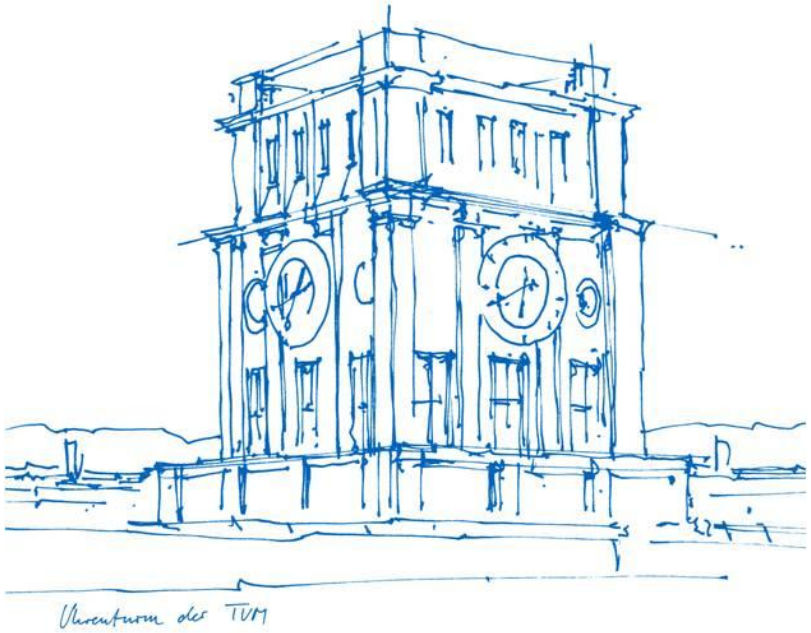
Multi-Modal Perception for Mobile Robotics

Simon Boche, Dr. Xingxing Zuo

Smart Robotics Lab

Technical University of Munich

WS 2022 / 23



Outline

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

How is the seminar organized?

- Slides / Material: seminar webpage
  - https://mlr.in.tum.de/teaching/w22/seminar_mmpmr
  - Password: mmpmr_w22 − Material page will go online soon

- Questions / Meeting arrangement: contact organizers
  - mmpmr-ws22@srl.in.tum.de

How is the seminar organized?

- Seminar meetings: talks and discussion
  - Time: Tuesdays, 10:00-12:00
  - Room: MI 02.09.023
  - Starting date: TBA (web page)
  - Number of meetings: TBA
  - **Attendance is mandatory!**

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

What about the presentation?

- General setup:
  - Duration: 20-25 minutes talk + 10-15 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
  - 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 mmpmr-ws22@srl.in.tum.de
  - 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 “Multi-Modal Perception for Mobile Robotics”

- Step 2: Personal registration via email
  - In the list of topics, select your **three** favorites
  - Write an email ranking these three favorites to mmpmr-ws22@srl.in.tum.de
  - Email subject: “MMPMR 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: July 27, 2022

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

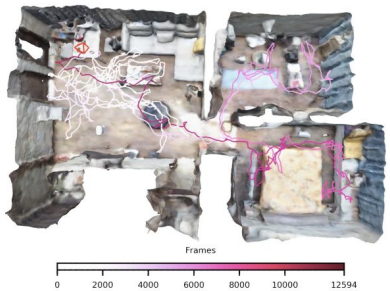
- Topic assignment
  - Topics are assigned after the participant list is finalized
  - We give our best to accommodate your preference list in the assignment

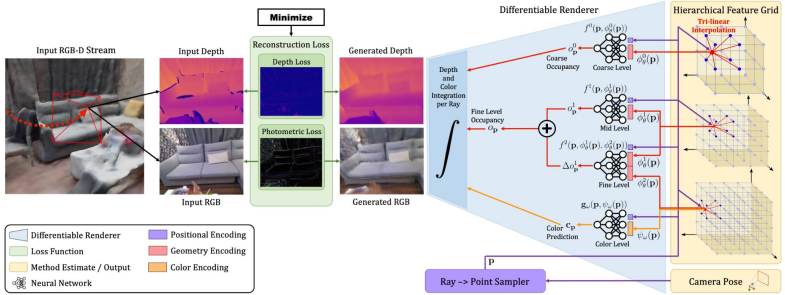
Outline

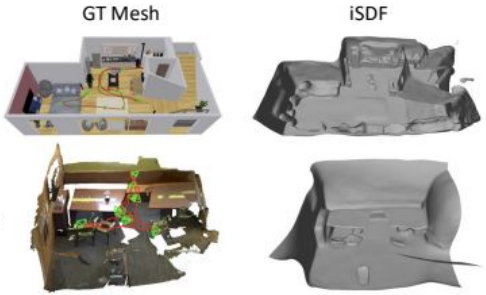
- General Information
  - About the seminar
  - Registration
- Topics
- Questions

Neural Implicit Representations for Robotic Perception

Advisor: Simon Boche





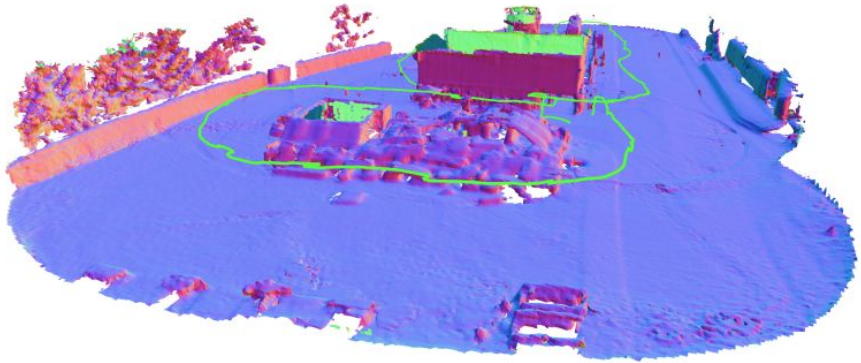


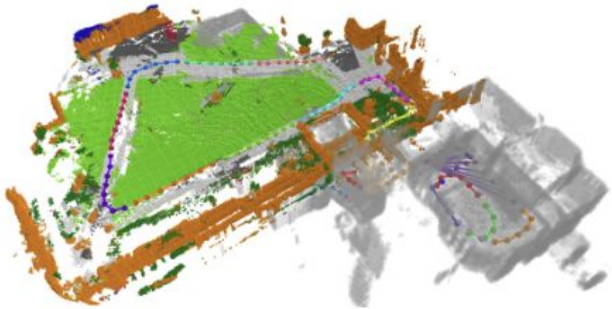
Related Papers:

- Ortiz, Joseph, et al. "iSDF: Real-Time Neural Signed Distance Fields for Robot Perception." *arXiv preprint arXiv:2204.02296* (2022).
- Zhu, Zihan, et al. "Nice-slam: Neural implicit scalable encoding for slam." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.
- Sucar, Edgar, et al. "iMAP: Implicit mapping and positioning in real-time." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Large-Scale LiDAR Reconstruction

Advisor: Simon Boche



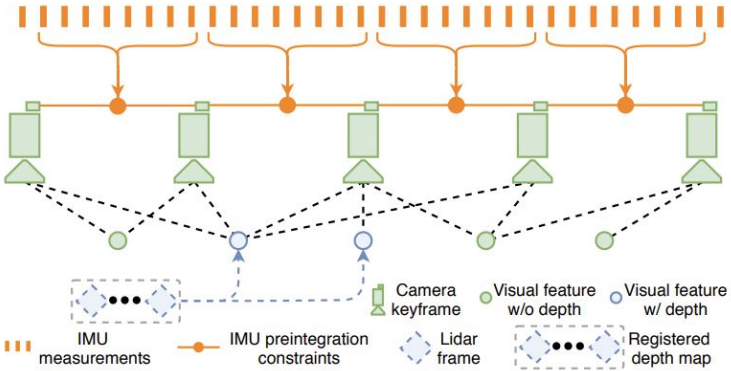


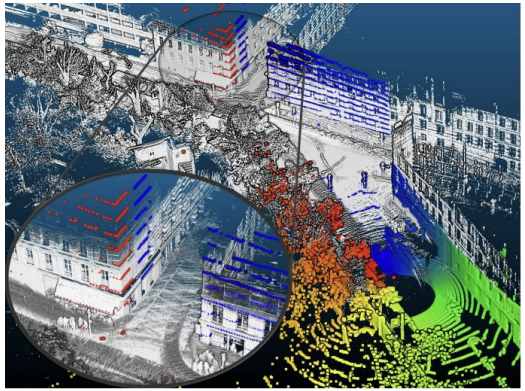
Related Papers:

- 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
- Reijgwart, Victor, et al. "Voxgraph: Globally consistent, volumetric mapping using signed distance function submaps." *IEEE Robotics and Automation Letters* 5.1 (2019): 227-234.

LiDAR-based Odometry / SLAM

Advisor: Simon Boche



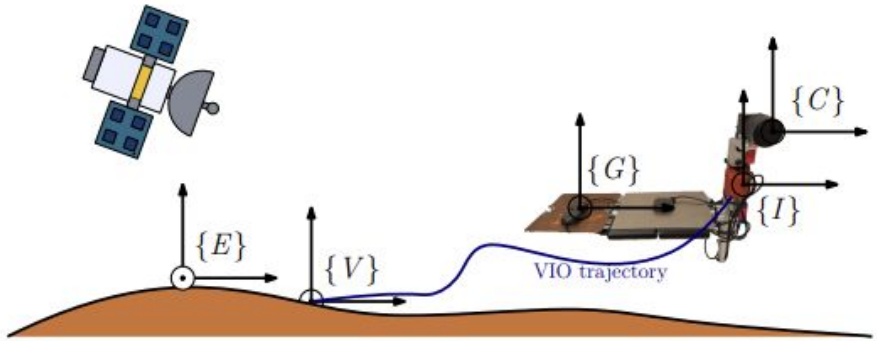


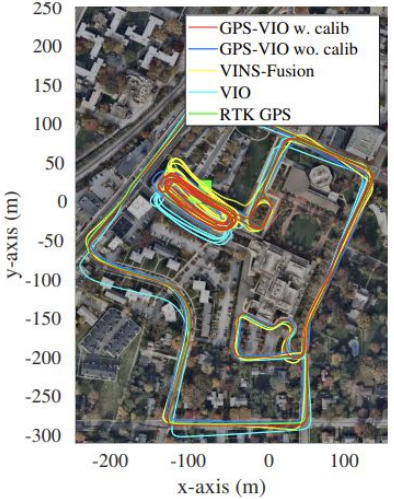
Related Papers:

- Shan, Tixiao, et al. "Lvi-sam: Tightly-coupled lidar-visual-inertial odometry via smoothing and mapping." *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2021.
- Dellenbach, Pierre, et al. "Ct-icp: Real-time elastic lidar odometry with loop closure." *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022.

GPS-aided Visual-Inertial Odometry

Advisor: Simon Boche



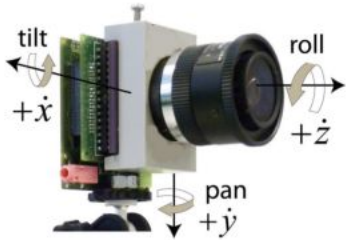


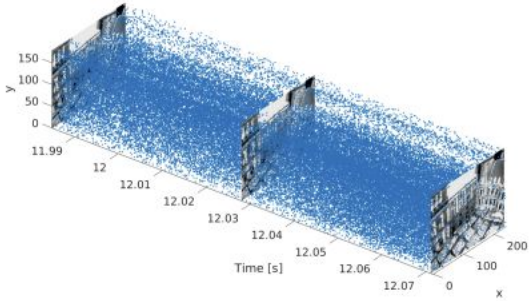
Related Papers:

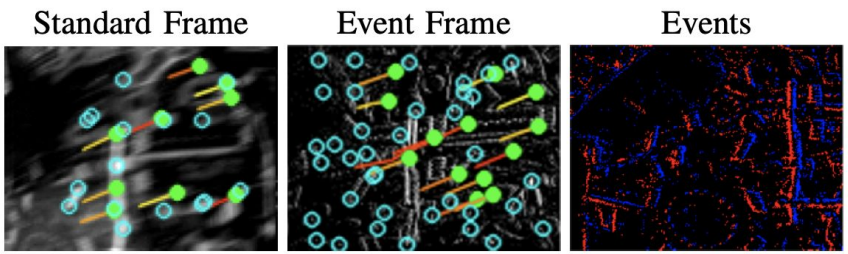
- Lee, Woosik, et al. "Tightly-coupled GNSS-aided Visual-Inertial Localization." *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022.
- Lee, Woosik, et al. "Intermittent gps-aided vio: Online initialization and calibration." *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020.

Event-based Odometry / SLAM

Advisor: Simon Boche





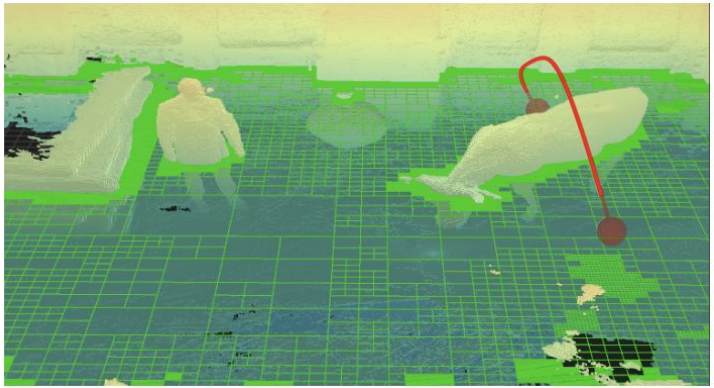


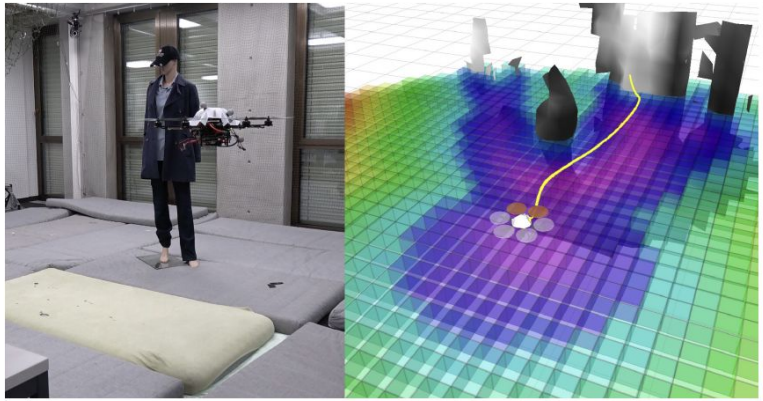
Related Papers:

- Vidal, Antoni Rosinol, et al. "Ultimate SLAM? Combining events, images, and IMU for robust visual SLAM in HDR and high-speed scenarios." *IEEE Robotics and Automation Letters* 3.2 (2018): 994-1001.
- Mueggler, Elias, et al. "Continuous-time visual-inertial odometry for event cameras." *IEEE Transactions on Robotics* 34.6 (2018): 1425-1440.

Real-Time Volumetric Mapping

Advisor: Simon Boche



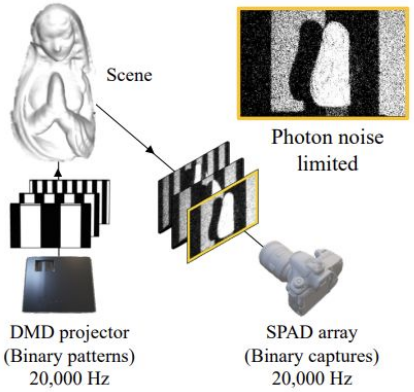


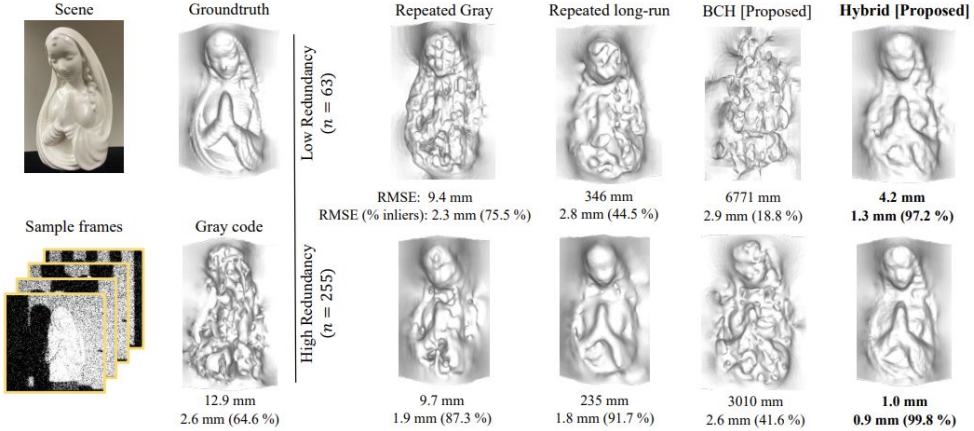
Related Papers:

- 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.
- Oleynikova, Helen, et al. "Voxblox: Incremental 3d euclidean signed distance fields for on-board mav planning." *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.

Single Photon Imaging

Advisor: Simon Boche



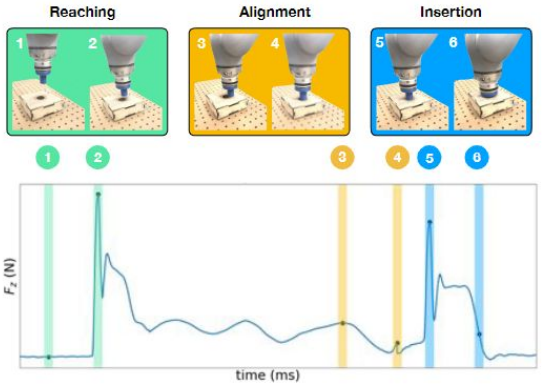


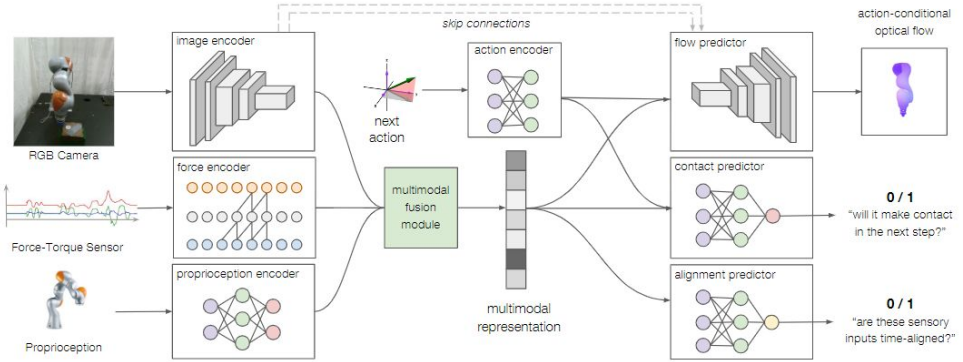
Related Papers:

- Sundar, Varun, et al. "Single-Photon Structured Light." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Vision and Touch Fusion for Perception and Manipulation

Advisor: Dr. Xingxing Zuo



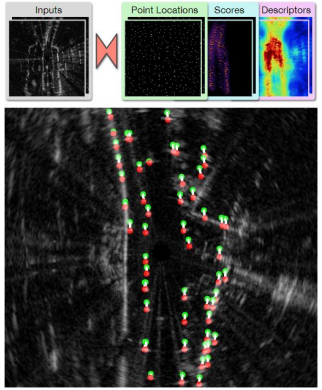


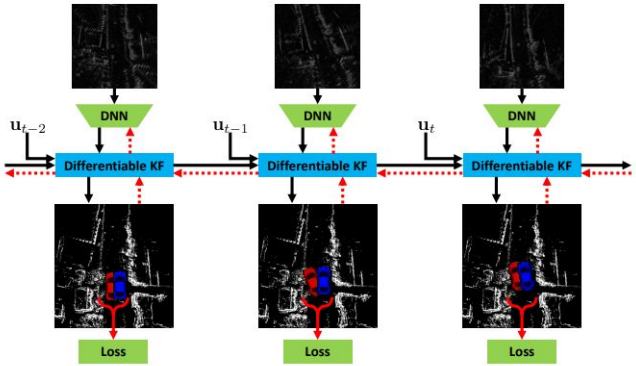
Related Papers:

- Lee, Michelle A., et al. "Making sense of vision and touch: Learning multimodal representations for contact-rich tasks." *IEEE Transactions on Robotics* 36.3 (2020): 582-596.

Radar Perception

Advisor: Dr. Xingxing Zuo



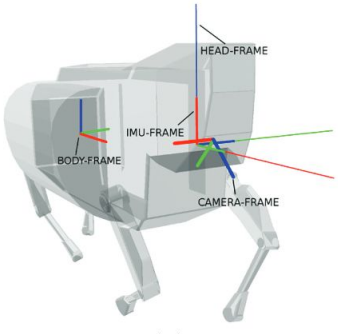


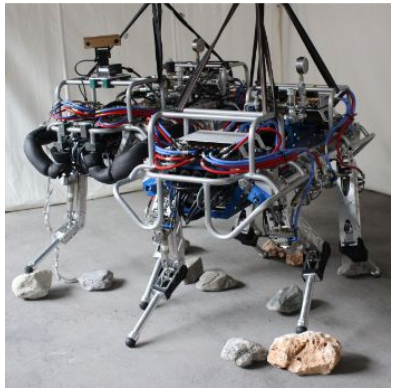
Related Papers:

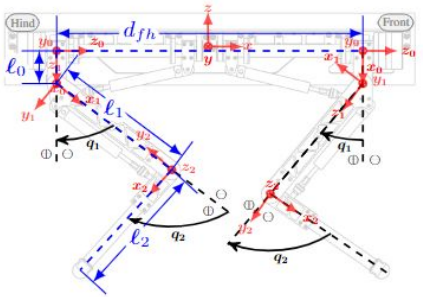
- Barnes, Dan, and Ingmar Posner. "Under the radar: Learning to predict robust keypoints for odometry estimation and metric localisation in radar." *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020.
- Yin, Huan, et al. "Rall: end-to-end radar localization on lidar map using differentiable measurement model." *IEEE Transactions on Intelligent Transportation Systems* (2021).

State Estimation for Quadrupedal Robot

Advisor: Dr. Xingxing Zuo





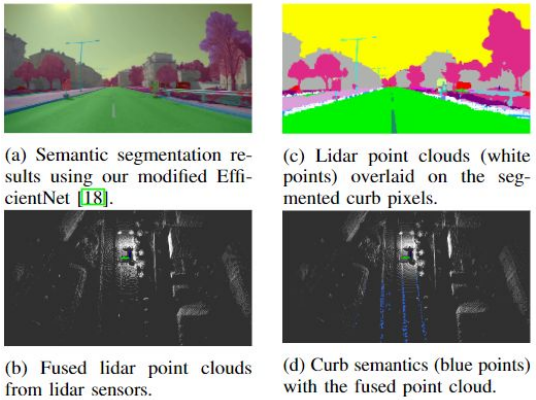


Related Papers:

- Camurri, Marco, et al. "Probabilistic contact estimation and impact detection for state estimation of quadruped robots." *IEEE Robotics and Automation Letters* 2.2 (2017): 1023-1030.
- Fink, Geoff, and Claudio Semini. "Proprioceptive sensor fusion for quadruped robot state estimation." *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020.
- Ma, Jeremy, et al. "Real-time pose estimation of a dynamic quadruped in GPS-denied environments for 24-hour operation." *The International Journal of Robotics Research* 35.6 (2016): 631-653.

Learning-based Multi-modal Perception

Advisor: Dr. Xingxing Zuo

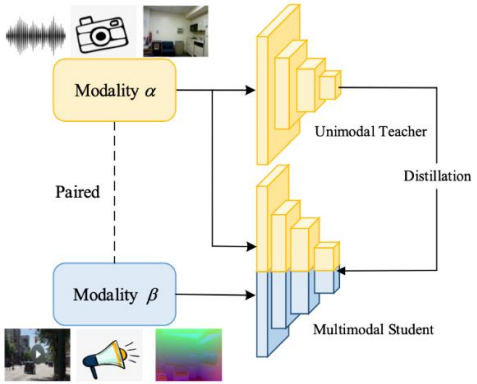


(a) Semantic segmentation results using our modified EfficientNet [18].

(c) Lidar point clouds (white points) overlaid on the segmented curb pixels.

(b) Fused lidar point clouds from lidar sensors.

(d) Curb semantics (blue points) with the fused point cloud.

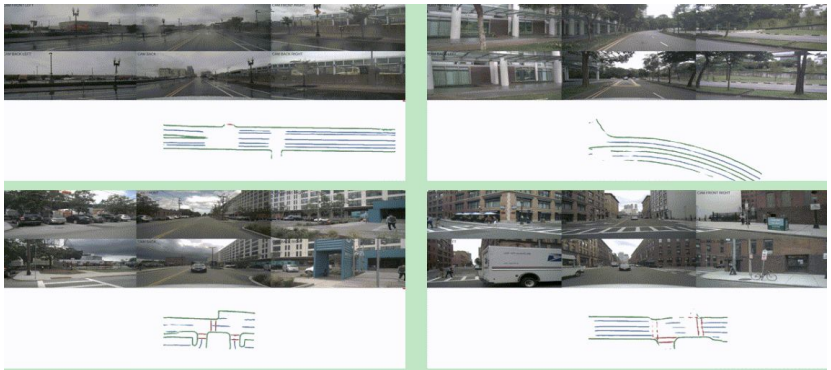


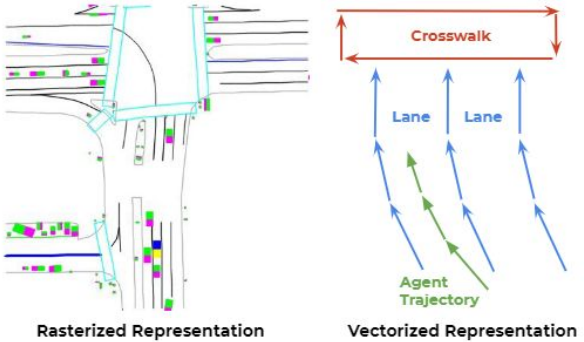
Related Papers:

- Das, Sandipan, et al. "Multi-modal curb detection and filtering." *arXiv preprint arXiv:2205.07096* (2022).
- Xue, Zihui, et al. "Multimodal knowledge expansion." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.
- Liu, Huayao, et al. "CMX: Cross-Modal Fusion for RGB-X Semantic Segmentation with Transformers." *arXiv preprint arXiv:2203.04838* (2022).

Learning-based Vector Map Reconstruction for Autonomous Driving

Advisor: Dr. Xingxing Zuo



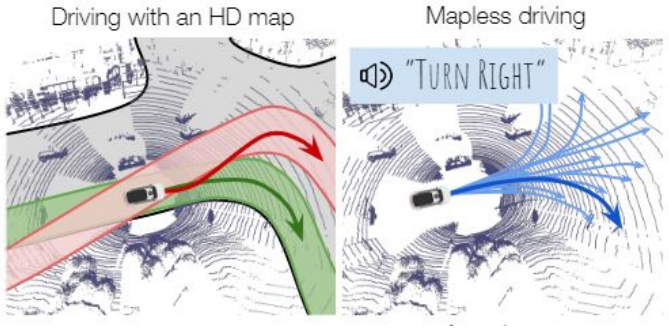


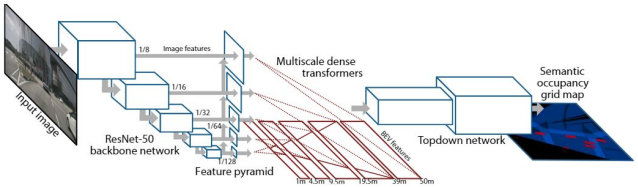
Related Papers:

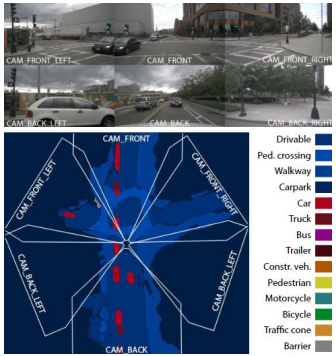
- Gao, Jiyang, et al. "Vectornet: Encoding hd maps and agent dynamics from vectorized representation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.
- Liu, Yicheng, et al. "VectorMapNet: End-to-end Vectorized HD Map Learning." *arXiv preprint arXiv:2206.08920*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

BEV Map Based Perception for Autonomous Driving

Advisor: Dr. Xingxing Zuo





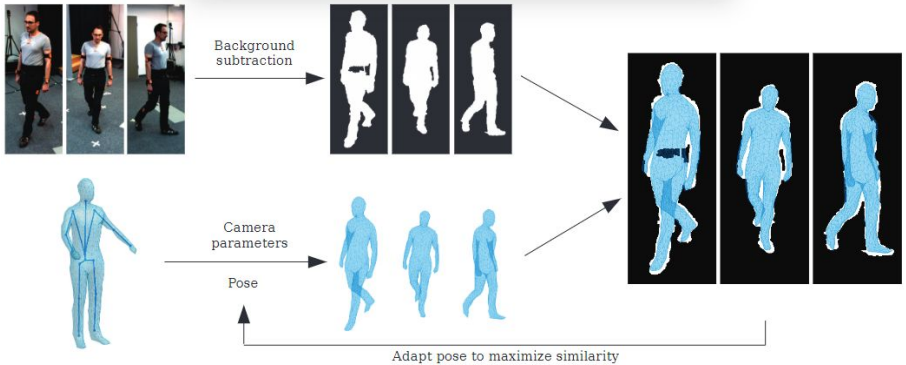


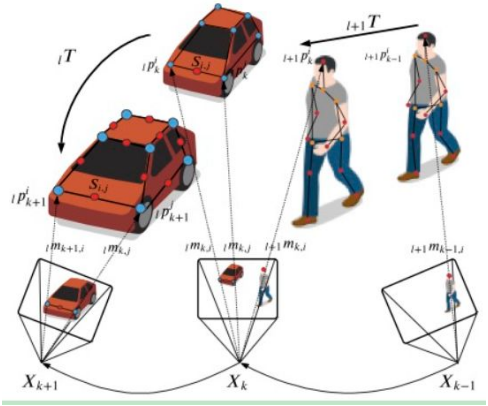
Related Papers:

- 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.

Dynamic Human Pose Tracking

Advisor: Dr. Xingxing Zuo





Related Papers:

- Yu, Ri, Hwangpil Park, and Jehee Lee. "Human dynamics from monocular video with dynamic camera movements." *ACM Transactions on Graphics (TOG)* 40.6 (2021): 1-14.
- Von Marcard, Timo, Gerard Pons-Moll, and Bodo Rosenhahn. "Human pose estimation from video and imus." *IEEE transactions on pattern analysis and machine intelligence* 38.8 (2016): 1533-1547.
- Qiu, Yuheng, et al. "Airdos: Dynamic slam benefits from articulated objects." *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022.

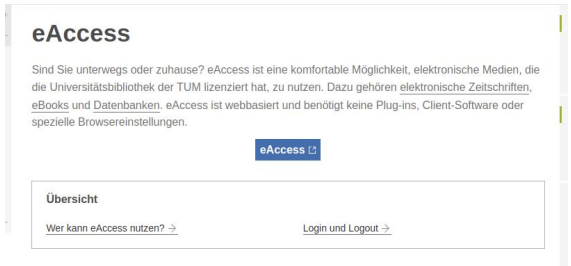
Where can I find the papers?

arxiv.org

IEEE Xplore

sign in with your TUM account

TUM eAccess



Questions?

- Web page: https://mlr.in.tum.de/teaching/w22/seminar_mmpmr
- Password: mmpmr_w22
- Contact: mmpmr-ws22@srl.in.tum.de
- **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.