CS231A Project Proposal: Tracked Based Semi Supervised Learning

Brian Chung Stanford University bpchung@stanford.edu

Abstract

In this paper, we tackle the issue of accurate object classification using semi-supervised learning methods based on tracking information. Based on the earlier work of Teichman and Thrun, we modify the approach using LIDAR data to extract useful information from a single fixed camera source. Then we develop a tracking algorithm using semi-supervised learning to prepare for the fine input of an object classifier. Finally the classifier gives the classification of objects in the camera using important tracking information.

Our track data comes from background subtraction and segmentation of camera video data. Consequently, a few objects as well as their tracks are labeled and that data is used in training the object classifier.

The classifier will find a few positive examples, and the resulting objects are further segmented and fed forward to training the classifier. Thus, the classifier and tracking work iteratively.

Future Distribution Permission

The author(s) of this report give permission for this document to be distributed to Stanford-affiliated students taking future courses.

1. Introduction

There is currently a great need for high throughput classification of video data. Rather than manually label individual video frames, we seek to automate the process through segmentation and track classification free of large training sets. In particular, we separate the process of learning into three separate tasks: background subtraction/segmentation, tracking/labeling, and object classification.

Background subtraction can be done using the new background methods of the updated OpenCV library. Segmentation can be done using several state-of-the-art computer graphical libraries. After background subtraction and segmentation, we will design algorithm and develop code to implement the tracking of the segmented objects. Jinjian James Zhai Stanford University jameszjj@stanford.edu

Then a user will hand label a few of those tracks. The set of tracks are fed into a classifier which will also classify further objects as being part of that object class. This new information feeds back into the segmentation/tracking step to glean further class data.

As seen in Teichman's work, tracking based semi supervised learning is surprisingly resilient to noise and has an uncanny ability to learn new useful instances of the classes. Thus difficulties will arise in optimizing over intersections of tracks.

1.1. Background Subtraction/Segmentation

Accurate background subtraction will be vital to identifying proper tracks. Due to the fixed camera, we can obtain a background model easily from the data. We will begin with using Gaussian modeling for quick background subtraction. If needed, Shah's method of Bayesian object detection will be used.

1.2. Track Identification

Pending background subtraction, tracks of objects can be gleaned from each sample. These tracks will be given their appropriate labels.

1.3. Object Classification

The classifier will be an off-the-shelf classifier such as Felzenszwalb¹. Once new frames of the class are identified, the rest of the track is obtained.

1.4. Goal

We hope to obtain results similar to fully supervised learning, if not slightly worse. The benefit of this project will be in much quicker results and the ability to learn new instances.

¹ P. Felzenszwalb, D. McAllester, D. Ramaman. A Discriminatively Trained, Multiscale, Deformable Part Model. Proceedings of the IEEE CVPR 2008.