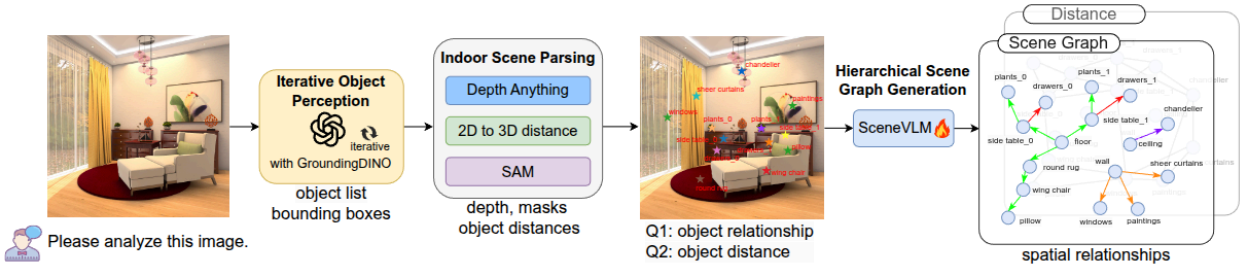
MSc Thesis Topic

Towards More Informative 3D Scene Graphs for Visual Reasoning

Supervisor: Panagiotis Petropoulakis | panagiotis.petropoulakis@tum.de
Examiner: Prof. Dr.-Ing. habil. AC Knoll. Chair of Robotics, Artificial Intelligence and Real-Time Systems
Start Date: Winter Semester 2025
Application Deadline: 24.10.2025



Source: ROOT, Wang et al., 2024

Project Overview

Scene graphs provide a structured representation of objects and their relationships within a scene, forming a crucial intermediate step for visual reasoning and embodied perception. However, current 3D scene graph generation methods often produce **dense and ambiguous relationship sets**, limiting their interpretability and practical value for downstream reasoning or robotic applications.

In this thesis, you will explore methods to **enhance the expressiveness and relevance** of 3D scene graphs by leveraging **vision–language foundation models** and **multi-modal reasoning techniques**.

Your work will build upon recent frameworks in **open-vocabulary visual understanding** (e.g., BLIP-2, PRISM-0, HOV-SG, ROOT, Panoptic Scene Graph Generation) and investigate how these can be adapted or extended for **more structured and interpretable 3D representations**.

Requirements

- Strong programming skills in **Python**; experience with **PyTorch** or other deep learning frameworks.
- Familiarity with **computer vision and visual–language models**.

How to Apply

Send your CV and a short motivation paragraph explaining your interest in **3D visual understanding and graph-based reasoning** to: panagiotis.petropoulakis@tum.de