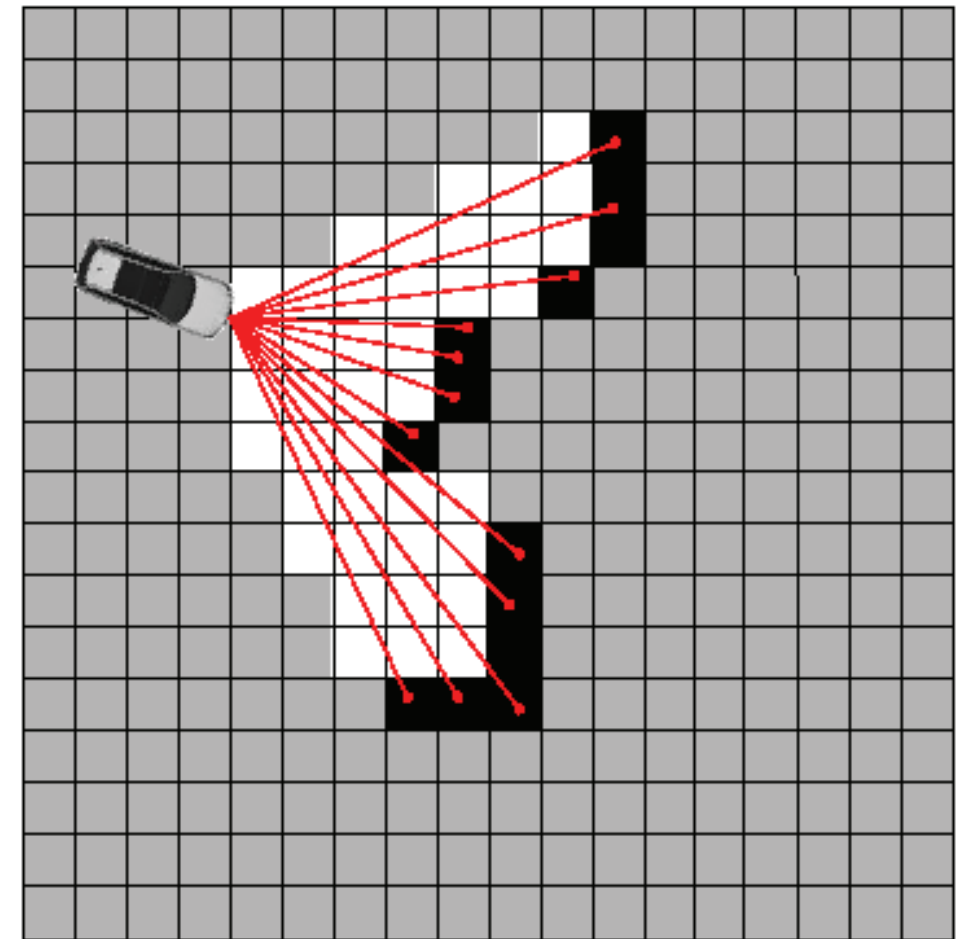
A Gentle Start: 2D Occupancy Grid Maps



Bayes rule (omitting “x” for simplicity):

$$p(m_i \mid z_{1:t+1}) = \frac{p(z_{t+1} \mid m_i) p(m_i \mid z_{1:t})}{p(m_i)}$$

Uninformative Prior

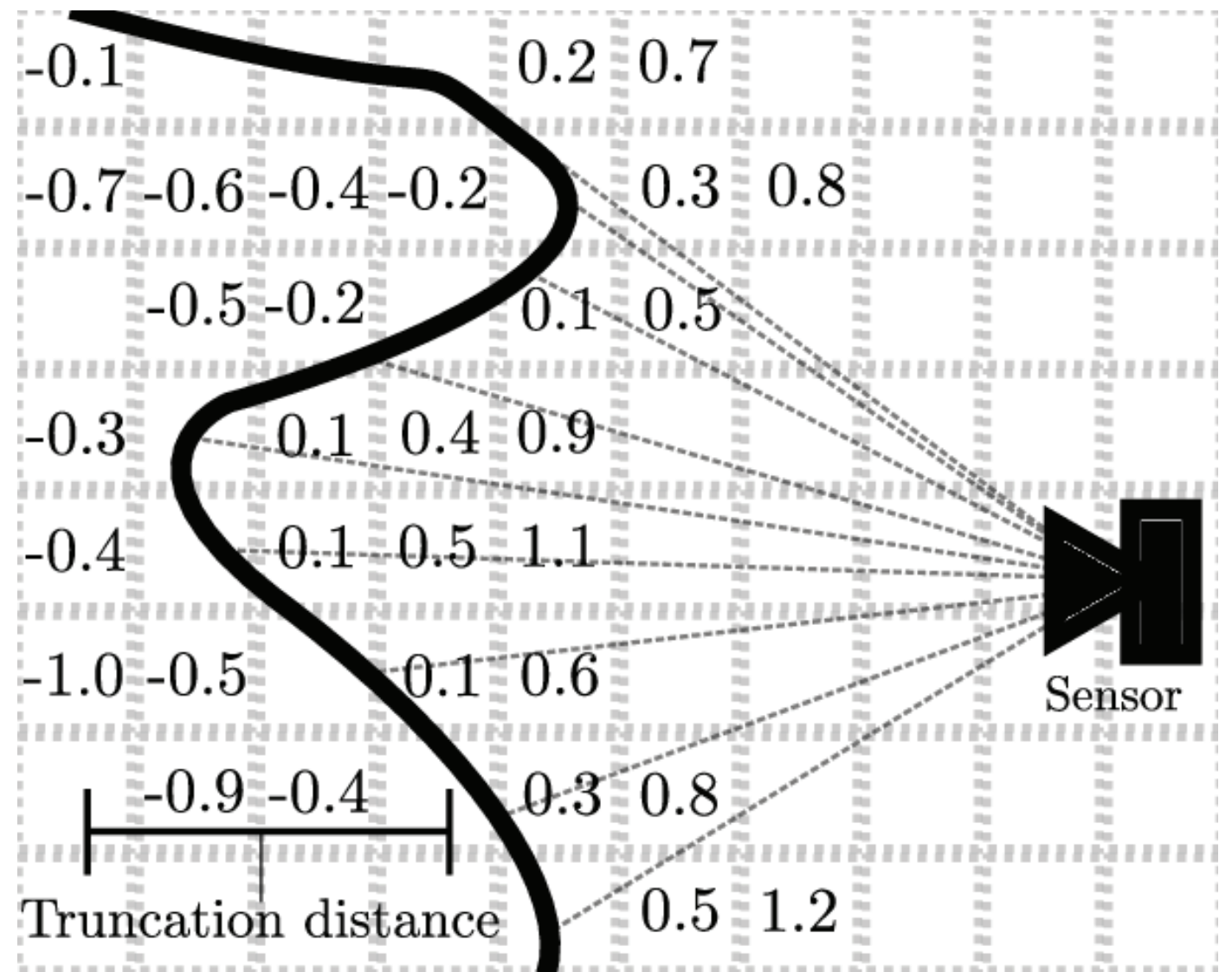


Log-odd representation is typically used to avoid numerical instabilities

$$\frac{p(\mathbf{m}_i \mid z_{1:t}, x_{1:t})}{1 - p(\mathbf{m}_i \mid z_{1:t}, x_{1:t})} \rightarrow l_{t,i} = \log \frac{p(\mathbf{m}_i \mid z_{1:t}, x_{1:t})}{1 - p(\mathbf{m}_i \mid z_{1:t}, x_{1:t})}$$

Truncated Signed Distance Function (SDF)

- Store distance to nearest obstacle (with sign)
 - Only update around obstacle itself
- (implicit surface model)



Update rule:

$$d(\mathbf{x}, \mathbf{p}, \mathbf{s}) = \|\mathbf{p} - \mathbf{x}\| \text{sign}((\mathbf{p} - \mathbf{x}) \bullet (\mathbf{p} - \mathbf{s})) \quad (1)$$

$$w_{\text{const}}(\mathbf{x}, \mathbf{p}) = 1 \quad (2)$$

$$D_{i+1}(\mathbf{x}, \mathbf{p}) = \frac{W_i(\mathbf{x})D_i(\mathbf{x}) + w(\mathbf{x}, \mathbf{p})d(\mathbf{x}, \mathbf{p})}{W_i(\mathbf{x}) + w(\mathbf{x}, \mathbf{p})} \quad (3)$$

$$W_{i+1}(\mathbf{x}, \mathbf{p}) = \min(W_i(\mathbf{x}) + w(\mathbf{x}, \mathbf{p}), W_{\text{max}}) \quad (4)$$

© ACM. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

[Curless and Levoy, "A Volumetric Method for Building Complex Models from Range Images", 2007]

Kinect Fusion (2011)

SIGGRAPH Talks 2011

KinectFusion:

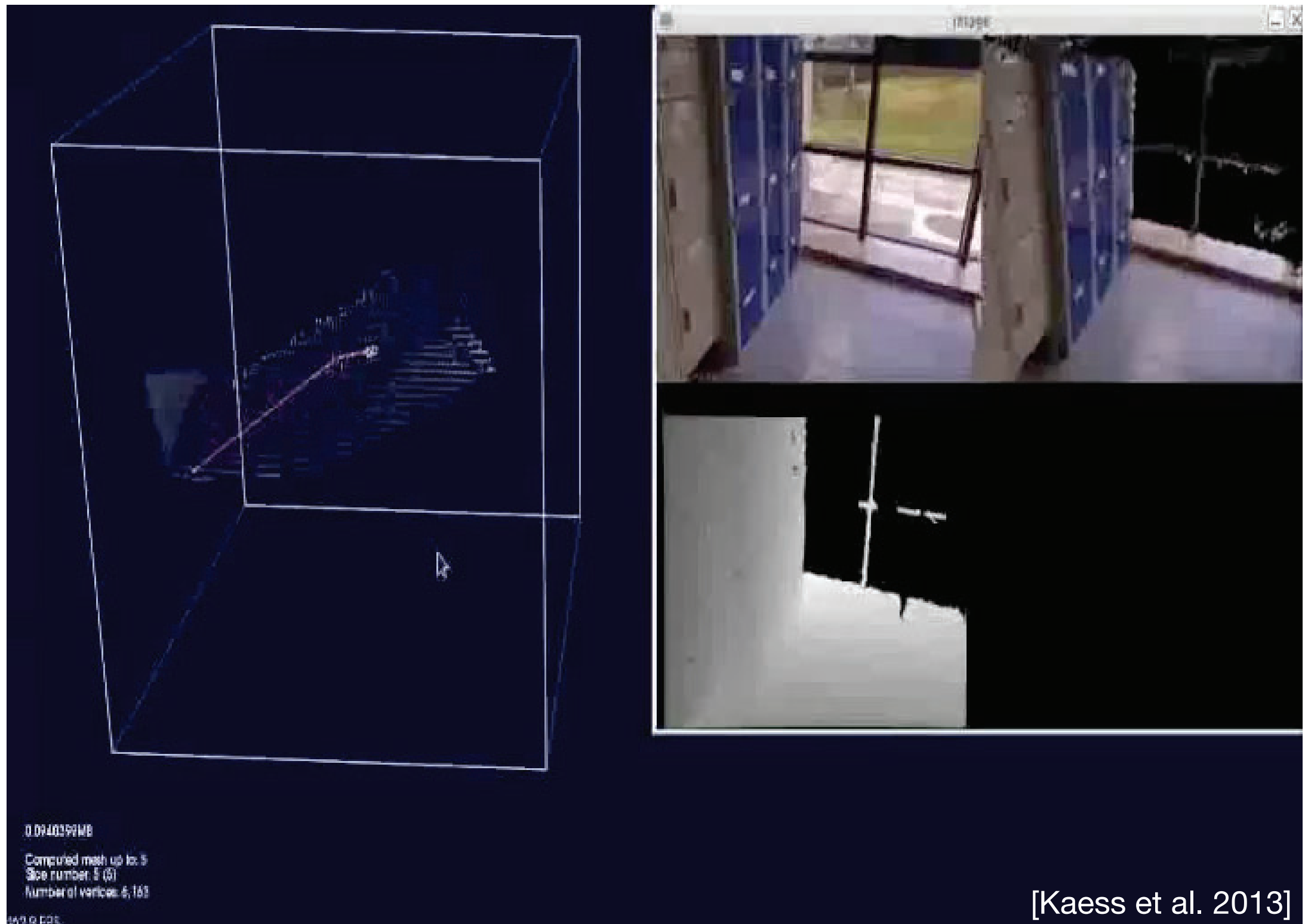
**Real-Time Dynamic 3D Surface
Reconstruction and Interaction**

**Shahram Izadi 1, Richard Newcombe 2, David Kim 1,3, Otmar Hilliges 1,
David Molyneaux 1,4, Pushmeet Kohli 1, Jamie Shotton 1,
Steve Hodges 1, Dustin Freeman 5, Andrew Davison 2, Andrew Fitzgibbon 1**

**1 Microsoft Research Cambridge 2 Imperial College London
3 Newcastle University 4 Lancaster University
5 University of Toronto**

GPU, memory ...

Kintinuous (2013)



GPU, bounded memory ...

VoxBlox (2017)

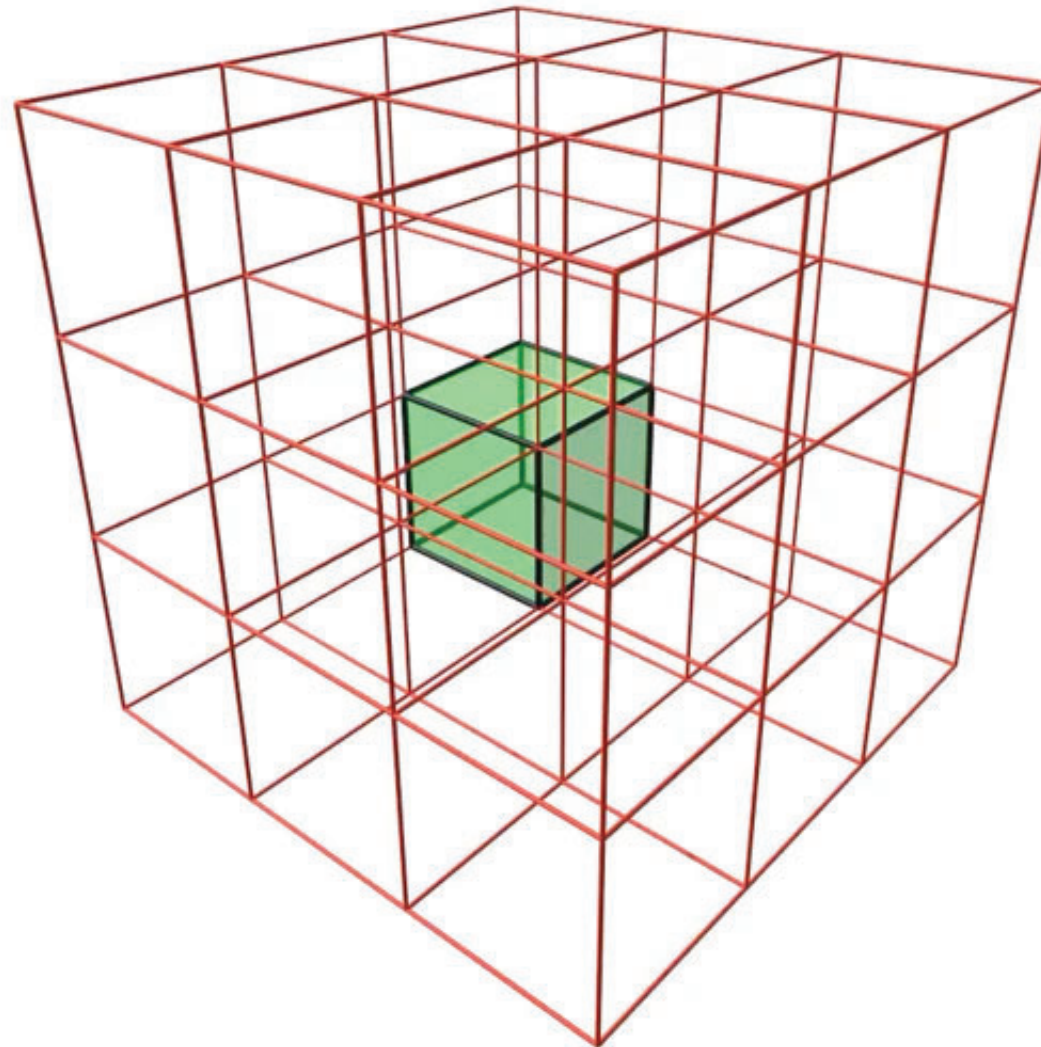


© ETH-Zurich. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

CPU, memory

From Voxels to Meshes

Marching cubes



© Algorithms Visualized. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

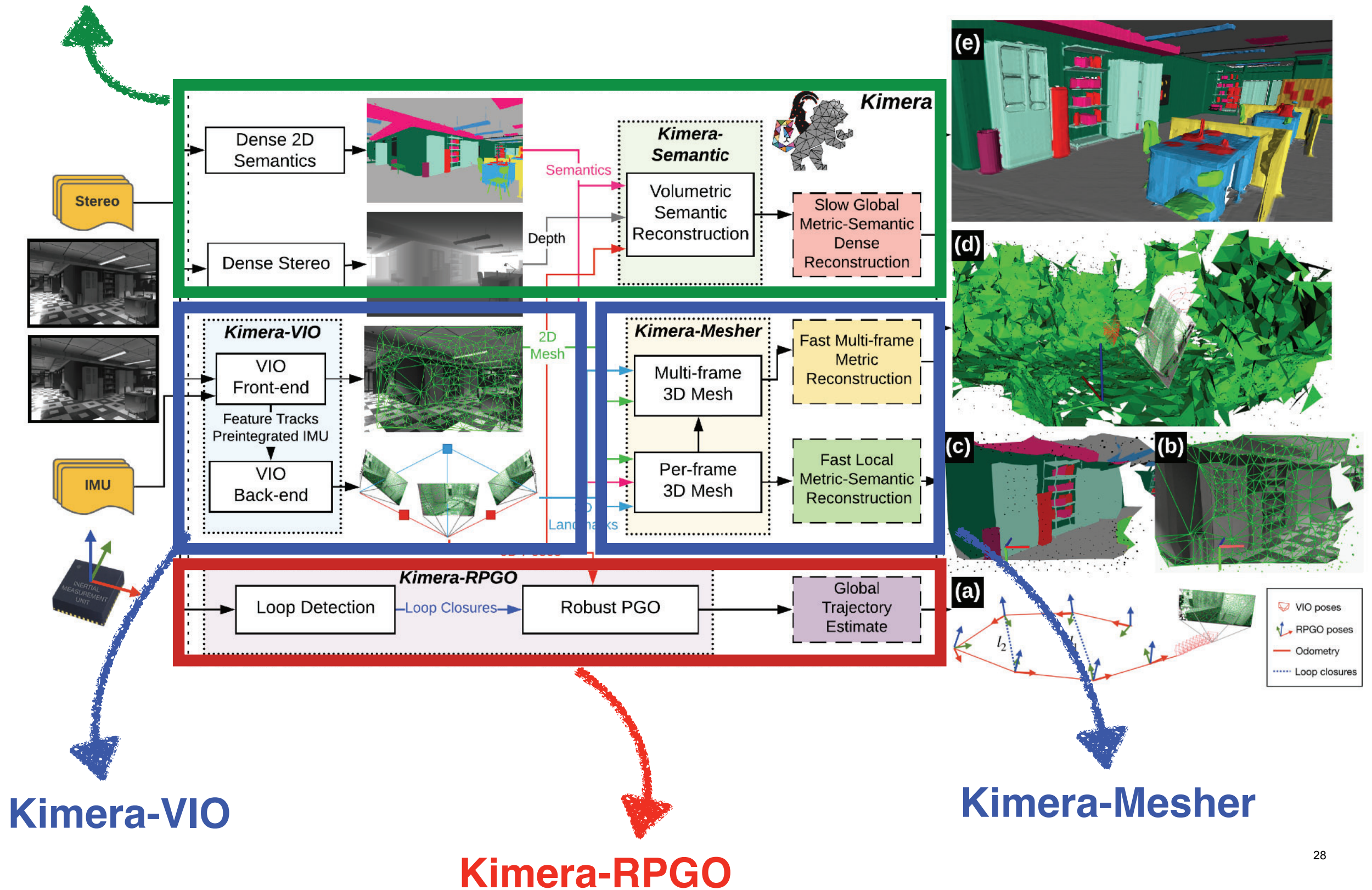
https://www.youtube.com/watch?v=B_xk71YopsA

Kimera (2020)

Kimera-VIO tracks sparse 3D landmarks for fast and accurate state estimation

Metric-semantic 3D Reconstruction

Kimera-Semantics



Kimera-VIO

Kimera-RPGO

Kimera-Mesher

Today

- Dense Reconstruction
 - 3D representations
 - (Some) Multi-view Stereo
 - Depth fusion
- Final thoughts

Multi-View Stereo: A 2015 Tutorial

Yasutaka Furukawa
Washington University in St. Louis
furukawa@wustl.edu

Carlos Hernández
Google Inc.
carloshernandez@google.com

ElasticFusion: Dense SLAM Without A Pose Graph

Thomas Whelan*, Stefan Leutenegger*, Renato F. Salas-Moreno[†], Ben Glocker[†] and Andrew J. Davison*

*Dyson Robotics Laboratory at Imperial College, Department of Computing, Imperial College London, UK

[†]Department of Computing, Imperial College London, UK

{t.whelan,s.leutenegger,r.salas-moreno10,b.glocker,a.davison}@imperial.ac.uk

2016

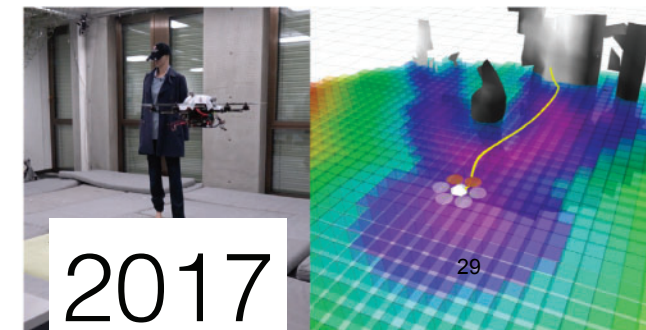
Figure 1 in H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart and J. Nieto, "Voxblox: Incremental 3D Euclidean Signed Distance Fields for on-board MAV planning," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, 2017, pp. 1366-1373, doi: 10.1109/IROS.2017.8202315. © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

Voxblox: Incremental 3D Euclidean Signed Distance Fields for On-Board MAV Planning

Helen Oleynikova, Zachary Taylor, Marius Fehr, Roland Siegwart, and Juan Nieto
Autonomous Systems Lab, ETH Zürich

Abstract—Micro Aerial Vehicles (MAVs) that operate in unstructured, unexplored environments require fast and flexible local planning, which can replan when new parts of the map are explored. Trajectory optimization methods fulfill these needs, but require obstacle distance information, which can be given by Euclidean Signed Distance Fields (ESDFs).

We propose a method to incrementally build ESDFs from Truncated Signed Distance Fields (TSDFs), a common implicit surface representation used in computer graphics and vision. TSDFs are fast to build and smooth out sensor noise over many observations, and are designed to produce surface meshes. Meshes allow human operators to get a better assessment of the robot's environment, and set high-level mission goals.



2017

29

KinectFusion: Real-Time Dense Surface Mapping and Tracking*

Richard A. Newcombe
Imperial College London

Shahram Izadi
Microsoft Research

Otmar Hilliges
Microsoft Research

David Molyneaux
Microsoft Research
Lancaster University

David Kim
Microsoft Research
Newcastle University

Andrew J. Davison
Imperial College London

Pushmeet Kohli
Microsoft Research

Jamie Shotton
Microsoft Research

Steve Hodges
Microsoft Research

Andrew Fitzgibbon
Microsoft Research



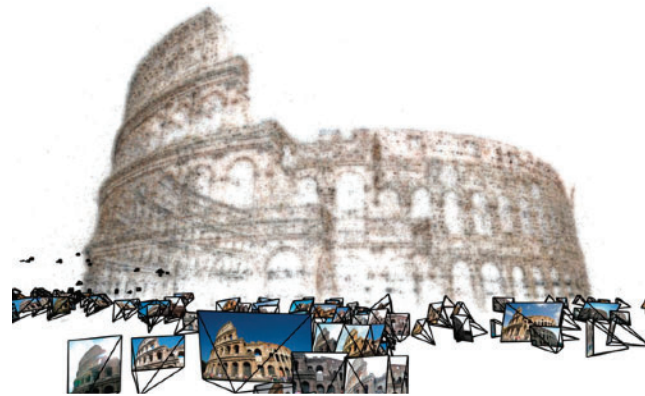
2011

Figure 1: Example output from our system, generated in real-time with a handheld Kinect depth camera and no other sensing infrastructure. Normal maps (colour) and Phong-shaded renderings (greyscale) from our dense reconstruction system are shown. On the left for comparison is an example of the live, incomplete, and noisy data from the Kinect sensor (used as input to our system).

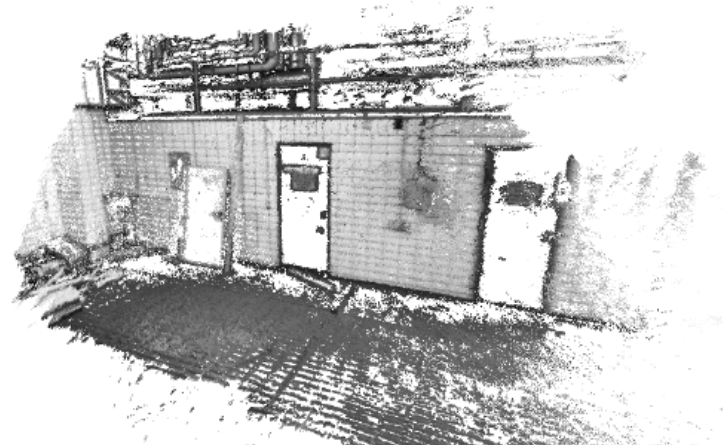
Robot Perception or Computer Vision?

Computer vision

.. “a day on a cluster with 500 compute cores”



Robotics



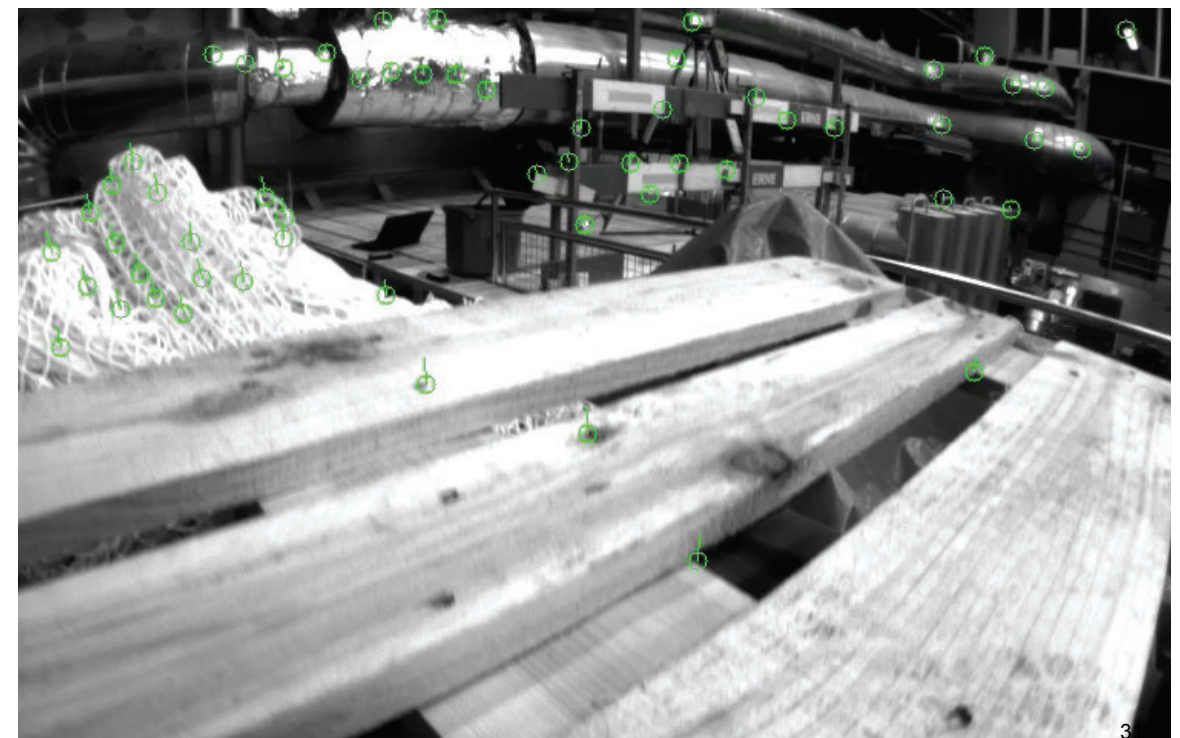
50-100ms latency,
embedded,
incremental

No longer a dichotomy for many vision applications!

Robot Perception or Computer Vision?



Unordered
Vs
Sequential



MIT OpenCourseWare
<https://ocw.mit.edu/>

16.485 Visual Navigation for Autonomous Vehicles (VNAV)
Fall 2020

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.