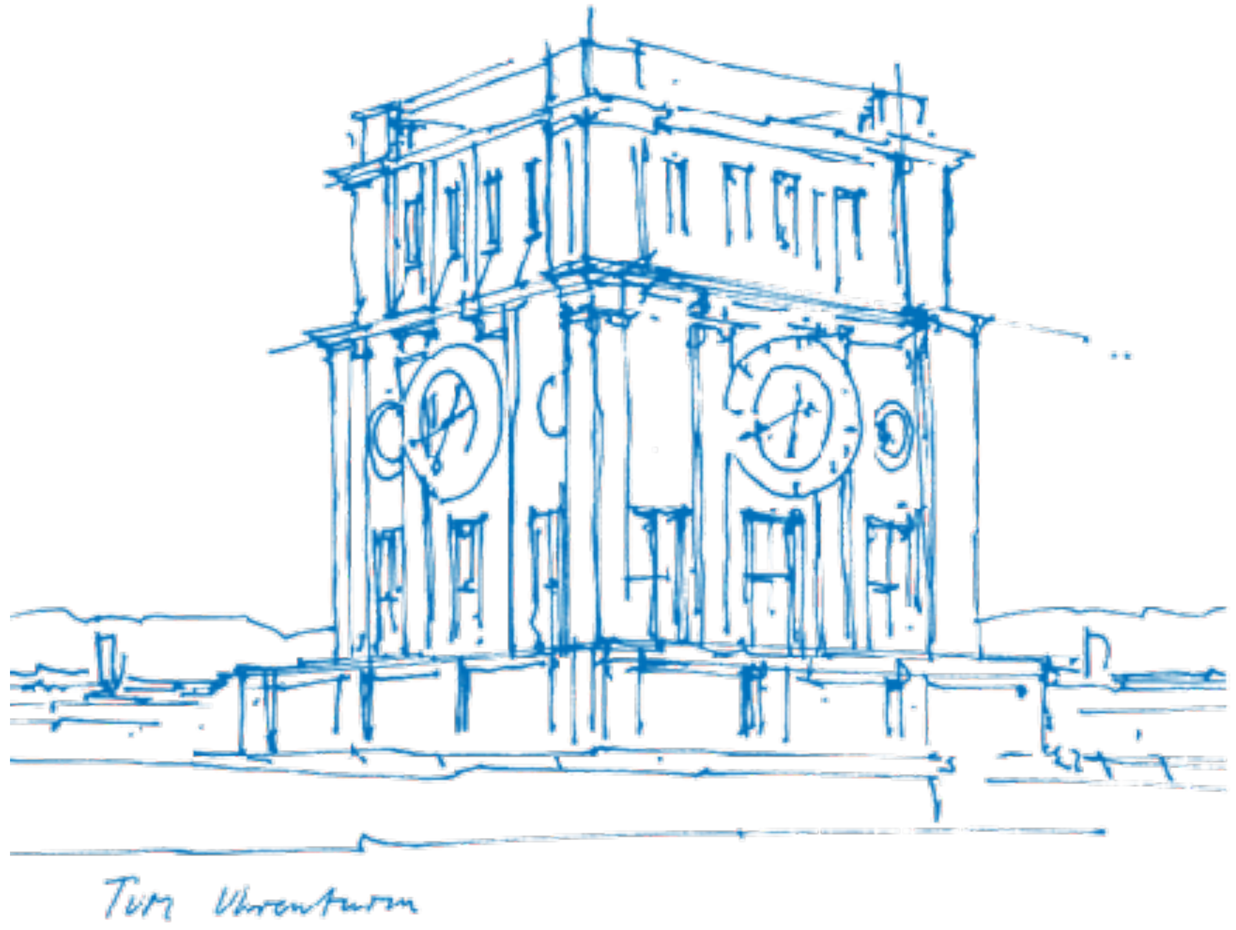
Geometric Scene Understanding
Preliminary meeting

Jason Chui, Linus Härenstam-Nielsen, Dominik Muhle, Lu Sang,
Abhishek Saroha, Dominik Schnaus, Nikita Araslanov
08.02.2023

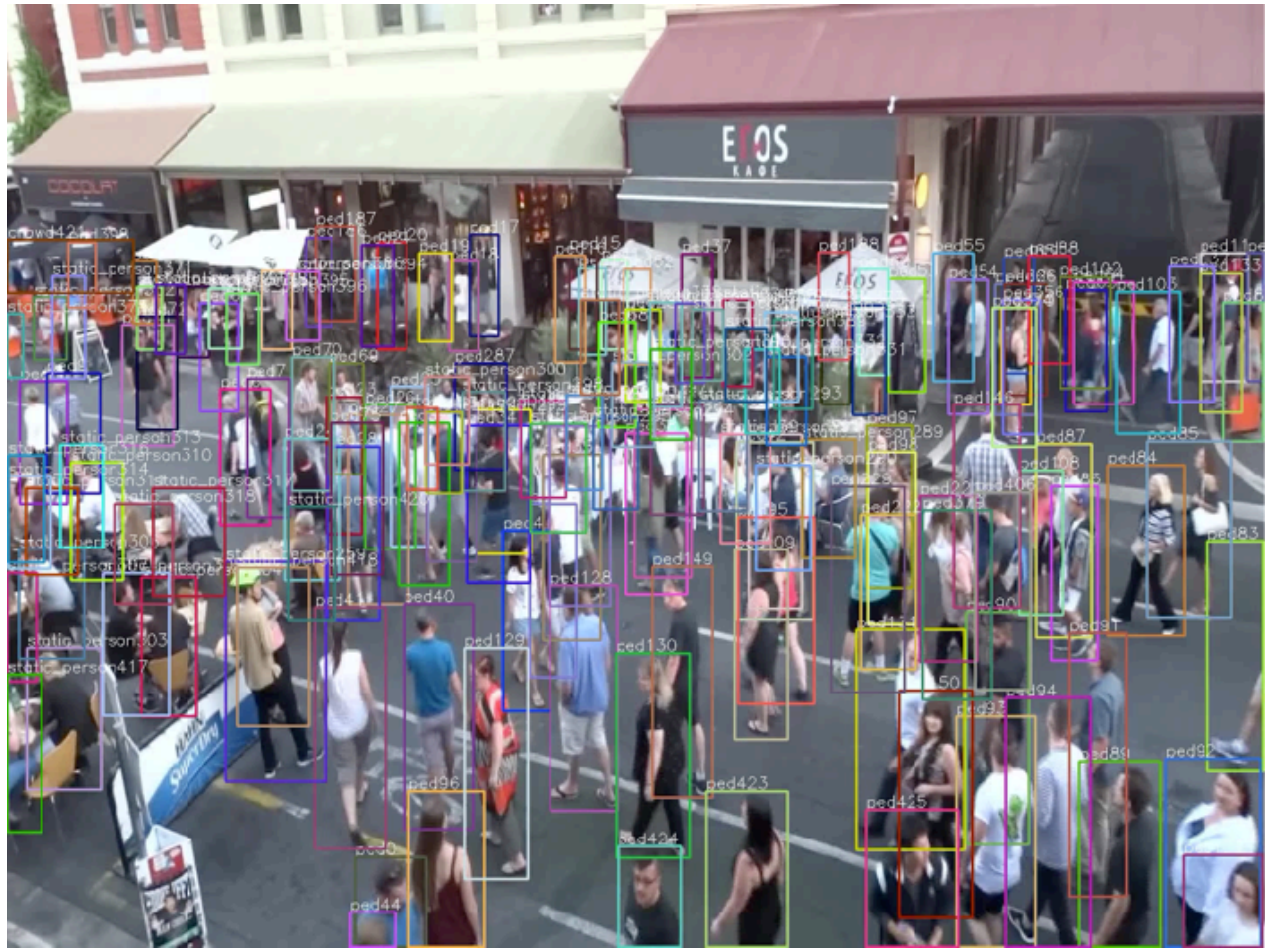


Motivation

- What is scene understanding?
  - Algorithms that describe an image in terms familiar concepts:
    - Content: what is the underlying 3D structure and its properties?
    - Relationship: how are two images related?
    - Temporal evolution: how does the camera/object move?

- What does it mean geometric?
  - Leveraging or inferring 3D geometry.

Real-world scene understanding



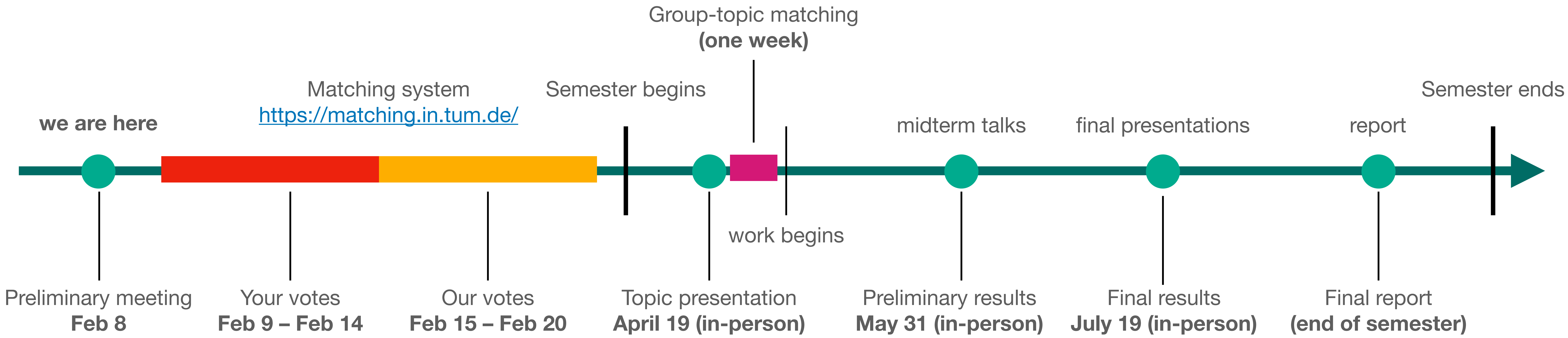
MOTChallenge

Cityscapes

What we expect

- Prerequisites:
  - Introduction to Deep Learning (IN2346)
  - Computer Vision II: Multi-view Geometry (IN2228)
- or any other relevant courses:
  - Computer Vision III (IN2375)
  - Machine Learning for 3D Geometry (IN2392)
- unsure → send us an email (mailing list)

Timeline



- Send us your application (transcripts) by February 14.
- Projects are assigned to groups of 3 people.
- Reports are due by the end of the semester.

Summary

What you get:

- Open research problems.
- Teamwork in a group of 3.
- Regular group meetings (e.g. weekly) with an advisor.
- Access to a GPU cluster.

What we get:

- two presentation (midterm and final);
- written project report.

Your potential advisors

Jason Chui

Linus Härenstam-Nielsen

Dominik Muhle

Lu Sang

Dominik Schnaus

Abhishek Saroha

Nikita Araslanov

You reach us over the mailing list: gsu-ss23@vision.in.tum.de

Nikita Araslanov

- Leverage geometric constraints to infer semantic-level concepts without supervision.

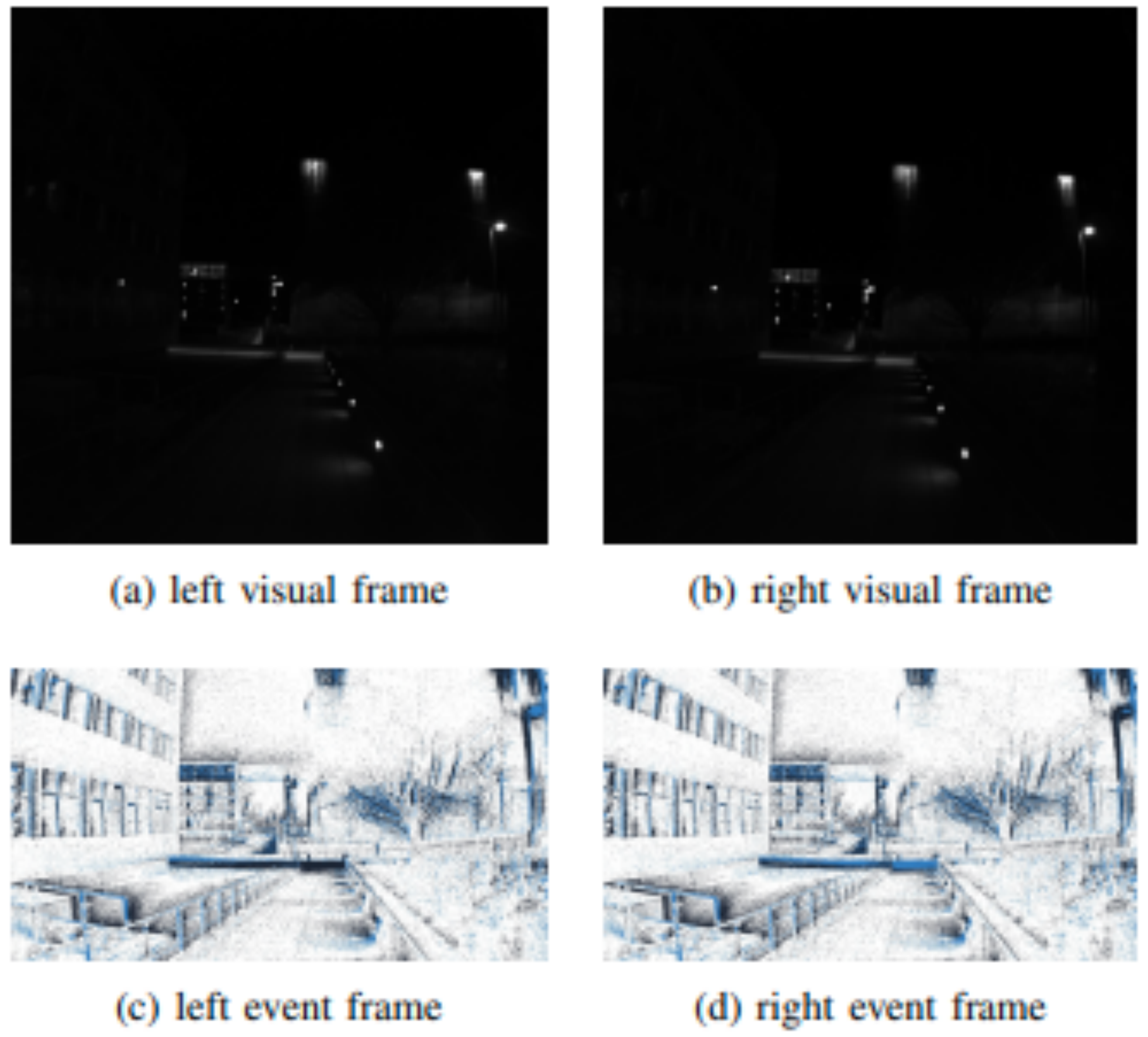


- Example scenario: Can we leverage temporal coherence (in the 3D world) to disentangle dynamic from static objects?

Jason Chui

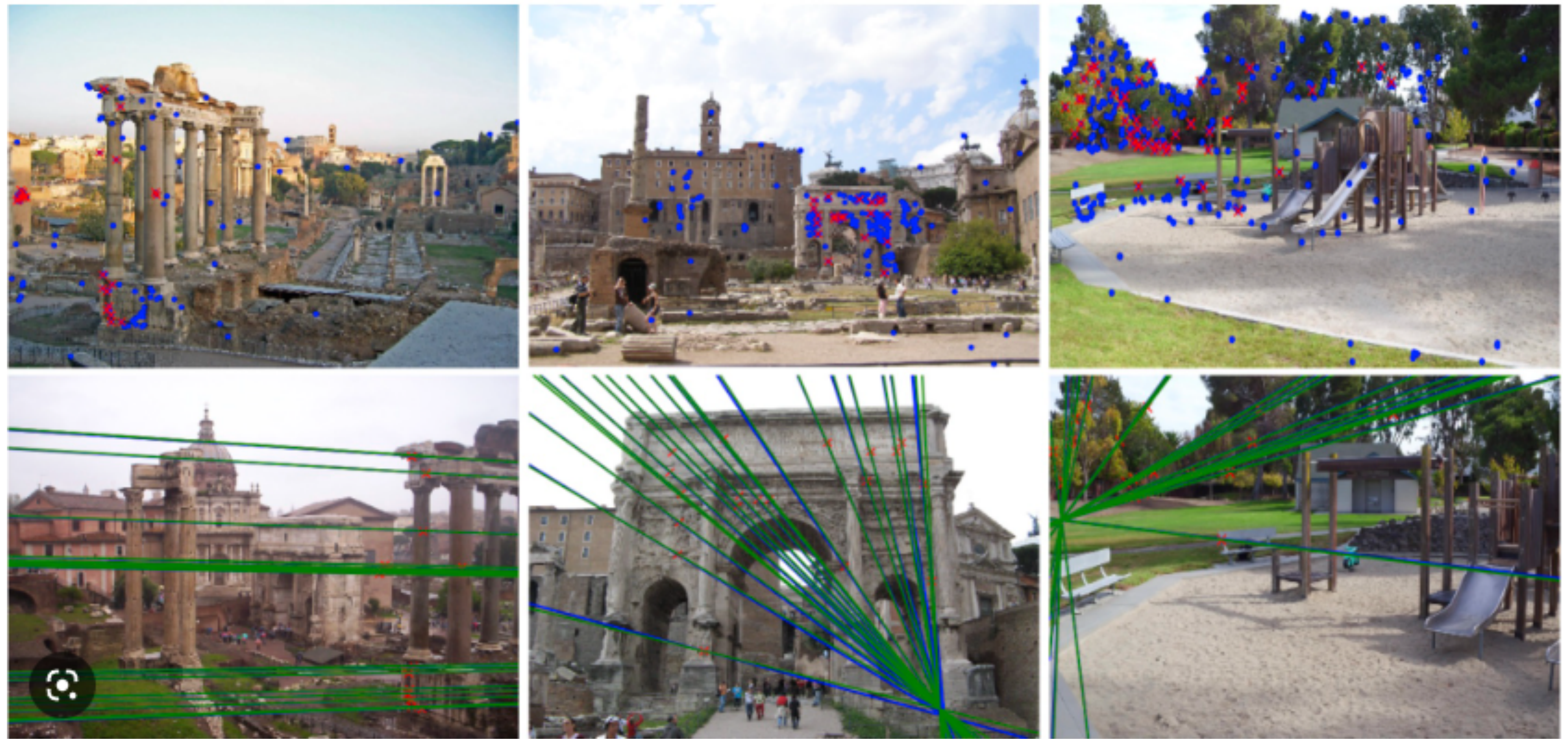
- Research interests:
  A. Sensor fusion (with event camera) (https://vision.in.tum.de/_media/spezial/bib/klenk2021tumvie.pdf)
  B. Deep feature detection and matching (https://arxiv.org/pdf/1905.03561.pdf)
  C. Deep fundamental matrix estimation (http://vladlen.info/papers/deep-fundamental.pdf)

A



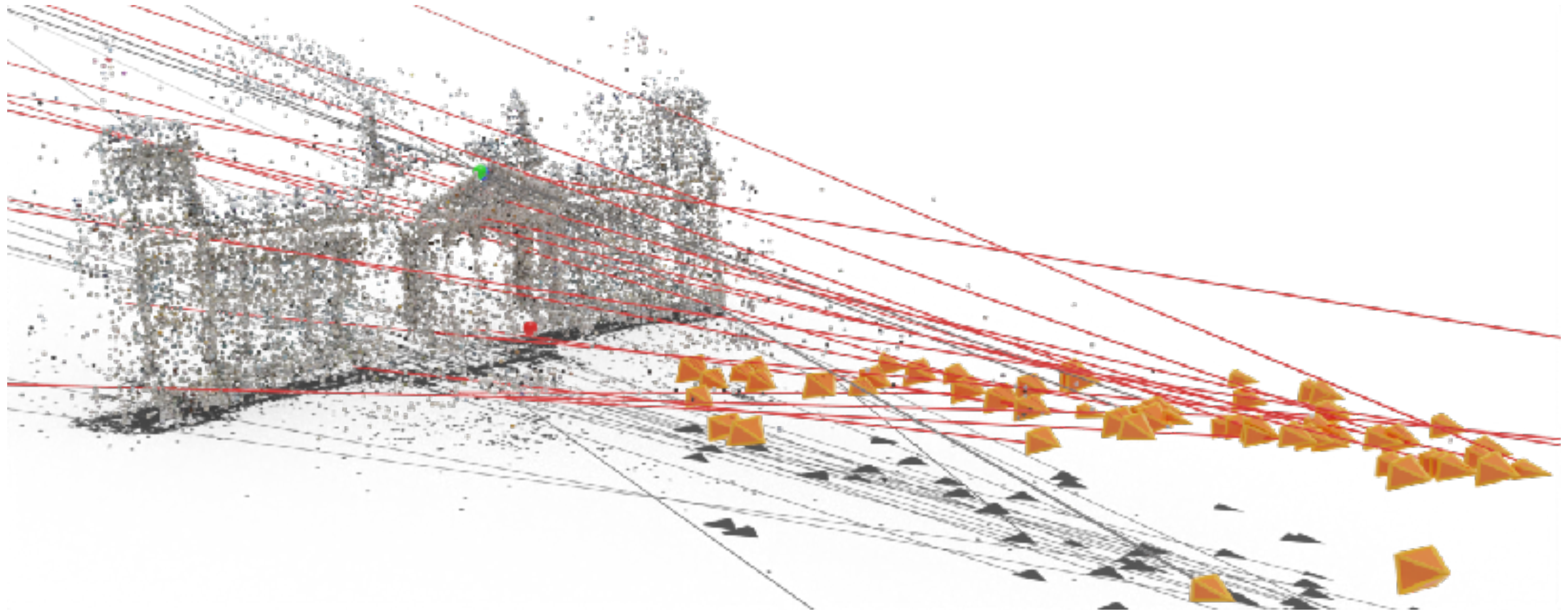
B

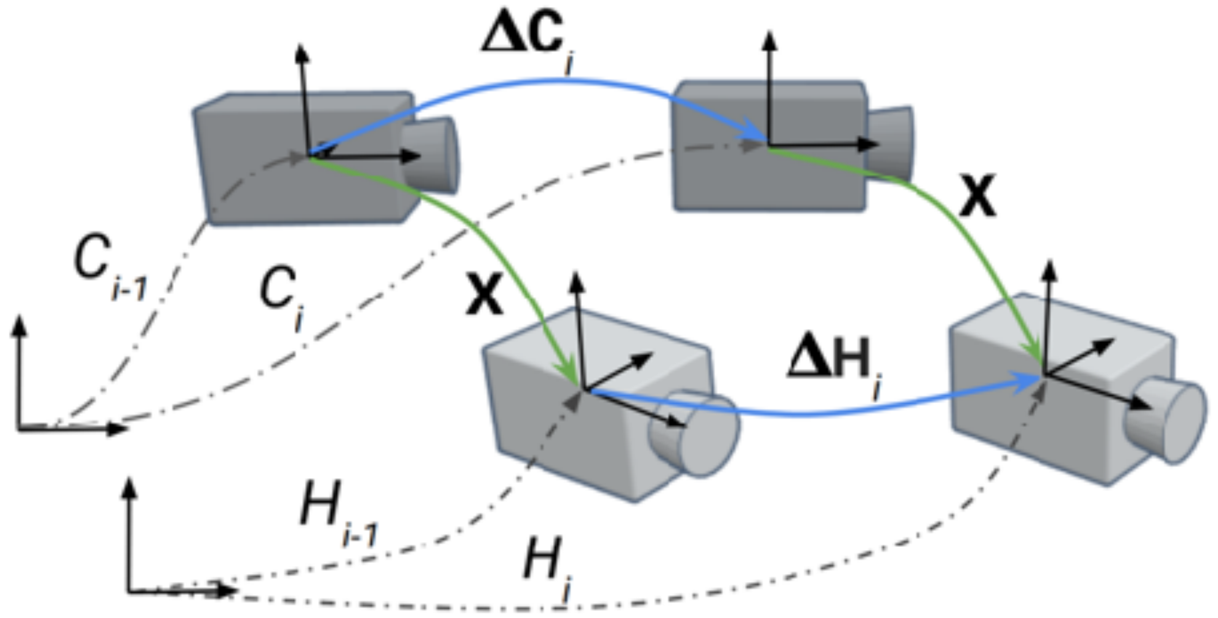
C

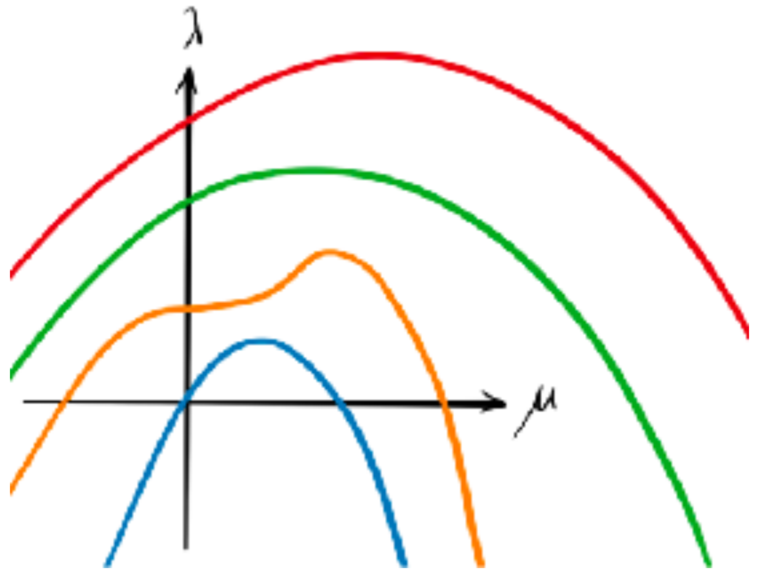


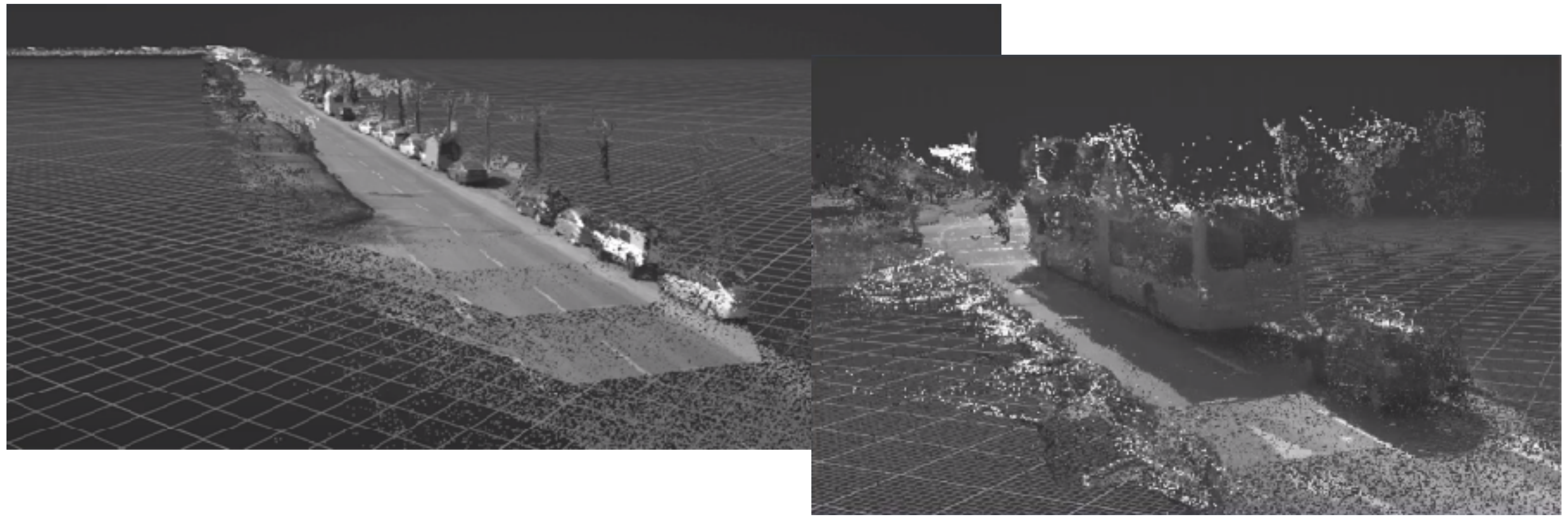
Linus Härenstam-Nielsen

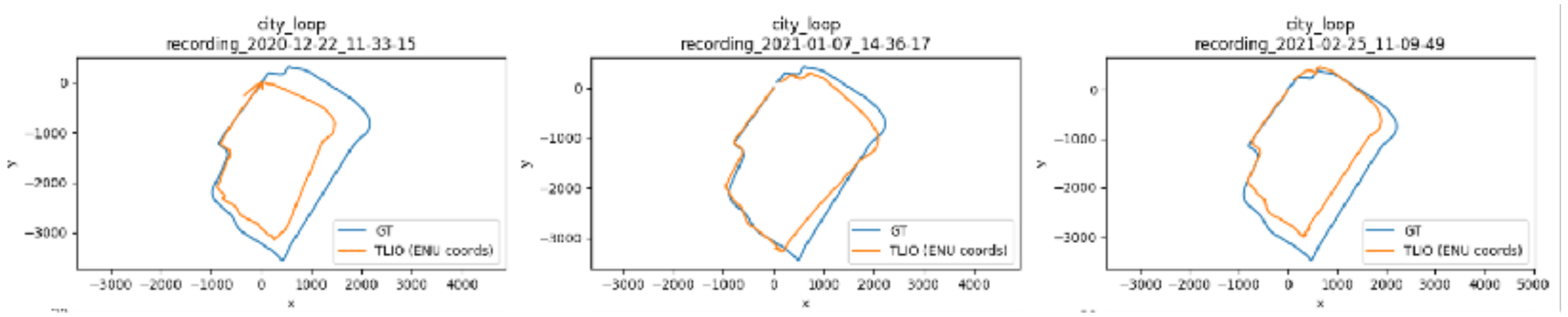
- Globally optimal solvers for geometry problems
- Deep learning for 3D reconstruction and tracking







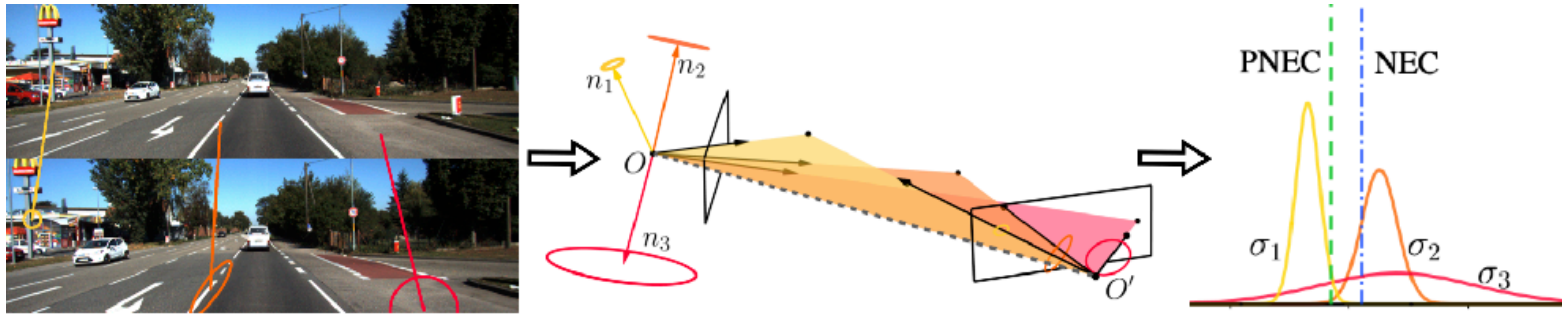




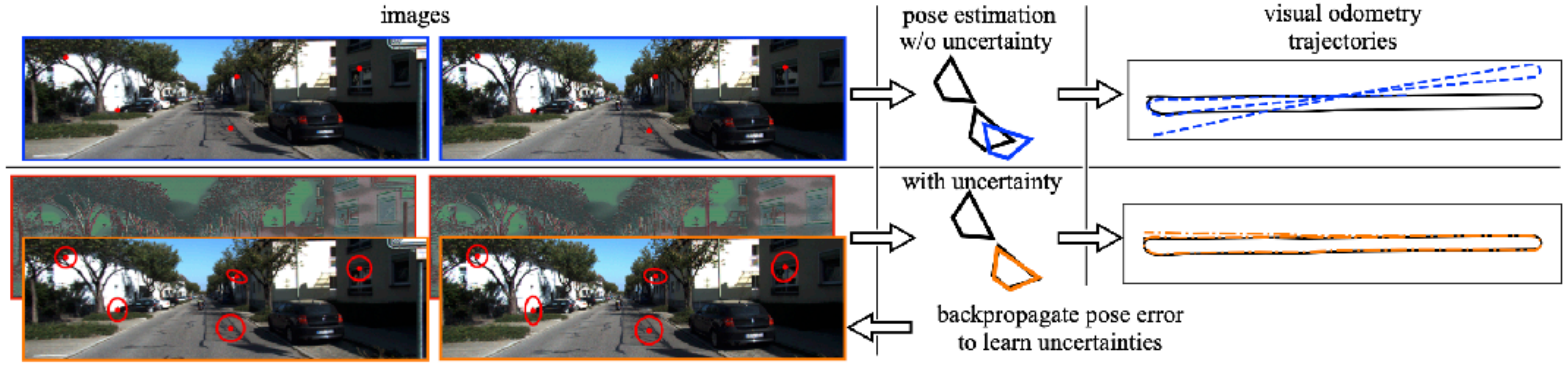
https://arxiv.org/abs/2301.11431
https://arxiv.org/abs/2002.10838

Dominik Muhle

Previous Research in Visual Odometry/Camera Pose Estimation



integrating feature position uncertainty into camera pose estimation

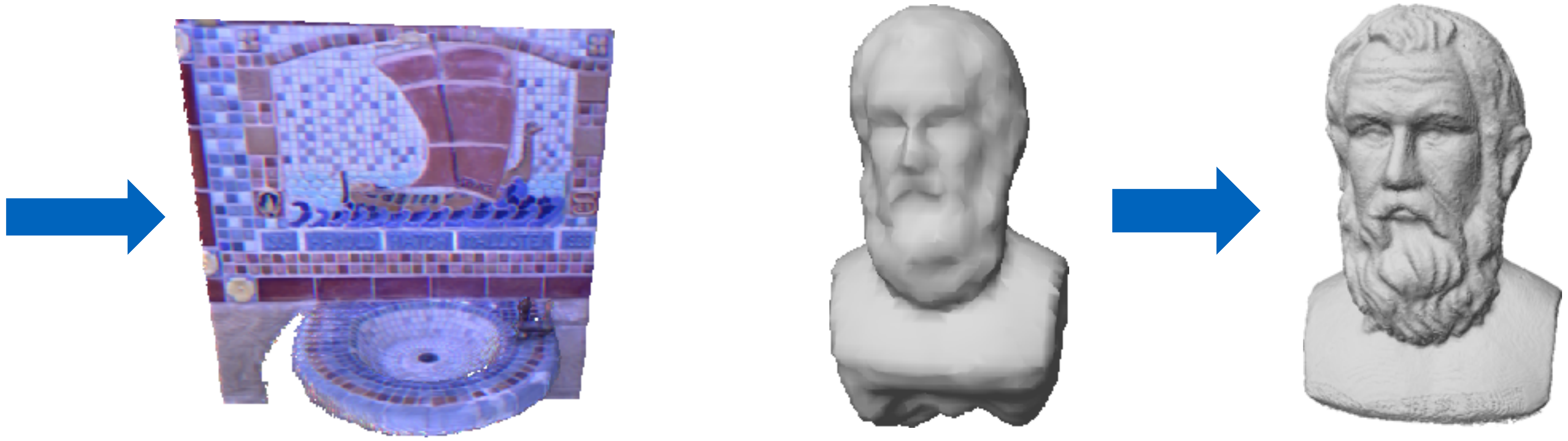


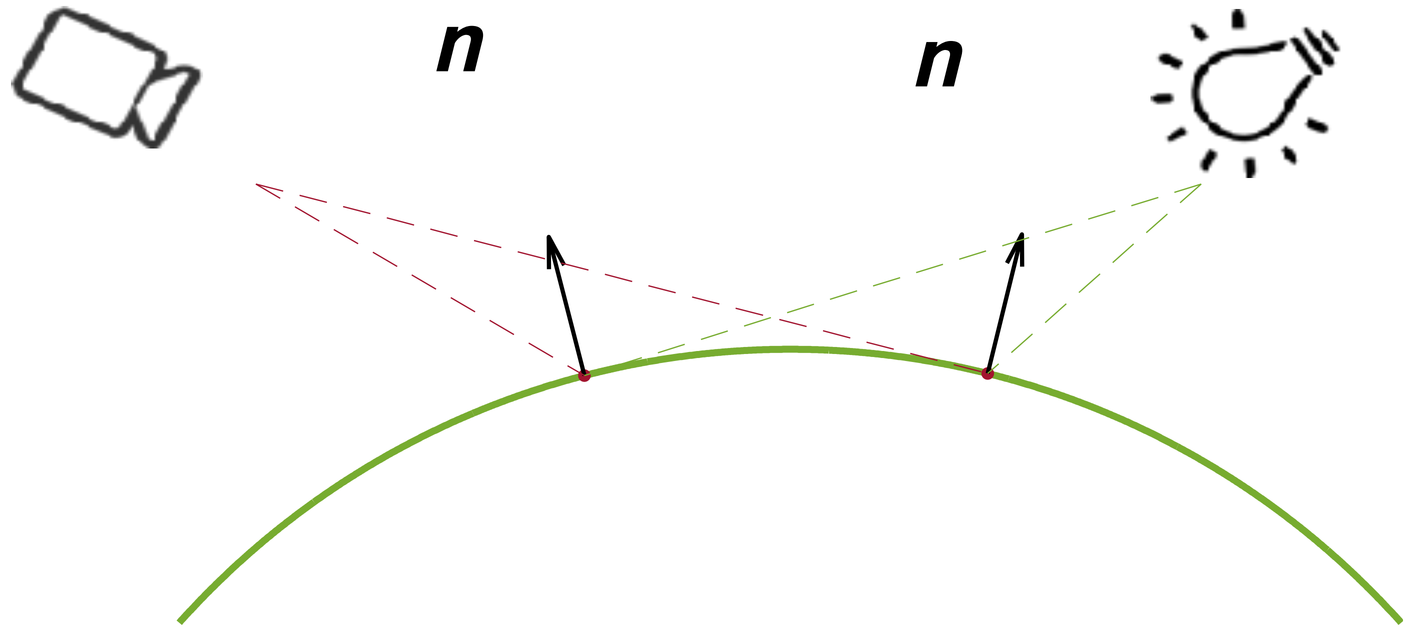
learning feature position uncertainty through differentiable non-linear least squares

GSU Topic: Camera re-localization through a shared 2D-3D embedding space

Lu Sang

Goal: Recover a complete & detailed 3D model of objects or scenes





photometrics stereo

Related topics & papers:

Classical surface representation and geometry refinement

Gradient-SDF: A Semi-Implicit Surface Representation for 3D Reconstruction, Sommer and Sang et al.

Neural network based 3D reconstruction

Multiview Neural Surface Reconstruction by Disentangling Geometry and Appearance, Yariv et al.

Classical method for photometric based surface refinement

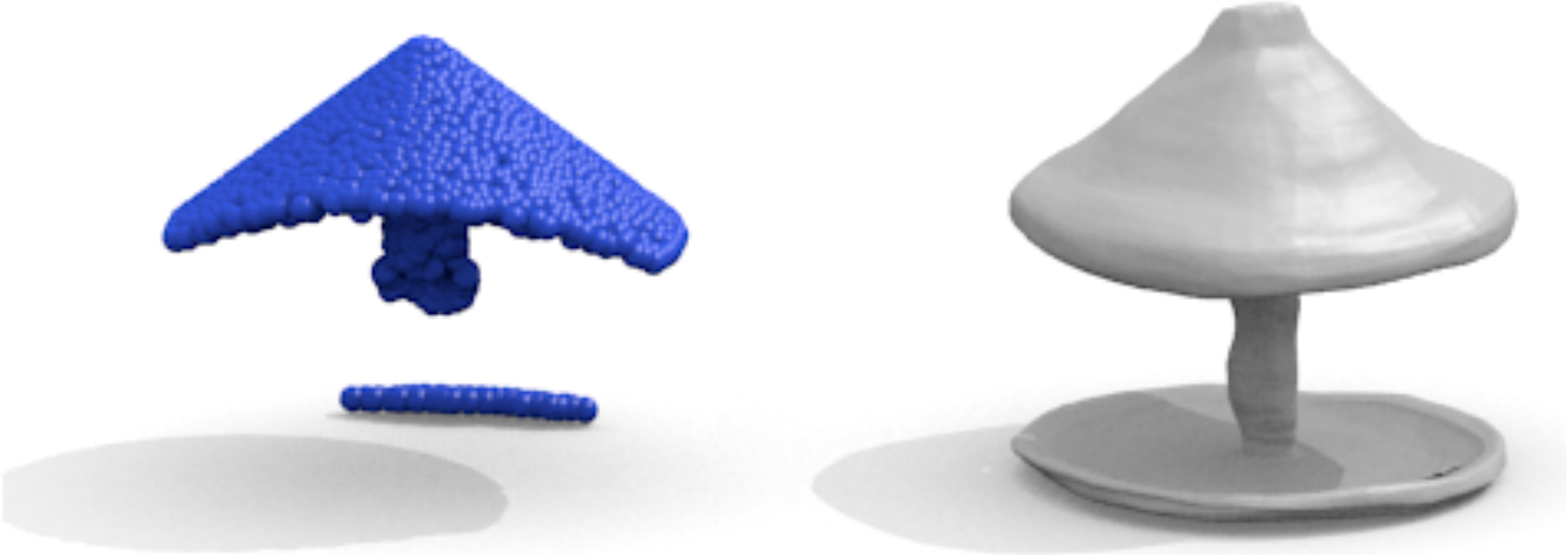
High-Quality RGB-D Reconstruction via Multi-View Uncalibrated Photometric Stereo and Gradient-SDF, Sang et al.

Neural network based photometric based surface refinement

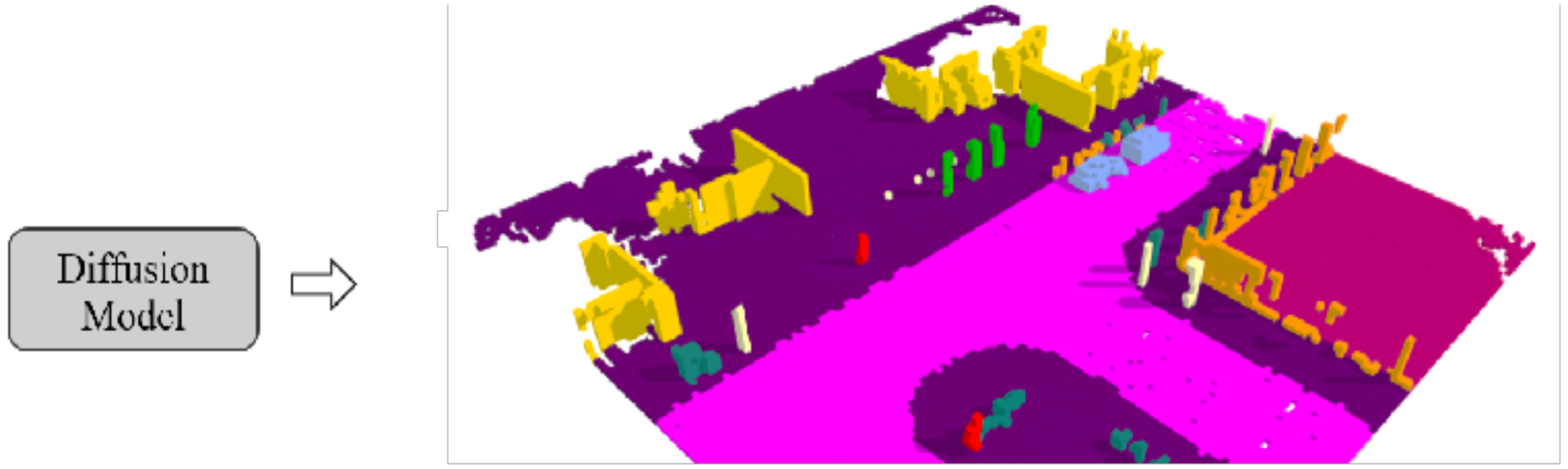
PhySG: Inverse Rendering with Spherical Gaussians for Physics-based Material Editing and Relighting, Zhang and Luan et al.

Abhishek Saroha

- Areas of Interest
  - 3D Shape Completion
  - 3D Reconstruction on Scenes
  - Latent Diffusion



GSU Topic: Generating Novel Scenes using Latent Diffusion



Building	Barrier	Other	Pedestrian	Pole
Road	Ground	Sidewalk	Vegetation	Vehicles

Image Source: https://github.com/zoomin-lee/scene-scale-diffusion

Dominik Schnaus

Goal: Segment images <u>without</u> labeled data using self-supervised learning

Image

Ground truth

Prediction

Related topics & papers:

- Unsupervised Semantic Segmentation by Contrasting Object Mask Proposals, Van Gansbeke et al.
- Deep spectral methods: A surprisingly strong baseline for semantic segmentation and localization, Melas-Kyriazi et al.
- Unsupervised Semantic Segmentation by Distilling Feature Co unsupervised rrespondences, Hamilton et al.
- Unsupervised Semantic Segmentation with Self-supervised Object-centric Representations, Zadaianchuk et al.
- Bridging the Gap to Real-World Object-Centric Learning, Seitzer et al.

Questions?

Contact: gsu-ss23@vision.in.tum.de