RBE 549: Homework Assignment 8

Aditya Mehrotra
amehrotra@wpi.edu

November 4, 2021

Problem 1

Included in .pdf below .

Problem 2

$| I_1 - I_2 |= \sum(I_{1_{ij}} - I_{2ij})^2$

where,

$(\mathrm{I}_{1_{ij}} - I_{2ij})^2 = \begin{cases} 0 & \text{for} \quad I_{1_{ij}} = I_{2ij} \\ 1 & \text{for} \quad I_{1_{ij}} \neq I_{2ij} \end{cases}$

$\therefore \sum(I_{1_{ij}} - I_{2ij})^2 = 1 + 1 + ... + 1(\text{where } \mathrm{I}_{1_{ij}} \neq I_{2ij}) = \#$ of pixels where $\mathrm{I}_{1_{ij}} \neq I_{2ij}$

Problem 3

a $\overrightarrow{\overline{X}} = \frac{1}{M}\sum_i X_i$

$\mathrm{Q} = \frac{1}{M}\sum_i (X_i - \overrightarrow{\overline{X}})(X_i - \overrightarrow{\overline{X}})^T$

So, for $Q.\overrightarrow{\overline{X}}$

Impossible to solve.

Problem 4

There are a total of $| P |$ pairings $= | E || M |$. Now, each of these pairings can lie in or out of the interpretations. Thus, there are a total of $2^{|P|} = 2^{|E||M|}$ interpretations.

3D Reconstruction through Incremental Structure from Motion

Aditya Mehrotra amehrotra@wpi.edu
Samarth Shah sshah5@wpi.edu
Anujay Sharma asharma6@wpi.edu
Aditya Patil apatil@wpi.edu

November 4, 2021

1 Introduction

Understanding the geometry of the scene is essential for various problems in autonomous mobile robotics. One such challenge for an autonomous mobile robot is to perceive an unstructured environment. The perception and 3D reconstruction of a scene can be achieved through a visual technique called Structure from motion. Structure from motion (SfM) is the process of estimating the 3D structure of a scene from a set of 2D images taken from different viewpoints. SfM is used in many applications, such as 3D scanning, augmented reality (AR), and visual simultaneous localization and mapping (vSLAM). This technique can be combined with data from other peripheral robotic hardware for other various applications.

2 Background & Significance

The groundbreaking work in [1] first introduced the point-based linear correspondences approach to solve SfM, which came to be formally known as the eight point algorithm. In particular, this problem encapsulates two aspects: The first is resolving a 3D structure from a set of overlapping frames that have an offset, and the latter being estimation of the relative camera pose giving the rotation and translation between frames. This method does not require a-prioris and is solved with a highly redundant bundle adjustment method by extracting relevant features from overlapping images [2].

The images are supplied in a certain order, such as the frames of a video or photographs recorded by a robot traversing around a room, in one frequent multi-image situation. If corresponding points between consecutive images can be matched, a simple way to reconstruct a sequence is to use the eight-point algorithm to recover the poses of the first two frames and triangulate their common points, then resection the next frame, triangulate any new points, and repeat: a form of visual ground truthing. This approach can work effectively with properly-calibrated cameras, especially on short image sequences. The method of incrementally growing the reconstruction with each image input is the standard and robust approach [3] which essentially is explored in this work.

3 Proposed Methodology & Technical Pipeline

This work aims to employ incremental SfM. Incremental SfM involves the four main stages namely, (1) Extraction and Matching, (2) Geometric Verification, (3) Reconstruction, and (4) Adjustment. The pipeline for building an SfM application is as follows:

1. Feature Extraction: The method should generate a tensor track that defines the matches across various views given a set of photos. (SIFT + Nearest Neighbours Algorithm)

2. Estimating the Camera Pose: Given an essential matrix, the function computes and finds eight camera position combinations, then chooses the optimal one based on cheirality. (8-point Algorithm)

3. Perspective-n-Point algorithm: The function should register a new picture utilizing 3D-2D correspondences using a linear perspective-n-point approach with RANSAC given a 3D reconstruction of points.
4. Reconstructing 3D Points: The function will reconstruct 3D points that have not previously been reconstructed given a newly registered image. Triangulation can be used to discover the spots that will be newly added and reconstruct them.
5. Running Bundle adjustment: This function will use nonlinear least squares optimization to refine the camera pose and reconstructed points while minimizing re-projection errors.

Although this method is robust, there are some scalability issues for large data sets.

4 Discussion & Conclusion

Based on the pipeline observed from the previous section, it can be stated that a solid groundwork has been settled to continue the development of this project. Currently, there is no significant analysis that can be performed given the lack of subjectivity in the judgment that can be inferred from the processes described in the earlier sections.The modules to be implemented are vital to the success of this project, and need to work in a reliable manner. However, at a later point, we will be able to carry out an analyses on the performance of our SfM code. Overall, the goal of this project is actively coming to fruition. Despite not being totally finished with the development of the software to carry out SfM, there is a solid groundwork in place that will certainly streamline future testing.

References

[1] H Christopher Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(5828):133–135, 1981.

[2] Keith N Snavely. *Scene reconstruction and visualization from internet photo collections*. University of Washington, 2008.

[3] Changchang Wu. Towards linear-time incremental structure from motion. In *2013 International Conference on 3D Vision-3DV 2013*, pages 127–134. IEEE, 2013.