Report On

3D reconstruction using Structure from motion(SFM)

Submitted in partial fulfillment of the requirements of the Course project in
Semester VII of Fourth Year Artificial Intelligence and Data Science

by
Karan Patra(Roll No. 20)
Khushi Tiwari (Roll No.28)

Supervisor
Dr. Tatwadarshi Nagarhalli

University of Mumbai

Vidyavardhini's College of Engineering & Technology

Department of Artificial Intelligence and Data Science

(2023-24)

Vidyavardhini's College of Engineering & Technology
Department of Artificial Intelligence and Data Science

CERTIFICATE

This is to certify that the project entitled “Title of the project” is a bonafide work of "Karan Patra (Roll No. 20) and Khushi Tiwari (Roll No. 28)" submitted to the University of Mumbai in partial fulfillment of the requirement for the Course project in semester VII of Second Year Artificial Intelligence and Data Science engineering.

Dr. Tatwadarshi P. N.
Head of Department

Table of Contents

Problem Statement

The project aims to reconstruct three-dimensional (3D) models of objects or scenes from a set of two-dimensional (2D) images frames. It has applications in various fields, including robotics, augmented reality, cultural heritage preservation, and autonomous navigation. Despite significant progress in SFM algorithms, there are still several challenges and problems that need to be addressed. SFM algorithms often perform well in controlled environments with good lighting and texture-rich scenes. However, they may fail when dealing with challenging conditions such as low-light environments, repetitive textures, or occlusions. The problem is to enhance the robustness of SFM algorithms to deal with these challenging scenarios. Solution to these problems in the domain of 3D reconstruction using Structure from Motion has the potential to impact a wide range of fields, from robotics and archaeology to entertainment and urban planning

Description

SFM algorithms reconstruct the underlying 3D structure and camera poses, enabling us to visualize the environment in a spatial context. This process mimics how humans perceive the world through stereoscopic vision and motion parallax.

- Feature Detection and Matching: In the first step, distinctive visual features, such as keypoints or interest points, are detected in the images. These features are then matched across different images to establish correspondences.

- Camera Pose Estimation: Given the matched features, SFM calculates the relative camera poses (positions and orientations) for each image in the dataset. This is often done using techniques like the Perspective-n-Point (PnP) algorithm or the Essential Matrix decomposition.

- Structure Reconstruction: Using the camera poses and feature correspondences, the algorithm estimates the 3D positions of the detected features in the scene. This forms the initial sparse 3D point cloud.

- Bundle Adjustment: To refine the accuracy of the reconstruction, the Bundle Adjustment step optimizes the camera poses and the 3D point positions simultaneously. It minimizes the reprojection error, ensuring that the 2D image points match their corresponding 3D points as closely as possible.

The output of SFM with Bundle Adjustment is a 3D point cloud representing the structure of the scene, camera poses for each image, and camera intrinsic parameters. These elements collectively create a detailed 3D model of the environment, which can be used in various applications, including robotics, augmented reality, archaeology, and more.

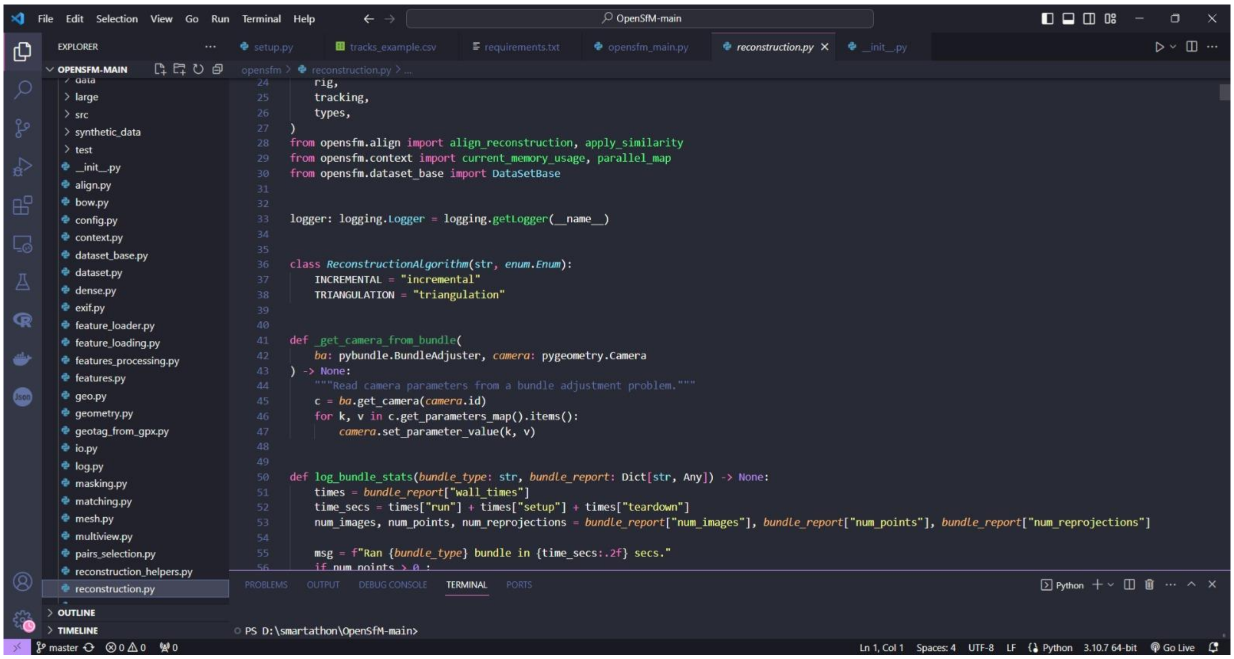
Module Description:

- Feature Detection and Matching:
  These features matched across different images to establish correspondences.
- Camera Pose Estimation:
  SFM calculates the relative camera poses (positions and orientations) for each image in the dataset
- Structure Reconstruction:
  This forms the initial sparse 3D point cloud.
- Bundle Adjustment:
  It optimizes the camera poses and the 3D point positions simultaneously.

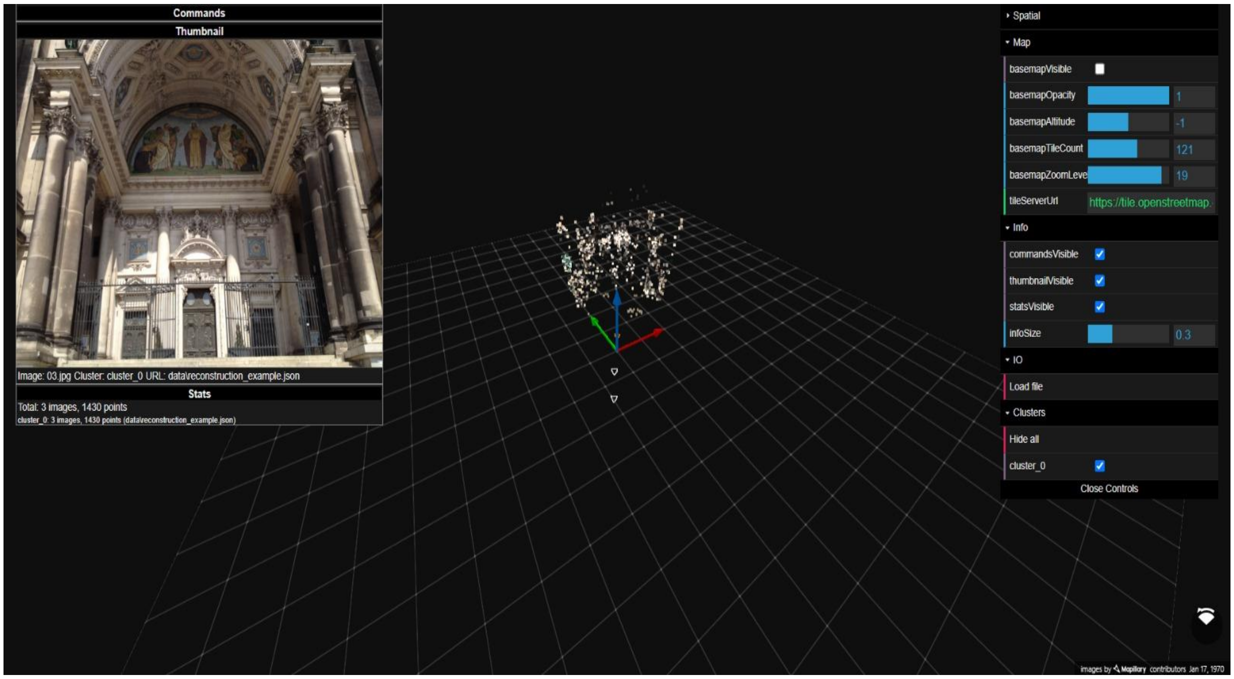
Brief Description of Software & Hardware

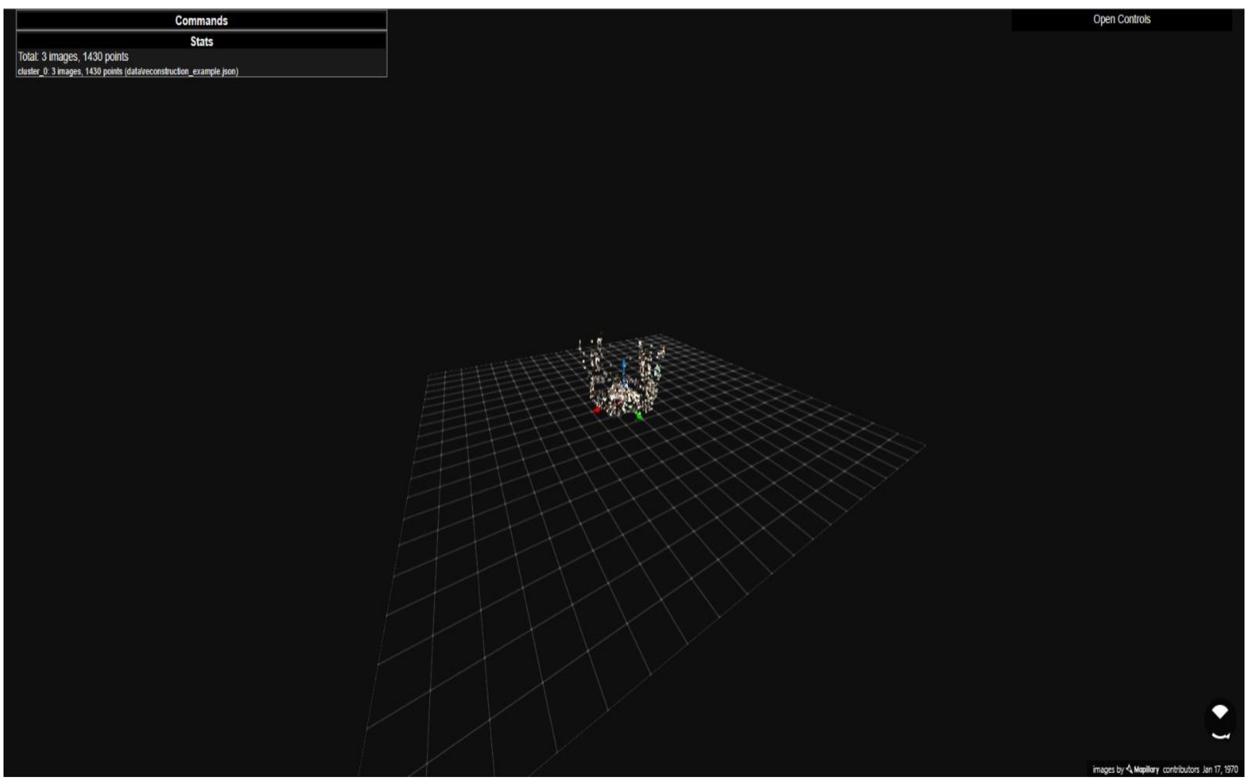
- Software: COLMAP and Bundler are proficient at creating both sparse and dense reconstructions, while VisualSFM offers an accessible interface for large-scale projects. CMVS-PMVS and MeshLab are useful for generating dense 3D models and post-processing.
- Hardware: The project was developed on a standard computer with a CPU and GPU for accelerated training.

Code



Results


Commands
Thumbnail
Image: 03.jpg Cluster: cluster_0 URL: data\reconstruction_example.json
Stats
Total: 3 images, 1430 points
cluster_0: 3 images, 1430 points (data\reconstruction_example.json)
Spatial
Map
basemapVisible
basemapOpacity
1
basemapAltitude
-1
basemapTileCount
121
basemapZoomLeve
19
tileServerUrl
https://tile.openstreetmap.
Info
commandsVisible
thumbnailVisible
statsVisible
infoSize
0.3
IO
Load file
Clusters
Hide all
cluster_0
Close Controls


Commands
Stats
Total: 3 images, 1430 points
cluster_0: 3 images, 1430 points (data\reconstruction_example.json)
Open Controls

Conclusion

In conclusion, 3D reconstruction using Structure from Motion (SFM) stands as a pivotal technology in computer vision and photogrammetry, offering a powerful means to transform two-dimensional images into detailed three-dimensional models. The process involves a series of essential steps, with the SFM algorithm, particularly the Bundle Adjustment phase, playing a central role in refining camera poses and scene structure to achieve highly accurate reconstructions. SFM has found applications in diverse fields, including robotics, augmented reality, cultural heritage preservation, and urban planning. While it has made significant strides, ongoing challenges, such as scalability, robustness in challenging conditions, real-time processing, and multi-modal data integration, continue to motivate research and innovation in this field.

References

[1]Chen, H., Xu, F., Liu, W., Sun, D., Liu, P. X., Menhas, M. I., & Ahmad, B. (2021). 3D reconstruction of unstructured objects using information from multiple sensors. IEEE Sensors Journal, 21(23), 26951-26963

[2] Dipanda A and Woo S 2005 Towards a real-time 3D shape reconstruction using a structured light system Pattern recognition 38(10) 1632-1650

[3] Prakoonwit S and Benjamin R 2007 3D surface point and wireframe reconstruction from multiview photographic images Image and Vision Computing 25(9) 1509-1518

[4] Park J S 2005 Interactive 3D reconstruction from multiple images: A primitive-based approach Pattern recognition letters 26(16) 2558-2571

[5] Hwang J T and Ting-Chen C 2016 3D Building Reconstruction By Multiview Images and the Integrated Application With Augmented Reality The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 41 1235

[6] Scharstein D and Szeliski R 2002 A taxonomy and evaluation of dense two-frame stereo correspondence algorithms International journal of computer vision