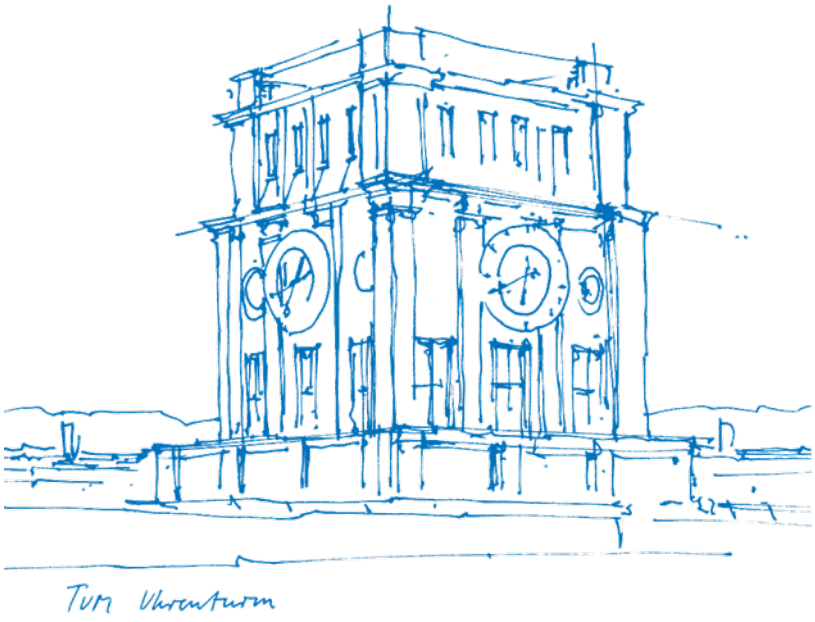
Seminar: The Evolution of Motion Estimation and Real-time 3D Reconstruction

Christiane Sommer, Lukas Köstler
Computer Vision Group
Technical University of Munich



How can I access these slides?

- Option 1 (preferred): seminar web page
  - `https://vision.in.tum.de/teaching/ss2021/seminar_realtime3d`
  - Password for material page: `ss21-realtime3d`
  - Material page will go online after this pre-meeting
- Option 2: contact organizers
  - `realtime3d-ss21@vision.in.tum.de`
  - **Only use this option if you forgot the password**

Outline

- General Information
  - About the Seminar
  - Registration
- Possible Papers
  - Depth and RGB-D Sensors
  - Monocular Cameras
- Questions

Outline

- General Information
  - About the Seminar
  - Registration
- Possible Papers
  - Depth and RGB-D Sensors
  - Monocular Cameras
- Questions

How is the seminar organized?

- Seminar meetings: Talks and discussion
  - Day: Tuesday, approximately 8 sessions (TBA)
  - Time: 14:00–16:00
  - Location: virtual (on Zoom)
  - Each session will consist of two talks which are held in English
  - **Attendance is mandatory!**
- Talk preparation / contact with supervisor
  - Day: Tuesday
  - Time: one hour between 16:00 and 18:00
  - Four weeks before talk: meet supervisor for questions (optional, but recommended)
  - Two weeks before talk: meet supervisor to go through slides (optional, but recommended)
  - One week before talk: upload slides to submission system (mandatory)
  - Two weeks after talk: upload report to submission system (mandatory)

What about the presentation?

- General set-up:
  - Duration: 20–25 minutes talk + 10–15 minutes discussion
  - Make sure to finish on time - not too early and not too late!
  - Rule of thumb: 1–2 minutes per slide $\rightarrow$ 10–20 slides
  - Do not put too much information on the slides!
- Recommended structure (talk):
  - Introduction
  - Overview / Outline
  - Method description
  - Experiments and results
  - Personal comments
  - Summary

What about the discussion after each talk?

- Discussion afterwards **will** influence your grade
- Ask questions!
- There are **no** stupid questions!

What about the final report?

- General set-up:
  - Use LaTeX template provided on web page
  - Length: 3-4 pages
  - Upload final report as pdf
  - Submission deadline: **Two weeks after talk**
- Recommended structure (main text only):
  - Introduction
  - Method description
  - Experiments and results
  - Discussion of results
  - Summary

Outline

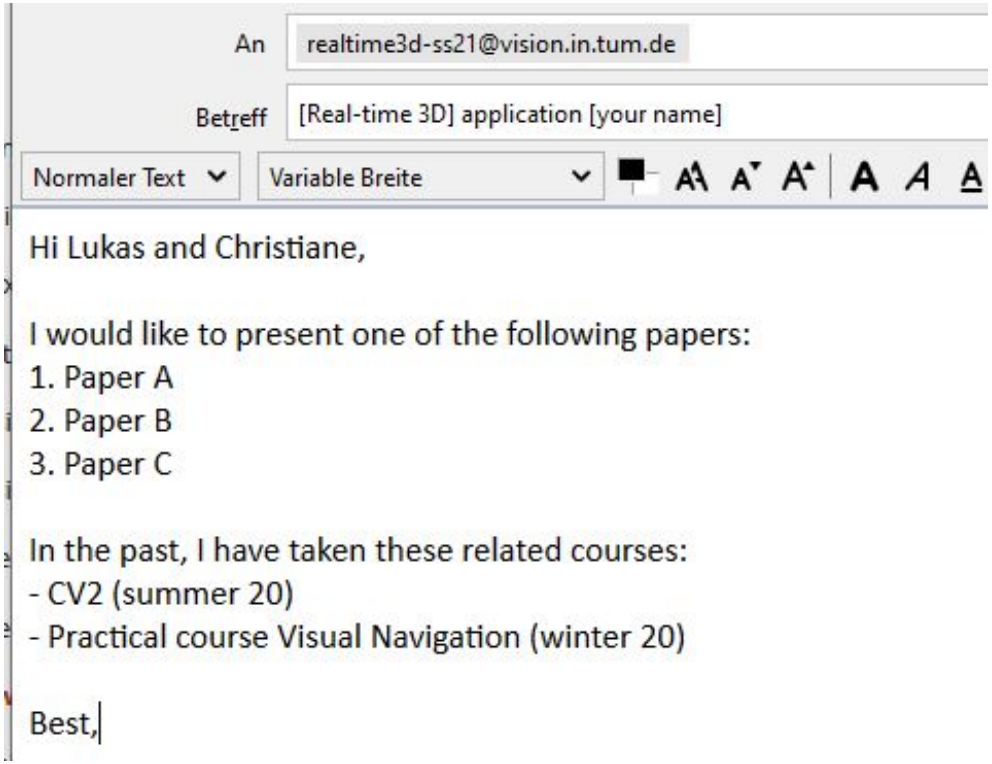
- General Information
  - About the Seminar
  - Registration
- Possible Papers
  - Depth and RGB-D Sensors
  - Monocular Cameras
- Questions

How do you register for the seminar?

- Step 1: Official registration via TUM matching system
  - Go to `https://matching.in.tum.de`
  - Register for seminar with the title *The Evolution of Motion Estimation and Real-time 3D Reconstruction*
- Step 2: Personal registration via email
  - In the list of papers on the web page, select your three favorites
  - Write an email ranking these three favorites to the seminar email address
  - Email subject: "[Realtime3D] application [your name]"
  - Include information about related lectures / courses you have taken so far
  - We do **not** need your CV or a motivation letter!
  - Registrations without email / emails with missing information will be ignored!
- Deadline for both registrations: February 16th, 2021

How do you register for the seminar?

Example registration email:



How do we select candidates and assign papers?

- Candidate selection
  - Only students registered in the matching system **AND**

    emails containing all required information will be considered
  - Among students meeting the formal criteria, selection will be random
  - Note that if you have not taken any related course, you must be willing to invest a lot of work to learn the required basics
  - You will get notified by the matching system about the decision (February 25th, 2021)
- Paper assignment
  - Papers are assigned after the participant list is finalized
  - We give our best to accommodate your preference list in the assignment

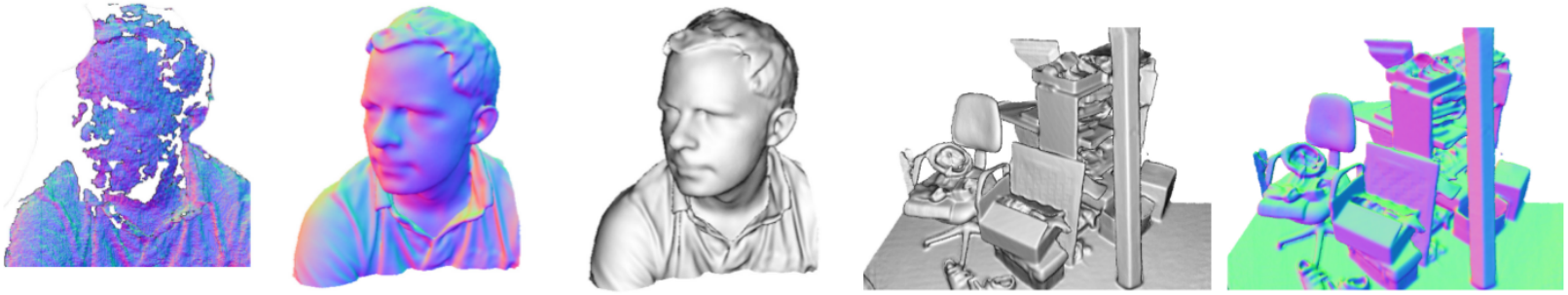
Outline

- General Information
  - About the Seminar
  - Registration
- Possible Papers
  - Depth and RGB-D Sensors
  - Monocular Cameras
- Questions

KinectFusion

R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohli, J. Shotton, S. Hodges, and A. W. Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking.

In *10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2011, Basel, Switzerland, October 26-29, 2011*, pages 127–136. IEEE Computer Society, 2011

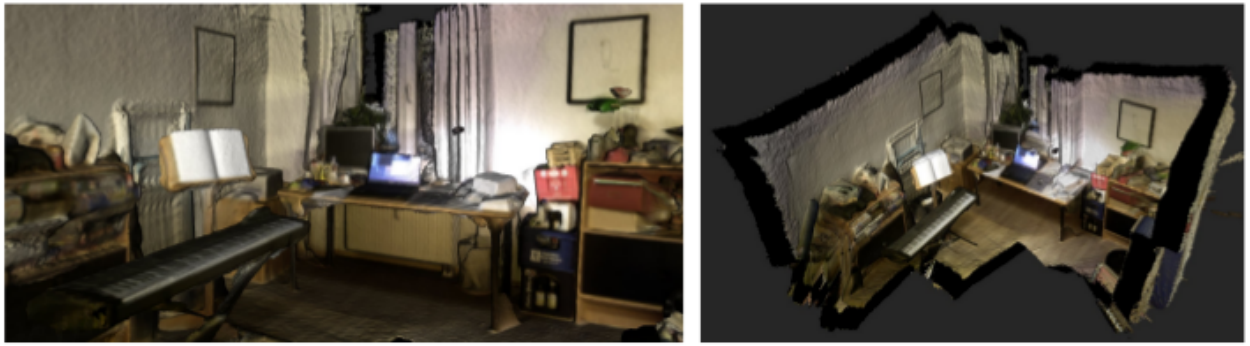


- First paper to generate dense 3D models in real-time using depth sensor and GPU
- Highly cited, impactful, baseline method for 3D reconstruction using RGB-D cameras

SDF-Tracking

E. Bylow, J. Sturm, C. Kerl, F. Kahl, and D. Cremers. Real-time camera tracking and 3d reconstruction using signed distance functions.

In P. Newman, D. Fox, and D. Hsu, editors, *Robotics: Science and Systems IX, Technische Universität Berlin, Berlin, Germany, June 24 - June 28, 2013*, 2013

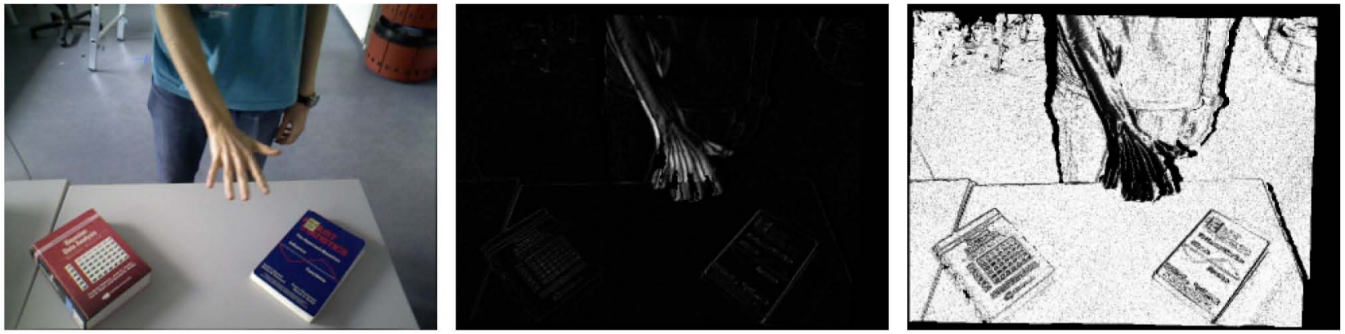


- Nice introduction to SDFs using RGB-D cameras

Dense Visual Odometry

C. Kerl, J. Sturm, and D. Cremers. Robust odometry estimation for RGB-D cameras.

In *2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, May 6-10, 2013*, pages 3748–3754. IEEE, 2013



- Odometry method that minimizes photometric cost using depth measurements
- Improved weighting function for robustness in the presence of outliers

Voxel Hashing

M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger. Real-time 3d reconstruction at scale using voxel hashing.

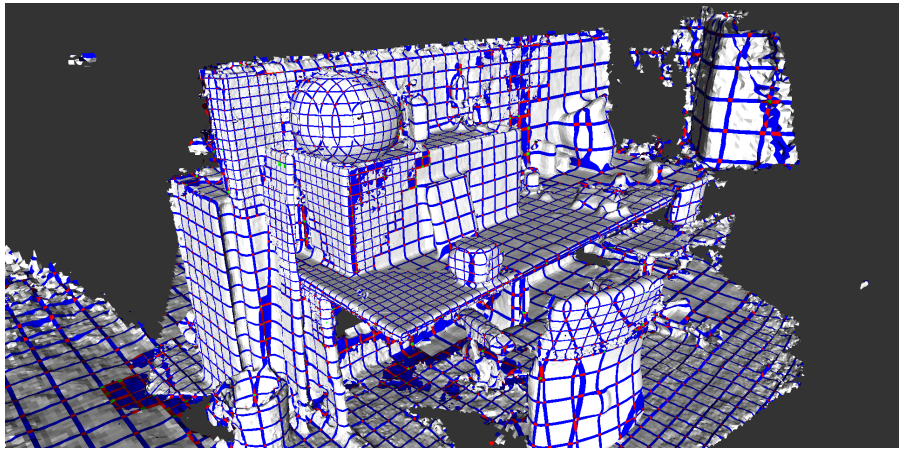
ACM Trans. Graph., 32(6):169:1–169:11, 2013

- Uses hashing to store TSDF grid efficiently
- Used within many state-of-the-art voxel-based reconstruction methods

Octree Mapping

F. Steinbrücker, J. Sturm, and D. Cremers. Volumetric 3d mapping in real-time on a CPU.

In *2014 IEEE International Conference on Robotics and Automation, ICRA 2014, Hong Kong, China, May 31 - June 7, 2014*, pages 2021–2028. IEEE, 2014

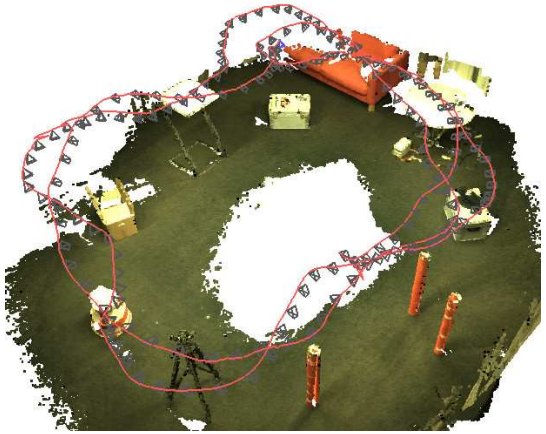


- Uses Octree to store TSDF grid efficiently
- Can run in real-time without GPU

BAD SLAM

T. Schöps, T. Sattler, and M. Pollefeys. BAD SLAM: bundle adjusted direct RGB-D SLAM.

In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 134–144. Computer Vision Foundation / IEEE, 2019



- Published on CVPR 2019
- Perform bundle adjustment on surfels to get a high quality pose

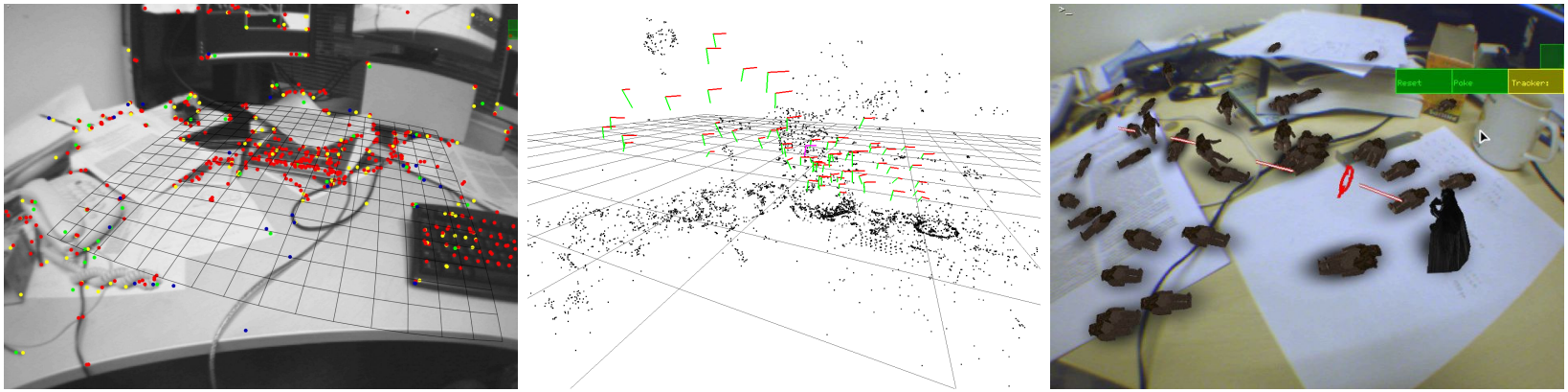
Outline

- General Information
  - About the Seminar
  - Registration
- Possible Papers
  - Depth and RGB-D Sensors
  - Monocular Cameras
- Questions

PTAM

G. Klein and D. W. Murray. Parallel tracking and mapping for small AR workspaces.

In *Sixth IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR 2007, 13-16 November 2007, Nara, Japan*, pages 225–234. IEEE Computer Society, 2007

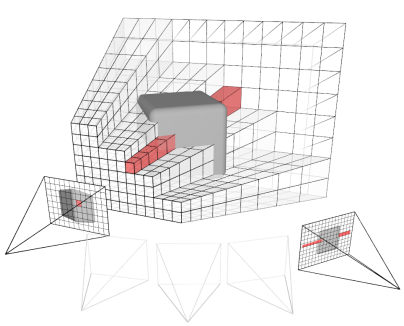


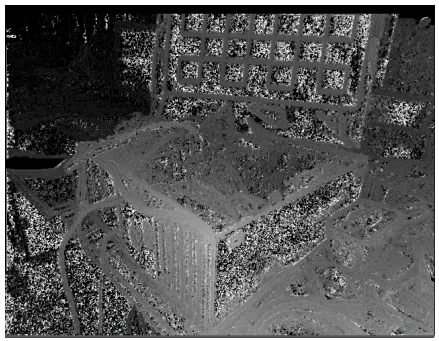
- One of the first systems capable of estimating both pose and geometry in real-time for handheld cameras
- Simple AR applications

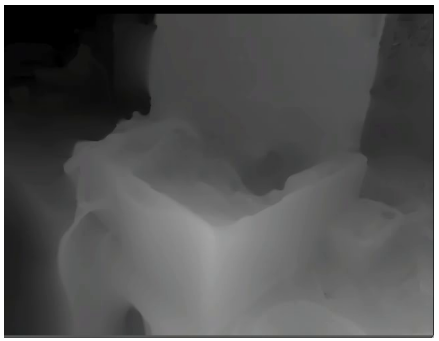
DTAM

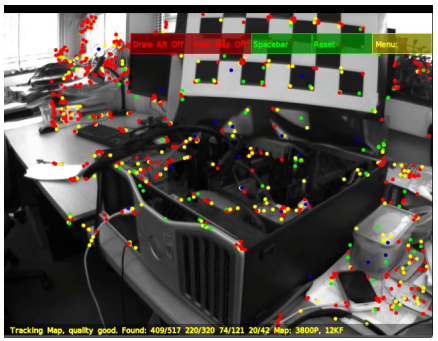
R. A. Newcombe, S. Lovegrove, and A. J. Davison. DTAM: dense tracking and mapping in real-time.

In D. N. Metaxas, L. Quan, A. Sanfeliu, and L. V. Gool, editors, *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*, pages 2320–2327. IEEE Computer Society, 2011







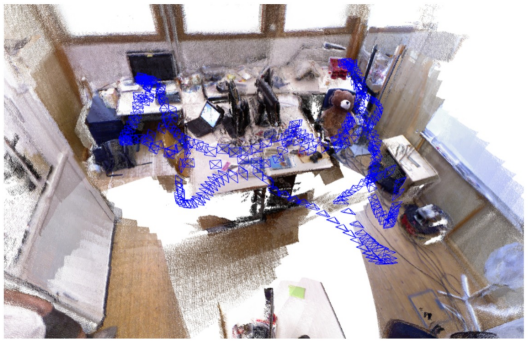


- One of the first monocular systems to create dense 3D models

ORB-SLAM

R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. ORB-SLAM: a versatile and accurate monocular SLAM system.

CoRR, abs/1502.00956, 2015

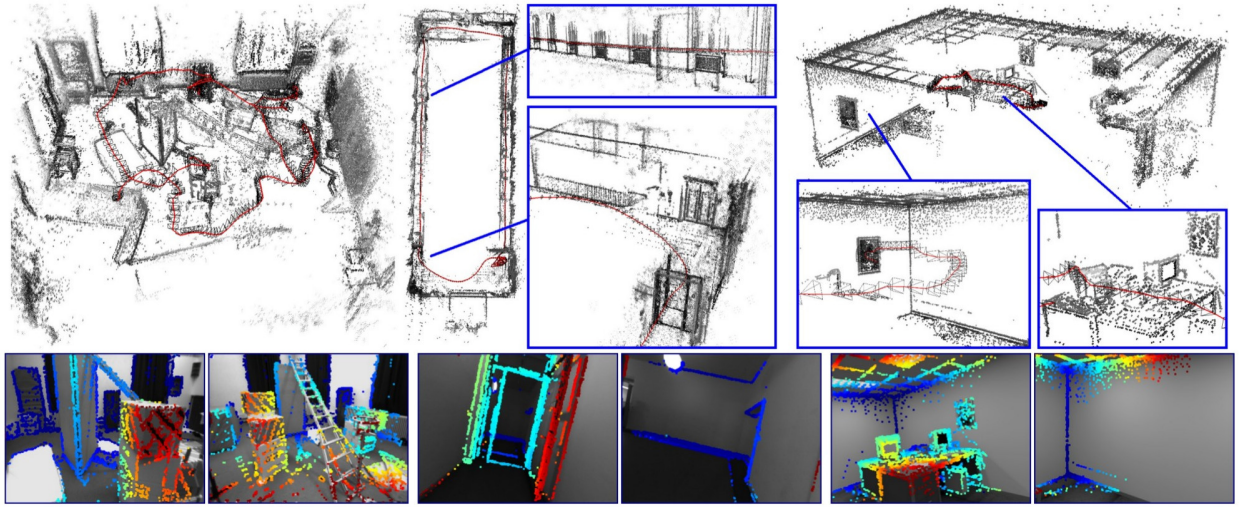


- Use all depth and color data to obtain consistent mapping

DSO

J. Engel, V. Koltun, and D. Cremers. Direct sparse odometry.

IEEE Trans. Pattern Anal. Mach. Intell., 40(3):611–625, 2018

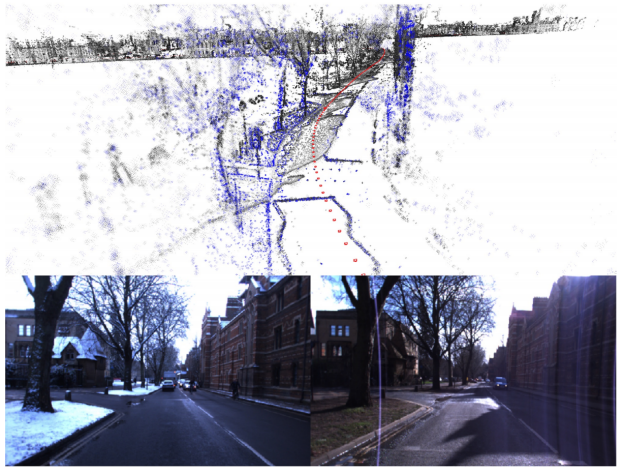


- Large-scale odometry
- Does not rely on keypoint detections

GN-Net

L. von Stumberg, P. Wenzel, Q. Khan, and D. Cremers. Gn-net: The gauss-newton loss for multi-weather relocalization.

IEEE Robotics Autom. Lett., 5(2):890–897, 2020

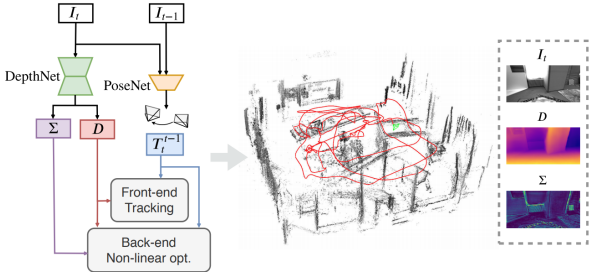


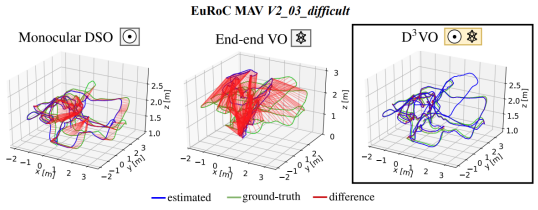
- Use feature-metric Bundle Adjustment for multi-weather relocalization
- Propose the Gauss-Newton loss to train NNs which generate feature maps that are suitable for direct image alignment

D3VO

N. Yang, L. von Stumberg, R. Wang, and D. Cremers. D3VO: deep depth, deep pose and deep uncertainty for monocular visual odometry.

In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 1278–1289. IEEE, 2020





- Monocular visual odometry framework that uses deep-learning on three levels: deep depth, pose and uncertainty estimation
- Shows impressive performance improvements in comparison to traditional methods (DSO, ORB)

Questions?

Reminder:

- Web page: `https://vision.in.tum.de/teaching/ss2021/seminar_realtime3d`
- Password: `ss21-realtime3d`
- Contact: `realtime3d-ss21@vision.in.tum.de`

References I

- E. Bylow, J. Sturm, C. Kerl, F. Kahl, and D. Cremers.
  Real-time camera tracking and 3d reconstruction using signed distance functions.
  In P. Newman, D. Fox, and D. Hsu, editors, *Robotics: Science and Systems IX, Technische Universität Berlin, Berlin, Germany, June 24 - June 28, 2013*, 2013.
- J. Engel, V. Koltun, and D. Cremers.
  Direct sparse odometry.
  IEEE Trans. Pattern Anal. Mach. Intell., 40(3):611–625, 2018.
- C. Kerl, J. Sturm, and D. Cremers.
  Robust odometry estimation for RGB-D cameras.
  In *2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, May 6-10, 2013*, pages 3748–3754. IEEE, 2013.
- G. Klein and D. W. Murray.
  Parallel tracking and mapping for small AR workspaces.
  In *Sixth IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR 2007, 13-16 November 2007, Nara, Japan*, pages 225–234. IEEE Computer Society, 2007.

References II

- R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós.
  ORB-SLAM: a versatile and accurate monocular SLAM system.
  CoRR, abs/1502.00956, 2015.
- R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohli, J. Shotton, S. Hodges, and A. W. Fitzgibbon.
  Kinectfusion: Real-time dense surface mapping and tracking.
  In *10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2011, Basel, Switzerland, October 26-29, 2011*, pages 127–136. IEEE Computer Society, 2011.
- R. A. Newcombe, S. Lovegrove, and A. J. Davison.
  DTAM: dense tracking and mapping in real-time.
  In D. N. Metaxas, L. Quan, A. Sanfeliu, and L. V. Gool, editors, *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*, pages 2320–2327. IEEE Computer Society, 2011.
- M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger.
  Real-time 3d reconstruction at scale using voxel hashing.
  ACM Trans. Graph., 32(6):169:1–169:11, 2013.

References III

- T. Schöps, T. Sattler, and M. Pollefeys.
  BAD SLAM: bundle adjusted direct RGB-D SLAM.
  In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 134–144. Computer Vision Foundation / IEEE, 2019.
- F. Steinbrücker, J. Sturm, and D. Cremers.
  Volumetric 3d mapping in real-time on a CPU.
  In *2014 IEEE International Conference on Robotics and Automation, ICRA 2014, Hong Kong, China, May 31 - June 7, 2014*, pages 2021–2028. IEEE, 2014.
- L. von Stumberg, P. Wenzel, Q. Khan, and D. Cremers.
  Gn-net: The gauss-newton loss for multi-weather relocalization.
  IEEE Robotics Autom. Lett., 5(2):890–897, 2020.
- N. Yang, L. von Stumberg, R. Wang, and D. Cremers.
  D3VO: deep depth, deep pose and deep uncertainty for monocular visual odometry.
  In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 1278–1289. IEEE, 2020.