Background Subtraction in various situations

Sarthak Behera - 2018CS10384
Shreyans Nagori - 2018CS10390

September 2021

1 Introduction

Background subtraction is a very common task to detect moving objects in sequence of frames. Background subtraction helps in object detection which is a very popular application of computer vision. In recent years deep learning has transformed this paradigm but in this assignment, we use classical approaches to obtain the foreground mask for the sequence of frames. In order the subtraction to be robust, we propose various algorithms that are applicable in wide variety of conditions that may affect the frame sequence. Our aim is to have algorithms which are robust to various changes in camera position and other environmental changes. In order to accommodate all these invariance we use a set of algorithms for each specific category to alleviate common issues of that particular category.

2 Our approach

We have algorithms for each of the categories of the dataset that aims to address the unpredictability associated with each of the categories. We use novel ideas to having bitwise AND and OR in order to combine various disparate approaches and each of the approach complements the other through this method. We had also implemented SubSense which took a lot of time to run due to LBSP computation (we used for loops as numpy indexing was not possible) thus it was relegated later and instead we implemented ViBe which has ideas of SubSense except that it doesn't use local binary patterns and unlike SubSense, parameters are fixed. We have innovated on the ideas of these papers and have made models that build upon it.

3 Baseline

For the baseline of static background, we have used the KNN implementation of OpenCV along with the techniques of opening and closing transformations to reduce false negatives. We have also used Gaussian blur to remove noise in the image. These filters are suitably placed so that there effect is maximized. In order to further reduce false positives, we take bitwise AND with flow methods to only generate foreground mask where some flow is generated. Our mean IOU = 0.7496

4 Illumination

To make our model robust to illumination changes, we use optical flow methods to identify relevant areas where motion occurs and also use the KNN and MOG2 methods already implemented in OpenCV to get general shape of people moving on the foreground. Then optical flow removes the false negatives from the background models. This removal is implemented through a bitwise AND. The only issue was the shadows associated with each person otherwise the model is very robust and accurate. Our mean IOU = 0.4236

5 Jitter

We use OpenCV methods and use various filters and bitwise AND and OR to get to a model which is far more effective to get the foreground. We have also used Gaussian filters to reduce noise in images and to blur out some redundant edges. Morphological transformations like closing and opening has also been performed. Our mean IOU = 0.6899

6 Dynamic BGs

We take inspiration from ViBe which maintains background models and does random updates on models based on the foreground mask. We improve upon this idea by adding in our ideas of bitwise AND and OR to do custom background update based on the foreground mask that is generated through some other method. This minimizes false positive rate and we get more accurate foreground masks. Our mean IOU = 0.8001

7 PTZ

For pan-tilt-zoom category, we use ViBe-like approach to reduce the false positives and get a coherent foreground mask. Still the performance is not satisfactory as the foreground mask IOU has low value so further improvements should be possible. This category in my opinion is one of the more difficult scenarios to design robust algorithms. Our mean IOU = 0.1132

8 Our improvement on ViBe's approach

The approach in the algorithm ViBe is to maintain a fixed number of background models. We then perform random updates on randomly chosen background model if the the pixel is background and a random neighbour is also updated. This leads to temporal collection of information in the background models. Our improvement was to get the model which updates the background from other methods like KNN, MOG and flow instead of using the same result mask generated from ViBe. This has significant advantage whenever some pixel location is classified as foreground that pixel never gets updated but our approach mitigates that issue. We have tried this approach in PTZ and got some encouraging results still this model has large scope of improvement.

Link to result masks:- Here