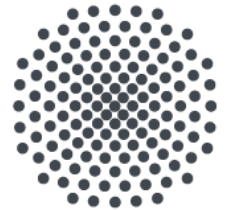


Advanced Seminar



3D Scene Understanding for Robotics and Augmented Reality



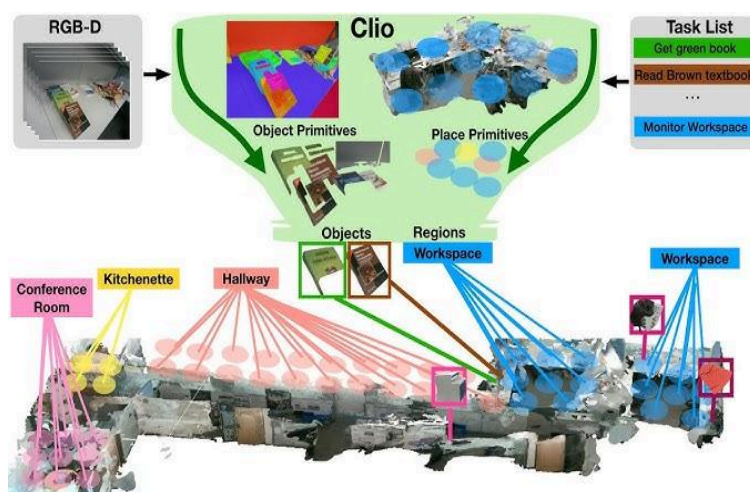
Understanding and reasoning about complex 3D environments is a key challenge for modern computer vision systems used in robotics, augmented reality (AR), and embodied AI. Recent advances in 3D perception and foundation models now allow machines not only to see geometry, but to **understand scenes semantically**, recognizing objects, their relationships, and the context in which they appear.

This seminar focuses on **3D scene understanding** from a computer vision perspective. We explore approaches that combine 3D geometry with visual and semantic information to build meaningful representations of real-world scenes. Topics include 3D object detection and semantic segmentation, object-centric and scene-level representations, spatial reasoning in 3D, and multimodal methods that integrate vision, geometry, and language. These ideas form the foundation for applications such as robot navigation and manipulation, human-robot interaction, and context-aware AR.



Figure 1: Meta Orion

The seminar will also discuss open research questions in semantic 3D understanding, such as how to obtain robust scene interpretations from incomplete and noisy sensor data, how to scale semantic understanding to large and dynamic environments, how to evaluate scene understanding methods, and how structured scene representations can support reasoning and generalization.



Participants will work with current research papers in the field, discuss open challenges and explore how structured scene representations support reasoning and generalization. A curated paper collection will be provided, and students are encouraged to propose additional papers based on their interests. Each participant will present a selected topic and write a short article reflecting on the key ideas and contributions.

Figure 2: Clio: Real-time Task-Driven Open-Set 3D Scene Graphs, D. Maggio et al., 2024 [1]

Target Group:

Master's students in the field of Computer Science

Language:

English

Room/Location:

Details about the time and place will be announced in C@MPUS or via email.

Organizers:

Prof. Dieter Schmalstieg

Michael Pabst

Deniz Bickici

References:

- [1] D. Maggio, et al. "Clio: Real-time Task-Driven Open-Set 3D Scene Graphs." IEEE Robotics and Automation Letters (2024).