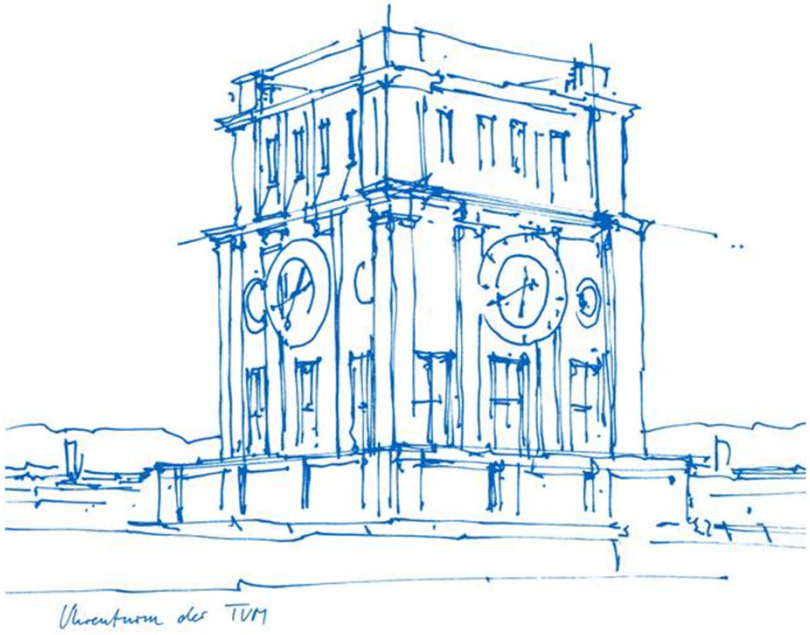
Deep-learning-based 3D scene representation and understanding for mobile robots

Dr. Xingxing Zuo

Smart Robotics Lab

Technical University of Munich

SS 2022



Outline

- General Information
  - About the seminar
  - Registration
- Topics/Papers
  - 3D Scene Representation
  - 3D Scene Understanding
  - ……
- Questions

Outline

- General Information
  - About the seminar
  - Registration
- Topics/Papers
  - 3D Scene Representation
  - 3D Scene Understanding
  - ……
- Questions

How is the seminar organized?

- Slides / Material: seminar webpage
  - https://mlr.in.tum.de/teaching/s22/seminar_dlr
  - Password: dlr_ss22 − Material page will be posted on the webpage!

- Questions / Meeting arrangement: contact organizers
  - xingxing.zuo@tum.de
  - Email subject: “DLR seminar – xx”

How is the block seminar organized?

- Seminar Days: talks and discussion
  - Time: 14:00 - 17:00, Monday, **May. 30**, 2022
    09:00 - 12:00, 14:00 - 17:00, Tuesday, **May. 31**, 2022
  - Room: MI 02.07.014 (Monday); MI 01.13.007 (Tuesday); or online with Zoom
  - Kick-off meeting date: TBA (webpage), online with Zoom
  - **Attendance is mandatory!**

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

What about the presentation?

- General setup:
  - Duration: 20-30 minutes talk + 5-10 minutes discussion
  - Focus on a topic with 1-3 relevant papers!
  - Make sure to **finish on time** **(maximum 40 mins)**!
  - Rule of thumb: 1-2 minutes per slide → 10-25 slides
  - Do not put too much text on a single slides!
- Recommended structure (talk only):
  - Introduction
  - Overview / Outline
  - Methodologies and main experimental findings
  - Personal comments
  - Summary and Future work

What about the final report?

- General setup:
  - Use LATEX template provided on course webpage
  - Review and comment on **more papers** relevant to this topic/paper
  - Length: 5-6 pages in English
  - Use some figures or tables for illustration
  - Send final report as pdf by email to xingxing.zuo@tum.de
  - Submission deadline: **two weeks after talk**
- Recommended structure:
  - Introduction
  - Related work
  - Methodologies and main experimental findings
  - Discussion of results
  - Summary and Future work (important)

Summary: how will the seminar be graded?

- The final grade will be based on
  - Presentation
  - Final Report and slides
  - Contributions to seminar discussions

    ⇒ Ask questions! There are no stupid questions!

  - Understanding of the topic and deep thinking of the future research

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 “Deep-learning-based 3D scene representation and understanding for mobile robots”

- Step 2: Personal registration via email
  - In the list of topics/papers on the webpage, select your **three** favorites
  - Write an email ranking three favorites (the first with the highest priority) to xingxing.zuo@tum.de
  - Email subject: “DLR seminar - application [your name]”
  - Include information about related lectures / courses (**transcript**) you have taken so far.
  - **Registrations without Emails or with missing information will be ignored**!

- Deadline for both registrations: February 15, 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 (February 24, 2022)

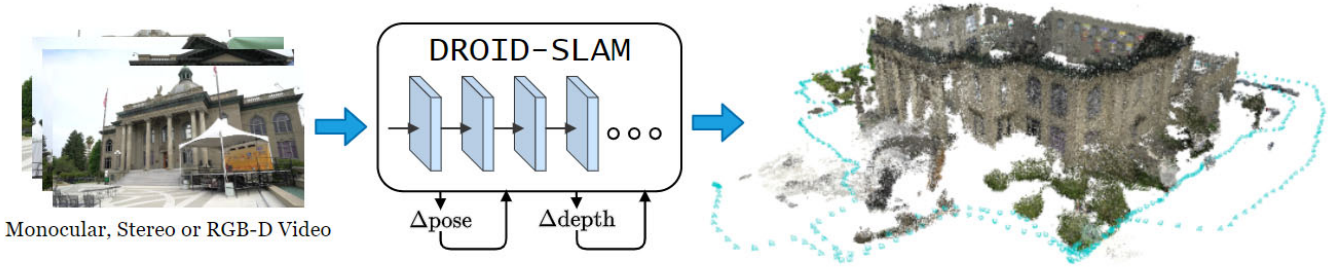
- Topics/papers assignment
  - Topics/papers are assigned after the participant list is finalized
  - We suggest only one paper as the start point, and you are supposed to find more relevant papers!
  - We give our best to accommodate your preference list in the assignment

Outline

- General Information
  - About the seminar
  - Registration
- Topics/Papers
  - 3D Scene Representation
  - 3D Scene Understanding
  - ……
- Questions

DROID-SLAM: Deep Visual SLAM for Monocular, Stereo, and RGB-D Cameras

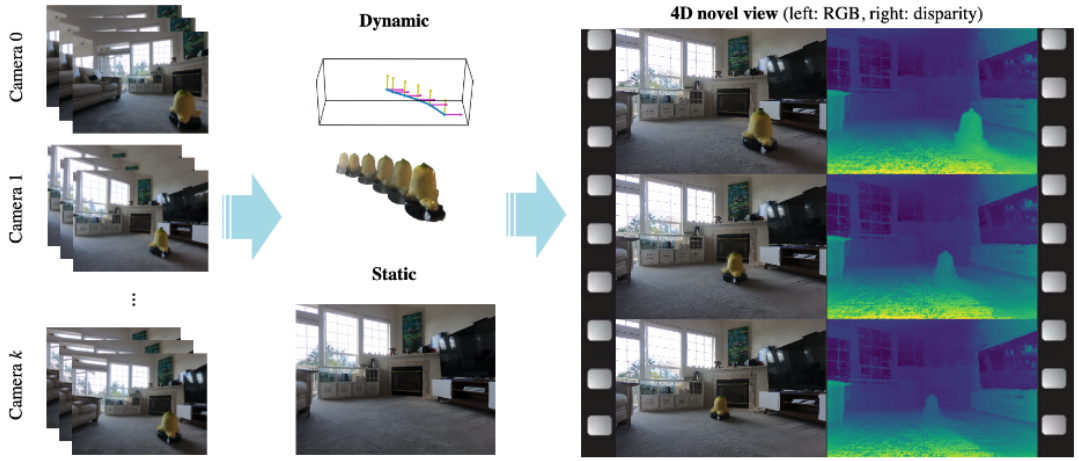
Deep learning based SLAM system， recurrent neural network, end-to-end neural architecture for visual SLAM， differentiable Dense Bundle Adjustment， outperforms many classic/prior works by large margins



Teed, Zachary and Deng, Jia (Advances in neural information processing systems, 2021)

STaR: Self-supervised Tracking and Reconstruction of Rigid Objects in Motion with Neural Rendering

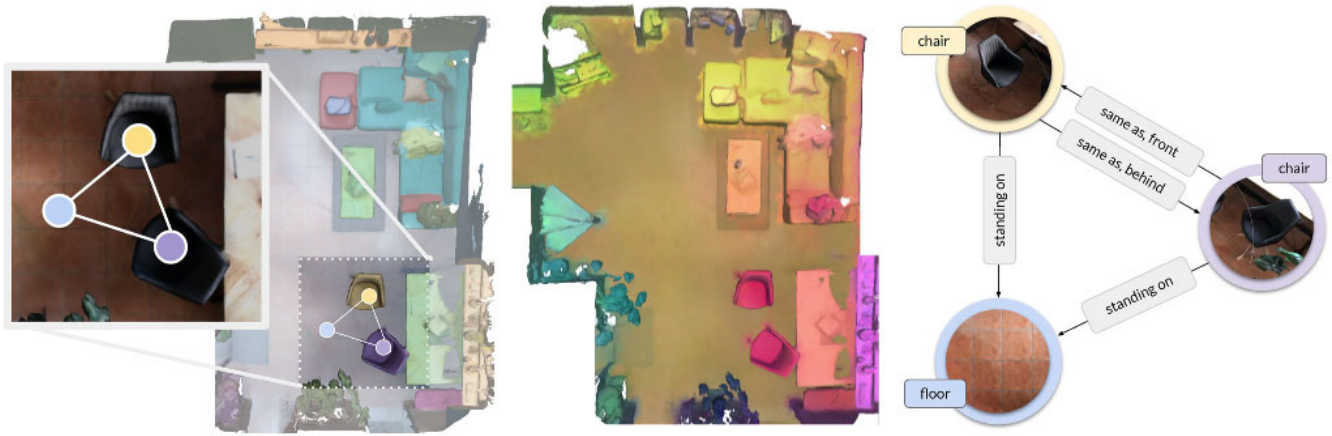
Neural Radiance Fields (NeRF), dynamic objects representation, neural volume rendering, reconstruction of dynamic scene, new view synthesis



Wentao Yuan, Zhaoyang Lv, Tanner Schmidt, Steven Lovegrove (CVPR, 2021)

Learning 3D Semantic Scene Graphs with Instance Embeddings

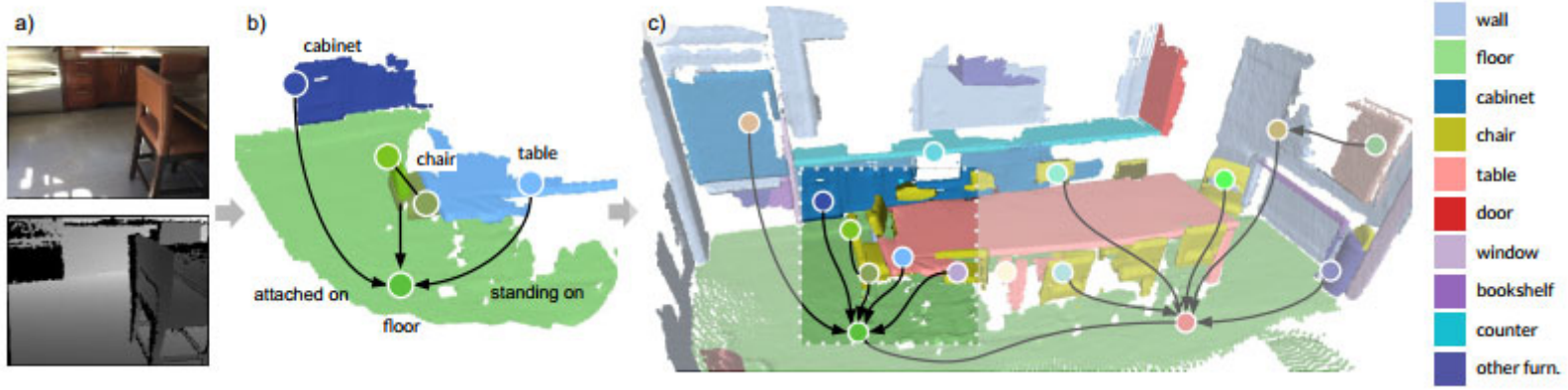
Scene graphs, 3D scene understanding, Semantic segmentation, Scene context, Instance embeddings alongside a scene segmentation



Johanna Wald, Nassir Navab, Federico Tombari (IJCV, 2021)

SceneGraphFusion: Incremental 3D Scene Graph Prediction from RGB-D Sequences

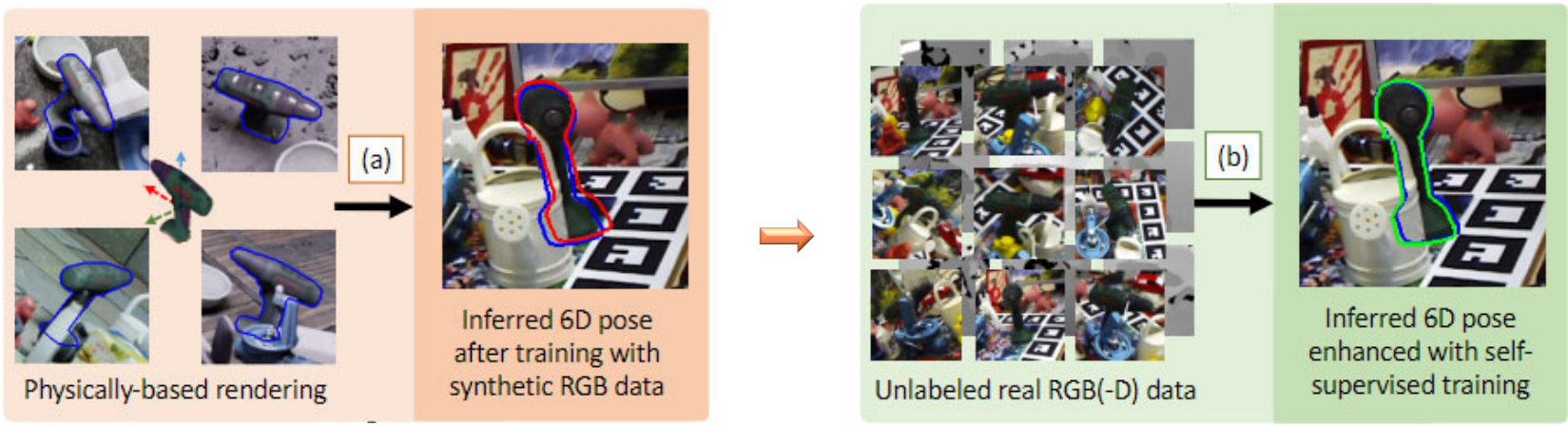
Scene graphs, 3D scene understanding, Semantic segmentation, Attention mechanism



S. Wu, J. Wald, K. Tateno, N. Navab, F. Tombari, (CVPR, 2021)

Occlusion-Aware Self-Supervised Monocular 6D Object Pose Estimation

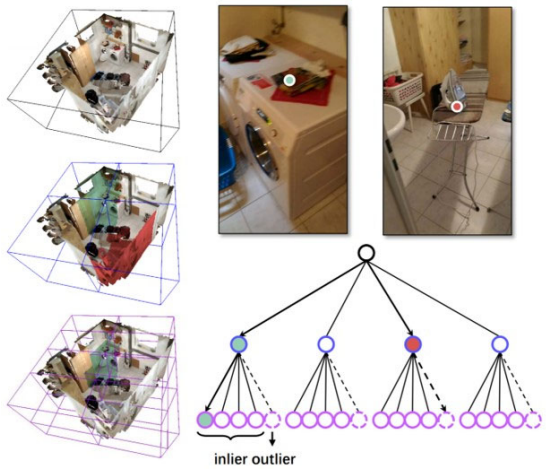
6D Object Pose Estimation, Self-Supervised Learning, Differentiable Rendering, Domain Adaptation

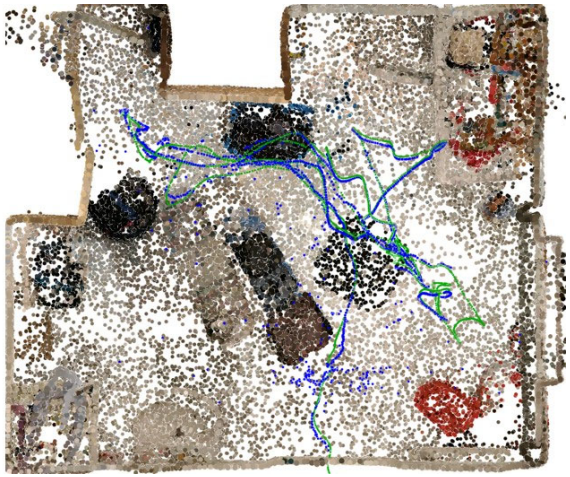


Gu Wang, Fabian Manhardt, Xingyu Liu, Xiangyang Ji, Federico Tombari (T-PAMI, 2021)

Robust Neural Routing Through Space Partitions for Camera Relocalization in Dynamic Indoor Environments

Dynamic scenes, Robust localization, Outlier aware neural tree

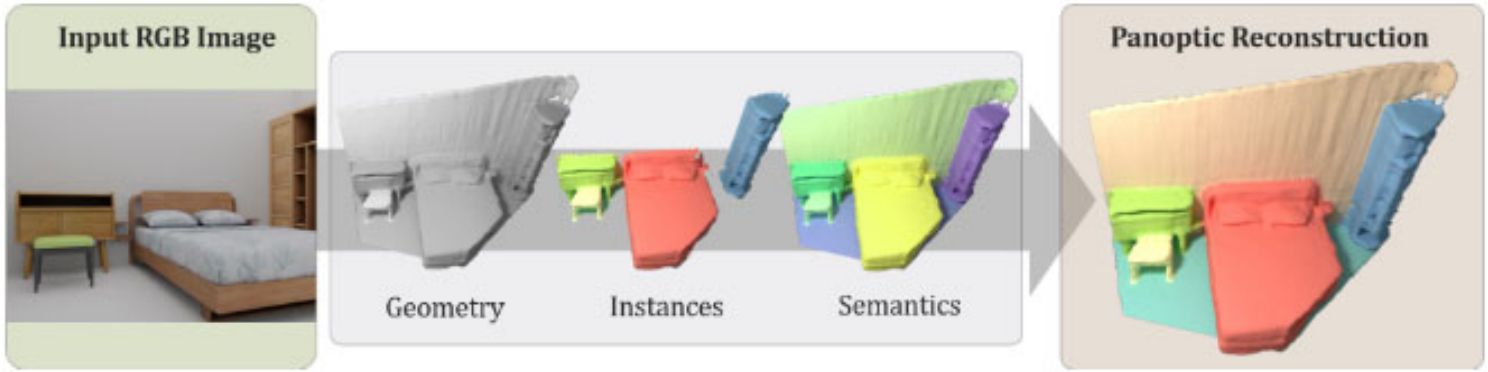




Siyan Dong, *Qingnan Fan, He Wang, Ji Shi, Li Yi, Thomas Funkhouser, Baoquan Chen, Leonidas J. Guibas (CVPR, 2021)

Panoptic 3D Scene Reconstruction From a Single RGB Image

3D Reconstruction, 3D instance segmentation，3D volumetric scene representation



Manuel Dahnert Ji Hou Matthias Nießner Angela Dai (NeurIPS, 2021)

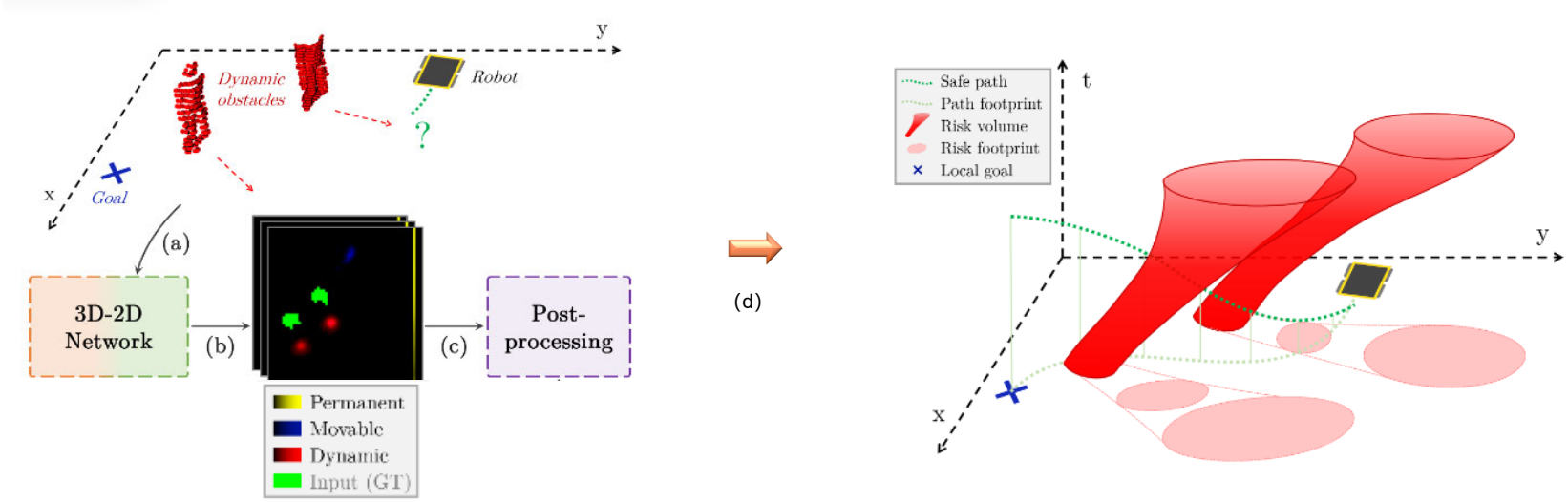
Learning To Recover 3D Scene Shape From a Single Image

3D Reconstruction, 3D scene shape, Depth prediction, Point-cloud networks

Wei Yin, Jianming Zhang, Oliver Wang, Simon Niklaus, Long Mai, Simon Chen, Chunhua Shen (CVPR, 2021)

Learning Spatiotemporal Occupancy Grid Maps for Lifelong Navigation in Dynamic Scenes

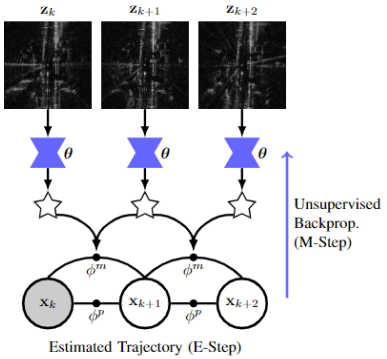
Dynamic scene representation, Self-supervised learning, Lifelong learning, Forecasting the motion of dynamic obstacles

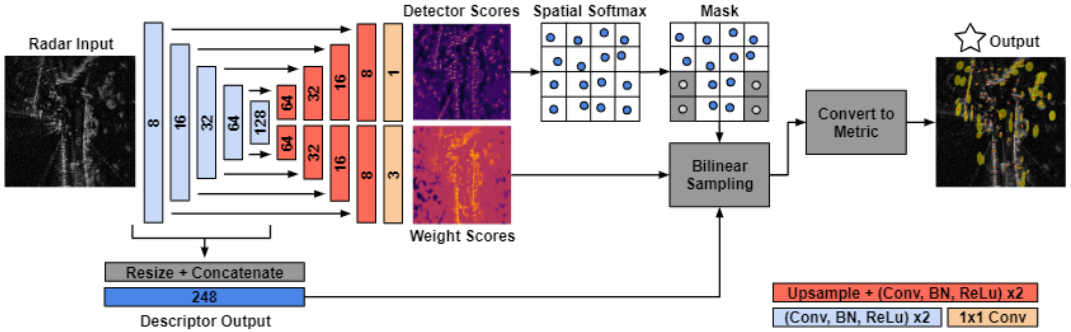


Hugues Thomas, Matthieu Gallet de Saint Aurin, Jian Zhang, Timothy D Barfoot (RAL, 2021)

Radar Odometry Combining Probabilistic Estimation and Unsupervised Feature Learning

Radar odometry, Robust localization, Feature learning, Probabilistic inference, Exactly Sparse Gaussian Variational Inference





Keenan Burnett, David J Yoon, Angela P Schoellig, Timothy D Barfoot (Arxiv, 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.

Object Goal Navigation using Goal-Oriented Semantic Exploration

Semantic exploration, Autonomous navigation, Long-term planning

Object Goal: dining table

Semantic Scene Understanding

Object detection

Learning Semantic Priors

Where is 'dining table' more likely to be found?

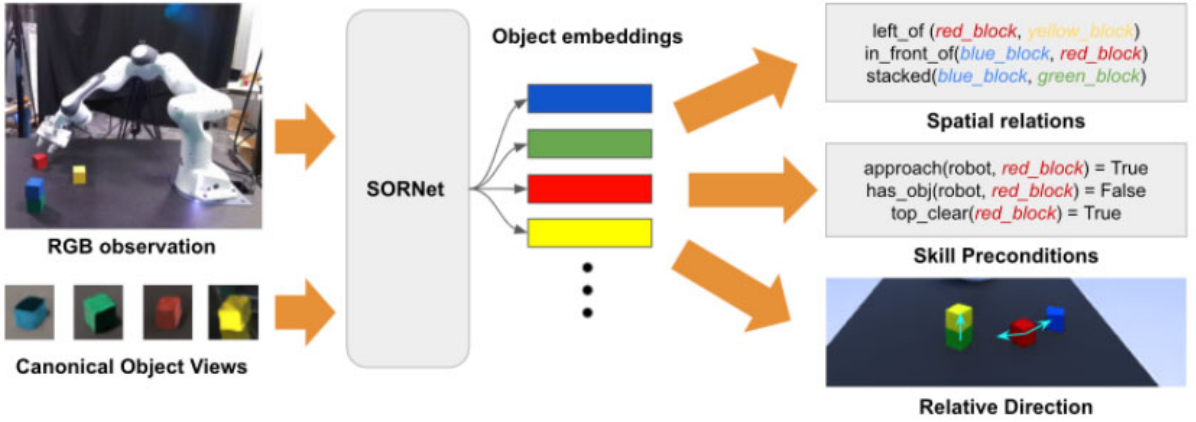
Episodic Memory

Keeping track of explored and unexplored areas

Devendra Singh Chaplot, Dhiraj Gandhi, Abhinav Gupta, Ruslan Salakhutdinov (NeurIPS, 2020)

SORNet: Spatial Object-Centric Representations for Sequential Manipulation

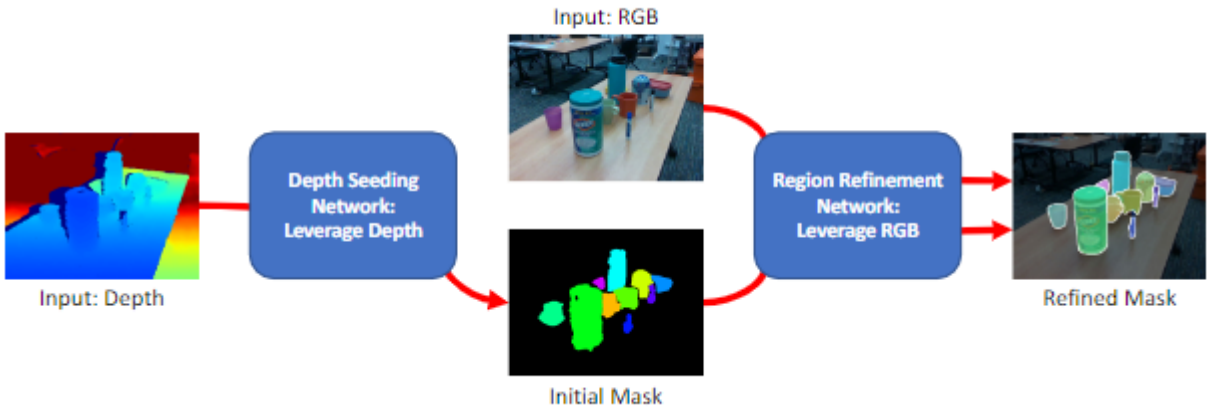
Object representation, Object manipulation, Reason about spatial relationships among object



Wentao Yuan, Chris Paxton, Karthik Desingh, Dieter Fox (CoRL, 2021)

Unseen object instance segmentation for robotic environments

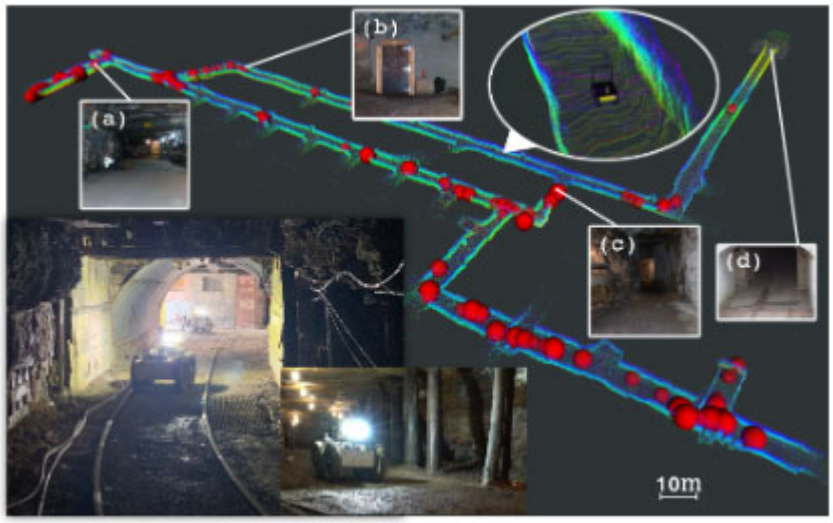
Instance segmentation, Robot perception, Sim-to-real



Christopher Xie, Yu Xiang, Arsalan Mousavian, Dieter Fox (T-RO, 2021)

Unsupervised Online Learning for Robotic Interestingness with Visual Memory

Unsupervised learning, Online learning, Visual memory, Robotic interestingness



Chen Wang, Yuheng Qiu, Wenshan Wang, Yafei Hu, Seungchan Kim, Sebastian Scherer (T-RO, 2021)

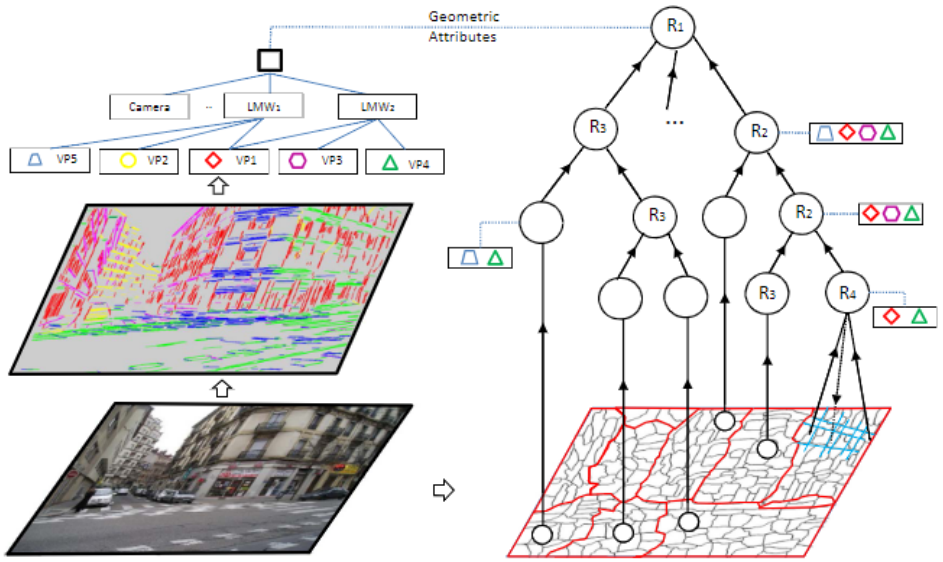
Learning High-Speed Flight in the Wild

Robot perception, Map noisy sensory observations to collision-free trajectories

A. Loquercio, E. Kaufmann, R. Ranftl, M. Mueller, V. Koltun, D. Scaramuzza (Science Robotics, 2021)

Single-View 3D Scene Reconstruction and Parsing by Attribute Grammar

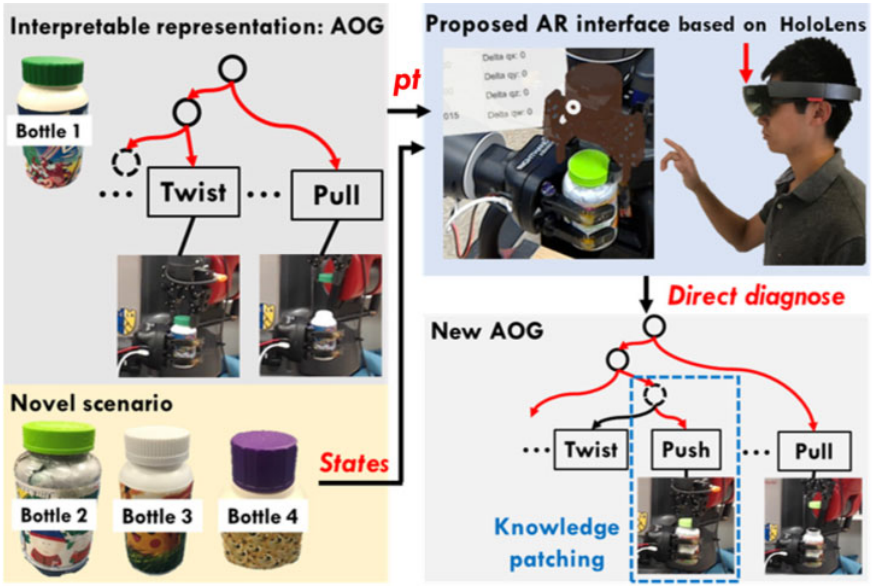
3D Scene Reconstruction, Scene Parsing, Attribute Grammar



Liu, Xiaobai, Yibiao Zhao, and Song-Chun Zhu. "Single-view 3D scene reconstruction and parsing by attribute grammar." T-PAMI, 40.3 (2017): 710-725.

Interactive Robot Knowledge Patching using Augmented Reality

Interactive knowledge structure patching, Augmented Reality



Liu, Hangxin, et al. "Interactive robot knowledge patching using augmented reality.“ ICRA, 2018.

Where can I find the papers?

- arxiv.org
- IEEE Xplore
- Google search
- The references in the recommended paper
- Google Scholar, papers that the recommended paper is “Cited by”
- Proceedings of top Computer Vision/Robotics conferences, like CVPR, ICCV, ECCV, ICRA, IROS, RSS, CoRL, etc.
- Top-tier journals in Computer Vision/Robotics, like Science Robotics, T-PAMI, IJCV, T-RO, IJRR, RA-L, J-FR, etc.

Outline

- General Information
  - About the seminar
  - Registration
- Topics/Papers
  - 3D Scene Representation
  - 3D Scene Understanding
  - ……
- Questions

Questions?

- Webpage (up to date): https://mlr.in.tum.de/teaching/s22/seminar_dlr
- Password for course material: dlr_ss22
- Contact: xingxing.zuo@tum.de , with Email subject: “DLR seminar – xxx”
- **Can I present other topics out of the given list?** You can also suggest topics/papers that you are interested in! We are always open for suggestions, please send Emails for negotiation!
- **Can I present methods without learning or neural networks?** Sure!