EE367 Term Project
Design of a Structure From Motion System

Author: Hans Magnus Ewald [hmewald@stanford.edu]

Project Idea and Goals

For this term project, the idea is to use a single mobile camera to obtain 3D information on a given scene. This information will be obtained using visual odometry combined with structure from motion techniques and stored or displayed as required. The larger scope of this idea is to eventually use it in a robotics application for exploration missions. Alternatively it could see deployment in a mobile application to capture 3D maps by moving a cell phone camera around.

For this project we will specifically aim for a final implementation using a camera connected directly to a laptop with a processing system implemented in MatLab. The goal is to reach a final system that can run in real time, generating an incrementally updated 3D map as images from new perspectives are added.

Proposed Methodology

For the theory behind implementing a structure from motion system we will draw primarily on existing literature and toolkits. The texts below are examples of material on the subject. [2], [3] and [4] are all textbooks that cover the topics of structure from motion and 3D reconstruction. These will serve as a good first place to look for further insight into the problem.

To update our 3D reconstruction in real-time, the algorithm we will use here is incremental structure from motion estimation. This means that each new image will lead to a round of optimization, before another image is added, and further optimization is performed. The principle of structure from motion is to recognize and keypoints from the different images before estimating a projection matrix for each image in the series. Firstly, keypoints can be found and matched using SIFT/SURF descriptors, and we will also make use of robust statistics in the form of the RANSAC algorithm for outlier rejection. Secondly, to optimize the projection matrix estimates we employ the standard homography estimation technique to begin with. We will seek begin our work with a pre-calibrated image set to partition the problem a little. If time permits, we will also make an attempt to use bundle adjustment to estimate structure and motion for larger groups of images.

Suggested Milestones:

- Implement structure from motion for 2 frames at a time from a standard image set
- Calibrate and use own camera
- Stretch Goal: Use full-on bundle adjustment with larger sets of images

References

[1] F. Dellaert, et.al.: "Structure from Motion without Correspondence" from the IEEE Conference on Computer Vision and Pattern Recognition, 2000. Available: `http://www.ri.cmu.edu/pub_files/pub2/dellaert_frank_2000_1/dellaert_frank_2000_1.pdf`

[2] L. Quan, T. Kanade:"Image-Based Modelling". Springer US, 2010

[3] R. Szeliski:"Computer Vision: Algorithms and Applications". Springer-Verlag London, 2011

[4] R. Hartley, A. Zisserman:"Computer Vision: Algorithms and Applications". Cambridge University Press, 2004