Review of Optical Flow Technique for Moving Object Detection

Anshuman Agarwal
HBTI Kanpur
Email: anshuman.hbtik@gmail.com

Shivam Gupta
HBTI Kanpur
Email: sadam9099@gmail.com

Dushyant Kumar Singh
Assistant Professor
MNNIT Allahabad
Email: dushyant@mnnit.ac.in

***Abstract*—Object detection in a video is a challenging task in the field of image processing. Some applications of the domain are Human Machine Interaction (HMI), Security and Surveillance, Supplemented Authenticity, Traffic Monitoring on Roads, Medicinal Imaging etc. There happens to be a number of methods available for object detection. Each of the method has some constraints on the kind of application it has been used for. This paper presents one of such method which is termed as Optical Flow technique. This technique is found to be more robust and efficient for moving object detection and the same has been shown by an experiment in the paper. Applying optical flow to an image gives flow vectors of the points corresponding to the moving objects. Next part of marking the required moving object of interest counts to the post processing. Post processing is the legitimate contribution of the paper for moving object detection problem. This here is discussed as Blob Analysis. It is tested on datasets available online, real time videos and also on videos recorded manually. The results show that the moving objects are successfully detected using optical flow technique and the required post processing.**

***Index Terms*—Optical flow, moving object detection, tracking, morphological operation, blob analysis.**

I. Introduction

Object detection and tracking is one of the most challenging tasks in digital image processing and it has many applications in Computer Vision [1]. In this paper the concept of optical flow [2, 3] for the motion detection presents an apparent change of moving object location between two frames. It insulates the moving objects from the static background objects. Optical flow estimation yield a two-dimensional vector field i.e. motion field that represent velocities of each point of an image sequence [4]. Optical flow estimation is useful in many applications. Some examples are Vehicles Navigation [4], Video Image reconstruction, Traffic Surveillance and object tracking [5]. Due to higher detection accuracy of optical flow method, motion parameters of moving objects are generated which results in avoiding any overlapping of different moving objects. The proposed algorithm initially takes the video frames as input one by one estimates the average flow vectors from them which results in Optical flow vectors. Noise filtering is done to remove the unwanted motion in the background. Then thresholding is done to achieve binary image. There are some uneven boundaries in threshold image which are rectified by morphological operations. Connected components are analyzed to evenly patch the generated white blobs in binary image. Finally, marking of moving object is done with a box which indicates the motion of the objects individually. Optical flow method has been preferred because of its low complexity and high accuracy [6].

II. Technical Review

Images are the combination of pixels which are spread around on the window in a regular pattern and that each point in a pixel has an intensity value that comprises an image. People can observe the image by many characteristics of it for detecting the object in image. For machine, an image is a two dimensional array of pixel intensities. So techniques are devised to achieve this goal of object detection.

Many numbers of techniques has been proposed for object detection in literature. Many researches discuss the problem of object detection specifically human detection and its usage for event categorization and other tasks. Here, study is restricted to concept of detecting objects those are in motion with respect to the background.

There were many algorithms proposed for the above tasks which are listed below:

- Frame differencing approach
- Viola Jones algorithm
- Skin color modeling

In an image a distinct boundary that separates two homogenous regions is taken as an edge. Frame differencing [7] and Edge Detection algorithm [8] subtracts the two consecutive frames based on these edges. If the difference comes out to be non-zero values, it is considered to be moving. But it has some limitations that during capturing the video due to the movement in air or any other source might cause the disturbance in the position of the camera resulting into the false detection of the non-moving objects [7].

The Viola-Jones algorithm [9] uses Haar-like features that are scalar product between the image and some Haar-like templates. Although it can be trained to detect a variety of object classes, it was motivated primarily by the problem of face detection [10]. But it has some limitations like the detector is most effective only on frontal images of faces and it is sensitive to lighting conditions.

The preliminary steps in skin detection [11] are the representation of image pixels in colour spaces, suitable distribution of skin and non-skin pixels, and after that skin colour [10] modelling. According to skin colours distribution characteristics on colour space, skin colour pixels can be detected quickly with skin colour model. But it has obvious disadvantage like skin colour also varies from one person to another belonging to different ethnic groups and from persons across different regions.

III. Optical Flow Technique

The optical flow method [12-14] describes the detection of moving objects by using the concept of flow vectors to detect motion in two subsequent frames of a video. It is known that vectors both have a magnitude and direction so the moving parts of the two consequent images are segmented from the static background based on zero and non-zero values of flow vectors. The cardinal surroundings are in a three-dimensional coordinate system with time as an extra variable. But for simplicity it is converted into a two-dimensional coordinate system with time as the third variable.

The intensity of a pixel in the coordinate plane is represented as I(x, y, t). This is brightness function of a point at particular time instance. The key assumptions that help in calculating the velocity vector of any point are Brightness Constancy and Spatial Coherence. Bright Constancy means that the brightness/intensity of a point doesnt change. Spatial Coherence means that there is very small movement of a point in two consecutive frames. These considerations derive an expression given as:

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (1)$$

Using Taylor series for the right hand part of equation (1) obtain:

$$I(x+\delta x, y+\delta y, t+\delta t) = I(x, y, t) + \frac{dI}{dx}\delta x + \frac{dI}{dy}\delta y + \frac{dI}{dt}\delta t + H.O.T \quad (2)$$

From (1) and (2), with neglecting higher order terms (H.O.T.) and after modifications, we get

$$I_x.V_x + I_y.V_y = -I_t \quad (3)$$

Or in formal vector representation

$$\nabla I^T.\overrightarrow{V} = -I_t \quad (4)$$

Where∇I is so-called the spatial gradient of brightness intensity and $\overrightarrow{V}$ is the optical flow velocity vector of the image pixel and It is the time derivative of the brightness intensity [7].

Equation (4) is most important for optical flow calculation and is called 2-D Motion Constraint equation or gradient Constraint. It represents one equation with two unknown quantities (the aperture problem) [12].

There are two most used methods to solve the problem are given below:

- Lucas-Kanade [1, 3]
- Horn-Schunck [1, 3]

IV. Optical Flow based Object Detection

The image with optical flow vectors obtained after applying the Optical Flow algorithm is used for marking the detected object in the image/video frame. This whole process is termed as post processing. The flow diagram in figure 1 shows the complete human detection with later part of post processing. Steps of the later part i.e. post processing are detailed in subsections.

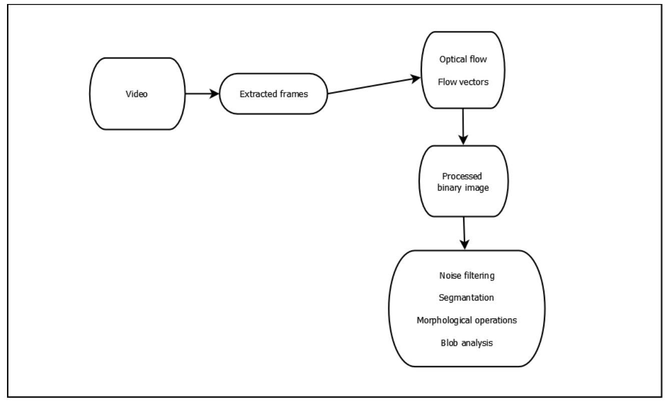

Fig. 1. Process of Object Detection using Optical Flow Technique

A. Noise Filtering

Video captured with a camera could have variety of noises that may not be visible in the original video but are clearly seen in the processed binary image. This is an obstruction in detecting moving objects.
Therefore the need emerges for implementing smoothing techniques that eliminate these noises out of the binary frame for image segmentation. To remove the noises, the sources of these should be known which are:

1) Slight motion of the camera due to which the still background seems to be moving.
2) Motion in the background due to environmental conditions like wind, rain, etc.

A class of filters that fulfills these requirements is called signal filters, so filtering is used to remove these noises. The major objective of these algorithms is to identify areas of homogeneity and eliminate them.
Signal filter belong to the class which are helpful to detect and suppress these noises. It is wide range and one of its types is the median filters. The median filter replaces the value of the processing pixel by the median value of the block of pixels that is declared previously. Median filters are an excellent choice as the effectively remove the unnecessary noises and smoothen the binary image which can be processed easily further.

B. Image Segmentation

To detect an object, image is normally segmented into blobs or regions using some common segmentation techniques such as background subtraction mean shift clustering and graph cuts. Segmented regions are then grouped together to represent an object based on some deterministic rules. In object detection algorithm, the content of each frame is read and the background is estimated. The interested objects are

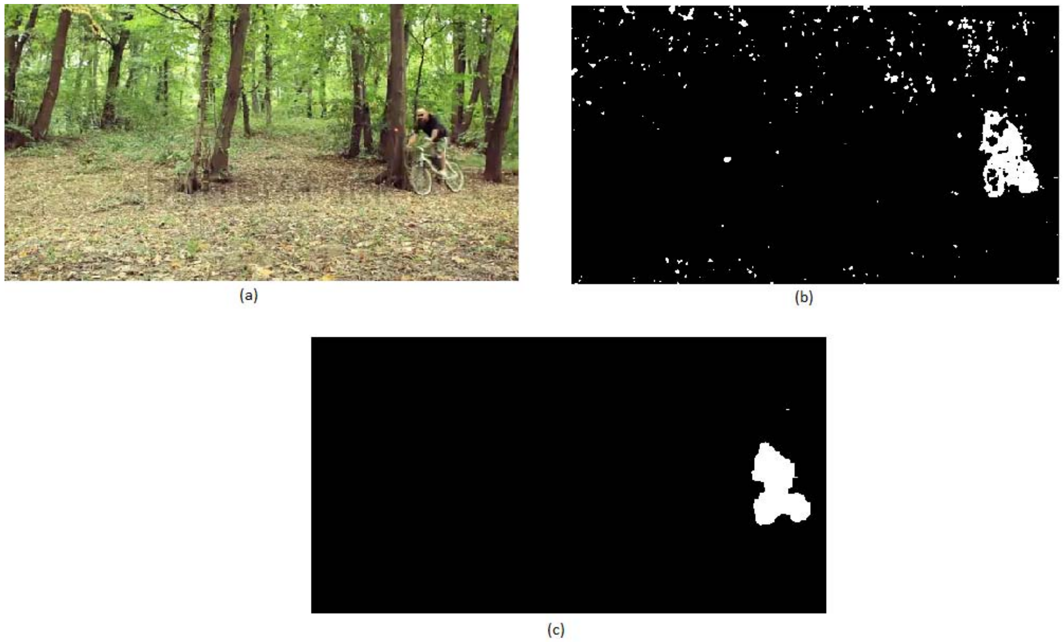

Fig. 2. Image results (a) Original Image (frame) (b) Binary image without morphological close(c) Binary image after applying dilation and erosion.

tagged by eliminating the background.
Thresholding function is applied for the conversion of the gray image to binary and consequently the moving objects are separated distinctly. So after applying filtering and the thresholding function, the moving parts of the background detected as objects are removed.

C. Morphological operations

In the binary image of the moving object obtained there are some improper boundaries, so Morphological operations are performed to extract significant features from those images that are useful in the representation and description of the shapes in those region. It is widely used in image segmentation and pattern recognition. Morphological closing and eroding is done respectively to remove portions of the objects which are unnecessarily detected as moving.

D. Blob Analysis

Blob analysis is done to find the exact object in the processed binary image. There can be white patches in the binary image which do not corresponds to the actual object. Such patches are due to the unwanted noise occurred due to movements in background. So, only those blobs which satisfy the moving objects characteristics are considered and all remaining unnecessary blobs are removed. This is performed by placing the threshold on size and aspect ratio of the blob. Size is the total connected pixels in the blob. Aspect ratio is the ratio of height to the width of the blob.

V. Experimental Results

Experiment is performed on a video taken from web, of a man is riding a cycle. Algorithm is as well tested on real-time videos and other manually recorded videos. The figure 2 & 3 well presents the image results of the experiment.

Before applying optical flow estimation on frames, the image format is converted from RGB to gray. Figure 2 shows morphological operations result for getting more appropriate shape of the object. Figure 2(a) is the frame extracted from the video. Figure 2 (b) is the median filtered binary image generated optical flow. Figure 2(c) is result after erosion and dilation.
The algorithm continues with blob analysis to detect specific objects of interest, and remove unnecessary blobs based on the size and aspect ratio. The size here is interpreted as the density or number of pixels in the blob. For aspect ratio connected component analysis is done, which the number of pixels that are connected to each other. The blobs after this stage are surrounded by box. Images in figure 3 present the results at different stages in the process.

A. Result Analysis

The performance analysis of the optical flow based object detection is done using slandered measurement parameters of True Positive Rate (TPR), False Positive Rate (FPR), Precision and Recall. However recall is nothing but TPR and FPR is measure for Fall-out.

$$TPR = TP/(TP + FN) \quad (5)$$
$$FPR = FP/(FP + TN) \quad (6)$$
$$Precision = TP/(TP + FP) \quad (7)$$

Where TP, FP, TN & FN are defined in compliance with problem as,

- TP is number of frames in which object detected when actually it is present

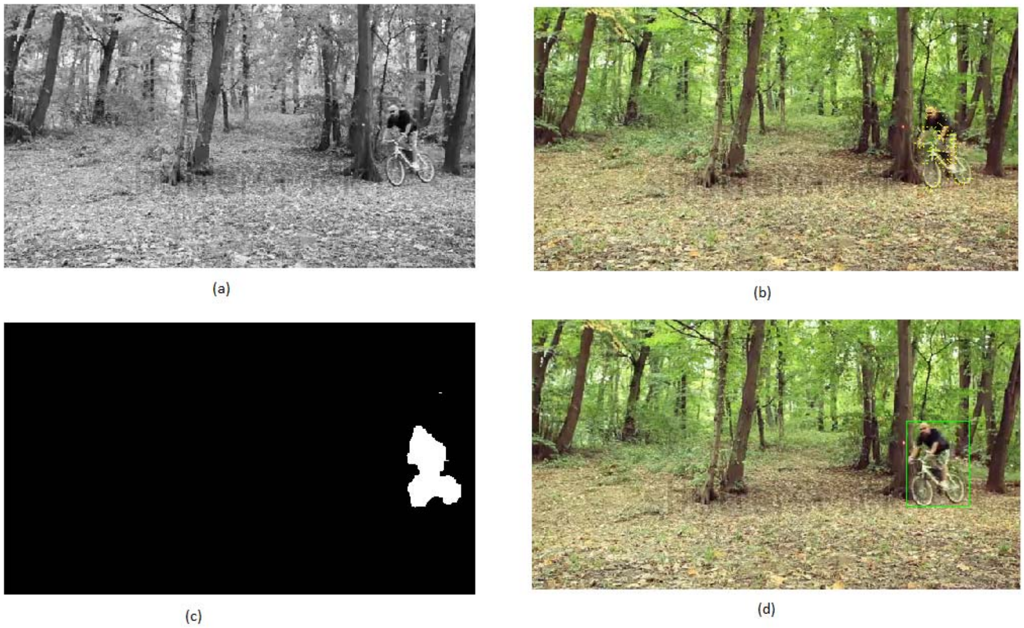

Fig. 3. (a) Gray Scale Image (b) Optical Flow Motion Vectors (c) Binary Image with Blobs (d) Tracking Moving Object with Bounding Box

TABLE I
TPR, FPR AND PRECISION CALCULATION

Video name	Number of frames	True Positive(%)	False Positive(%)	False negative(%)	True Negative(%)	Precision	TPR(%)	FPR(%)
Record.mp4	75	88.89	6.11	11.11	93.89	93.47	88.89	6.11
Car.mp4	100	78.75	9.58	21.25	90.42	89.15	78.75	9.58
Vissiontraffic.mp4	200	86.73	7.50	13.27	92.50	92.04	86.73	7.50

- FP is number of frames in which object detected when actually it is not present
- TN is number of frames in which object not detected when actually it is not present
- FN is number of frames in which object not detected when actually it is present

The experiment was done for 3 different videos and performance results for the same have been tabulated in Table 1.

VI. CONCLUSION

The algorithm in this paper is able to detect and track the moving object in the sequence of video frame taken from the static camera in any kind of background and terrain. In every subsequent frame initially the average flow vectors are estimated and then the generation of optical flow vectors takes place. For the better accuracy of the detection morphological erosion and dilation is performed. Lucas-Kanade has been chosen for the estimation of optical flow because of its high accuracy and its basic principle that uses the change of intensity between two consecutive video frames for motion detection. Now the filtering is done to smooth out the boundaries of the moving object using median filters. Ultimately, the algorithm will detect only those moving objects that will satisfy the restrictions applied on the blob areas rest will remain as undetected. The paper work can be extended using the identified centroid of the moving object the distance and velocity can be calculated which helps in finding the speed of the moving object from the video sequence.

REFERENCES

[1] Revathi, R., and M. Hemalatha. "Certain Approach of Object Tracking using Optical Flow Techniques." International Journal of Computer Applications53.8 (2012).

[2] Klappstein, Jens, et al. "Moving object segmentation using optical flow and depth information." Pacific-Rim Symposium on Image and Video Technology. Springer Berlin Heidelberg, 2009.

[3] Kale, Kiran, Sushant Pawar, and Pravin Dhulekar. "Moving object tracking using optical flow and motion vector estimation." Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), 2015 4th International Conference on. IEEE, 2015.

[4] Chaple, Manisha, and S. S. Paygude. "Vehicle Detection and Tracking from Video Frame Sequence." International Journal of Science & Engineering Research 4.3 (2013): 1-7.

[5] Singh, Dushyant Kumar, and Dharmender Singh Kushwaha. "Tracking Movements of Humans in a Real-Time Surveillance Scene." Proceedings of Fifth International Conference on Soft Computing for Problem Solving. Springer Singapore, 2016.

[6] Han, Hong, and Minglei Tong. "Human detection based on optical flow and spare geometric flow." Image and Graphics (ICIG), 2013 Seventh International Conference on. IEEE, 2013.

[7] Zhan, Chaohui, et al. "An improved moving object detection algorithm based on frame difference and edge detection." Image and Graphics, 2007. ICIG 2007. Fourth International Conference on. IEEE, 2007.
[8] Kumar, S. Hari, and P. Sivaprakash. "New approach for action recognition using motion based features." Information & Communication Technologies (ICT), 2013 IEEE Conference on. IEEE, 2013.
[9] Ikizler-Cinbis, Nazli, and Stan Sclaroff. "Object, scene and actions: Combining multiple features for human action recognition." European conference on computer vision. Springer Berlin Heidelberg, 2010.
[10] Zhou, Dongxiang, and Hong Zhang. "Modified GMM background modeling and optical flow for detection of moving objects." 2005 IEEE International Conference on Systems, Man and Cybernetics. Vol. 3. IEEE, 2005.
[11] Yamaguchi, Koichiro, Takeo Kato, and Yoshiki Ninomiya. "Vehicle ego-motion estimation and moving object detection using a monocular camera."18th International Conference on Pattern Recognition (ICPR'06). Vol. 4. IEEE, 2006.
[12] Aslani, Sepehr, and Homayoun Mahdavi-Nasab. "Optical flow based moving object detection and tracking for traffic surveillance." International Journal of Electrical, Electronics, Communication, Energy Science and Engineering 7.9 (2013): 789-793.
[13] Denman, Simon, Clinton Fookes, and Sridha Sridharan. "Improved simultaneous computation of motion detection and optical flow for object tracking." Digital Image Computing: Techniques and Applications, 2009. DICTA'09.. IEEE, 2009.
[14] Shirgeri, S., Naik, P.U., Udupi, G.R. and Bidkar, G.A., 2013. Design and development of Optical flow based Moving object detection and tracking (OMODT) system. Int. J. Comput. Eng. Res, 3(4).
[15] Gonzalez, Rafael C., and E. Richard. "Woods, digital image processing." ed: Prentice Hall Press, ISBN 0-201-18075-8 (2002).