600.461 Computer Vision Progress Report

Noah Belcher and Juneki Hong and Roger Xu

I. INTRODUCTION

Our group has been working on the "Object Detection and Classification in a Video Sequence" project as the final project for this course. We have made some basic progress so far, but we still remain in the initial stages of this project. In other words, we have background subtraction working, and we are trying to stabilize the images in our video to improve our background subtraction.

November 14, 2011

A. Background Subtraction

From the image stack of the video provided, have gotten background subtraction working. It is a naive approach without us doing any sort of image stabilization, so the produced video is somewhat noisy; the buildings in the back and the crosswalk painted lines frequently appear. However, the vehicles passing by are pretty clearly defined and I think even from here, we could identify all the vehicles.

B. Image Stabilization

In order to get some nicer results, it would be nice if we could achieve some form of image stabilization. Currently, we plan on taking the sift features from every image and computing all the homographies required to warp every image to a single frame of reference.

We have currently computed the sift features for every image and we can compute the homographies between each of the images (using our implementations from homework 2). We can also compute and store all the homographies. With a little bit of more work, we should be able to properly warp each image with the corresponding homography. Hopefully, we will soon be able to produce an image stack full of stabilized images. If we can run our background subtraction on this, then we will have nicer and cleaner images to work with.

We have talked with a few other groups, and apparently they have had some trouble trying to get this approach to work. We are going to try out this method of image stabilization anyway because it might work out for us. However, if it doesn't, we will then try to stabilize our images using affine transformations.

Noah Belcher, `nbelche1@jhu.edu`
Juneki Hong, `jhong29@jhu.edu`
Roger Xu, `rxu8@jhu.edu`

C. Detection, Tracking, and Classification

Whether or not we sufficiently stabilize our images, we should be able to produce background subtraction footage, so that we can use it to detect, track, and classify moving objects.

1) Detection: After background subtraction, we should then begin trying to detect all the moving images in our footage. Because moving images in our background subtraction is white while the background is black, we would like to use Window Based Object Detection to detect moving objects and draw bounding boxes around them.

2) Tracking: After detection, we should then begin trying to track the moving objects in the video. Once again, we think it may be possible to use a modified version of Window Based Object Detection. We could assume that the bounding boxes around moving objects shouldn't move too much in between two frames of the video. We could say that if the bounding boxes of an object in two frames are really close together, then we could infer that the object is the same.

3) Classification: After detection and tracking, we should then begin trying to identify the moving objects. For classification, we would like to use the "Bag of Words" approach. We would like to maintain a database of categorized images, which is a hash table storing SIFT features paired with a classified image. While we are tracking an object, we could also query our Bag of Words and categorize it.

II. CONCLUSION

So far, we have only gotten background subtraction working and we also have some functionality for image stabilization. Ahead of us, we will still need to do object detection, tracking, and classification, but I think we do have an idea of how we are going to go about accomplishing these tasks.

REFERENCES

[1] Vedaldi, Andrea, and Brian Fulkerson. *VLFeat*, http://www.vlfeat.org/.