



Multiple View Reconstruction: A Short Introduction

Prof. Dr. Daniel Cremers

Chair of Computer Vision & Artificial Intelligence
School of Computation, Information & Technology
Technical University of Munich

Tutorials



Regine Hartwig



Mohammed Brahim



Zhenzhang Ye

AI Panel, Paris, April 5, 2024

PR[AI]RIE

Paris Artificial Intelligence Research Institute
PSL

France  Digitale

Future of AI



**Francis
Bach**
Inria



**Daniel
Cremers**
TU Munich



**Michael
Jordan**
Inria



**Yann
LeCun**
Meta AI



**Cordelia
Schmid**
Inria

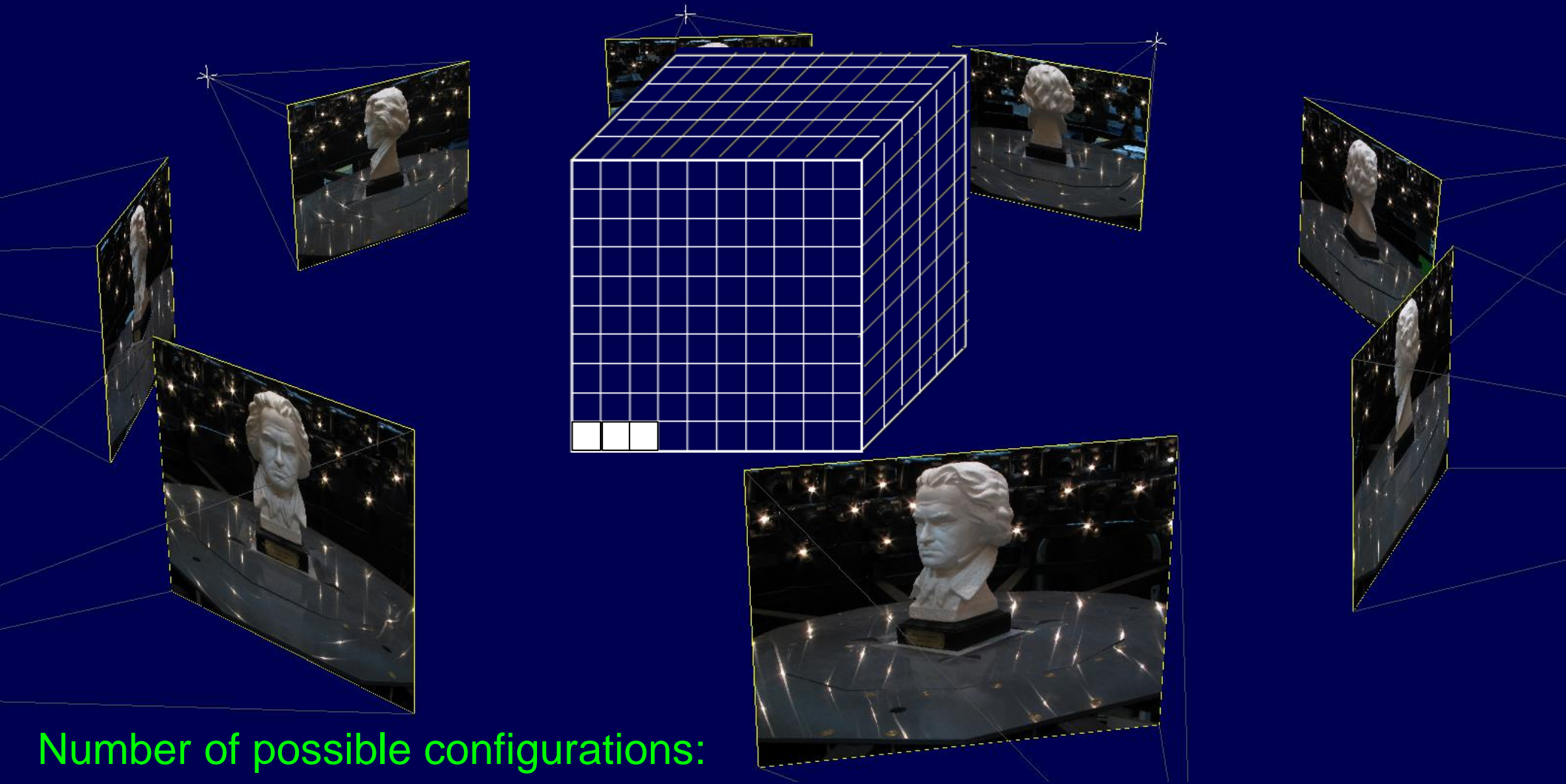


3D Reconstruction from Images



infinite-dimensional optimization

3D Reconstruction from Images



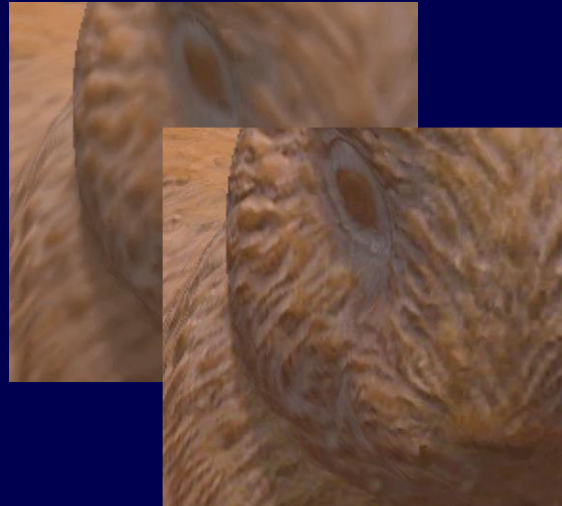
Number of possible configurations:

$$2^N = 2^{512 \times 512 \times 512} \approx 10^{40} \text{ Mio}$$

Overview



Multiview reconstruction



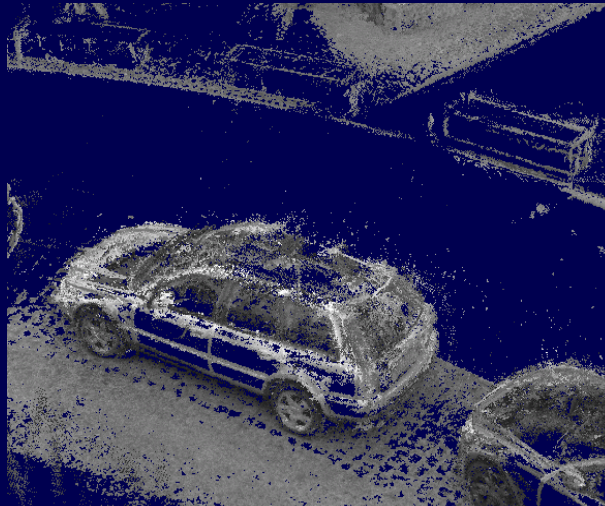
Super-res. textures



Free-viewpoint TV



Realtime dense scanning

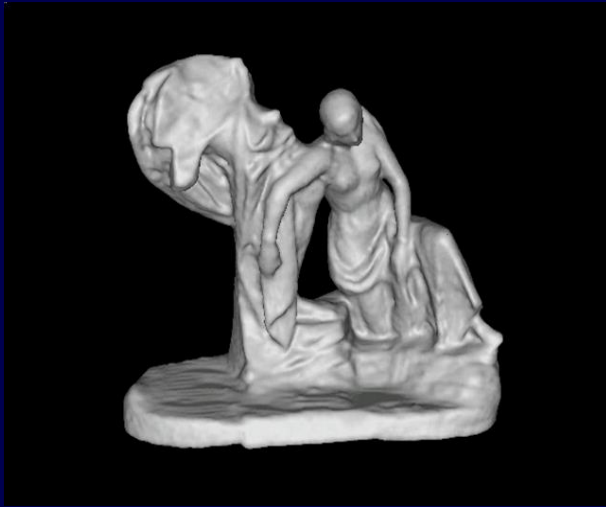


Visual SLAM



Deep Networks for SLAM

Overview



Multiview reconstruction



Super-res. textures



Free-viewpoint TV



Realtime dense scanning



Visual SLAM

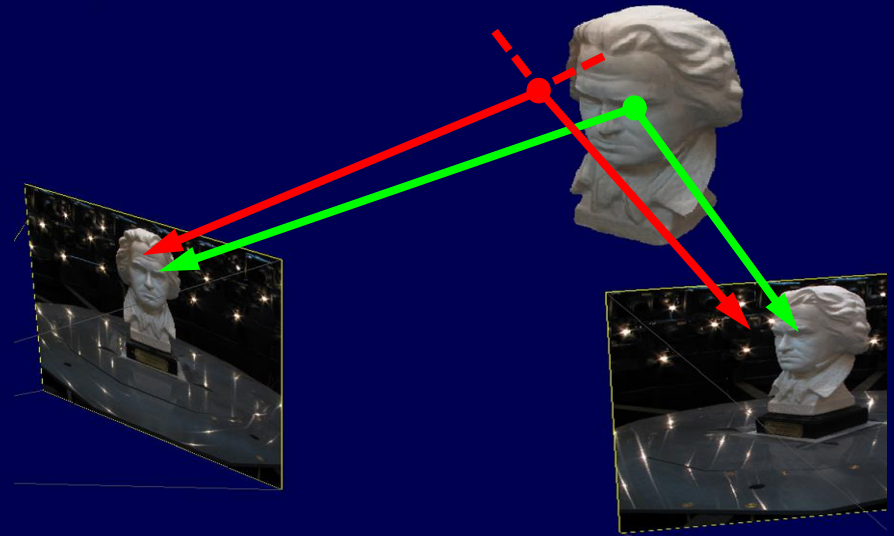


Deep Networks for SLAM

Solutions via Energy Minimization

Photoconsistency function:

$$\rho : \mathbb{R}^3 \rightarrow [0, 1]$$

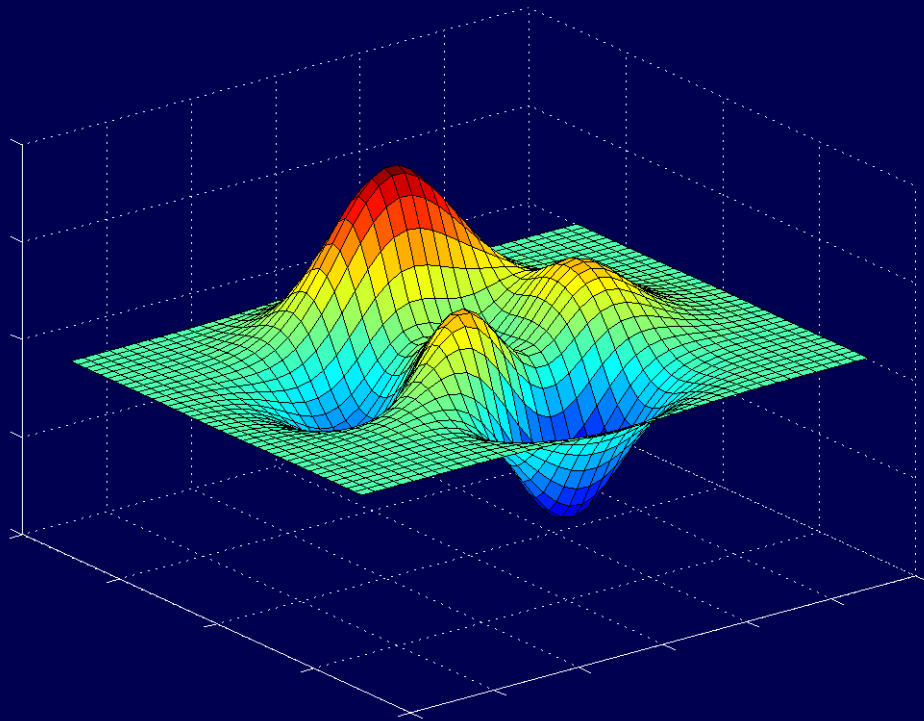


Determine a surface S of optimal photoconsistency by minimizing

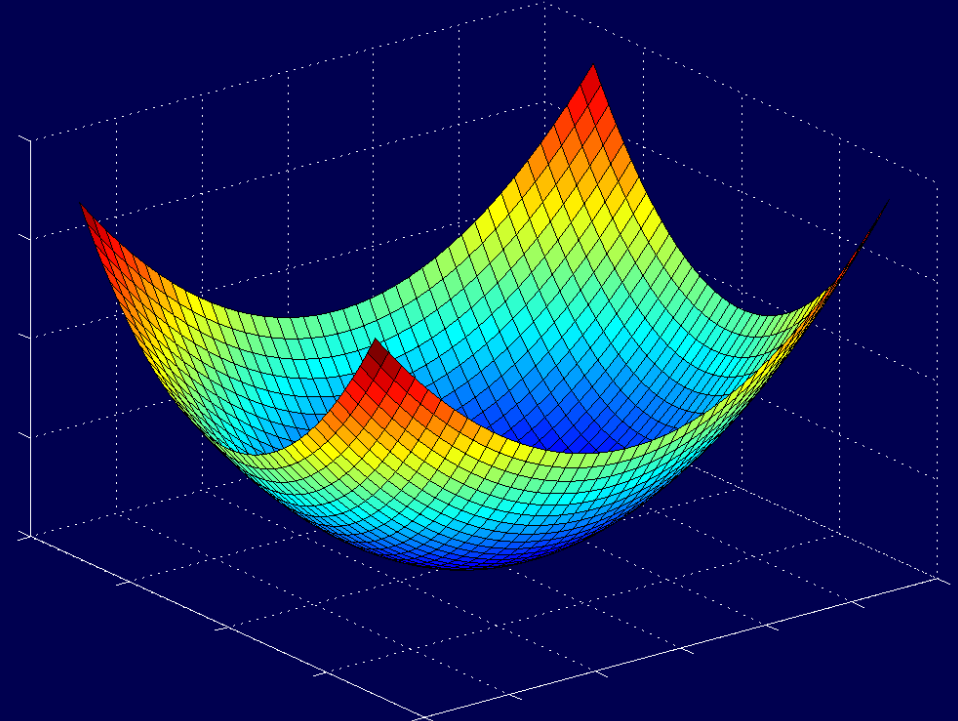
$$E(S) = \int_S \rho \, dA$$

Kolev, Klodt, Brox, Cremers, Int. J. of Computer Vision '09:

Theorem: Globally optimal surfaces can be computed via convex relaxation.

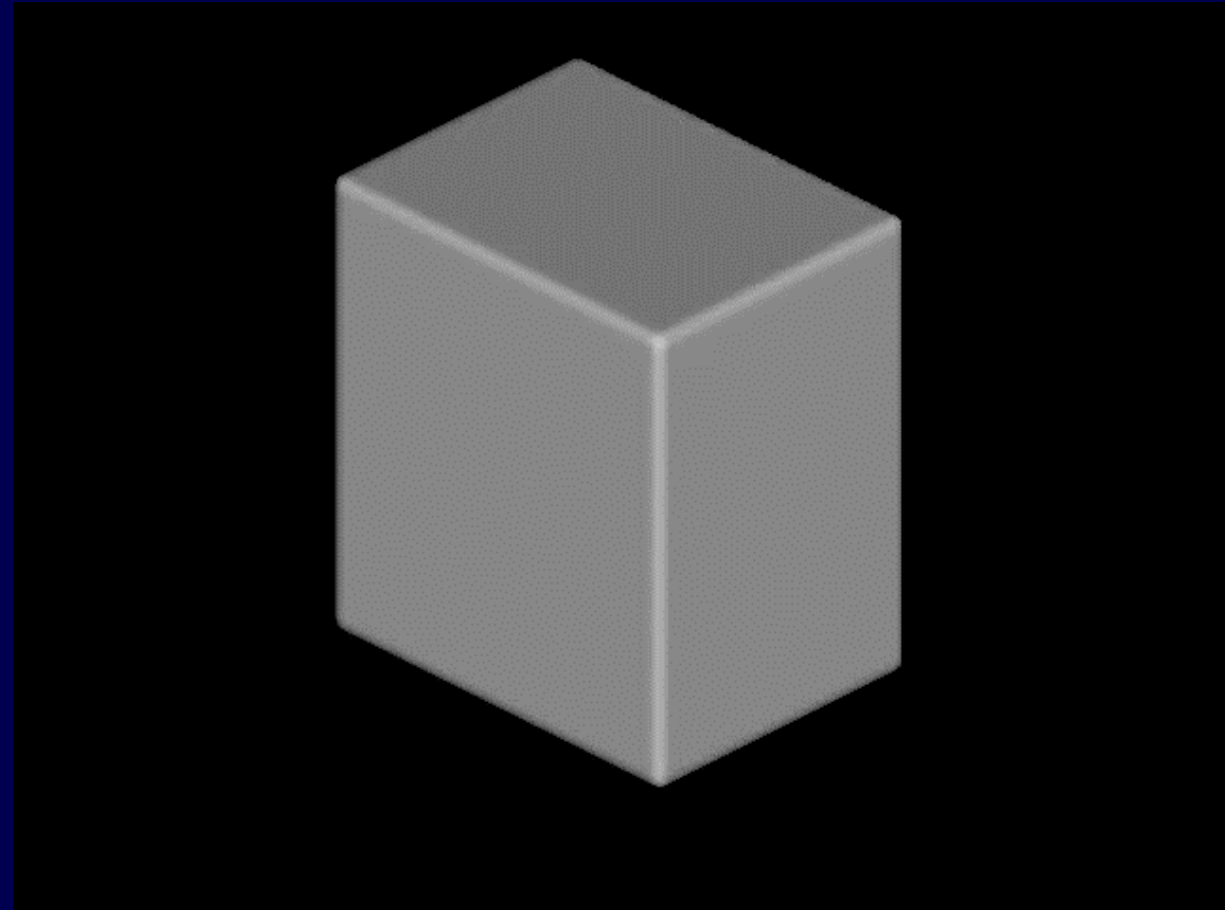
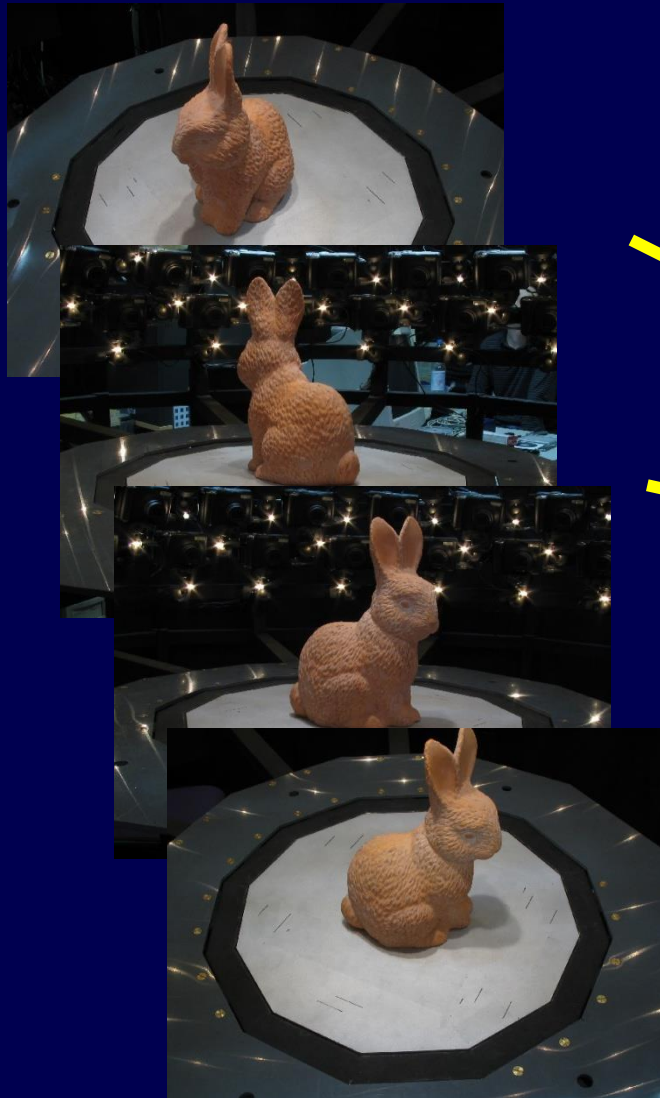


Non-convex energy



Convex energy

Evolution to Global Optimum

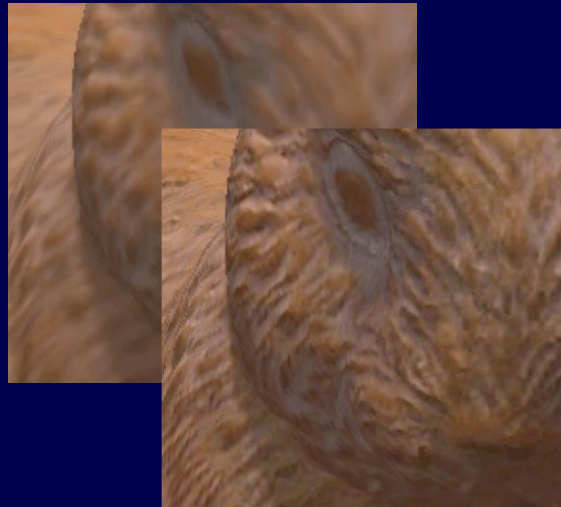


Kolev, Klodt, Brox, Cremers, IJCV 2009

Overview



Multiview reconstruction



Super-res. textures



Free-viewpoint TV



Realtime dense scanning



Visual SLAM



Deep Networks for SLAM

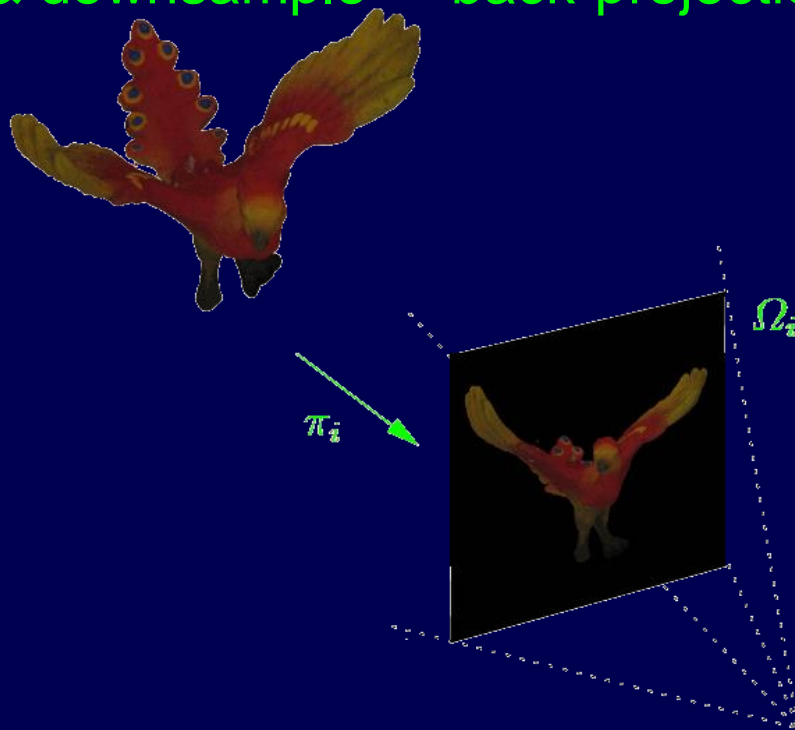
Super-Resolution Texture Map

Given all images $\mathcal{I}_i : \Omega_i \rightarrow \mathbb{R}^3$, determine the surface color $T : S \rightarrow \mathbb{R}^3$

$$\min_T \sum_{i=1}^n \int_{\Omega_i} \left(b * (T \circ \pi_i^{-1}) - \mathcal{I}_i \right)^2 dx + \lambda \int_S \|\nabla_S T\| ds$$

blur & downsample

back-projection



Goldlücke, Cremers, ICCV '09, DAGM '09*, IJCV '13

* Best Paper Award

Super-Resolution Texture Map



Goldlücke, Cremers, ICCV '09, DAGM '09, IJCV '13*

** Best Paper
Award*

Super-Resolution Texture Map



Closeup of input image



Super-resolution texture

Goldlücke, Cremers, ICCV '09, DAGM '09, IJCV '13* * Best Paper Award

Overview



Multiview reconstruction



Super-res. textures



Free-viewpoint TV



Realtime dense scanning

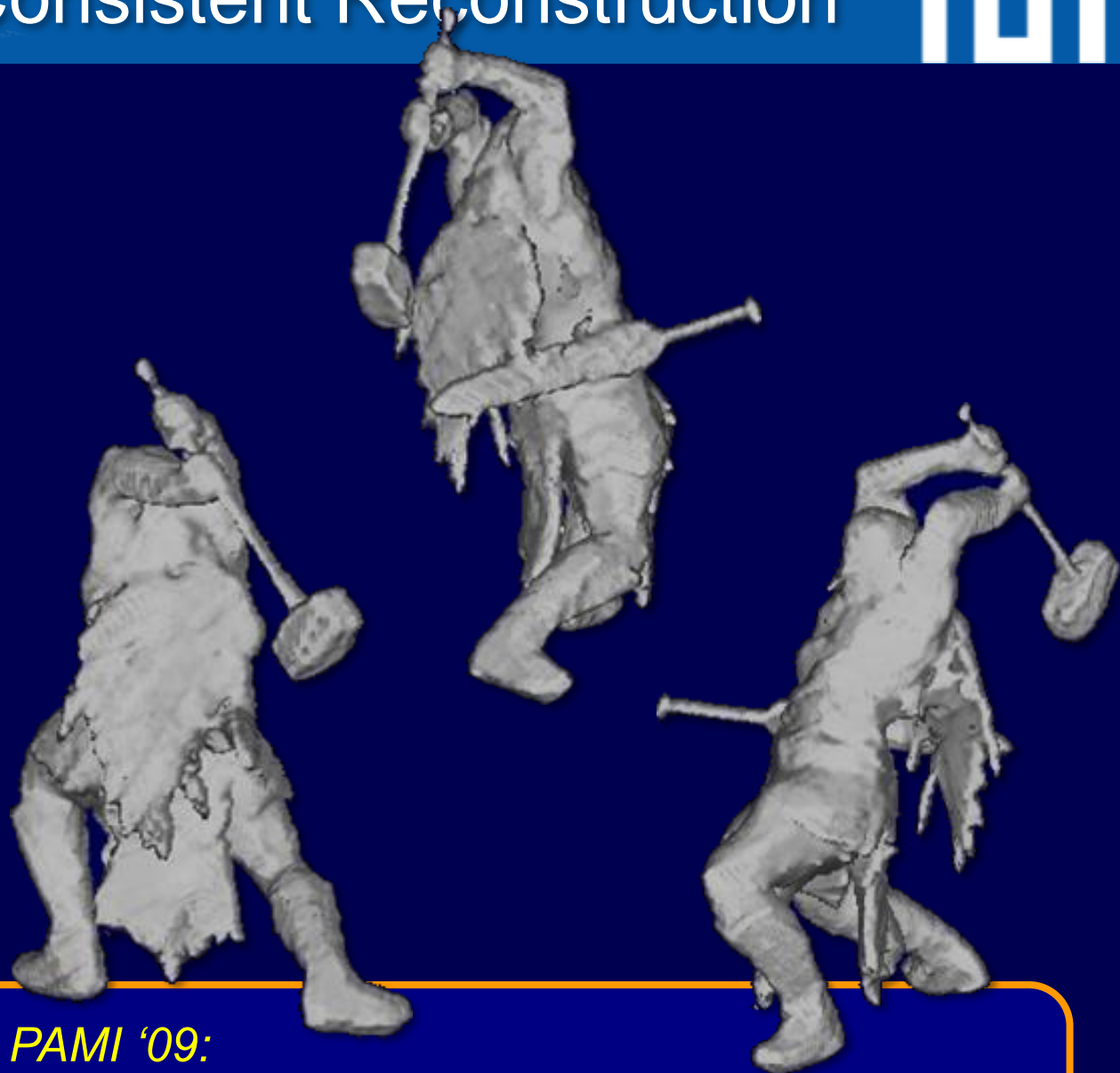


Visual SLAM



Deep Networks for SLAM

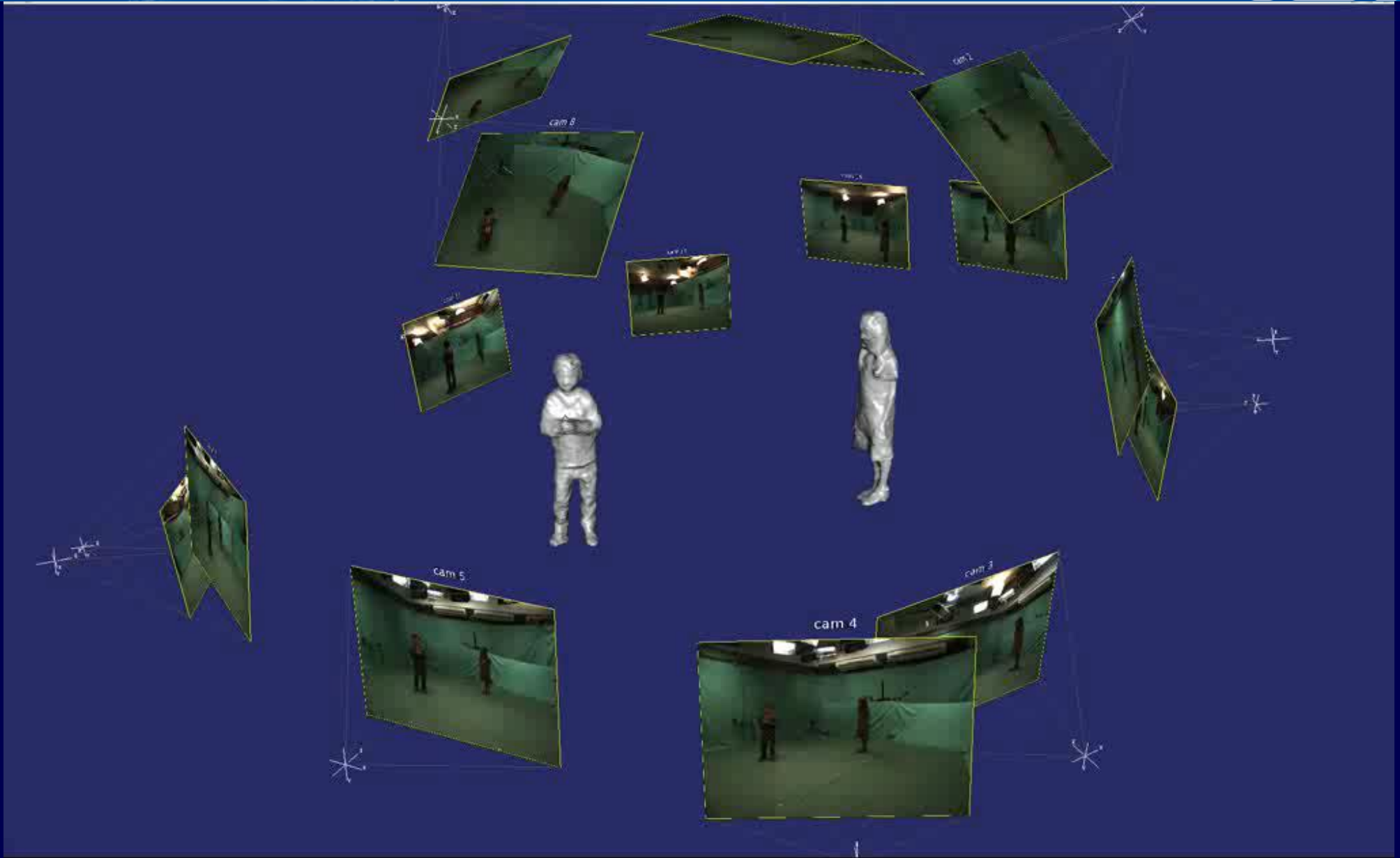
Silhouette-Consistent Reconstruction



Kolev, Cremers, ECCV '08, PAMI '09:

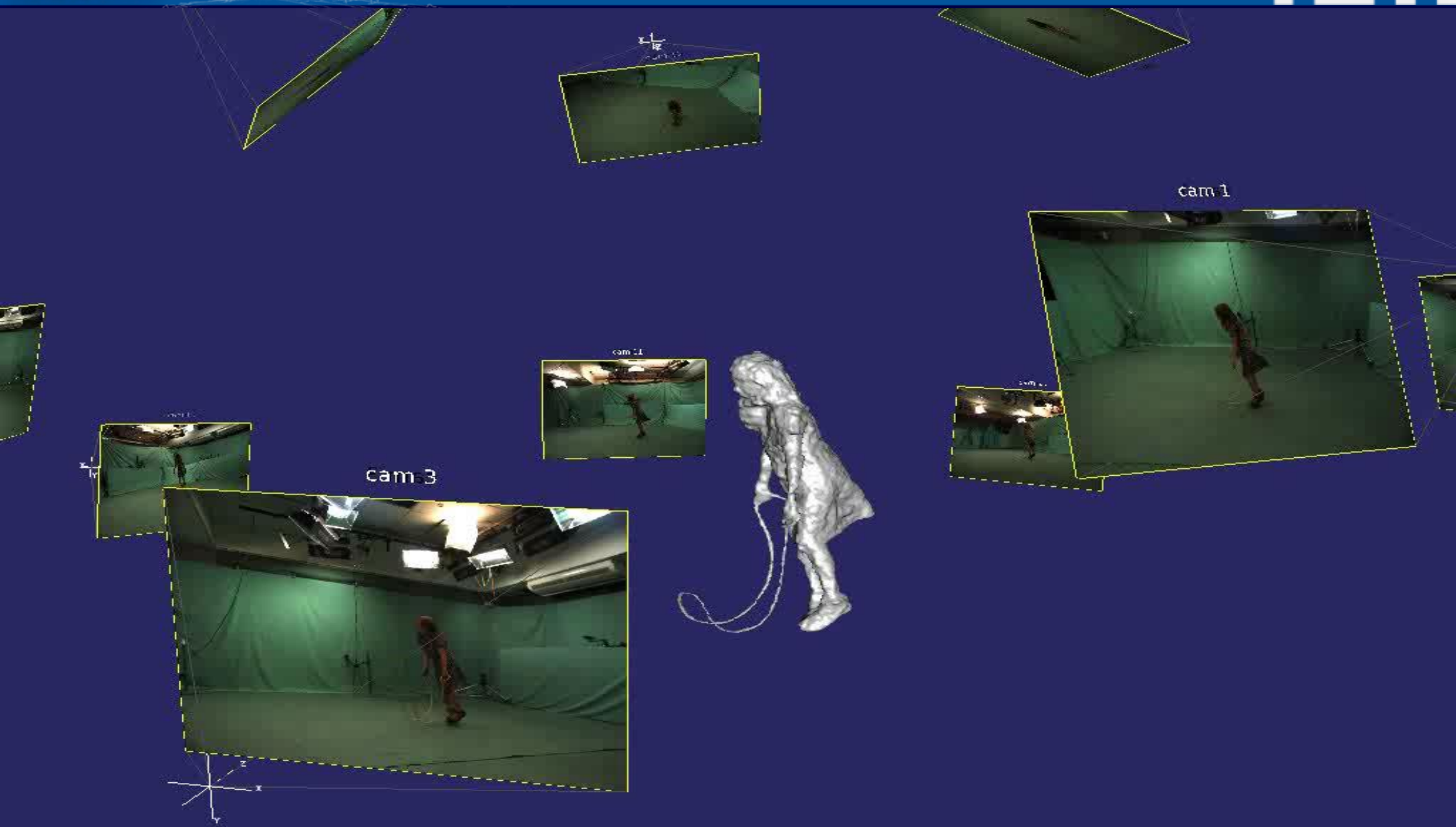
Theorem: Provably silhouette-consistent reconstructions can be computed by convex optimization over convex domains.

Action Reconstruction



Oswald, Cremers, ICCV '13 4DMoD Workshop

Reconstructing Dynamic Scenes



Oswald, Stühmer, Cremers, ECCV '14

Multiview Reconstruction



Can we do realtime dense reconstruction
from a handheld camera?

Overview



Multiview reconstruction



Super-res. textures



Free-viewpoint TV



Realtime dense scanning



Visual SLAM



Deep Networks for SLAM

From Realtime Optical Flow...



Input video



Optical flow field

Wedel, Pock, Bischof, Cremers, ICCV '09

From Realtime Optical Flow...



Input video



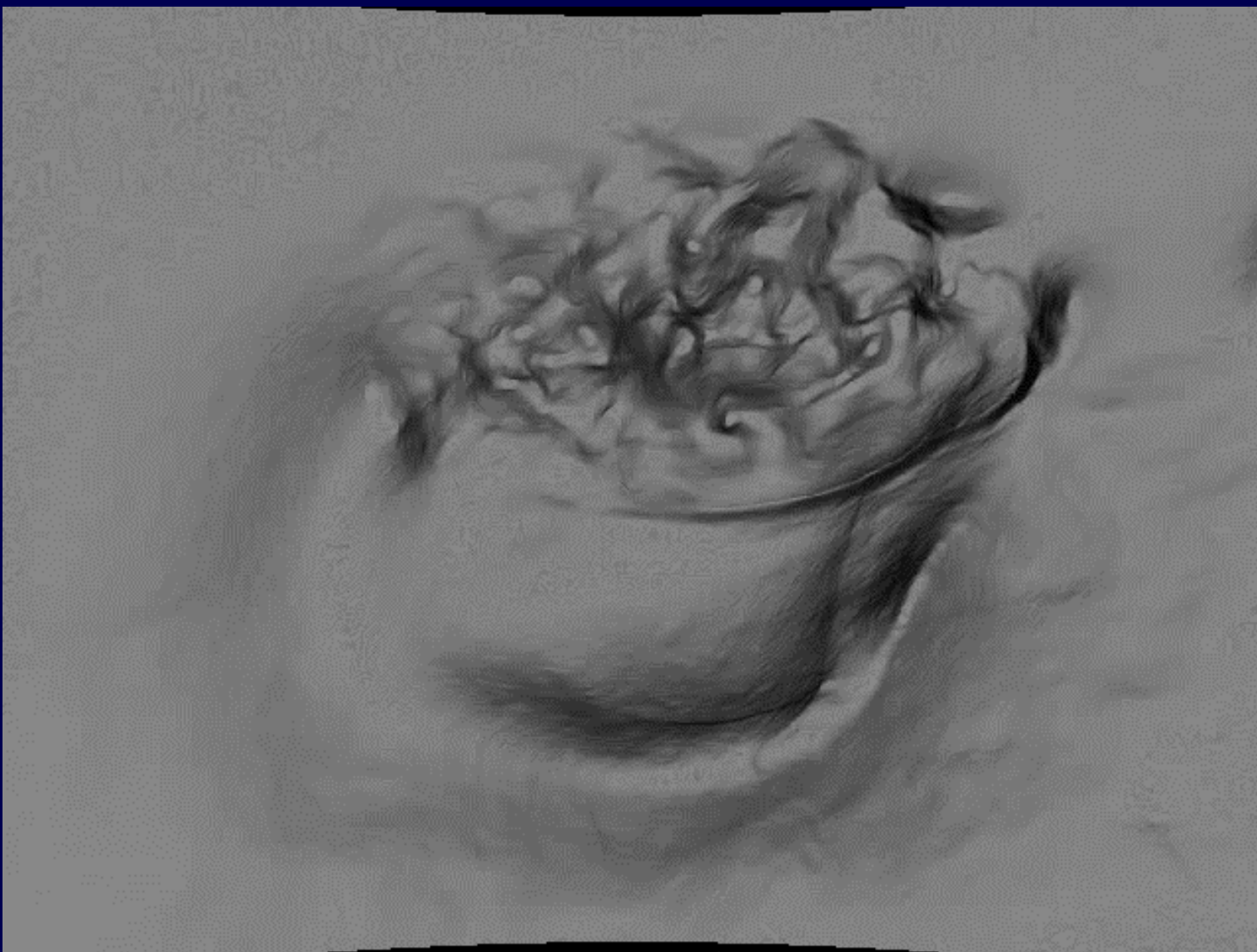
Optical flow field^{*}

* 60 fps at 640 x 480 resolution

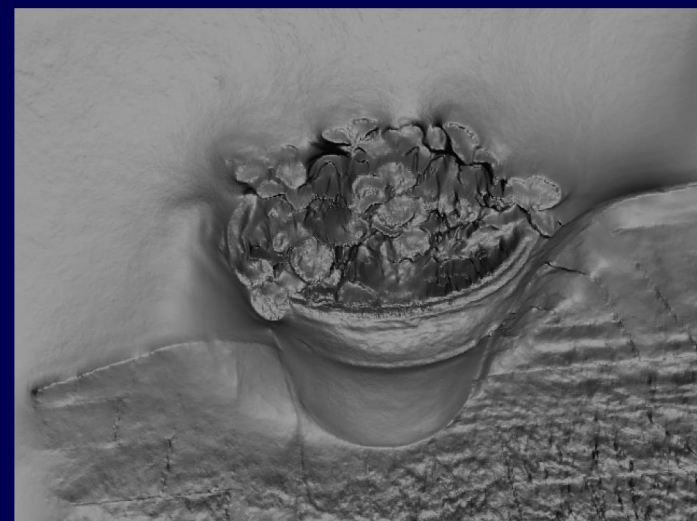
Wedel, Pock, Bischof, Cremers, ICCV '09



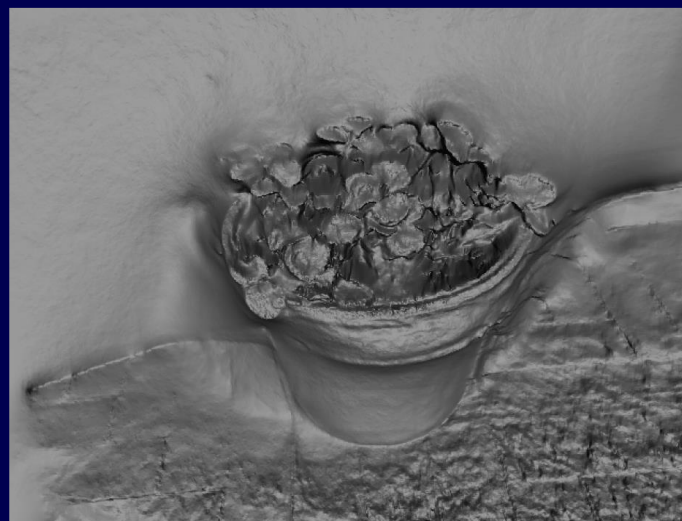
Stuehmer, Gumhold, Cremers, DAGM '10



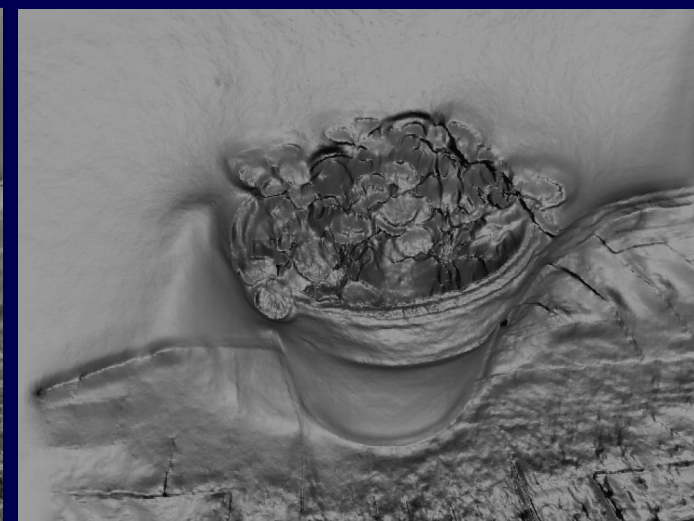
Stuehmer, Gumhold, Cremers, DAGM '10



16.0 fps



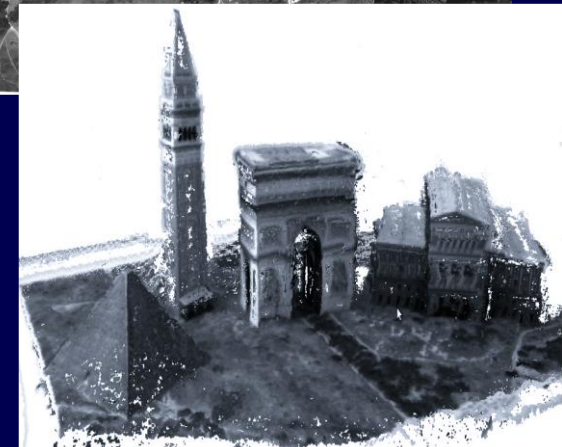
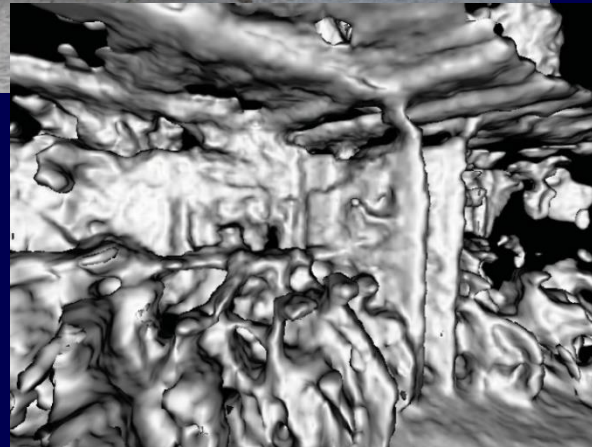
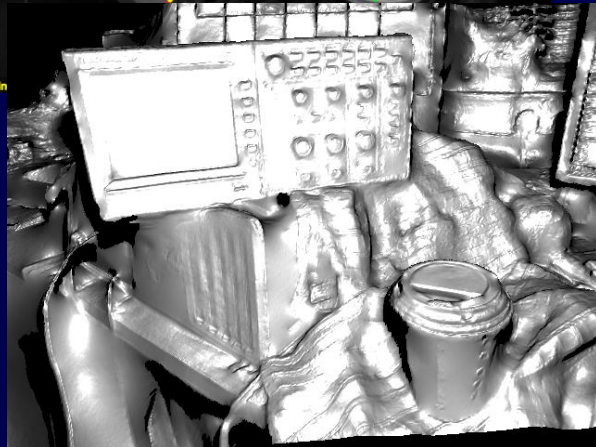
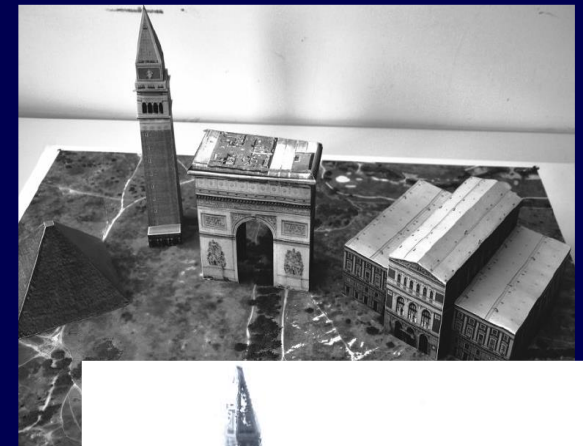
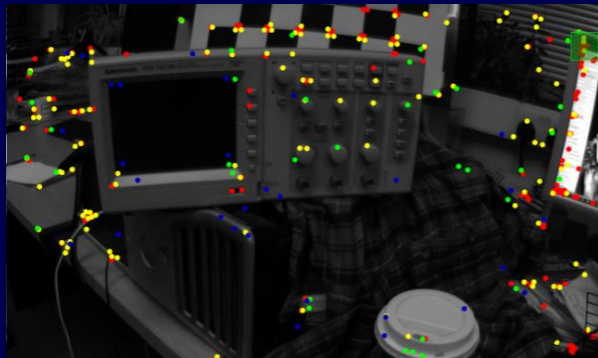
22.0 fps



41.1 fps

Stuehmer, Gumhold, Cremers, DAGM '10

Realtime Dense Reconstruction



Newcombe et al. ICCV '11

Wendel et al. CVPR '12

Pizzoli et al. ICRA '14



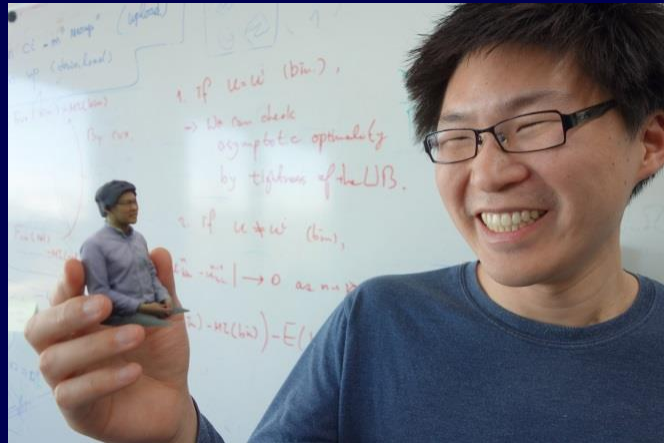
S Sturm, Bylow, Kahl, Cremers, GCPR '13

Realtime 3D Modeling





Realtime 3D Modeling



Sturm, Bylow, Kahl, Cremers, GCPR '13

Realtime 3D Modeling





Realtime 3D Modeling

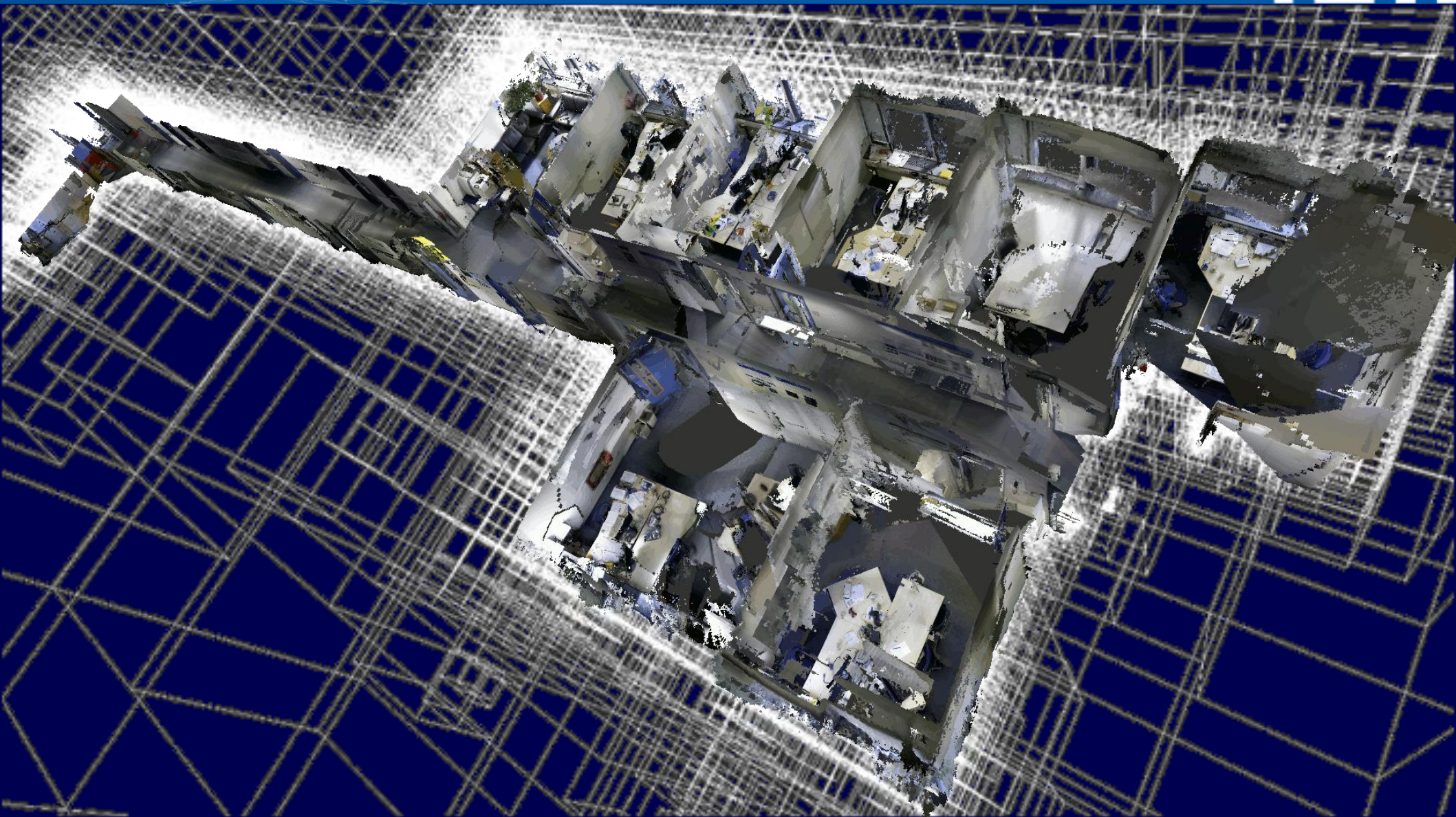


Reconstruction on the Fly

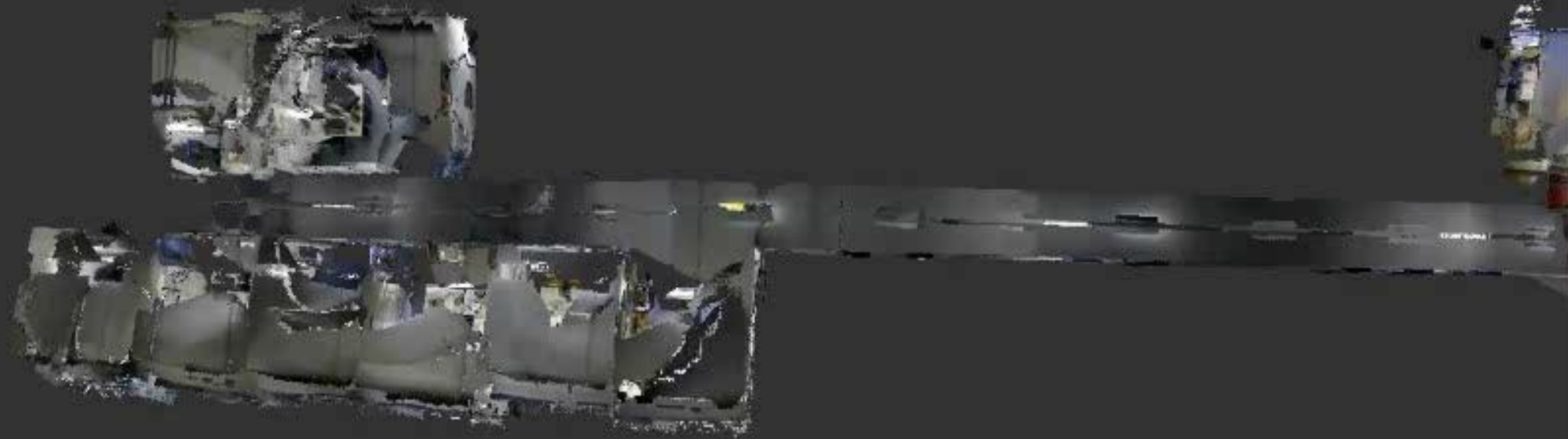


Bylow, Sturm, Kerl, Kahl, Cremers RSS '13

Large Scale: Octrees



Steinbrücker, Kerl, Sturm, Cremers ICCV '13



Large-Scale Reconstruction

Steinbrücker, Kerl, Sturm, Cremers ICCV '13, ICRA '14

Overview



Multiview reconstruction



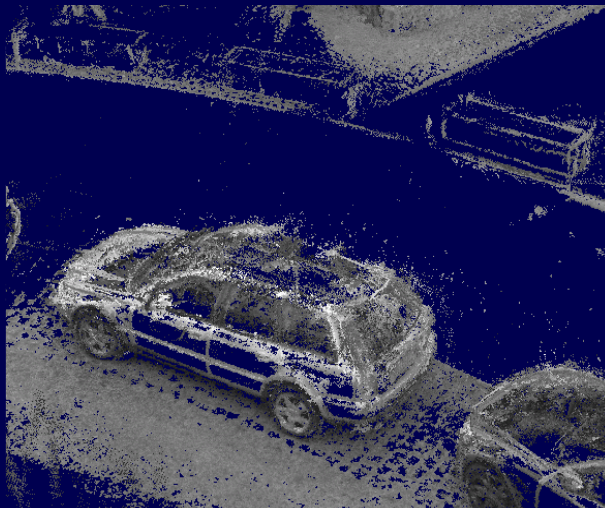
Super-res. textures



Free-viewpoint TV



Realtime dense scanning



Visual SLAM



Deep Networks for SLAM

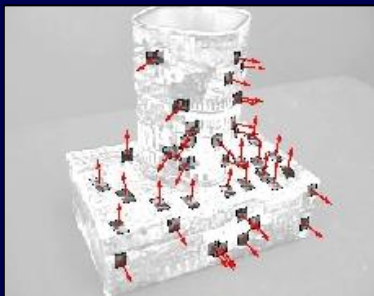


Erwin Wilhelm Kruppa (1885 - 1967)

Kruppa 1913:

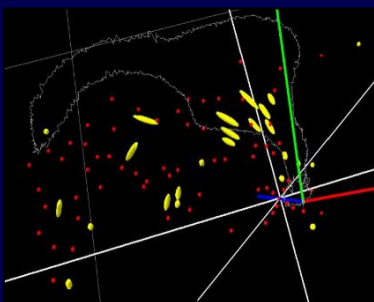
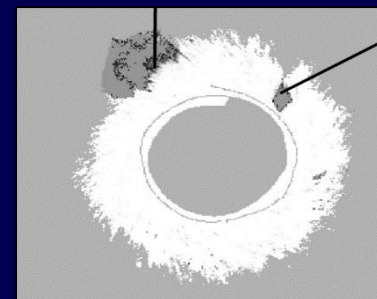
Two views of five points are sufficient to determine both the camera motion and the 3D location of these points up to finitely many solutions.

Real-time Visual SLAM



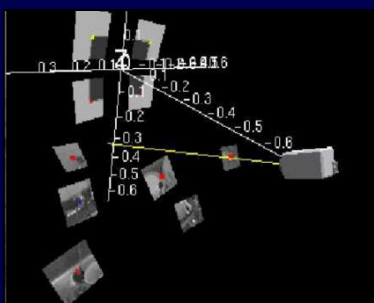
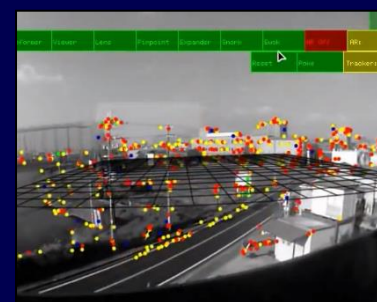
Structure from Motion Causally Integrated Over Time.
Chiuso, Favaro, Jin, Soatto; PAMI '02.

Visual Odometry.
Nistér, Naroditsky, Bergen; CVPR '04.



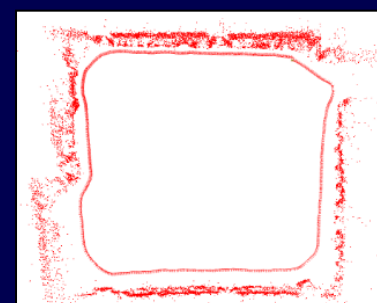
Scalable monocular SLAM.
Eade, Drummond; CVPR '06.

Parallel Tracking and Mapping for Small AR Workspaces.
Klein, Murray; ISMAR '07.



MonoSLAM: Real-time single camera SLAM.
Davison, Reid, Molton, Stasse; PAMI '07.

Scale Drift-Aware Large Scale Monocular SLAM.
Strasdat, Montiel, Davison; RSS '10.



DTAM: Dense Tracking and Mapping in Real-Time.
Newcombe, Lovegrove, Davison; ICCV '11.

REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time.
Pizzoli, Forster, Scaramuzza; ICRA '14.

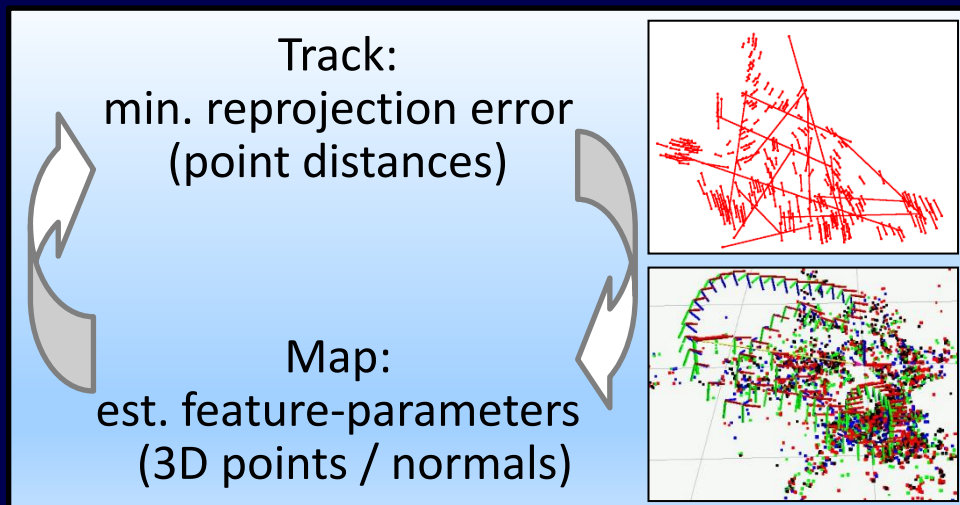


Real-time Visual SLAM

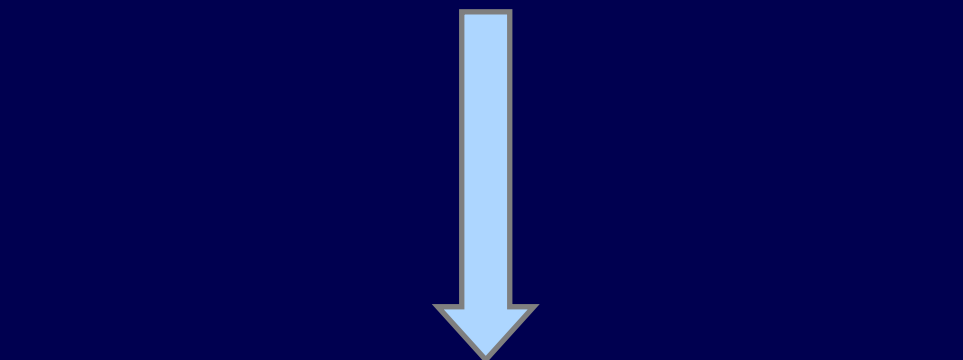
Keypoint-Based



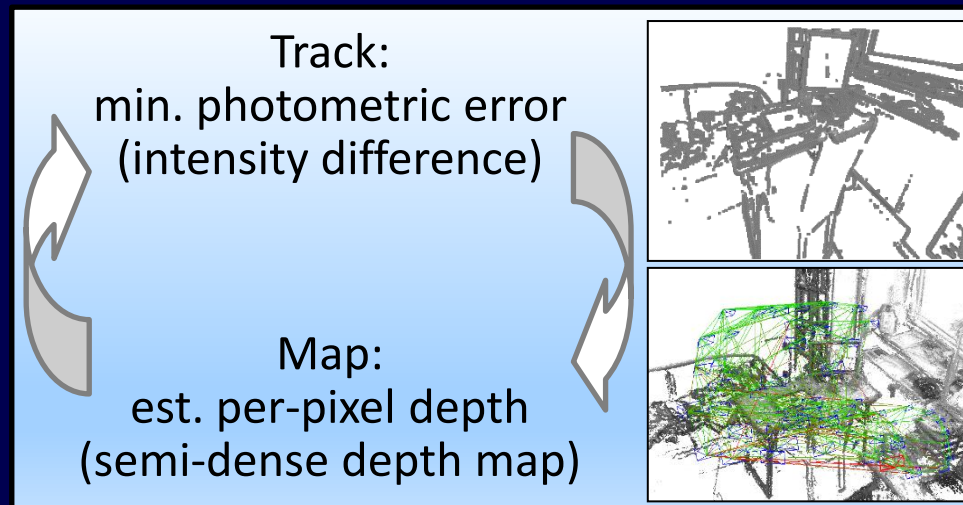
abstract images to feature observations



Direct (LSD-SLAM)



keep full image





Engel, Sturm, Cremers, ICCV '13, Engel, Schöps, Cremers, ECCV '14

Camera-Based Navigation of a Low-Cost MAV using LSD-SLAM



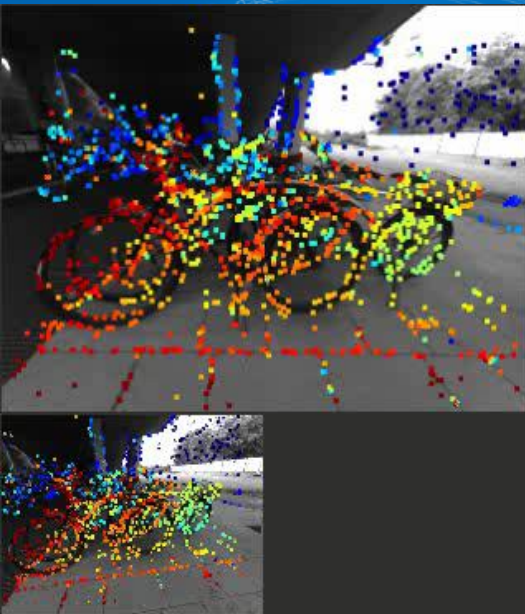
Lukas von Stumberg, Jakob Engel, Jörg Stückler, Daniel Cremers



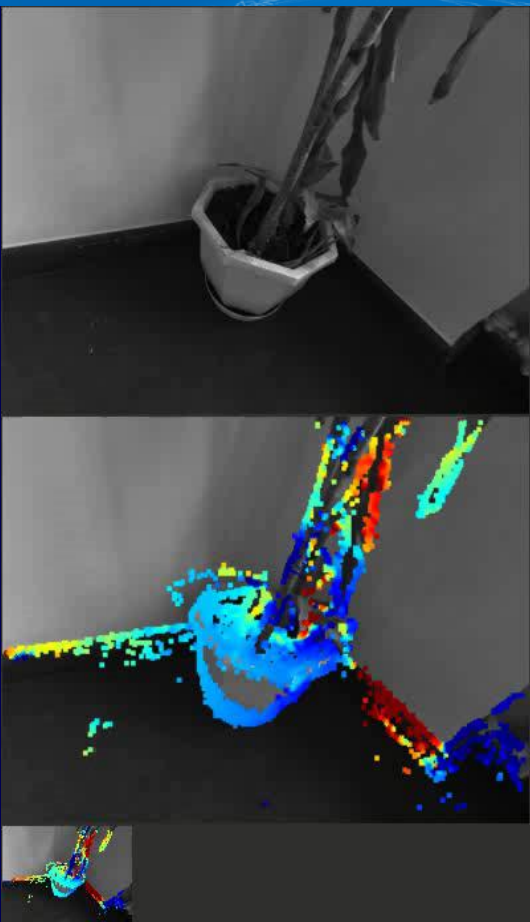
Computer Vision Group
Department of Computer Science
Technical University of Munich



Von Stumberg, Engel, Stückler, Cremers '15



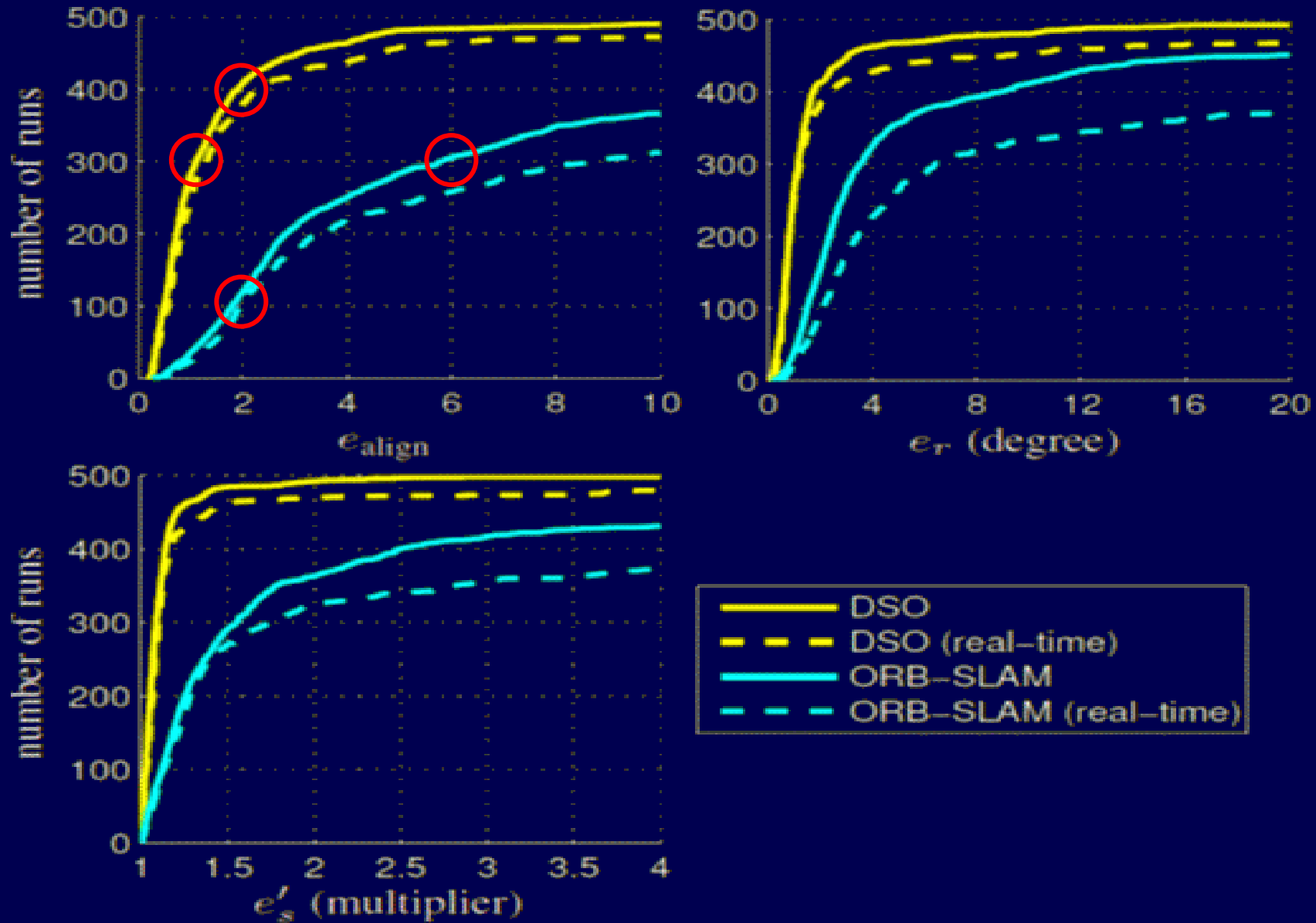
Engel, Koltun, Cremers, "Direct Sparse Odometry", PAMI '18



Engel, Koltun, Cremers, "Direct Sparse Odometry", PAMI '18



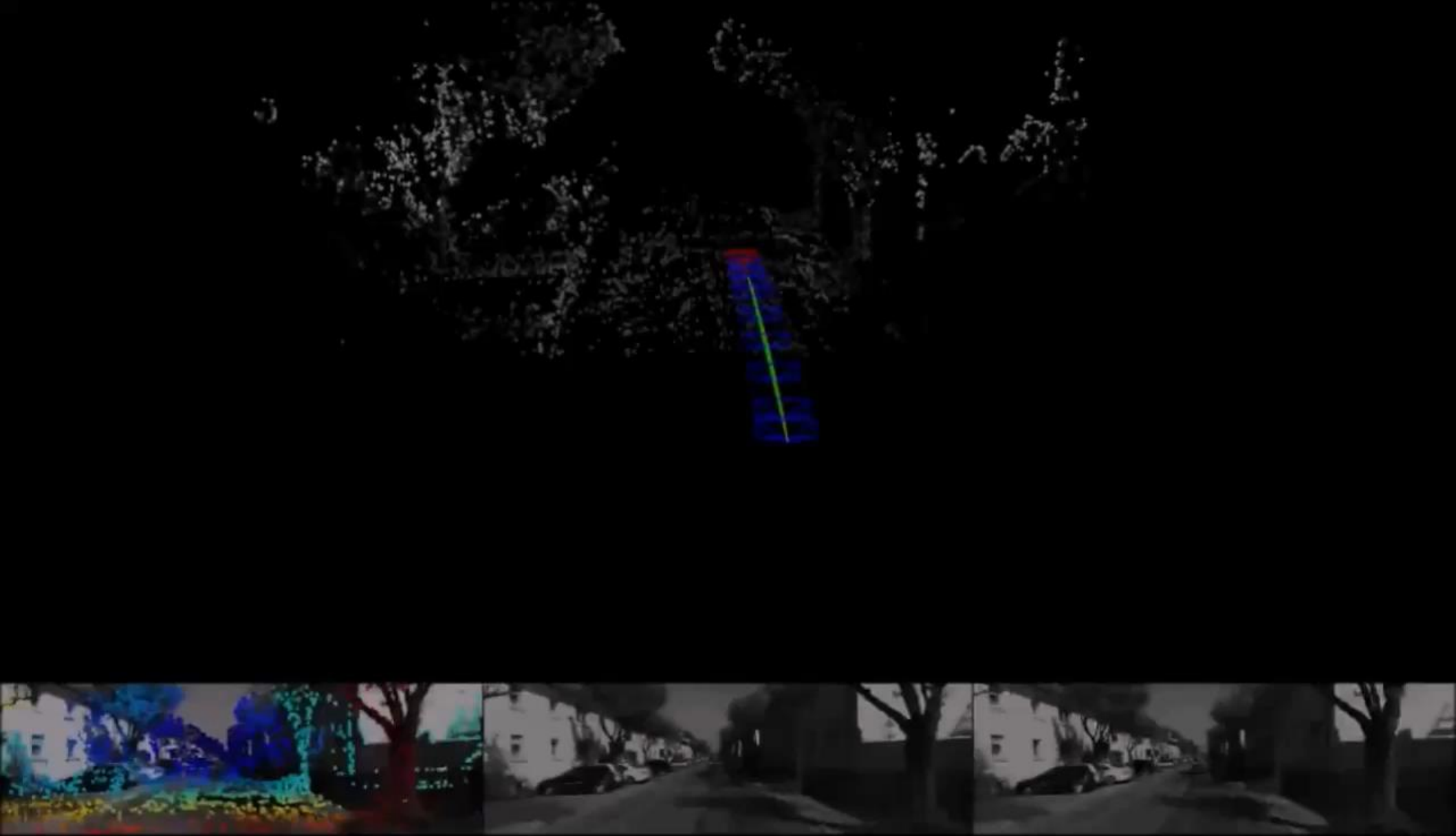
50 sequences
105 minutes
190k frames



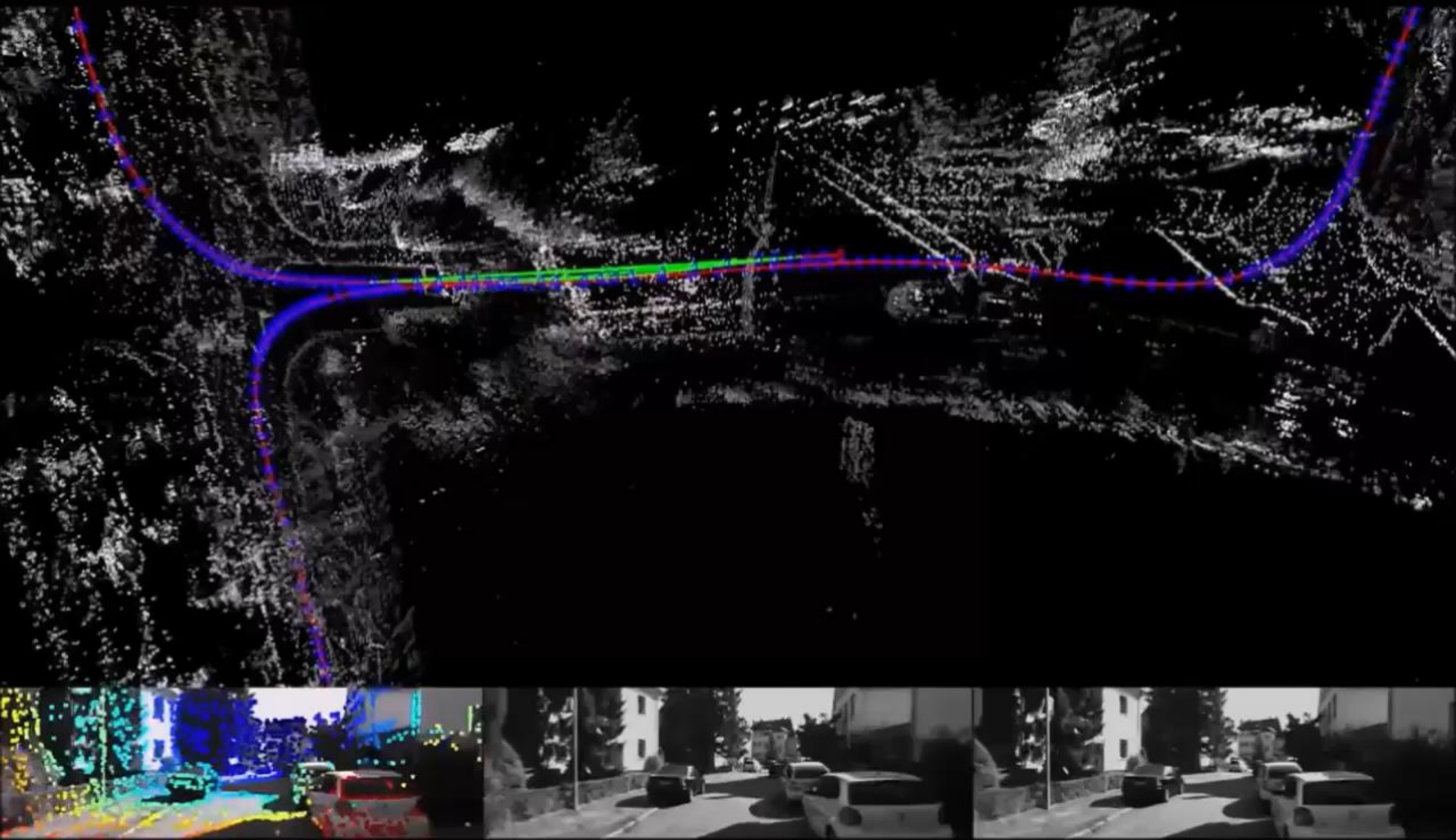
Engel, Koltun, Cremers, IEEE PAMI '18



*Engel et al. 2023:
Project Aria: A New Tool for Egocentric Multi-Modal AI Research*



Wang, Schwörer, Cremers, "Stereo DSO", ICCV '17



Wang, Schwörer, Cremers, "Stereo DSO", ICCV '17



Von Stumberg, Cremers, "DM-VIO", ICRA '22

Overview



Multiview reconstruction



Super-res. textures



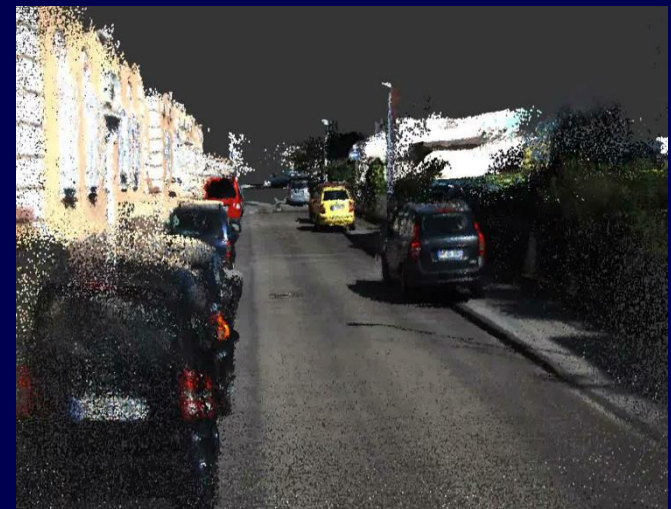
Free-viewpoint TV



Realtime dense scanning

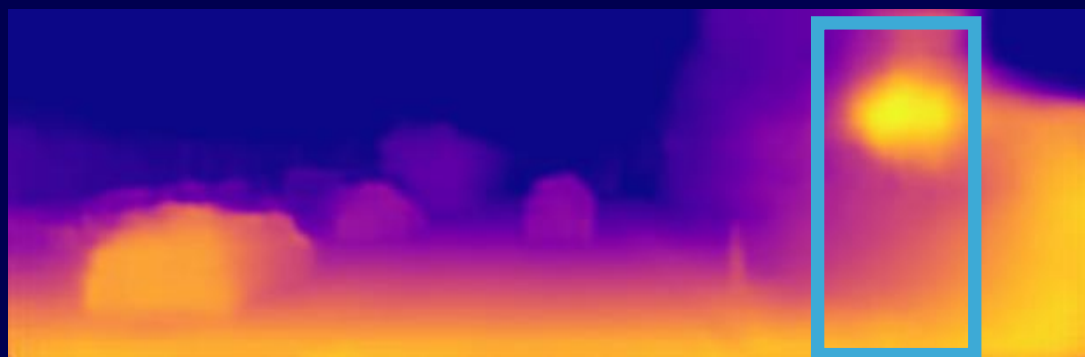


Visual SLAM

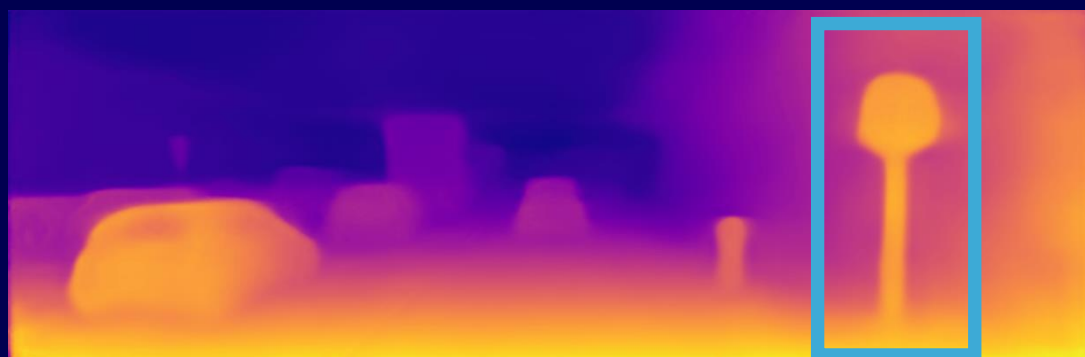


Deep Networks for SLAM

Depth Prediction from a Single Image



Kuznietsov et al. CVPR 2017



Yang, Wang, Stueckler, Cremers, ECCV 2018

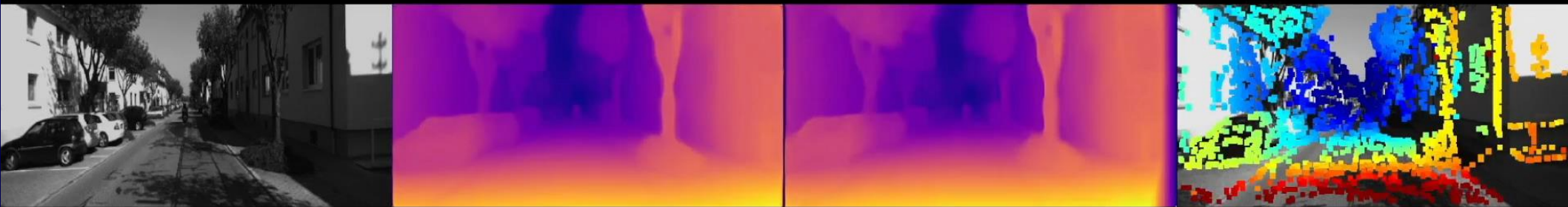
Deep
Neural
Network

Deep Virtual Stereo Odometry (DVSO)

Yang et al., “Deep Virtual Stereo Odometry”, ECCV 2018

Deep Virtual Stereo Odometry

KITTI 09

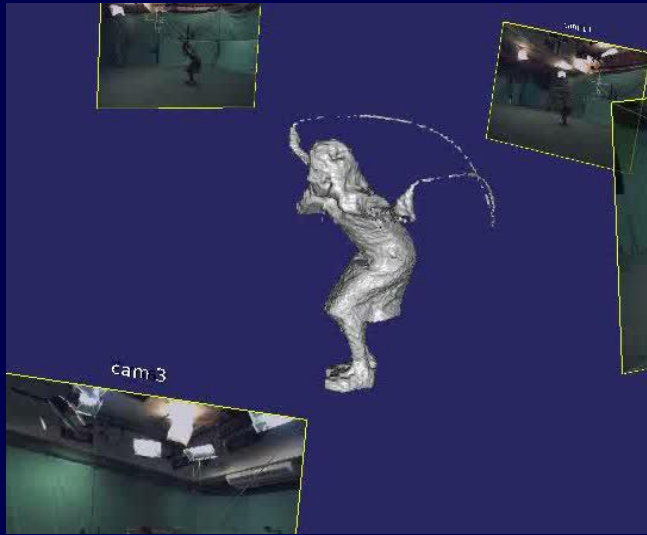


Yang et al., ECCV 2018, CVPR 2020



Wimbauer et al., "MonoRec: Monocular Dense Reconstruction", CVPR '21

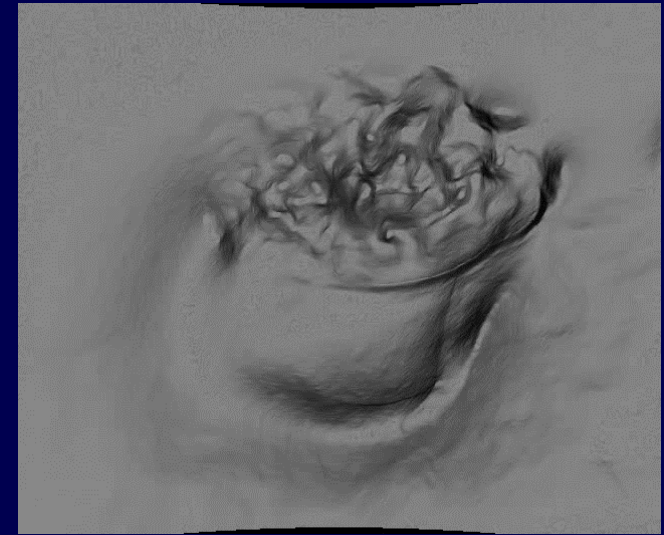
Summary



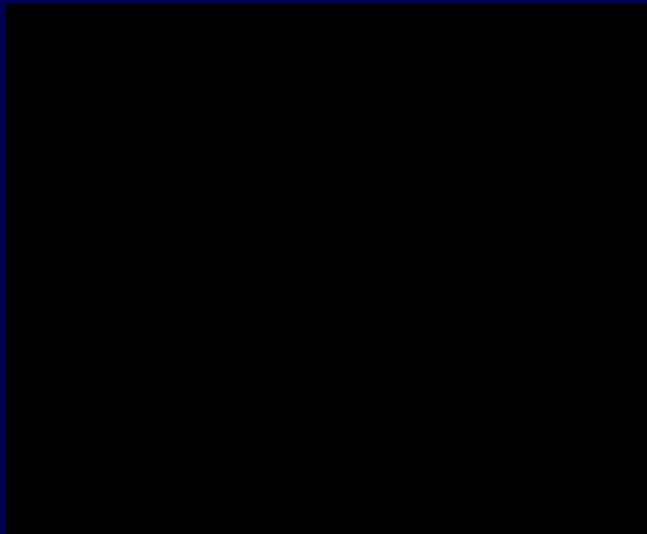
3D reconstruction



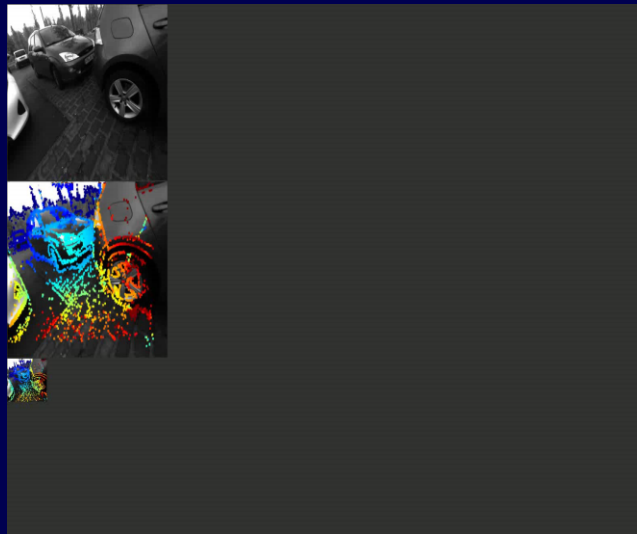
Super-res. textures



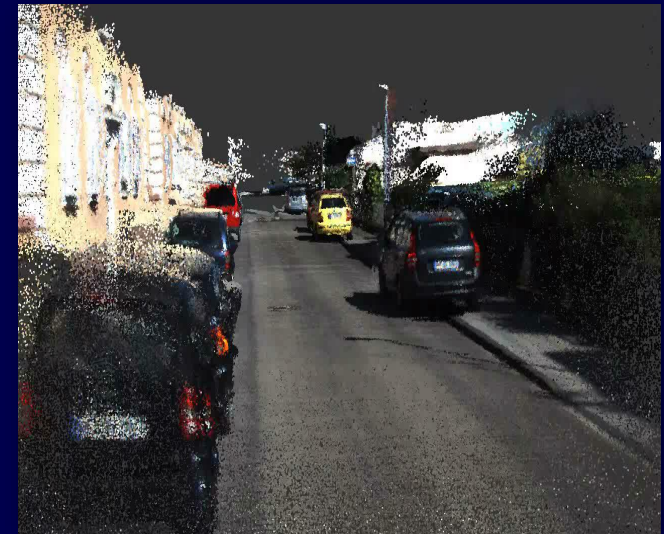
Realtime dense geometry



RGB-D Scanning



Direct Sparse Odometry



Mono dense reconstruction