Machine Learning Project Report

Topic: Foreground-Background Segmentation for Human tracking in Real-Time video.

1401001 - Deep Parekh
1401011 - Jeet Parekh
1401035 - Krima Doshi
1401061 - Maharshi Bhavsar

***Abstract*— The feature of detecting and separating moving objects from the background is required by cameras in many areas. The basic technique used for separating moving and static objects is background separation. Background separation will subtract the current frame from the reference frame of static objects.we had used a Gaussian filter, which further could be developed into a MOG system.Further the process is optimized through the concept of optical flow.**

I. INTRODUCTION

Motion is one of the major challenge faced by the researchers in the field of Computer Vision.

For background-foreground segmentation, we mainly tried 3 approaches-

Foreground-background segmentation is required when motion of objects need to be detected. The moving objects can be considered as foreground and the stationary objects can be considered as background.

Next, we used MOG filter for the initial separation. Then we took a reference frame and subtracted current frame from the reference frame. This approach works for the stable camera and a stable background and only a moving foreground. For the moving background-foreground segmentation we used a complex algorithm called Optical Flow.

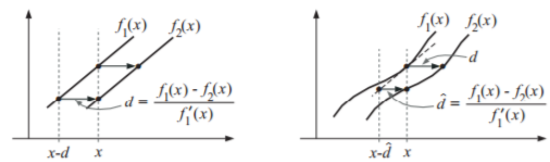
Optical Flow is one of the tool at rescue to solve many problems such as 3D shape acquisition, oculomotor control, perceptual organization, object recognition and scene understanding. Our problem statement is concerned with real time video sequence in which objects(Human) or the camera maybe moving according to 3D path. Optical flow provides visual perception. It takes a set of points in a frame and find those same points in another frame, and then find the vectors for each point and those vectors represents the direction and the magnitude of the change of that particular point with reference to the previous frame.

Applications include traffic supervision, gesture tracking, person tracking, model-based video coding, and content-based video retrieval. The process we followed to detect this motion is known as optical flow.

II. OPTICAL FLOW

A. Brief

Optical flow is implemented by studying the velocity of objects. The velocity is related to the space-time image derivatives at one image location using an equation often called the gradient constraint equation. n. If one has access to only two frames, or cannot estimate It, it is straightforward to derive a closely related gradient constraint, in which It(x, t) in (1.5) is replaced by I(x, t) I(x, t + 1) I(x, t).



The gradient constraint relates the displacement of the signal to its temporal difference and spatial derivatives (slope). For a displacement of a linear signal (left), the difference in signal values at a point divided by the slope gives the displacement. For nonlinear signals (right), the difference divided by the slope gives an approximation to the displacement.[3]

The question then arises what should be tracked from the image to compare the velocity. We need some characteristic of the object that can help identify it like features. A good feature has: Texture and Corners, i.e.

$$\begin{bmatrix} \sum_{neighbourhood} \left(\frac{\partial I}{\partial x}\right)^2 & \sum_{neighbourhood} \frac{\partial^2 I}{\partial x \partial y} \\ \sum_{neighbourhood} \frac{\partial^2 I}{\partial x \partial y} & \sum_{neighbourhood} \left(\frac{\partial I}{\partial y}\right)^2 \end{bmatrix}$$

The higher the eigen value obtained, the better is the feature detected. Optical flow can be implemented with the help of many such feature detection methods such as Lucas Kanade. The Lucas Kanade algorithm returns a high pixel difference score or a state of rapid change for the target.

B. The Algorithm

Optical flow can be implemented by following the below mentioned steps:

1) Compute the intensity of each pixel
2) For each pixel position compute the gradient matrix and store an eigenvalue of matrix
3) Separate the high scoring pixels by flag matrix F and region size k and flag region size f
4) Take the top n eigen values and use those for the trackable features
5) The Gaussian random distribution is applied for speed base

6) Next, warp one image, take derivatives of the other so you dont need to re-compute the gradient after each iteration
7) Repeat until complete

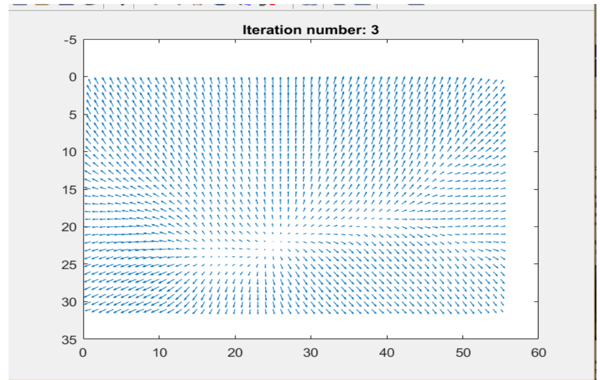
Note: Instead of eigenvalues, we will simply down scale the image so as to decrease the computational cost as well as matrix size so the there would be no need of selecting the higher eigenvalues, i.e. all eigenvalues will be considered.

III. Results

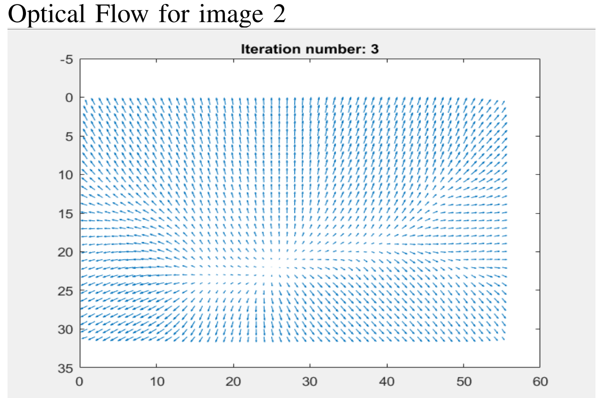
A. Human Tracking in real time video

These images above are snapshots taken from a real time surveillance video. The human figure detected is enclosed inside a box.

B. Optical Flow



Optical Flow for image 2



Optical Flow for image 3

IV. Literature Review

1) Sri Devi Thota, Kanaka Sunanda Vemulapalli, Kartheek Chintalapati, Phanindra Sai Srinivas Gudipudi - Comparison Between The Optical Flow Computational Techniques[2] This paper talks compares the results obtained by the application of two major optical flow algorithms on different sets of image sequences. It also discusses the applications of optical flow in vehicle detection and tracking. The basic method is Gradient method, but it cannot give a complete solution. So two other differential methods - Horn-Schuck and Lucas-Kanade algorithms are discussed.

2) David J. Fleet, Yair Weiss - Optical Flow Estimation[3] This paper gives a walkthrough for gradient-based optical flow estimation. It also discusses topics like least-squares and robust estimators, iterative coarse-to-fine refinement, different forms of parametric motion models, different conservation assumptions, probabilistic formulations, and robust mixture models.

3) Thomas Brox, Jitendra Malik, Fellow, IEEE - Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation[4] Optical flow estimation requires dense sampling in real time. Coarse-to-fone warping schemes somewhat relax this constraint. But still there is a dependency between the

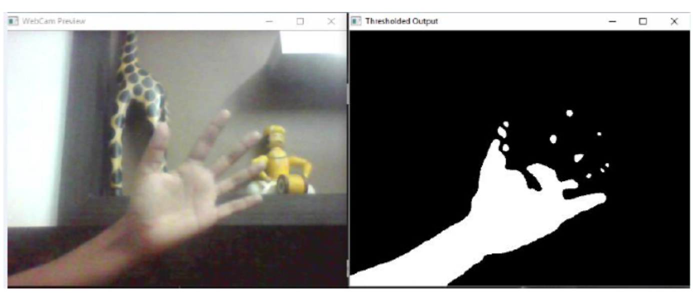
scale of structures and their velocity. This particularly renders the estimation of detailed human motion problematic, as small body parts can move very fast. In this paper, a way to approach this problem is discussed by integrating rich descriptors into the variational optical flow setting. This way a dense optical flow field can be estimated with almost the same high accuracy as known from variational optical flow, while reaching out to new domains of motion analysis where the requirement of dense sampling in time is no longer satisfied.

V. Discussion

Our initial attempts were in MATLAB to perform the foreground-background segmentation.But since the MATLAB camera module is not easily available, we started parallel attempts using python and Open CV.

In our previous approach, After each specific no. of iteration, a reference frame is being updated. And we are detecting the motion by differentiating the current frame and that reference frame. So once the object comes into the picture with some motion, it will get detected. But when it stops moving, the reference frame will be updated and eventually the current frame and the reference frame will be the same. So no motion will be detected.

Below is the web cam output when the moving object comes into frame and Filtered output which also detects only a moving object.



These methods are not capable of handling multiple backgrounds like moving water in a fountain. Hence, other approaches were taken into consideration.

Currently, we have implemented the concept of optical flow.It refers to the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer and a scene. Motion vectors, represents flow of a moving object, are obtained using Lucas - kanade optical flow algorithm for moving object detection with complex background. These flow vectors are quantized using a predefined threshold to decide whether a pixel is a part of an object or a background.We are successful in our attempt to improve the results obtained from the previous approaches.

VI. Further Planning

We are still trying to make our results more accurate by making the algorithms more efficient.The output of human detection can be optimized further by increasing the accuracy of the algorithms used.There are still some objects detected which are non-human from the surveillance video which need to be corrected further. We are also trying to implement different algorithms which can successfully detect the human figures.

References

[1] Kyungnam Kim, Thanarat H. Chalidabhongse, David Harwood, Larry Davis - Real-time foregroundbackground segmentation using codebook model $1077-2014/-r$ 2005 $Elsevier\ Ltd.\ All\ rights\ reserved.\ doi$: $10.1016/j.rti.2004.12.004$

[2] Sri Devi Thota, Kanaka Sunanda Vemulapalli, Kartheek Chintalapati, Phanindra Sai Srinivas Gudipudi - Comparison Between The Optical Flow Computational Techniques. [online] Available: http://ijettjournal.org/volume-4/issue-10/IJETT-V4I10P142.pdf

[3] David J. Fleet, Yair Weiss - Optical Flow Estimation. [online] Available: http://www.cs.toronto.edu/ fleet/research/Papers/flowChapter05.pdf

[4] Thomas Brox, Jitendra Malik, Fellow, IEEE - Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation. [online] Available: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/papers/bm$_pami2010.pdf\ DavidStavens, TheOpenCV\ Library$: $ComputingOpticalFlow.[online]http$: $//robots.stanford.edu/cs223b05/notes/CS\%20223-B\%20T1\%20stavens_opencv_optical_flow.pdf$

[5] C. Liu. Beyond Pixels: Exploring New Representations and Applications for Motion Analysis. Doctoral Thesis. Massachusetts Institute of Technology. May 2009.