Optical Flow Based Multiple Object Detection and Tracking System in Static and Dynamic Environment

Smitha Gayathri.D[1], Dattatreya K. A.[2]

1- M.E, Communication System student, Department of Electronics and Communication Engineering
2- Professor, Department of Electronics and Communication Engineering Adhiyamaan College, Hosur.

Abstract: *Estimation and tracking of multiple moving objects in a dynamic environment is a challenging task in computer vision. It demands real-time computation with reliable accuracy on embedded processors. This project focuses on the identification, analysis and implementation of suitable single and multiple moving object detection and tracking algorithm which makes use of spatiotemporal information.*

In this Thesis, the spatiotemporal information is obtained using Optic Flow computation. Optical Flow is a vector field which gives both magnitude and direction of the pixel movements with respect to consecutive frames. Horn and Schunck algorithm is used in this work for computation of Optical Flow vectors. Single and multiple objects based on the computed vectors are segmented using intensity based thresholding. Image enhancement techniques consisting of morphological operations and pixel connectivity are applied to enhance the segmented objects. The designed tracking algorithm is implemented MATLAB. The developed algorithm is first validated and its parameters tuned using real and virtual data with static background. For tracking in dynamic background the segmented objects are identified and the user is asked to choose one of the objects to track. A centroid based approach is used to track the chosen object.
Key Words: Optical Flow, Segmentation, Threshold and Centroid

Nomenclature

$I(x,y)$	Intensity
E_x	Intensity in x direction
E_y	Intensity in y direction
E_t	Intensity in t frame
u,v	Vectors obtained from Optical Flow
α	Regularization term

Abbreviations

HS	Horn and Schunck
CLG	Combined Local and Global Method
ROI	Region of Interest

I.   Introduction

The goal of computer vision is to make useful decisions about physical objects and scenes based on sensed images. In order to make decisions about real objects it is almost always necessary to construct some description or model of them from the image [1]. Computer vision is the combination of hardware and software algorithms, the vision obtained through hardware been analyzed through algorithms and the necessary information been extracted.

Multiple object detection and tracking involves in major constraints like object detection, motion estimation, object segmentation and object tracking.

An object detection algorithm must have the ability to recognize the variation in visual appearance [2]. Motion estimation is the process to find the motion vectors of a real world 3D scene captured projected onto 2D plane. Detection and motion estimation of object is based upon comparing two frames for finding moving objects. Segmentation refers to segregating the objects detected in the image in groups depending on its size, shape and various characteristics. By doing a segmentation on the object which has movement in the given frame that particular object been tracked.

There are two main problems to be solved, estimation of the object and data association with the object [3].Background subtraction needs background models to detect an object [4][5][6][7]. Optical Flow algorithms are used to track objects in static and dynamic background and are robust to noise [8][9] [10][11].Optical Flow method is implemented, through MACH filter, log r-θ mapping to track an object [6] and also polar-log images is proposed as an effective method to track multiple objects [7]. Colour and blob tracking been done through Bayesian tracking methods [12] and also by using multiple stationary and moving cameras [13].

They are complexities involved in tracking object in dynamic camera scenes. The complexity involved in object detection in dynamic environment is foreground and background image data changes in successive frames which makes difficult to detect object.

In this paper object detection is implemented using Optical Flow algorithm. The obtained vectors are segmented using intensity based threshold technique. The segmented object tracked based on centroid approach.

1.1 Optical Flow

Optical Flow is the distribution of apparent velocities of movement of brightness patterns in an image. Optical Flow can arise from the relative motion of objects and the viewer [10].

Optical Flow is an approximation of the local image motion based upon local derivatives in a given sequence of images. That is, in 2D it specifies how much each image pixel moves between adjacent images while in 3D it specifies how much each volume voxel moves between adjacent volumes [11]. The 2D image sequences used here are formed under perspective projection via the relative motion of a camera. Scene objects Optical Flow vectors are obtained based on the estimation of the pixels intensity motion in the corresponding frame and apparent velocity of the pattern [11]. The different methods in Optical Flow algorithm are differential method, frequency method and correlation method.

Horn and Schunck is a differential technique used to compute image velocity from spatiotemporal derivative of image intensities. There are mainly two differential techniques are local and global method. Global method finds out the vectors on entire image and does normalization on the vectors to obtain dense flow of vectors. On other hand local methods are compared in locality by fixing its neighborhood size and processed to obtain vectors. These algorithms give more dense flow vectors, robust to noise and sensitive for small intensity change in the successive images. The computation of differential Optical Flow is a two-step procedure:
1. Measure the spatio temporal intensity derivative (which is equivalent to measuring the velocities normal to the local intensity structures) [11].
2. Integrate normal velocities into full velocities, for example, either locally via a least squares calculation or globally via a regularization [11].

Optical Flow vectors been calculated through apparent velocities of movement of brightness pattern in an image. The change in pixel intensity in consecutive frames of an object is zero and equation given as [10]

$$\frac{dE(x,y,t)}{dt} = 0 \qquad (1)$$

Consider the intensity of a pixel at time t is $I(x,y,t)$ and at time Δt the intensity is $I(x+\Delta t, y+\Delta t, t+\Delta t)$. The rate of change of motion in x and y direction is given as [12] $(\frac{dx}{dt}, \frac{dy}{dt})$.

1.2 Horn and Schunck Algorithm

The intensity at time mathematically given by $I(x,y,t)$, where I is the intensity of object at position (x, y). The Horn and Schunck algorithm computes the velocity $\bar{v} = (u,v)$ for each pixel in an image by minimizing the functional given by [12].

$$F = \int_{\Omega} \left[\alpha^2 E_c^2 + E_b^2 \right] d\Omega, \qquad (2)$$

Where α is regularization parameter, E_b^2 and E_c^2 are the data and regularization term respectively given by [12]

$$E_b^2 = (I_x u + I_y v + I_t)^2, \qquad (3)$$

$$E_c^2 = (u_x)^2 + (u_y)^2 + (v_x)^2 + (v_y)^2 = \|\Delta \bar{v}\|^2 \qquad (4)$$

In (3), $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$ and $I_t = \frac{\partial I}{\partial t}$.

In (4),

$$u_x = \frac{\partial u}{\partial x}, \ u_y = \frac{\partial u}{\partial y}, \ v_x = \frac{\partial v}{\partial x} \text{ and } v_y = \frac{\partial v}{\partial y}.$$

A. Estimating the Partial Derivatives

We must estimate the derivatives of brightness from the discrete set of image brightness measurements available. Horn and Schunck method is a global Optical Flow method does smoothening on the computed flow field given by

$$E_{HK} = \iint ((f_x u + f_y v + f_t)^2 + \alpha(|\Delta v|^2 + |\Delta u|^2)) dx dy \quad (5)$$

It is important to estimate f_x, f_y and f_t be consistent. The values are obtained through partial derivatives of image as

$$f_x = \frac{1}{4}\{E_{i,j+1,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i+1,j,k}$$
$$+ E_{i,j+1,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i+1,j,k+1}\} \quad (6)$$

$$f_y = \frac{1}{4}\{E_{i+1,j,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i,j+1,k}$$
$$+ E_{i+1,j,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i,j+1,k+1}\} \quad (7)$$

$$f_t = \frac{1}{4}\{E_{i,j,k+1} - E_{i,j,k} + E_{i+1,j,k+1} - E_{i+1,j,k}$$
$$+ E_{i,j+1,k+1} - E_{i,j+1,k} + E_{i+1,j+1,k+1} - E_{i+1,j+1,k}\} \quad (8)$$

B. Estimation the Laplacian of the Flow Velocities

The Laplacian of u and v vectors are given by [1]

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \text{ and } \nabla^2 v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}$$

To obtain more dense flow fields the local averages of *u* and *v* vectors between the pixels are approximated using

$$\bar{u}_{i,j,k} = \frac{1}{6}(u_{i-1,j,k} + u_{i,j+1,k} + u_{i+1,j,k} + u_{i,j-1,k}$$
$$+ u_{i-1,j-1,k} + u_{i-1,j+1,k} + u_{i+1,j+1,k} + u_{i+1,j-1,k}) \quad (9)$$

$$\bar{v}_{i,j,k} = \frac{1}{6}(v_{i-1,j,k} + v_{i,j+1,k} + v_{i+1,j,k} + v_{i,j-1,k}$$
$$+ v_{i-1,j-1,k} + v_{i-1,j+1,k} + v_{i+1,j+1,k} + v_{i+1,j-1,k}) \quad (10)$$

The new set of velocity estimates are calculated from previous vector values to smoothen the obtained vectors values given by [1]

$$U^{n+1} = U_{avg} - \frac{E_x[E_x U_{avg} + E_y U_{avg} + E_t]}{(\alpha^2 + E_x^2 + E_y^2)} \quad (11)$$

$$V^{n+1} = V_{avg} - \frac{E_x[E_x V_{avg} + E_y V_{avg} + E_t]}{(\alpha^2 + E_x^2 + E_y^2)} \quad (12)$$

1.3 The Combined Local – Global (CLG) Method

Local and global differential methods have complementary advantages and short comings. It would be interesting to construct a hybrid technique that constitutes the advantages of two methods. It combines the robustness of local methods with the density of global approaches. Lucas kanade algorithm is designed based on the foundation considering the flow is essentially constant in a local neighborhood by least square criterion. Lucas and Kanade proposed to assume that the unknown optic flow vector is constant within some neighborhood of size ρ. In this case it is possible to determine the two constants 'u' and 'v' at some location (x, y, t) from a weighted least square fit by minimizing the function given by equation 5.14. Equation 5.11 represents the

$$W := (u, v, 1)^T, \quad (5)$$

$$|\nabla w|^2 := |\nabla u|^2 + |\nabla v|^2, \quad (5.12)$$

Equation 5.12 is the regulariser used to fill in the information from the neighbourhood where no reliable local estimate is possible. '*u*' and '*v*' are vectors where *w* is regulariser.

$$\nabla_3 f := \left(f_x, f_y, f_z\right)^T, \qquad (5.13)$$

In equation 5.13, f_x, f_y, f_z represents the partial derivate of the image *I* with respect to position *x*, *y* and *t*.

$$J_p\left(\nabla_3 f\right) := K_p * \left(\nabla_3 f \nabla_3 f^T\right) \qquad (5.14)$$

In equation $J_p\left(\nabla_3 f\right)$ is the second moment matrix derived from the gradient, K_p is the kernel mask for the size of the neighbourhood size of ρ.

$$E_{LK}(w) := W^T J_p\left(\nabla_3 f\right) W \qquad (5.15)$$

Equation 5.15 is the Lucas Kanade method quadratic form obtained from equation 5.11-5.14.

From Horn and Schunck it is evident that the intensity is given by equation

$$E_{HK} = \int_\Omega w^T J_0\left(\nabla_3 f\right) w^2 + \alpha \left(|\Delta u|^2 + |\Delta u|^2\right)) \, dx \, dy \qquad (5.16)$$

By replacing the matrix $J_0\left(\nabla_3 f\right)$ by structure tensor $J_\rho\left(\nabla_3 f\right)$ on combining both local and global intensity levels the obtained equation is

$$E_{CLG} = \int_\Omega w^T J_\rho\left(\nabla_3 f\right) w^2 + \alpha \left(|\Delta u|^2 + |\Delta u|^2\right)) \, dx \, dy \qquad (5.17)$$

1.4 Segmentation

The Optical Flow vectors obtained from the algorithm are segmented through threshold technique. The concurrent mean of the image and the pixel value is changed iteratively to the size of the image M x N.

$$\text{Mean} = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} k \qquad (21)$$

$$k(i, j) = \begin{cases} h = 1 \ge mean, i = 0,1....M-1 \\ h = 0 \le mean, j = 0,1,.....N-1 \end{cases} \qquad (22)$$

1.5 Centroid Calculation

The centroid of the segmented image of size w x h is been calculated from the following equation.

$$X_c = \sum_{x=0}^{w-1}\sum_{y=0}^{h-1} x * I(x, y) \qquad (23)$$

$$Y_c = \sum_{x=0}^{w-1}\sum_{y=0}^{h-1} y * I(x, y) \qquad (24)$$

The centroid of the identified object is been calculated as with primary implication the object intensity will be same in the consecutive frames so the segmented image of object will be almost equal to original segmented area. The centroid of the newly segmented image will be near to the identified object centroid value. The centroid of the interested object will be within the neighbourhood size of assumed value n. If the centroid falls within the neighbourhood then it is considered as the movement of the object selected.

$$abs((C_x, C_y) - (C_{x+\Delta t}, C_{y+\Delta t})) < windowsize(n) \quad (25)$$

II. PROBLEM STATEMENT

The ability to track a particular object in consecutive frames is crucial step in object tracking and classified applications. Object tracking has two main problems to be solved, estimation of the object and data association with objects such as occlusion effects. Tracking of an object in dynamic environment is challenging task as foreground and background pixel data changes for every instance. To develop vision based multiple object detection and tracking system.
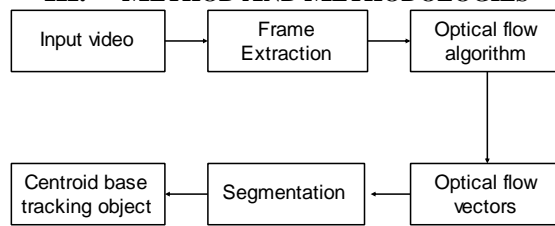
III. METHOD AND METHODOLOGIES

Fig. 1 Tracking of Multiple Objects Using Horn and Schunck

Figure 1 shows the design of multiple object detection and tracking system. It has the steps; frames extraction from the input video sequence and implementing Optical Flow algorithm. The Optical Flow vectors obtained are segmented using threshold technique. From the processed segmented binary image the centroid been calculated and the object with respect to the centroid values in consecutive frames the object been tracked.

IV. Results

The input video sequence is tracked using Horn and Schunck and Combined Local and global method. An outdoor video sequence is taken as input. Single object with static back ground, multiple objects with static background are considered. The experimental results obtained on static background are satisfactory. So, the experiment results for the multiple objects detection with dynamic background for two different images and resolution have been analysed.

Fig. 2 Input Image to Track Objects

Figure 2 shows the input video sequence with dynamic background of a race car to be tracked which is captured using a dynamic camera

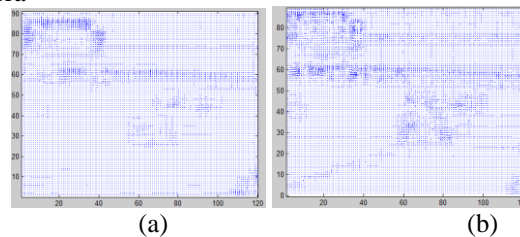(a)   (b)

Fig. 3 Optical Flow Vectors (A) Vectors Obtained from Horn and Schunck (B) Vectors Obtained from CLG Method

Figure 3 (a) shows the Optical vectors obtained by processing through Horn and Schunck method of Figure 2. Figure 3 (b) shows the vectors obtained from CLG method. The Optical Flow vectors obtained from CLG method has more dense flow fields; it locally finds Optical vectors in neighbourhood and does regularization on the vectors. Figure 4 shows segmentation region using a simple threshold technique.
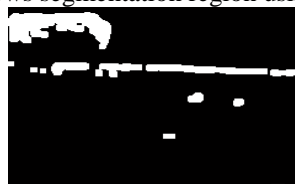

Fig. 4 Segmentation on Optical vectors

Figure 5 shows the centroid of the segmented regions as dots. The ROI segmented object centroid been compared with next frame centroid value, if centroid are near then the object is tracked.
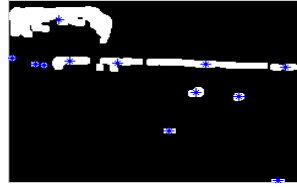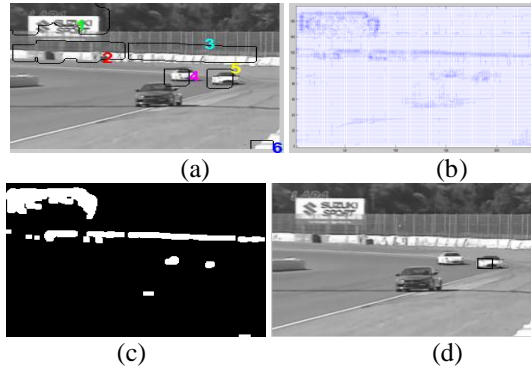
Fig. 5 Centroid of the Segmented Regions

(a) (b)

(c) (d)

Fig. 6 Results of the First Video Sequence (180x240)

Figure 6(a) shows the object movements labelled with respect object movements. The user selected object is been tracked in the consecutive frames.

Figure 6(b) shows the optical flow vectors obtained by implementing Horn and Schunck algorithm. The obtained optical flow vectors are segmented using threshold technique and the movements are shown in figure 6(c) Figure 6(d) shows the object tracked highlighted through boundary box.

The Input video sequence is a car moving from left to right direction in a dynamic background. Figure 9(a) shows the object movements labelled which is obtained through Horn and Schunck method.
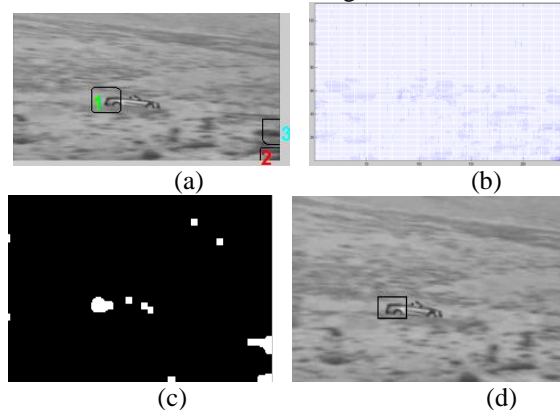

(a) (b)

(c) (d)

Fig. 7 Results of Second Video Sequence (135x240)

Figure 7 (b) shows the Optical flow vectors obtained from by implementing HS algorithm. The Segmented regions are obtained by implementing threshold technique. The ROI movement is tracked and highlighted by constructing boundary box.

Figure 8(a) and 9(a) shows the labeled image obtained through Combined Local and Global method. The Optical flow vectors are shown in Figure 8(b) and 9(b). The Optical Flow vectors are segmented by threshold technique. The ROI movement is tracked has been highlighted through boundary box as shown in 8(c) and 9(c).
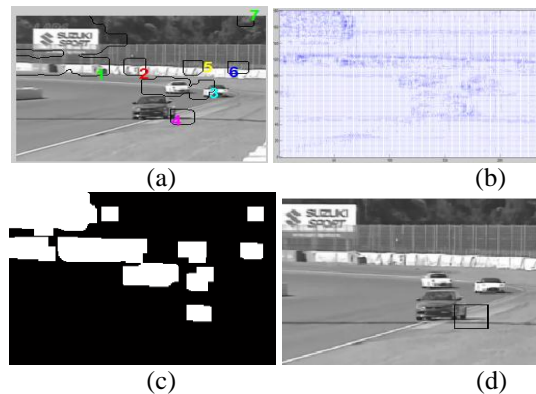
(a)   (b)

(c)   (d)

Fig. 8 Results obtained through CLG method (180x240)

The results obtained though Horn and Schunck and CLG method are shown in figure 6 and 8 respectively. The optical flow vectors obtained through CLG method are denser flow fields than Horn and Schunck. But Horn and Schunck highlights only the larger intensity changes which results in accurate tracking.
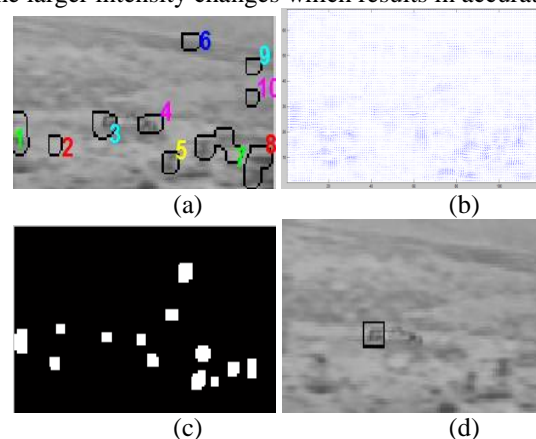

(a)   (b)

(c)   (d)

Fig.9 Results obtained through CLG method (68x120)

Examining the results of Horn and Schunck algorithm with combined local and global method it's clear that the vector fields obtained through combined method gives more accurate number of vectors due to its local processing technique.

V. Conclusion

Multiple object detection and tracking in a dynamic background video is tracked using horn and Schunck Optical Flow method along with CLG method. Estimation of the object movement and data association with objects in video plays a vital role in tracking of an object. So in Optical Flow methods the estimation of the object movement is been tracked using centroid. The execution time of Horn and Schunck and CLG method are nearly similar. The vectors obtained from combined method are more accurate. The execution time to track an object in Horn and Schunck is less comparatively with CLG method, but vectors are accurate. Implementing pyramidal techniques with these methods gives much less execution time. Centroid based tracking gives good results. The work can be extended to recognize the object and dynamic occlusion has to be solved.

References

[1] Linda Shapiro and George Stockman, *Computer Vision,* Prentice Hall, Jan 2001
[2] Henry Schneiderman and Takeo Kanade, *Object Detection Using the Statistics of Parts,* International Journal of Computer Vision, Issue no: 56, pp. 151–177, 2004
[3] Hue C., Le Cadre J.-P and P. Perez, *Tracking Multiple Objects With Particle Filtering*, IEEE transactions on Aerospace and Electronic Systems, Vol. 38, pp. 791-812, Dec. 10. 2002.
[4] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian, *Foreground object detectionin changing background based on color co-occurrence statistics*, In Proc. IEEE Workshop on Applications of Computer Vision, 2002.
[5] C. Stauffer and W. E. L. Grimson, *Adaptive background mixture models for real-time tracking*, In Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1999.
[6] Mittal and N. Paragios, *Motion-based background subtraction using adaptive kernel density estimation*, In Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 2, 2004, pp. 302–309.
[7] Elgammal, D. Harwood, and L. Davis, *Non-parametric model for background subtraction*, in Proc. European Conf. Computer Vision vol. II, May 2000, pp. 751–767.
[8] Waqar Shahid Qureshi and Abu-Baqar Nisar Alvi, Object Tracking using MACH filter and Optical Flow in Cluttered Scenes and Variable Lighting Conditions, World Academy of Science, 2009.

[9]   Hai-Yan Zhang, *Multiple Moving Objects Detection and Tracking Based On Optical Flow in Polar-Log Images,* College of Information, Beijing Forestry University, Proceedings of the Ninth Beijing, China, Proceedings of International Conference on Machine Learning and Cybernetics, Qingdao, 11-14 July 2010.
[10]   Berthold K. P. Horn and Brain G. Schunck*, Determining Optical Flow,* Artificial Intelligence, Vol. 17 (1981), pp. 185-203.
[11]   Rafael A. B. de Queiroz, Gilson A. Giraldi, Pablo J. Blanco and Raul A. Feijoo, *Determining Optical Flow using a Modified Horn and Schunck's Algorithm,* 17th International Conference on Systems, Signals and Image Processing, Brazil, June 2010.
[12]   Manjunath Narayana and Donna Haverkamp, A Bayesian algorithm for tracking multiple moving objects in outdoor surveillance video, IEEE, 2007
[13]   Jinman Kang, Isaac Cohen and Gerard Medioni, *Tracking Objects from Multiple Stationary and Moving Cameras,* IRIS, Computer Vision Group, University of Southern California, USA