Generative Synthetic Scene Expansion for Immersive Telepresence Using Diffusion Models

Background

Immersive telepresence applications and volumetric streaming demand rich 3D scene content. However, capturing real scenes at scale is costly. Recent advances in generative modelling (e.g., diffusion models) for images and 3D scenes open up new possibilities.

Problem Description

Design a pipeline that uses a **generative diffusion model** to expand or augment captured 3D/Light Field scenes for telepresence: e.g., filling missing viewpoints, synthesising scene variations (lighting, objects), or expanding content coverage for immersive experiences. The student will explore how to condition generative models on 3D/light-field inputs and output plausible augmented scenes.

Milestones and Extensions

- Review literature on image/scene diffusion models, synthetic data for XR.
- Build or adapt a diffusion-based architecture conditioned on sparse 3D/light-field inputs.
- Generate augmented scenes and perform a user study on immersive quality (VR headset).
- **Extension:** evaluate compression/transmission benefits of synthetic-augmented content vs real.

Tools, Qualifications, and Outcomes

- Skills: Python, deep learning, generative models, light-field/3D data handling.
- Tools: HuggingFace, stable-diffusion variants, light-field/volumetric datasets, VR display.
- Outcome: Pipeline and dataset of synthetic augmented immersive scenes, user-study results, potential publication.

Relevant Articles and Resources

- Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J. and Rombach, R., 2023. Sdxl: Improving latent

diffusion models for high-resolution image synthesis. *arXiv preprint arXiv:2307.01952*.

- Poole, B., Jain, A., Barron, J.T. and Mildenhall, B., 2022. Dreamfusion: Text-to-3d using 2d diffusion. *arXiv preprint arXiv:2209.14988*.

- Borkman, S., Crespi, A., Dhakad, S., Ganguly, S., Hogins, J., Jhang, Y.C., Kamalzadeh, M., Li, B., Leal, S., Parisi, P. and Romero, C., 2021. Unity perception: generate synthetic data for computer vision. *arXiv preprint arXiv:2107.04259*.