A Benchmark Study on Algorithms to Identify Ideal Models for Motion Detection

Si Chen & Bjoern Cheng Yi

Introduction

Today in computer vision, a lot of research power is focused on the applications behind self-driven automobiles. For example, Google has been involved in self-driving cars since 2009 and other companies such as Uber are following in its footsteps. Since the technology behind autonomous vehicles is relatively new, researchers face many challenges. One such problem is detecting incoming objects so that cars can slow down or turn away to avoid accidents.

This benchmark will focus on a key part of that challenge by evaluating algorithms' ability to detect motion by segmenting the image into clusters representing moving objects. In this project, we will extract the optical flow data from a video and then, segment the data to find the moving object using various algorithms. From those algorithms, the various algorithms will be analyzed with regards to their speed as well as their output's accuracy with respect to the ground truth annotations.

Algorithm

In this benchmark, we will compare different methods for tracking moving objects on optical flow data. The RGB optical flow frames will be used to compare the background subtraction and clustering-based segmentation algorithm.

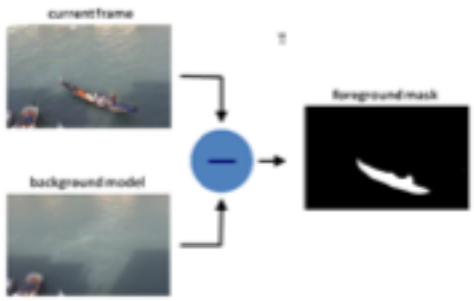
Background subtraction will allow for an image's foreground to be extracted. In this case, the foreground will be the moving object. The algorithm is able to do this by using initial frames of just the background to establish a model of the background that includes noise and slight shifts in a background. This model is then compared with new frames of the scene and subtraction occurs to show the difference. The prevalence of using background subtraction in video surveillance indicates its effectiveness as a method for identifying moving objects. However, this method has no understanding about the blobs discovered in the segmentation.

Since RGB optical flow will highlight areas of the frame that were moving at the same rate in the same color, k-means clustering based segmentation will allow us to segment the image based on color. Analysis of this algorithm will require us to test various k values and cluster initializations. These motion videos contain different moving objects and we are interested in seeing how the clusters will be able to update their location with respect to new objects in a scene. K-means clustering will cluster an image by color into different segments, but its adaptability is hindered by the need to choose a set number of clusters.

We will also test mean shift as a basis for segmentation. This algorithm approaches the problem of segmentation by homogenizing local groupings, replacing each pixel with the mean of the pixels in a diameter range that has a RGB value within a certain range. In this sense, it is very similar to k-means segmentation. The main differences lie in the fact that cluster locations do not need to be initialized and neither do the number of clusters, although they can be, by iterating over an image with mean shift until only a certain number of clusters remain.

The watershed algorithm is another common image segmentation algorithm that is commonly used for biomedical applications in identifying individual cells based on coloring. Watershed focuses on being able to separate overlapping segments into individual segments. It does so by identifying "ridges" in an image and separating basins created by the ridges into separate segments.

Our last approach will be a graph-based approach to image segmentation. This algorithm represents the problem of segmentation in terms of a graph where each node in the set of vertices corresponds to a pixel in the image and each edge connects pairs of neighboring pixels. Based on the property (ex. HSV value) of the pixels, the edges have varying weights, and pixels are classified based on the weighted edges into segments.

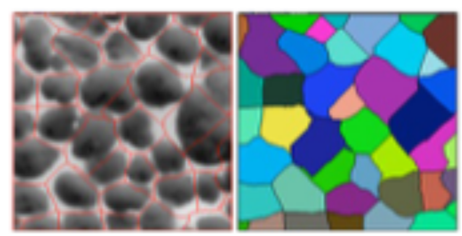


Background Subtraction

K-Means Segmentation

Mean Shift Segmentation

Graph-Based Segmentation



Watershed Segmentation

Dataset

The dataset that will be used in this project will be the video dataset from the Background Models Challenge (BMC). This dataset contains videos with manual annotations on given frames. It also provides over 120 initial frames of just the data, which will allow us to compare the background subtraction algorithm, which needs over 115 frames to build a model of the background, with the segmentation algorithms.

Performance Evaluation

For the clustering algorithms that require a number of clusters to be pre-chosen, we will evaluate them against one another with different numbers of clusters and different initializations of the cluster centers. From the best cluster number choice and initialization for the clustering algorithms, we will compare their accuracies and speeds to the other algorithms' accuracies and speeds. We will also evaluate the performance of the various algorithms on videos from high quality, which will be the videos from the dataset, to low quality, which will be the videos from the dataset that are compressed.

The BMC dataset we are using contains ground truth annotations that will allow us to calculate the accuracies of the motion clusters. We will use f-scores to evaluate the precision and recall of each algorithm in being able to identify the moving objects in the frame.