ETH zürich

Mixed Reality
Autumn Semester 2024

Real-time 3D Reconstruction Visualizer

Keita AZUMA[1], Aron Tse Rong CHOO[2], Yang LI[3], Jixian TANG[4]
[1]D-INFK (Mobility), ETH Zurich; [2]D-MAFT, ETH Zurich; [3]D-INFK, ETH Zurich; [4]IfI, UZH

Supervisors: Julia CHEN[5], Zador PATAKI[5]
[5]Computer Vision and Geometry Group (CVG), ETH Zurich

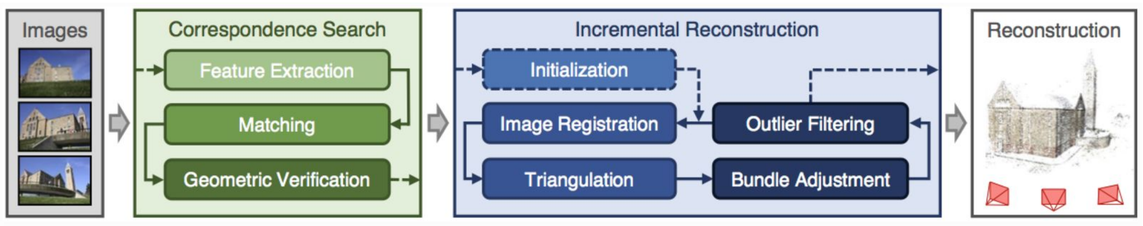
1. Introduction

The goal of this project is to enable interactive visualisation of 3D reconstructions so as to improve understanding of the complex pipeline involved in Structure-from-Motion (SfM) processes.

2. Background

COLMAP

- Popular incremental SfM pipeline used to produce a 3D reconstruction (point cloud and camera poses) from images
- Widely used by the computer vision research community as a preliminary step to generate ground truth
- Unclear how each stage of the pipeline and the order of image registration affects the resultant reconstruction



COLMAP's incremental Structure-from-Motion pipeline.

- ❖ **Image Registration**: Estimate camera position and orientation for each image using image features and correspondences
- ❖ **Triangulation**: Add new 3D points to the reconstruction through triangulation using feature matches from newly added images
- ❖ **Bundle Adjustment**: Perform nonlinear optimization on both point clouds and camera poses to minimize the sum of reprojection errors
- ❖ **Outlier Filtering**: Remove outliers identified during Bundle Adjustment and iterates Bundle Adjustment multiple times

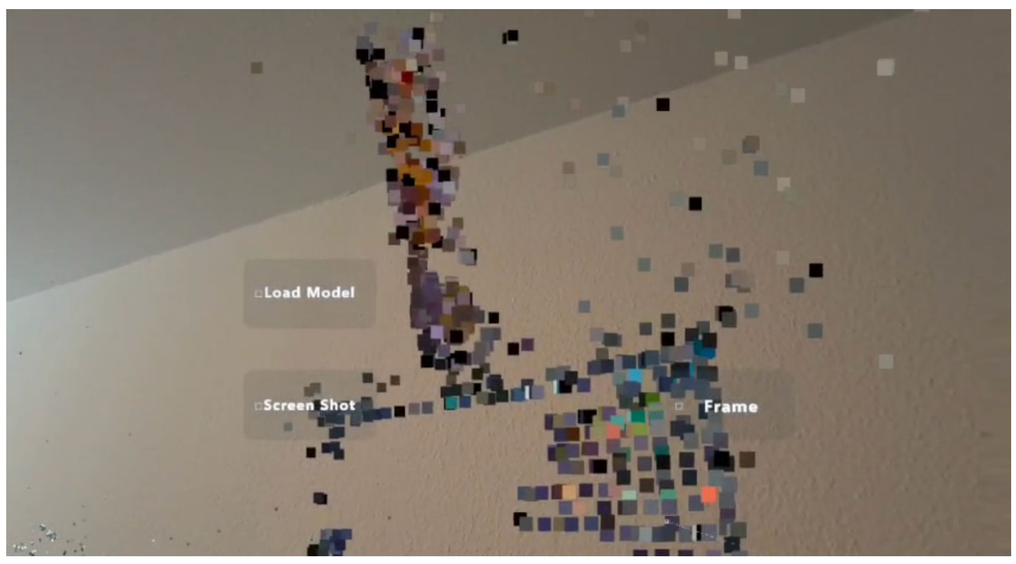
3. Method Overview

1. Add ability to add/remove images from reconstruction in user-specified order
   - Challenge: COLMAP designed for a fixed set of images, algorithmically decides order of image registration
   - Solution: Modify pipeline using pycolmap library to allow for custom incremental mapping
2. Add breakpoints to skip stages within COLMAP pipeline
   - Challenge: Need to support dynamic operations based on user input, such as desired pipeline stages and order of image registration
   - Solution: Multi-threaded server/COLMAP that pauses COLMAP while awaiting user input from client request

4. Results and Discussion

User-friendly Unity app on a Meta Quest 3 that acts as a "visual debugger" for COLMAP's incremental SfM pipeline.

- Collects images of the user's surroundings using Quest's onboard camera
- Offers real-time interactive visualization of the reconstruction after each stage of the pipeline for each image
- From the quality of the reconstructed point cloud, users can intuitively understand:
  - What types of images are suitable for SfM (e.g., avoiding translational image pairs).
  - The roles and importance of individual components, such as Bundle Adjustment, in achieving high-quality reconstructions.



5. Conclusion and Proposed Improvements

The visualizer serves as a valuable tool for researchers to form a deeper, more intuitive understanding into the complex COLMAP reconstruction process, crucial for thorough research.

Future Iterations

- Support user-defined conditions (e.g. min triangulation angle)
- Visualize dense reconstructions and additional information
- Support user-specified feature extractors and matchers

References

1. Schonberger, J. L., & Frahm, J. M. (2016). Structure-from-motion revisited. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4104-4113).
2. Sarlin, P. E., Cadena, C., Siegwart, R., & Dymczyk, M. (2019). From coarse to fine: Robust hierarchical localization at large scale. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12716-12725).

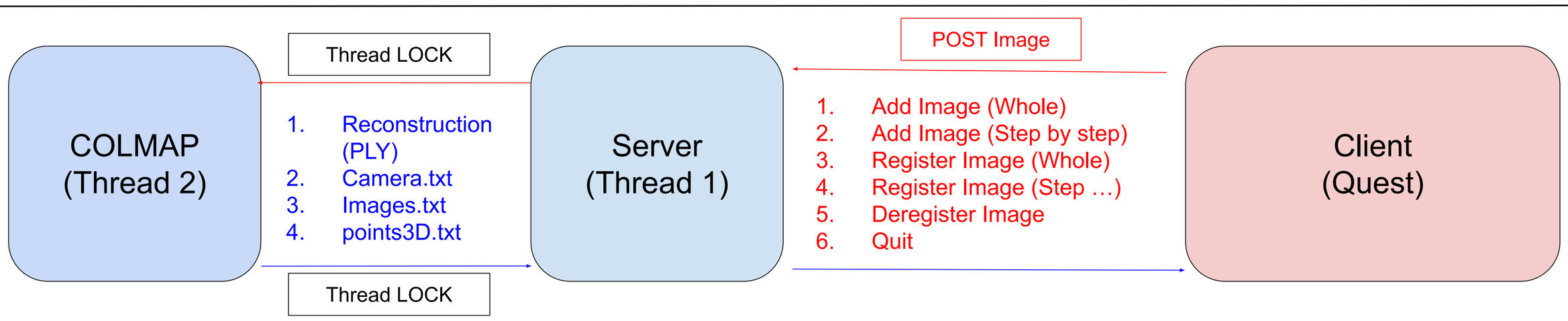


Figure of Method Pipeline. Data communication flow between COLMAP and the Unity application. The server employs thread-based synchronous processing to implement breakpoints within the pipeline. Reconstruction results up to the user-specified step are returned to the client and visualized interactively.