

## AI6121 Project on Structure from Motion

This project aims for a structure-from-motion technique that allows reconstructing 3D structures from multi-view 2D images or video frames of a scene. You are expected to select a scene on the NTU campus, collect multi-view 2D images or video frames by using your camera (like your smartphone camera), and reconstruct 3D structures (e.g., in a form of point clouds) of the selected scene. You may either leverage open-source codes for the development or implement the structure-from-motion algorithm step by step by yourself. In either case, you're expected to produce a project report that describes your development with detailed analysis and discussion.

You may evaluate the generalization of your development by working on multiple scenes. For example, you may first develop and evaluate your algorithm by using images of one scene, and then examine whether your developed algorithm can work for images of new scenes.

### Submission

You need to submit your project report in **PDF format**, and there are no standard templates for your report. Ensure you include the **names** and the matriculation numbers of **all group members** (2 – 5 group members if completed by groups) clearly in the cover page of your report. Your report will be evaluated according to both contents and report presentation.

Please submit your project report through **NTULearn** before the deadline on **Nov 21<sup>st</sup> 2024**. There will be penalty for late submissions.

### References

- [1] Schonberger J L, Frahm J M. Structure-from-motion revisited. CVPR, 2016: 4104-4113.
- [2] Schönberger J L, Zheng E, Frahm J M, et al. Pixelwise view selection for unstructured multi-view stereo. ECCV, 2016: 501-518.
- [3] Teed Z, Deng J. Droid-slam: Deep visual slam for monocular, stereo, and RGB-D cameras. NeurIPS, 2021, 34: 16558-16569.
- [4] Xu G, Wang X, Ding X, et al. Iterative geometry encoding volume for stereo matching. CVPR. 2023: 21919-21928.
- [5] Wang Y, Zeng Z, Guan T, et al. Adaptive patch deformation for textureless-resilient multi-view stereo. CVPR, 2023: 1621-1630.