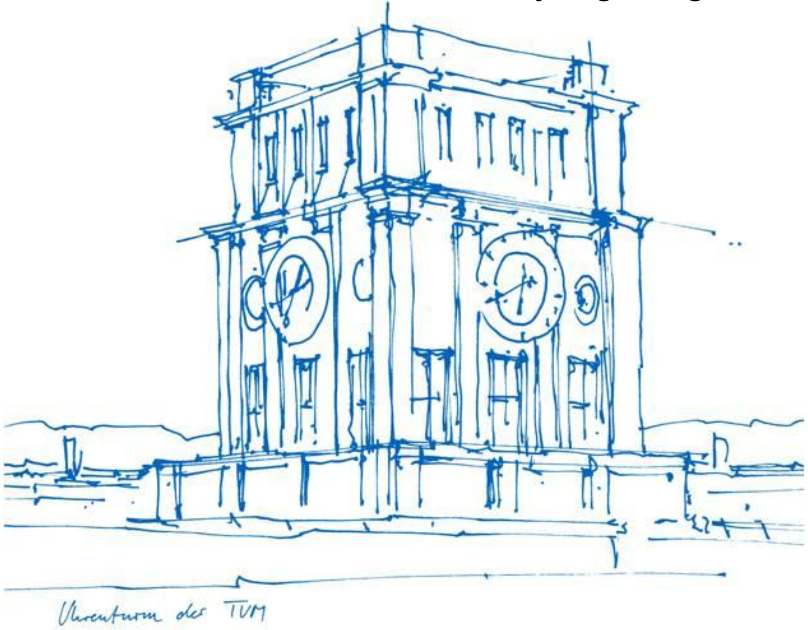
Robot Perception & Intelligence

Simon Schaefer, Simon Boche, Hanzhi Chen, Sebastian Barbas Laina, Jiaxin Wei, Dr. Jaehyung Jung

Smart Robotics Lab

Technical University of Munich

SS 2024



Outline

- General Information
    - About the seminar
    - Registration
- Topics
- Questions

How is the seminar organized?

- Slides / Material: seminar webpage
  - https://srl.cit.tum.de/teaching/s24/seminar_rpi
  - Password: rpi_s24 − Material page will go online soon

- Questions / Meeting arrangement: contact organizers
  - Use email address from website

How is the seminar organized?

- Seminar meetings: talks and discussion
  - Block Seminar
  - Time: 14:00 ~ 19:00 June 12th and 8:00 ~ 12:00 June 13th
  - Room: 00.08.053
  - **Attendance is mandatory!**

- Talk preparation / contact with supervisor
  - Read through your topic and related papers and write down what you don’t understand
  - **Up to one week before talk** (optional, but recommended) talk: meet supervisor for questions & feedback
  - **Two weeks after** talk: submit your report via email

What about the presentation?

- General setup:
  - Duration: 20-25 minutes talk + 5-10 minutes discussion
  - Make sure to **finish on time**!
  - Rule of thumb: 1-2 minutes per slide → 10-20 slides
  - Do not put too much information on the slides!

- Recommended structure (talk only):
  - Introduction / Motivation
  - Overview / Outline
  - Related Work(s)
  - Method description(s)
  - Experiments and results
  - Personal comments
  - Future work (important)
  - Summary

What about the final report?

- General setup:
  - Use LATEX template provided on web page
  - Length: 4-5 pages
  - Send final report as pdf by email to course email
  - Submission deadline: **two weeks after talk**

- Recommended structure (main text only; can be more comprehensive/extensive than your presentation):
  - Introduction
  - Related work
  - Method description(s)
  - Experiments and results
  - Discussion of results
  - Future work (important)
  - Summary

Summary: how will the seminar be graded?

- Presentation
- Final Report
- Contributions to seminar discussions

  ⇒ Ask questions!

How do you register for the seminar?

- Step 1: Official registration via TUM matching system
  - Go to matching.in.tum.de
  - Register for seminar named “Robot Perception & Intelligence”

- Step 2: Personal registration via email
  - In the list of topics, select your **three** favorites
  - Write an email ranking these three favorites to course email
  - Email subject: “RPI seminar application [your name]”
  - Include information about related lectures / courses you have taken so far (**Transcript should be attached**).
  - We do **not** need a CV or a motivation letter!
  - **Registrations without email / emails with missing information will be ignored**!

- Deadline for both registrations: 14.02.

How do we select candidates and papers?

- Candidate selection
  - Only students registered in the matching system **AND** with emails containing all required information will be considered
  - Among students meeting the formal criteria, selection will be random (matching system)
  - You will get notified by the matching system about the decision

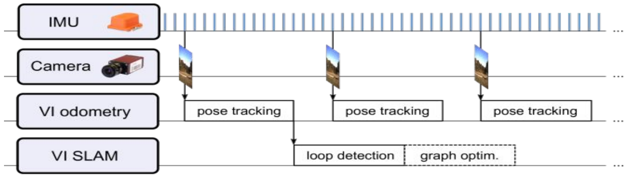
- Topic assignment
  - Topics are assigned after the participant list is finalized
  - We give our best to accommodate your preference list in the assignment

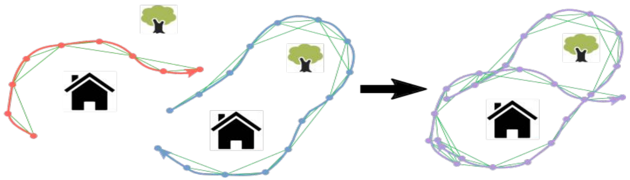
Outline

- General Information
  - About the seminar
  - Registration
- Topics
- Questions

Recent Advances in Visual Inertial SLAM

Advisor: Sebastián Barbas Laina



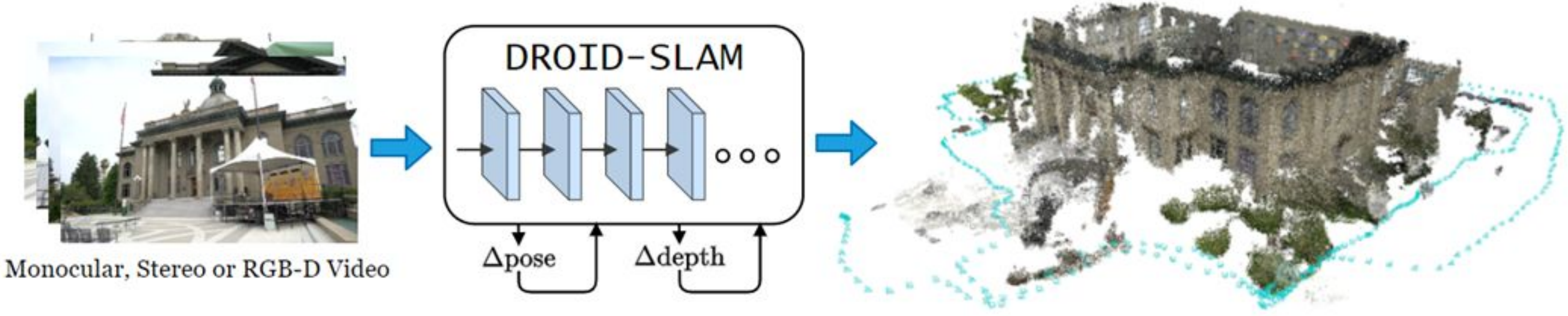


Related Papers:

- Leutenegger, Stefan. "Okvis2: Realtime scalable visual-inertial slam with loop closure." *arXiv preprint arXiv:2202.09199* (2022).
- Han, Liming, et al. "Deepvio: Self-supervised deep learning of monocular visual inertial odometry using 3d geometric constraints." *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019.
- Yang, Mingyu, Yu Chen, and Hun-Seok Kim. "Efficient deep visual and inertial odometry with adaptive visual modality selection." *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXVIII*. Cham: Springer Nature Switzerland, 2022.

Learning-Based Differentiable SLAM

Advisor: Sebastián Barbas Laina, Simon Schaefer

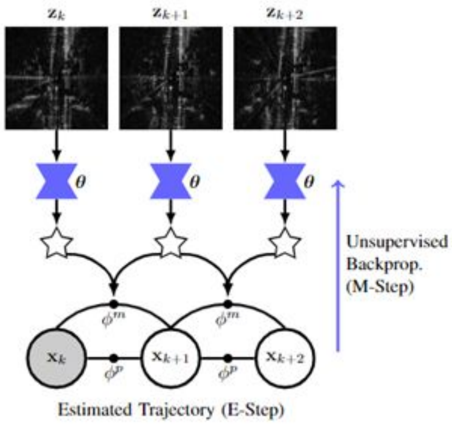


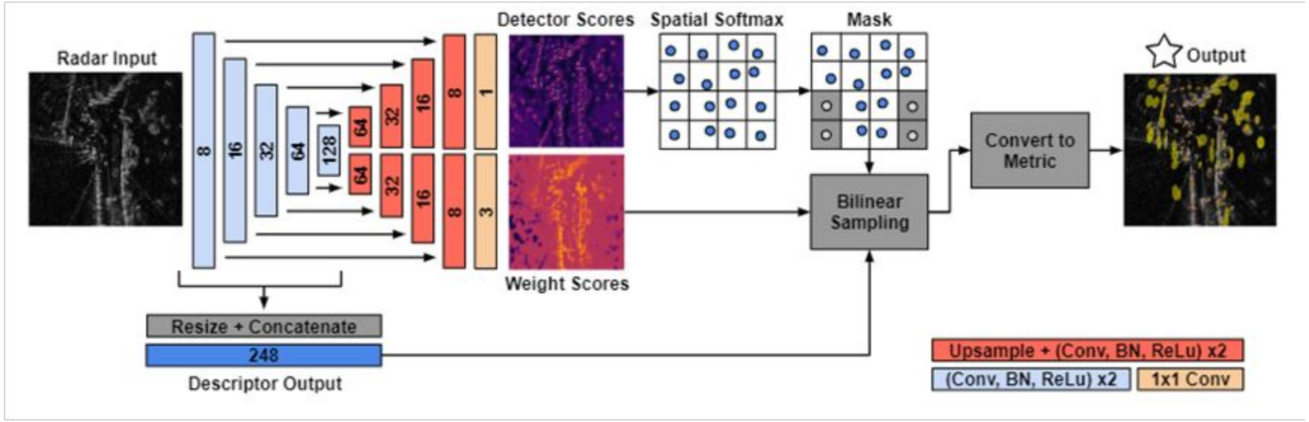
Related Papers:

- Teed, Zachary, and Jia Deng. "Droid-slam: Deep visual slam for monocular, stereo, and rgb-d cameras." *Advances in neural information processing systems* 34 (2021): 16558-16569.
- Lisus, Daniil, and Connor Holmes. "Towards Open World NeRF-Based SLAM." *arXiv preprint arXiv:2301.03102* (2023).

Unsupervised Feature Learning for Odometry

Advisor: Sebastián Barbas



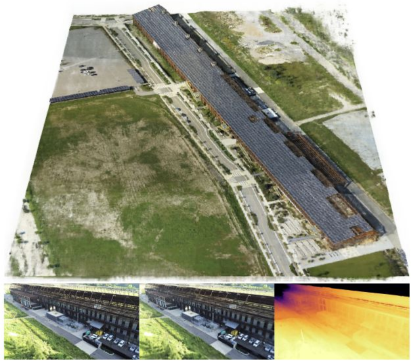


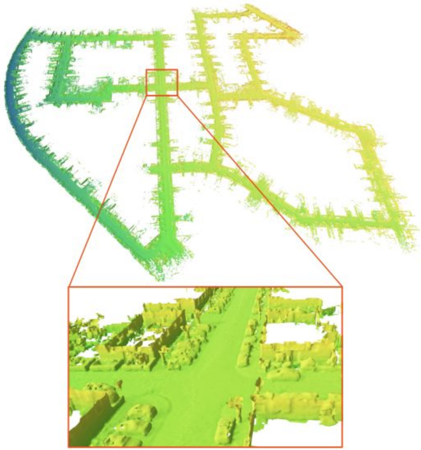
Related Papers:

- Burnett, Keenan, et al. "Radar odometry combining probabilistic estimation and unsupervised feature learning." *arXiv preprint arXiv:2105.14152* (2021).
- Another reference: Unsupervised Learning of Lidar Features for Use in a Probabilistic Trajectory Estimator, David J. Yoon, Haowei Zhang, Mona Gridseth, Hugues Thomas, Timothy D. Barfoot, RAL 2021.

Implicit Mapping at large scale

Advisor: Sebastián Barbas Laina



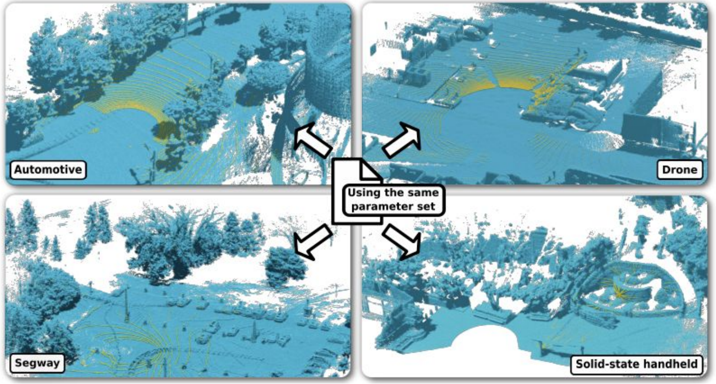


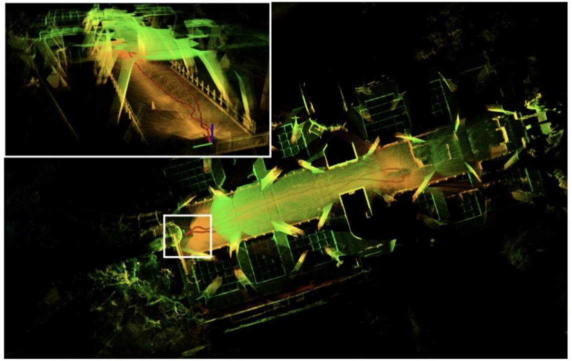
Related Papers:

- Zhong, Xingguang, et al. "SHINE-Mapping: Large-Scale 3D Mapping Using Sparse Hierarchical Implicit Neural Representations." *arXiv preprint arXiv:2210.02299* (2022).
- Turki, Haithem, Deva Ramanan, and Mahadev Satyanarayanan. "Mega-nerf: Scalable construction of large-scale nerfs for virtual fly-throughs." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Lidar-based Odometry

Advisor: Simon Boche,
Sebastián Barbas Laina



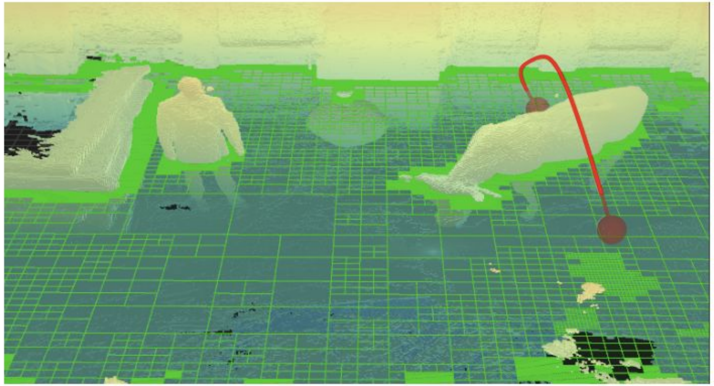


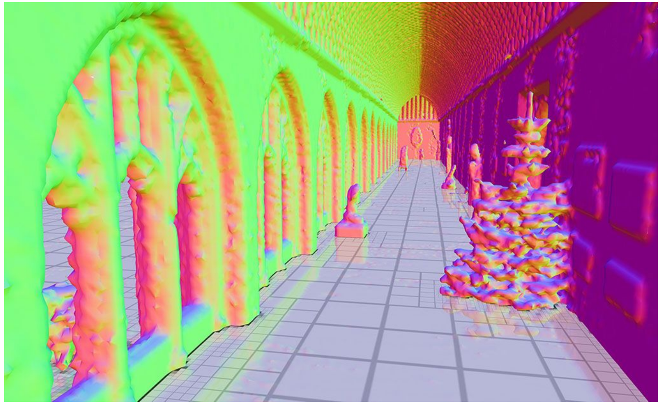
Related Papers:

- Vizzo, Ignacio, et al. "KISS-ICP: In Defense of Point-to-Point ICP Simple, Accurate, and Robust Registration If Done the Right Way." *IEEE Robotics and Automation Letters* (2023).
- Xu, Wei, et al. "Fast-lio2: Fast direct lidar-inertial odometry." *IEEE Transactions on Robotics* 38.4 (2022): 2053-2073.

Efficient / Real-Time Volumetric Mapping

Advisor: Simon Boche



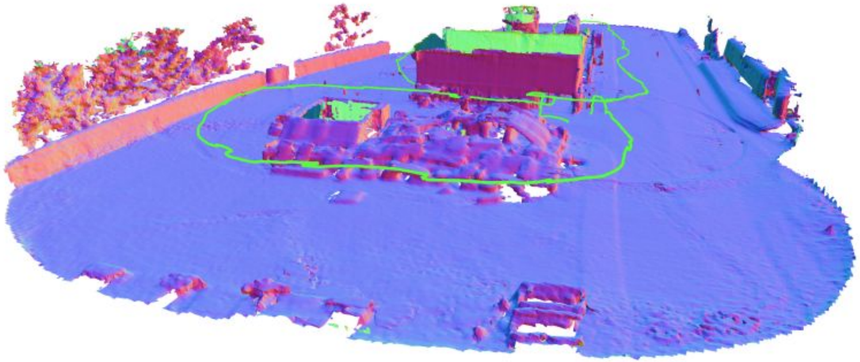


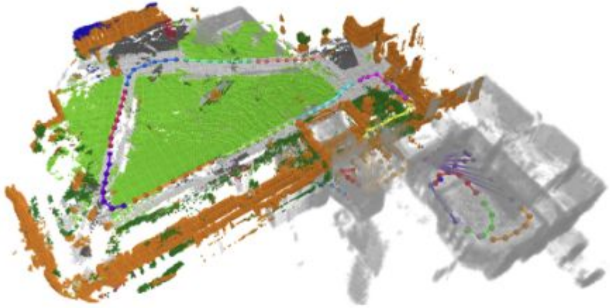
Related Papers:

- Reijgwart, Victor, et al. "Efficient volumetric mapping of multi-scale environments using wavelet-based compression." (2023).
- Funk, Nils, et al. "Multi-resolution 3D mapping with explicit free space representation for fast and accurate mobile robot motion planning." *IEEE Robotics and Automation Letters* 6.2 (2021): 3553-3560.
- Vizzo, Ignacio, et al. "Vdbfusion: Flexible and efficient tsdf integration of range sensor data." *Sensors* 22.3 (2022): 1296.
- Duberg, Daniel, and Patric Jensfelt. "UFOMap: An efficient probabilistic 3D mapping framework that embraces the unknown." *IEEE Robotics and Automation Letters* 5.4 (2020): 6411-6418.

Large-Scale Volumetric Mapping

Advisor: Simon Boche



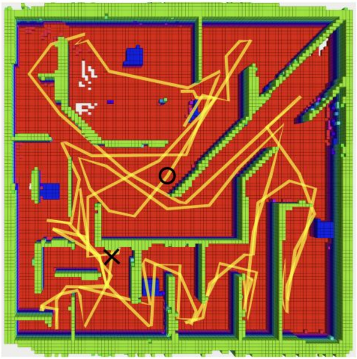


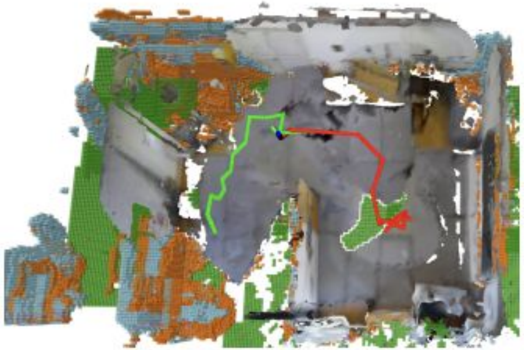
Selection of Related Papers:

- Reijgwart, Victor, et al. "Voxgraph: Globally consistent, volumetric mapping using signed distance function submaps." *IEEE Robotics and Automation Letters* 5.1 (2019): 227-234.
- Wang, Yiduo, et al. "Strategies for large scale elastic and semantic LiDAR reconstruction." *Robotics and Autonomous Systems* (2022). https://www.sciencedirect.com/science/article/pii/S0921889022001075

Autonomous Exploration

Advisor: Simon Boche



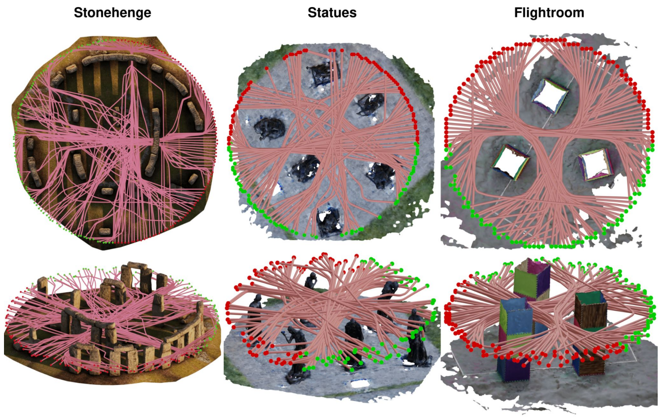


<u>Selection of Related Papers:</u>

- Bircher, Andreas, et al. "Receding horizon path planning for 3D exploration and surface inspection." *Autonomous Robots* 42 (2018): 291-306.
- Dai, Anna, et al. "Fast frontier-based information-driven autonomous exploration with an mav." *2020 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2020.
- Schmid, Lukas, et al. "Incremental 3d scene completion for safe and efficient exploration mapping and planning." *arXiv preprint arXiv:2208.08307* (2022).

Robot Navigation in Neural Radiance Fields

Advisor: Simon Boche

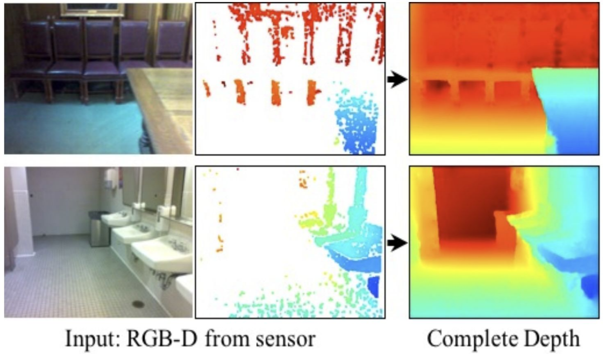


Selection of Related Papers:

- Adamkiewicz, Michal, et al. "Vision-only robot navigation in a neural radiance world." *IEEE Robotics and Automation Letters* (2022)
- Chen, Timothy, Preston Culbertson, and Mac Schwager. "CATNIPS: Collision Avoidance Through Neural Implicit Probabilistic Scenes." *arXiv preprint arXiv:2302.12931* (2023).

Depth Completion using RGB/RGB-D/LiDAR

Advisor: Sebastián Barbas Laina

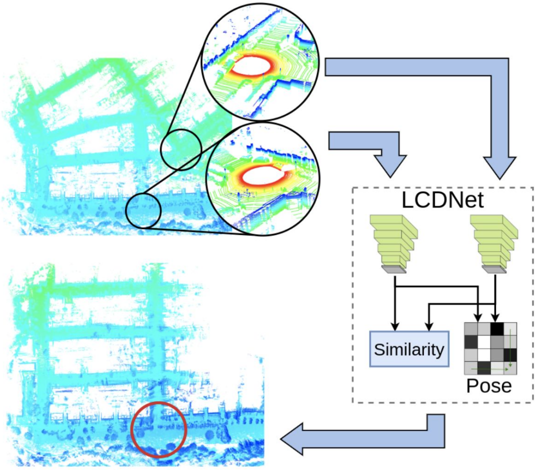


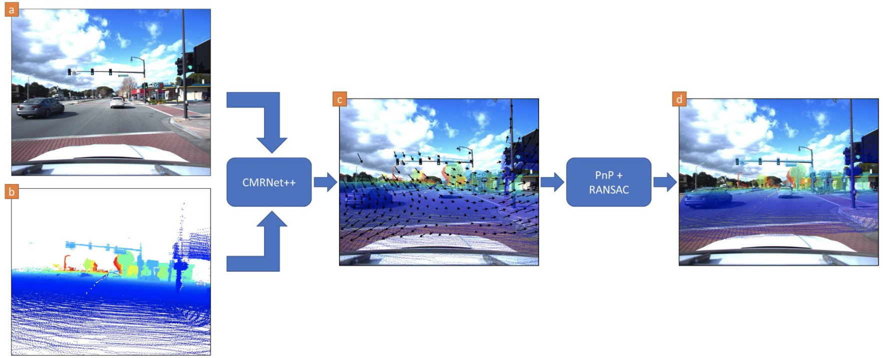
Related Papers:

- YAN, Zhiqiang, et al. RigNet: Repetitive image guided network for depth completion. In: *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXVII*. Cham: Springer Nature Switzerland, 2022. S. 214-230.
- Park, Jinsun, et al. "Non-local spatial propagation network for depth completion." *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16*. Springer International Publishing, 2020.
- Ma, Fangchang, Guilherme Venturelli Cavalheiro, and Sertac Karaman. "Self-supervised sparse-to-dense: Self-supervised depth completion from lidar and monocular camera." *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019.

Multimodal Learning for Images and LiDAR

Advisor: Simon Schaefer, Simon Boche



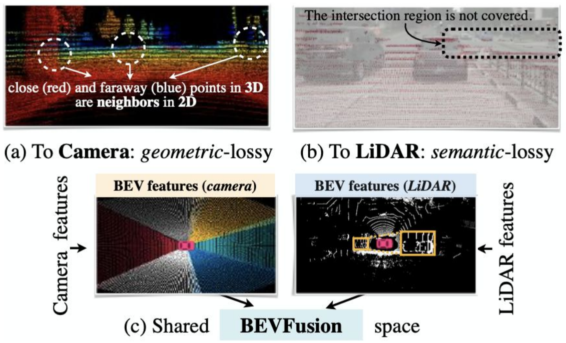


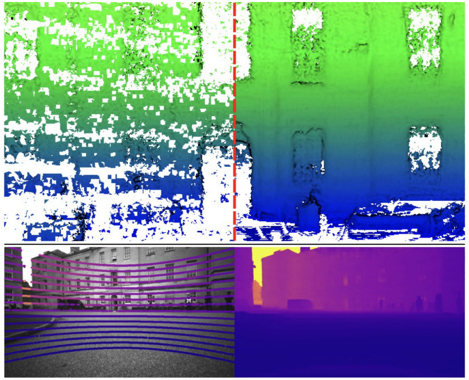
Related Papers:

- CMRNet: Camera to LiDAR-Map Registration, Cattaneo et al, 2019
- LCDNet: Deep Loop Closure Detection and Point Cloud Registration for LiDAR SLAM, Cattaneo et al, 2021

Learning-based multi-modal perception

Advisor: Dr. Jaehyung Jung





Related Papers:

- Zhijian Liu et al., "BEVFusion: Multi-task multi-sensor fusion with unified bird's-eye view representation," *ICRA 2023*.
- Yifu Tao, et al. "3D Lidar Reconstruction with Probabilistic Depth Completion for Robotic Navigation," *IROS 2022.*
- Yuxiang Sun et al., "RTFNet: RGB-thermal fusion network for semantic segmentation of urban scenes." *RAL 2019.*

Uncertainty-aware perception

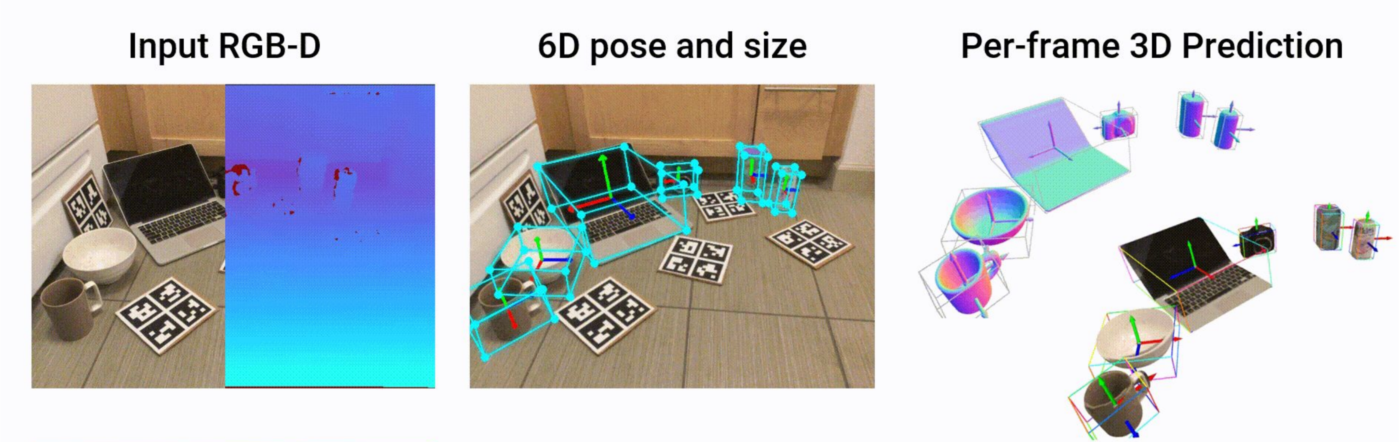
Advisor: Dr. Jaehyung Jung

<u>Related Papers:</u>

- Liyan Chen et al., "Learning the Distribution of Errors in Stereo Matching for Joint Disparity and Uncertainty Estimation," *CVPR 2023*.
- Gangwei Xu et al,. "Accurate and efficient stereo matching via attention concatenation volume," *TPAMI 2023.*
- Xuran Pan et al., "ActiveNeRF: Learning where to see with uncertainty estimation." *ECCV*, 2022.

Object-level Perception

Advisor: Hanzhi Chen

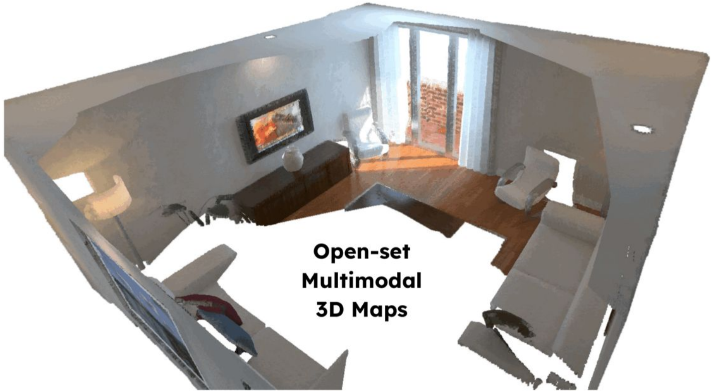


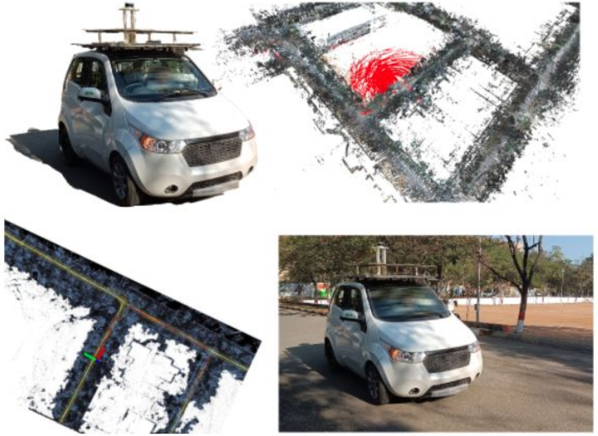
Related Papers:

- Irshad, Muhammad Zubair, et al. "Shapo: Implicit representations for multi-object shape, appearance, and pose optimization." ECCV 2022
- Landgraf, Zoe, et al. "Simstack: A generative shape and instance model for unordered object stacks." ICCV 2021
- Li, Guanglin, et al. "Generative Category-Level Shape and Pose Estimation with Semantic Primitives." CoRL 2022

Open Vocabulary 3D Scene Understanding

Advisor: Hanzhi Chen



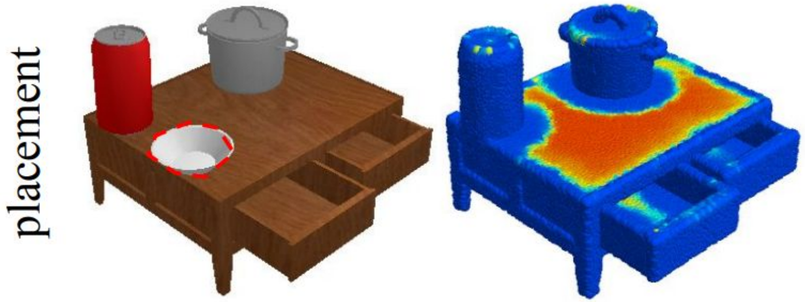


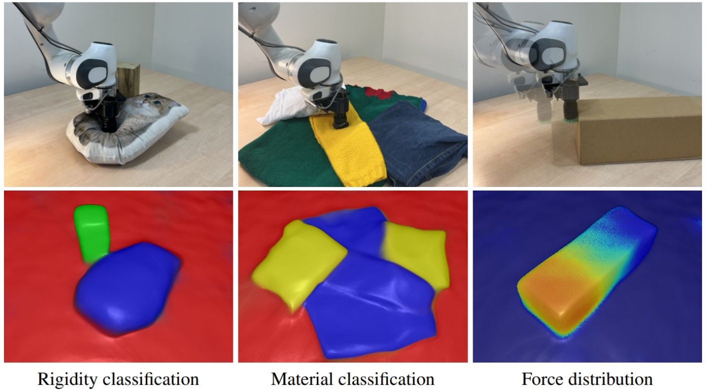
Related Papers:

- Peng, Songyou, et al. "3D Scene Understanding with Open Vocabularies." CVPR 2023.
- Jatavallabhula, Krishna Murthy, et al. "ConceptFusion: Open-set Multimodal 3D Mapping." RSS 2023
- Liu, Minghua, et"OpenShape: Scaling Up 3D Shape Representation Towards Open-World Understanding", Arxiv 2023

Mapping Beyond Geometry

Advisor: Hanzhi Chen





Related Papers:

- Mo, Kaichun, et al. "O2O-Afford: Annotation-free large-scale object-object affordance learning." CoRL 2022.
- Haughton, Iain, et al. "Real-time Mapping of Physical Scene Properties with an Autonomous Robot Experimenter." CoRL 2022

Computationally Efficient NeRFs

Advisor: Simon Schaefer, Hanzhi Chen

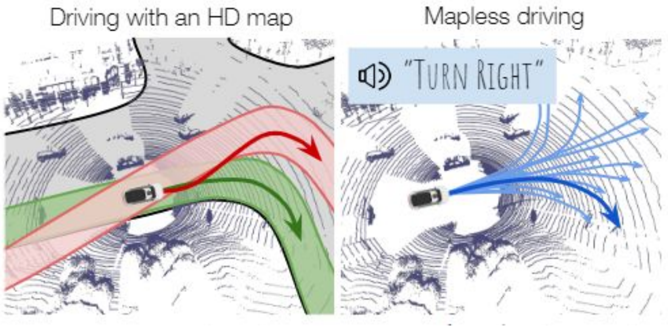
(a) None — 411k parameters, 10:45 (mm:ss)
(b) Multiresolution grid — 10k + 16.3M parameters, 1:26 (mm:ss)
(c) Frequency — 438k + 0 parameters, 13:53 (mm:ss)
(d) Hashtable ($T=2^{14}$) — 10k + 494k parameters, 1:40 (mm:ss)
(e) Hashtable ($T=2^{19}$) — 10k + 12.6M parameters, 1:45 (mm:ss)

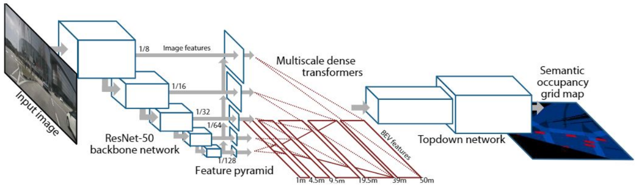
Related Papers:

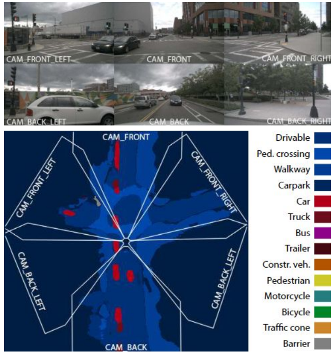
- Instant Neural Graphics Primitives with a Multiresolution Hash Encoding, Müller et al, 2022
- Plenoxels: Radiance Fields without Neural Networks, Fridovic-Keil et al, 2022

BEV Map Based Perception for Autonomous Driving

Advisor: Dr. Xingxing Zuo





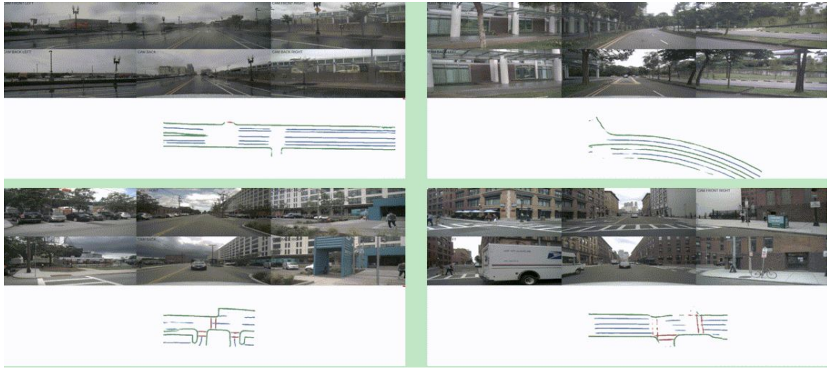


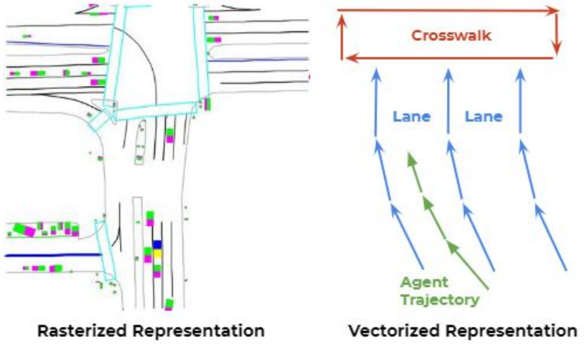
Related Papers:

- Roddick, Thomas, and Roberto Cipolla. "Predicting semantic map representations from images using pyramid occupancy networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.
- Ye, Maosheng, Shuangjie Xu, and Tongyi Cao. "Hvnet: Hybrid voxel network for lidar based 3d object detection." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.
- Casas, Sergio, Abbas Sadat, and Raquel Urtasun. "Mp3: A unified model to map, perceive, predict and plan." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.
- Saha, Avishkar, et al. "Translating images into maps." *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022.

Learning-based Vector Map Reconstruction for Autonomous Driving

Advisor: Dr. Xingxing Zuo



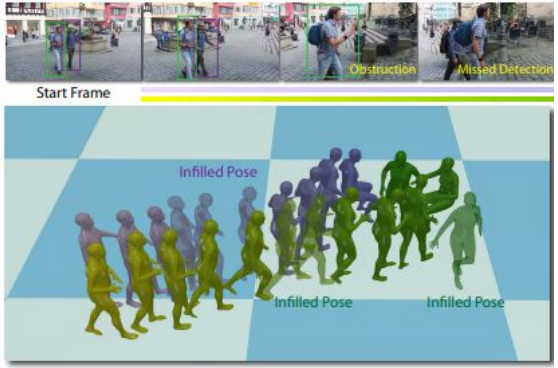


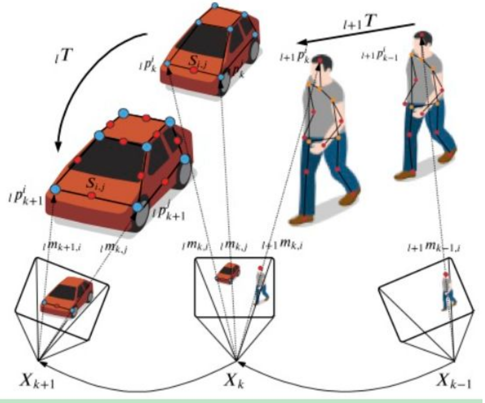
Related Papers:

- Gao, Jiyang, et al. "Vectornet: Encoding hd maps and agent dynamics from vectorized representation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.
- Liu, Yicheng, et al. "VectorMapNet: End-to-end Vectorized HD Map Learning." *arXiv preprint arXiv:2206.08920* (2022).
- Li, Qi, et al. "Hdmapnet: An online hd map construction and evaluation framework." *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022.

Human Pose Tracking with Dynamic Camera

Advisor: Simon Schaefer



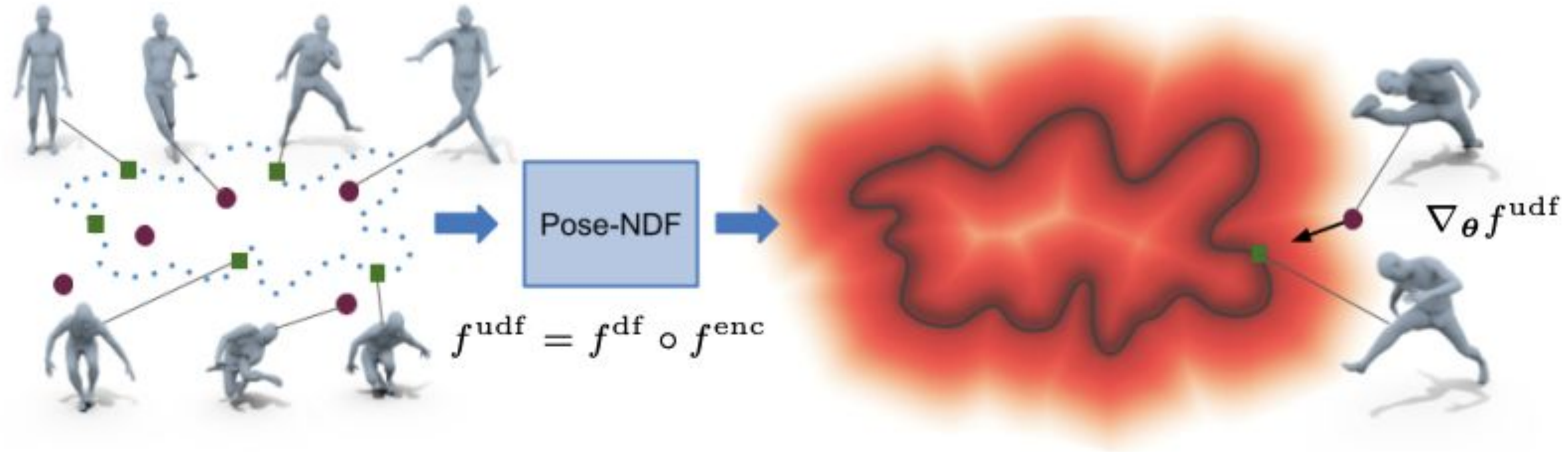


Related Papers:

- Yuan, Ye, et al. "GLAMR: Global occlusion-aware human mesh recovery with dynamic cameras." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.
- Qiu, Yuheng, et al. "Airdos: Dynamic slam benefits from articulated objects." *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022.

Learning the Human Distribution

Advisor: Simon Schaefer

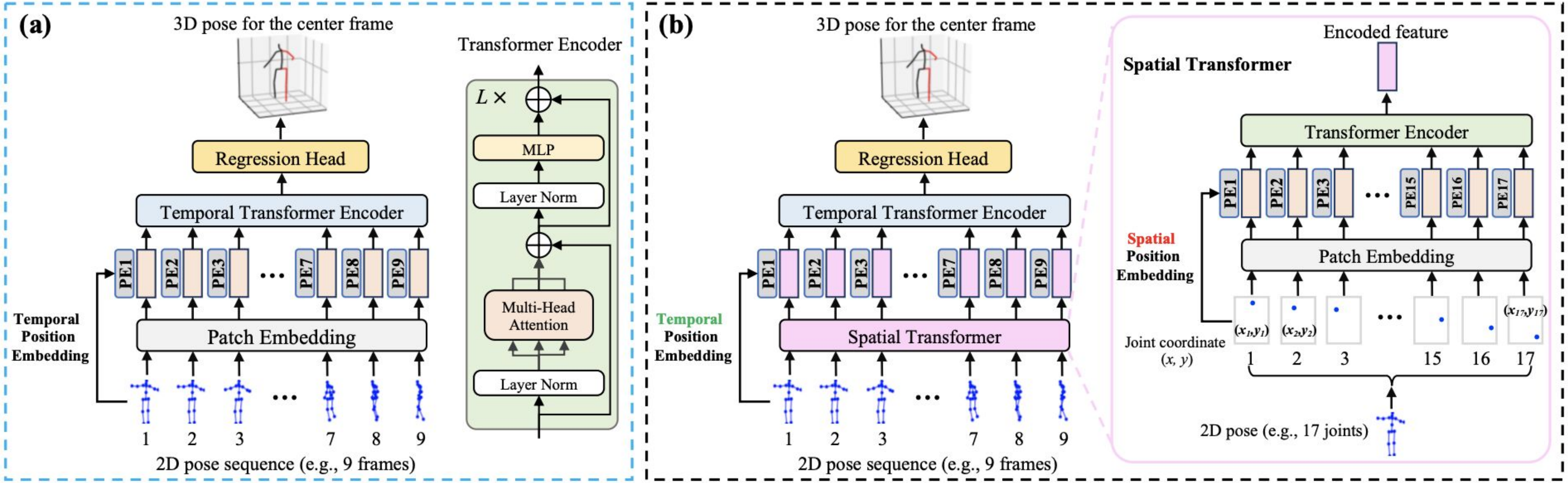


Related Papers:

- VPoser: Variational Human Pose Prior for Body Inverse Kinematics, Pavlakos et al, 2019
- Pose-NDF: Modeling Human Pose Manifolds with Neural Distance Fields, Tiwari et al, 2022

3D Human Pose Estimation

Advisor: Simon Schaefer

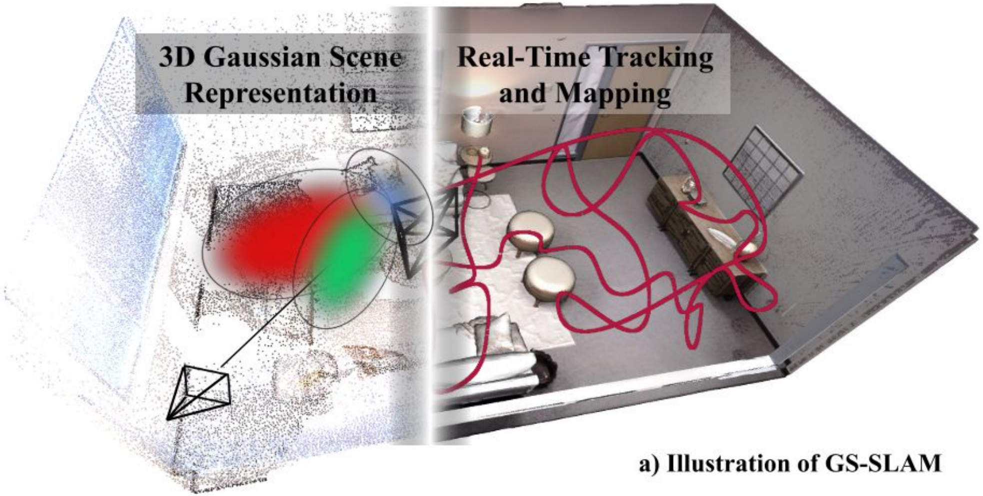


Related Papers:

- 3D Human Pose Estimation for Free-from and Moving Activities Using WiFi, Ren et al, 2022
- 3D Human Pose Estimation with Spatial and Temporal Transformers, Zheng et al, 2021

Gaussian Splatting SLAM

Advisor: Jiaxin Wei



a) Illustration of GS-SLAM

Related Papers:

- GS-SLAM: Dense Visual SLAM with 3D Gaussian Splatting
- SplaTAM: Splat, Track & Map 3D Gaussians for Dense RGB-D SLAM
- Gaussian Splatting SLAM
- Gaussian-SLAM: Photo-realistic Dense SLAM with Gaussian Splatting

Gaussian Splatting for Dynamic Scene Reconstruction

Advisor: Jiaxin Wei

Related Papers:

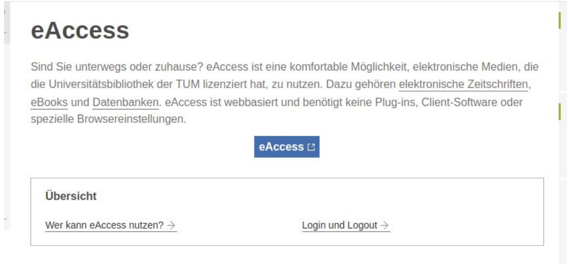
- Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis
- 4D Gaussian Splatting for Real-Time Dynamic Scene Rendering
- Deformable 3D Gaussians for High-Fidelity Monocular Dynamic Scene Reconstruction

Where can I find the papers?

arxiv.org

IEEE Xplore

TUM eAccess



sign in with your TUM account

Outline

- General Information
  - About the seminar
  - Registration
- Topics
- Questions

Questions?

- Web page: https://srl.cit.tum.de/teaching/s24/seminar_rpi
- Password: rpi_s24
- **Can I present another topic?** You can also suggest a topic / paper that you are interested in! If you have a topic in mind, that you are interested in and that is not in the list, we are always open for suggestions. In that case, attach it to your three favorite papers and we will decide whether it fits.