Chalmers University of Technology

SSY097 - Image Analysis

Sequential 3D Reconstruction

Li Weiming
Jiang Yumeng
Xie Yanni

March 16, 2017

Department of Signals and Systems

1 Introdution

3D models are essential to many hot topics like indoor navigation, self-drive, virtual reality and etc. However, constructing a fast and reliable 3D model usually requires specialized high cost hardware and thus 3D reconstruction introduces a cheap way to solve these problems. The limitation lies in that the calculation load is so high that it's almost impossible to calculate online. Our main task in this project is to reduce the calculation load, meanwhile ensure accuracy.

An efficient way to solve this problem is to explore the limit of computer using SIFT on GPU[1], Multicore Bundle Adjustment[2] and etc. Structure from motion (SFM) is another way to handle reconstruction of increasingly large uncontrolled photo collections[3]. Instead of using those complicated methods, we implement a very simple way to reduce the calculation load.

We follow the traditional way of 3D reconstruction: start from two views, match SIFT features, use RANSAC to estimate the camera pose, estimate 3D points and use feature tracks to track all useful information and then try to add new views. Match SIFT features of the new view and the last view, update all feature tracks, use the feature point to feature track correspondences and use feature group to estimate the pose of the new camera, and finally, rerun triangulation for each feature group and re-estimate 3D points. When estimating a new camera pose, since the approximate pose is predictable, that is , it should be close to the last camera pose, we add a limit region to the pose, if it is outside the region, just skip the following calculation. Also, to keep the accuracy, when adding new features to the feature track, we update 3D points in all feature groups considering the new camera pose and only keep those feature groups with 3D points within safety region. More details see our codes. Fig 1 shows the result of 3D reconstruction from all views (details see reconstruction.fig and reconstruction.mat), the total running time is 2.3h (in a Windows system computer with 2.5GHz CPU), which is 5 times less than before. We run 500 times RANSAC iterations to estimate each new camera pose, which is far not enough, so the result is not the best. To keep balance of speed and accuracy, a lot of work need to do.

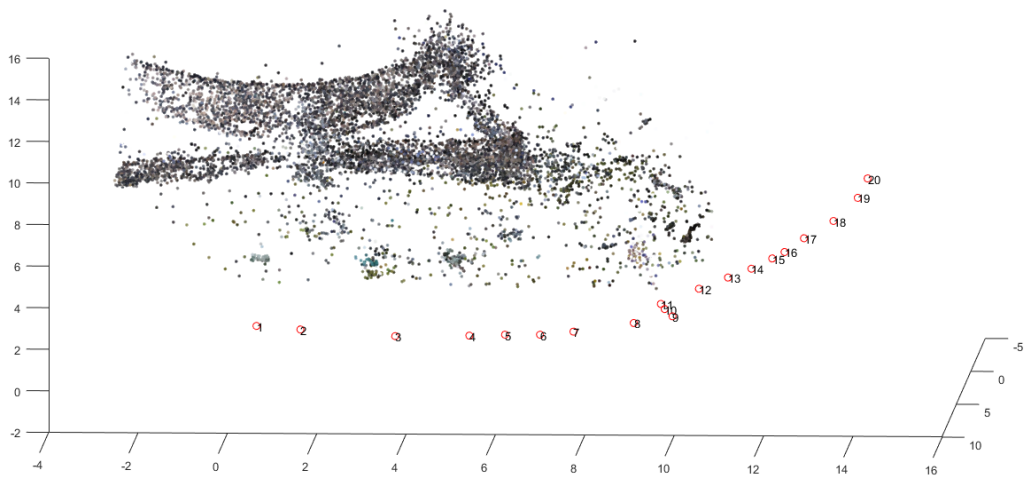

Figure 1: 3D reconstruction from all views

References

[1] Changchang Wu. *A GPU Implementation of Scale Invariant Feature Transform*. URL: `http://www.cs.unc.edu/~ccwu/siftgpu/`.

[2] Changchang Wu. *Multicore Bundle Adjustment*. URL: `https://grail.cs.washington.edu/projects/mcba/`.

[3] Changchang Wu. *Towards Linear-time Incremental Structure from Motion*. URL: http://ccwu.me/vsfm/vsfm.pdf.