Structure from Dense Paired Stereo Reconstruction

Davis Wertheimer, Leah Kim, James Cranston

{jamesc4, daviswer, leahkim}@stanford.edu

Abstract—Here's what this needs to include:

- **Title and authors**
- **Sec 1. intro: problem you want to solve and why**
- **Sec 2. technical part: how do you propose to solve it?**
- **Sec 3. milestones achieved so far**
- **Sec 4. remaining milestones (dates and sub-goals)**
- **References**

I. INTRODUCTION

We began our group project trying to use methods to create more accurate disparity maps, but changed our project focus to focus on an improved approaches to SFM stereo reconstruction. We are creating an algorithm to run SFM on stereo correspondences to build a reconstruction without prior information on point correspondences in the images. After meeting with Prof. Silvio Savarese on Thursday May 12, 2016 and getting advice, we agreed to proceed with the following three approaches. We will discuss these approaches in more detail in section II.

1) Given stereo pairs of cameras from fixed aperture cameras (rectified pairs) with known internal and external parameters, generate point clouds for each pair. These point clouds come from the correspondence algorithm with a sliding window. Then, use bundle adjustment on the point clouds to find 3D points and generate a reconstruction.
2) Find close pairs of images using minimum Euclidean distance. Then, rectify these pairs to make dense correspondences. With these correspondences, then run SFM on the dense correspondences to create a 3D reconstruction.
3) Take either option 1 or 2, then with these results, compare with the results of having known point correspondences from the start to see how well our algorithm performed absent of this information.

For this milestone, we are deciding to only complete option 2, which we will describe in more detail in the section.

II. TECHNICAL PART

Given a bunch of point clouds, it allows you to match them up. You may/may not need to know point correspondences

DAVIS: given two point clouds and a shared camera, just generate the the point correspondences. If we have this code and didn't want to do bundle adjustment, it wouldn't be hard to generate a best fit affine transformation to map the point clouds into each other.

III. MILESTONES ACHIEVED SO FAR

1) Construct pairs by forming a minimum spanning tree. (James)
2) Rectify the images for each pair. (Leah)
3) Build a dense correspondence for each pair, from lecture 6 page 23.
   - Triangulate the points
   - Given M, M', and two points in each image from the sliding window,
   - Need camera matrices for rectified images.
   - Output is a point in 3D
4) Run SFM: for every pair of pairs with a shared camera, define the point correspondences, do bundle adjustment using the point cloud, correspondences, and camera information.

IV. REMAINING MILESTONES

REFERENCES

[1] Aamir Khan , Muhammad Farhan, Asar Ali, *Speech Recognition: Increasing efficiency of SVMs*, 2011
[2] Dong Yu, Li Deng, Frank Seide, *The Deep Tensor Neural Network With Applications to Large Vocabulary Speech Recognition*, 2012
[3] Li, S. Z. *Content-based audio classification and retrieval using the nearest feature line method. IEEE Transactions on Speech and Audio Processing* 2000
[4] Lu, L., Zhang, H.-J., & Li, S. Z.*Content-based audio classification and segmentation by using support vector machines. Multimedia Systems*, 2003
[5] Huang, R., & Hansen, J. H. L. *Advances in unsupervised audio classification and segmentation for the broadcast news and NGSW corpora. IEEE Transactions on Audio, Speech and Language Processing* 2006
[6] Abdulrahman Alalshekmubarak *Towards a robust Arabic speech Recognition system based on reservoir computing*, 201
[7] http://recognize-speech.com/feature-extraction/mfcc