Professorship Big Geospatial Management
TUM School of Engineering and Design
Technichal University of München

Master Thesis

3D Dynamic Human-Scene Interaction Synthesis

Are you looking for a Master thesis in 3D computer vision and generative AI? Are you interested in combining state-of-the-art generative models with 3D scene understanding? Do you want to explore the intersection of video generation, 3D reconstruction, and human-scene interaction?

We look forward to you joining us as a Master Thesis student (d/f/m) within the Big Geospatial Data Management Group at the Department for Aerospace and Geodesy, TUM School of Engineering and Design. A Supervision in the School of Computation, Information and Technology is also possible.

Location: Ottobrunn/Munich/Remote

Duration: 6 months


[1]


[2]

Your topic:
Synthesizing realistic human-scene interactions in 3D remains a significant challenge. So far, most research has focused on static or overly simplified interaction scenarios. However, rich and diverse examples of interactions already exist in the form of 2D formats in datasets, films or everyday recordings. By leveraging recent advances in large video generative models and 3D scene reconstruction, we now have the opportunity to bring these 2D interactions to life within reconstructed 3D environments. In this thesis, you will explore the potential of generative models to enrich static 3D scenes with dynamic and realistic human-scene interactions. The central research question is: How can we synthesize and inpaint view-consistent and temporally coherent human-scene interactions in 3D environment using 2D generative models?

This includes:

- Literature Review of recent work on video generation models, 3D scene reconstruction and human-scene interaction modeling.
- Evaluate generative models for their ability to generate dynamic human-scene interactions.
- Design and Implement a pipeline for editing reconstructed scenes and generating dynamic interactions.
- Benchmark the quality, realism and consistency of the approaches in different scenes and with different objects.

Related Work:
[1] GenZI: Zero-Shot 3D Human-Scene Interaction Generation (Li and Dai 2024)
[2] CG-HOI: Contact-Guided 3D Human-Object Interaction Generation (Diller and Dai 2024)
[3] Veo and Imagen 3: Announcing new video and image generation models on Vertex AI. (Barkley 2024)

Qualifications:

- Interest in video or image generation using generative models (GANs, diffusion models, etc)
- Familiarity with 3D data representation (point clouds, meshes or voxel grids)
- Advanced programming skills in Python
- Experience with machine learning frameworks and libraries such as PyTorch or TensorFlow
- Interest and experience in literature-based work with a good scientific practice
- Enrolled full time student within Computer Science, Electrical Engineering or similar field of study
- Fluent English is mandatory; German would be an asset

Applications via Mail with CV and transcript to:

Advisor: M.Sc. Wejdene Mansour
Raum: 9377.01.109
Telefon: 089/289-555 57
Email: wejdene.mansour@tum.de