World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:7, No:9, 2013

Optical Flow Based Moving Object Detection and Tracking for Traffic Surveillance

Sepehr Aslani, Homayoun Mahdavi-Nasab

***Abstract*—**Automated motion detection and tracking is a challenging task in traffic surveillance. In this paper, a system is developed to gather useful information from stationary cameras for detecting moving objects in digital videos. The moving detection and tracking system is developed based on optical flow estimation together with application and combination of various relevant computer vision and image processing techniques to enhance the process. To remove noises, median filter is used and the unwanted objects are removed by applying thresholding algorithms in morphological operations. Also the object type restrictions are set using blob analysis. The results show that the proposed system successfully detects and tracks moving objects in urban videos.

***Keywords*—**Optical flow estimation, moving object detection, tracking, morphological operation, blob analysis.

I. Introduction

MOTION analysis and estimation is one of the most challenging tasks in digital video processing and computer vision [1]-[5]. Optical flow presents an apparent change of a moving object's location or deformation between frames. Optical flow estimation yields a two-dimensional vector field, i.e., motion field, that represents velocities and directions of each point of an image sequence [1]. As it is an ill-posed problem, so far a wide variety of constraints between frames have been introduced in optical-flow modeling. Such constraints are based on image brightness and velocity. In particular, assumption of image brightness constancy between frames is one of the mostly used constraints. Moreover, a motion smoothness constraint is incorporated to enhance the detection accuracy and stability of the estimation. However, it is known that many issues remain unsolved in a scene with ambient disturbance, noise and occlusion. In particular, the background can spatiotemporally change due to natural complexity such as no-uniform background texture change with illumination change or/and wind-driven motion. In addition, objects by themselves can show a wide variety of properties ranging from rigidity and elasticity to fluidity. Other properties are the objects different appearances; different sizes, shapes, speeds and different degrees of transparency.

Optical flow estimation is used in many applications. Vehicles navigation, video image reconstruction and object tracking are some examples [6]. Moving vehicle detection is an important part of Intelligent Transport System (ITS) [7]. The goal of moving vehicle detection is to separate moving vehicles from background, and its detection result has a great impact on post image processing. At present, moving vehicles detection method from video are mainly temporal difference between two consecutive frames [8], image subtraction with background [9] and optical flow estimation [10]. Due to the higher detection accuracy of optical flow, it is more suitable for multiobjective moving analysis in complex scenes. Through optical flow estimation, motion parameters of moving objects can be obtained and at the same time, phenomena of occlusion and overlapping of objects may be avoided as far as possible [11].

The main objectives of this work are the design and development of the tracking system. In this paper we propose a novel method which is in fact a combination of a number of well-known computer vision techniques to track cars in a video. As this system is implemented in a modular fashion with a step-by-step approach, it is easy to evaluate a particular algorithm's applicability for the system simply by substituting one sub-module for another. This allows attention to be focused on framework design rather than on intricacies within each sub-module.

After applying optical flow calculation for detecting motion vectors, the vector magnitudes threshold are used to segment objects from the background. Filtering process removes the speckle noise and finally blob analysis is employed to identify the cars for the tracking process.

This paper is organized as follows. Section II expresses the tracking algorithm. Section III discusses a theory of optical flow methods. Section IV describes noise filtering. Section V demonstrates the image segmentation process. In Section VI the experimental results are presented. Finally, conclusion is drawn in Section VII.

II. The Tracking Algorithm

The proposed algorithm is presented here. Before any operation a scene should be selected from a static camera. Our test movies are selected from urban surveillance videos. Some pre-processing operations have to be done for making the scene ready to process. Due to the camera's auto white balance and the effect of sudden environment intensity changes, the mean of every frame is calculated on gray-scale format. The optical flow estimation is the essential part of the algorithm which is executed next. Filtering speckle, impulse and general external noises induced due to weather conditions is one of the most important sections of the procedure. Median filter is used in our framework. During filtering operation, some holes are created in the frames. To fill these holes and prevent detection mistakes morphological closing is implemented. Now the motion objects

S. Aslani is with the Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Isfahan, Iran (e-mail: Aslani.sepehr@gmail.com)

H. Mahdavi-Nasab is with the Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Isfahan, Iran

are detected, but many of them are not interested. The pedestrians or waving flags are the example of these unwanted motions. Blob analysis helps us to cluster objects and filter out objects which cannot be cars based on blob sizes. Drawing Bounding boxes around the vehicles and counting the number of tracked cars are the last segments of the algorithm shown in Fig. 1.

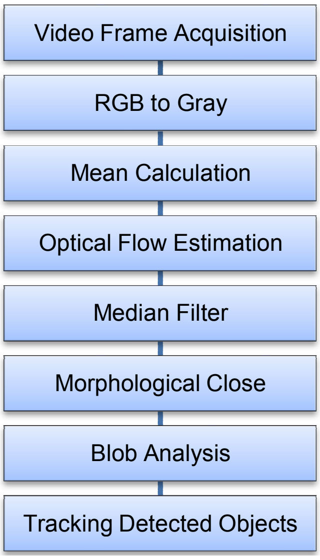


Fig. 1 Block diagram of the proposed algorithm

III. Optical Flow

The optical flow describes the direction and time rate of pixels in a time sequence of two consequent images. A two dimensional velocity vector, carrying information on the direction and the velocity of motion is assigned to each pixel in a given place of the picture.

For making computation simpler and quicker we may transfer the real world three dimensional (3-D+time) objects to a (2-D+time) case. Then we can describe the image by means of the 2-D dynamic brightness function of location and time I(x, y, t). Provided that in the neighborhood of a displaced pixel, change of brightness intensity does not happen along the motion field, we can use the following expression.

$$I(x,y,t) = I(x+\delta x, y+\delta y, t+\delta t) \tag{1}$$

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

$$I(x+\delta x, y+\delta y, t+\delta t) = I(x,y,t) + \frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\delta y + \frac{\partial I}{\partial t}\delta t + H.O.T. \tag{2}$$

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

$$I_x.v_x + I_y.v_y = -I_t \tag{3}$$

or in formal vector representation

$$\nabla I.\vec{v} = -I_t \tag{4}$$

where∇I is so-called the spatial gradient of brightness intensity and $\vec{v}$ is the optical flow(velocity vector) of the image pixel and I_t is the time derivative of the brightness intensity.

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).

Optical flow estimation is computationally demanding. At present there are several groups of methods for its calculation. All the methods come from (1) and consequently the presumption of conservation of brightness intensity. In this article our interest is concentrated in the differential methods. The optical flow determination is solved by the calculation of partial derivatives of the image signal. There are two most used methods, namely:

- Lucas-Kanade
- Horn-Schunck

A. Lucas-Kanade Method

Lucas and Kanade introduced the error term ρ_{LK} for each pixel [12]. This one, according to the following relation, is calculated as the sum of the weighted smallest squares of gradient constraint (4) in a close neighborhood of the pixel.

$$\rho_{LK} = \sum_{x,y\in\Omega} W^2(x,y)\left[\nabla I(x,y,t).\vec{v} + I_t(x,y,t)\right]^2 \tag{5}$$

where Ω is the neighborhood of the pixel; W(x,y) are weights allocated to individual pixels in Ω (typically2-D Gaussian coefficients).To find a minimal error it is necessary to compute derivation of the error term ρ_{LK} by individual components of velocity and putting the result equal to zero. Finally, a matrix form of the expression for the optical flow is as follows

$$\vec{v} = \left[A^T W^2 A\right]^{-1} A^T W^2 \vec{b} \tag{6}$$

For N pixels ($N = n^2$, for n × n of Ω neighborhood) and (x_i, y_i) ∈Ω at time t:

$$A = \left[\nabla I(x_1,y_1),\ldots,\nabla I(x_N,y_N)\right] \tag{7}$$

$$W = diag\left[W(x_1,y_1),\ldots,W(x_N,y_N)\right] \tag{8}$$

$$\vec{b} = -\left[I_t(x_1,y_1),\ldots,I_t(x_N,y_N)\right] \tag{9}$$

So we will obtain the resultant velocity for one pixel as the solution of the system (6). Instead of the calculation of the sums, the convolution by means of *Gaussian* or *difference* *temporal gradient* filter is used to reduce the algorithm complicacy [13].

B. Horn-Schunck Method

Horn and Schunck [14] issue from Lucas-Kanade method (5). In addition to the gradient constraint (4) they add another error term (called the global term of smoothing) for the limitation of too great changes of optical flow components (v_x, v_y) in the Ω. The minimization of the total error ρ_{HS} is then given by the relation:

$$\rho_{HS} = \int_D \left(\nabla I.\vec{v} + I_t\right) + \lambda^2 \left[\left(\frac{\partial v_x}{\partial x}\right)^2 + \left(\frac{\partial v_x}{\partial y}\right)^2 + \left(\frac{\partial v_y}{\partial x}\right)^2 + \left(\frac{\partial v_y}{\partial y}\right)^2\right] dxdy \quad (10)$$

where D (domain) is the region of the whole image, λ expresses relative effect of the second added error term (typically λ =1.0).

The relation (10) leads to the system of equations for which it is convenient to use Jacobi or Gauss-Seidel iterative methods [15].

Horn-Schunck method is more accurate [13], but with regard to relatively large number of iterations (in practice there are 10 to 100 steps) it is slower (as shown in Table I).

IV. Noise Filtering

Motion digital images are often interfered by a variety of noise distributions dependent on the prevalent conditions. The observed noise can be modeled either as additive white, impulsive, signal dependent or a combination of them [16]. Some of these noise distributions are very annoying when are involved in intensity changes in video frames. They randomly and sparsely corrupt pixels to two intensity levels: relative high or relative low, when compared to its neighboring pixels [17].

Therefore, the need emerges for implementing smoothing techniques that are able to treat different kinds of noise. Furthermore, a noise-free version of the corrupted image or sequence required by adaptive filtering algorithms during the training procedure is not always available. Moreover, it is well known that the main objectives of image filtering algorithms are: (a) the suppression of noise in homogeneous regions, (b) the preservation of edges (spatial or temporal) and (c)the removal of impulses (of constant and/or random value) [17]. A class of filters that fulfills these requirements is the so called signal filters. Standard median (SM) is a paradigm of this class. Median filter, as its name implies, replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel. Median filters are quite popular because, for certain types of random noise, they provide excellent noise-reduction capabilities, with considerably less blurring than linear smoothing filters of similar size [18].

V. 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 [19]; in this paper, the optical flow. In tracking algorithm the content of each frame is read and the background is estimated. The unwanted/interested objects are tagged by eliminating the background. Thresholding function is used to convert the gray image to binary so that the objects of interest can be highlighted by fixing a threshold limit [20].

A. Morphological Close

Morphological operations are performed to extract significant features from images that are useful in the representation and description of the shapes in the region; mostly used in image segmentation and pattern recognition [19]. In the proposed system we used both morphological close and erode, respectively, to remove portions of the road and unwanted objects. After morphological closing operation, on condition that vehicle's appearance is not destroyed, objects including many small holes and separated pixels may be connect into one big actual vehicle shape [21]. The following is the definition of morphological closing operation and the applied structural element B.

$$P \bullet B = \left(P \oplus B\right) \ominus B \quad (11)$$

where:

$$B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (12)$$

The matrix *P* which includes moving vehicle information is obtained through threshold segmentation.

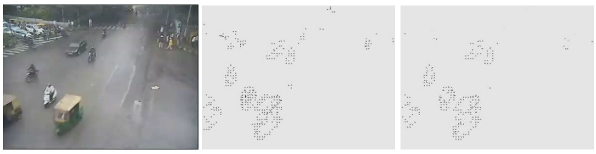
B. Blob Analysis

Blob analysis can be used to detect any kind of 2-dimensional shapes of an image. The detection is based on spatial characteristics using certain criteria. In many applications where the computation is time consuming, one can use blob analysis to eliminate blobs that are of no interest based on certain spatial characteristics and keep only the relevant blobs for further analysis [20]. The blobs which satisfied our system are vehicles. Other useless blobs are removed by setting limitations on the relative features in the algorithm.

VI. Experimental Results

In this section we show the experimental results using standard datasets from traffic closed-circuit TV (CCTV) to evaluate our system.

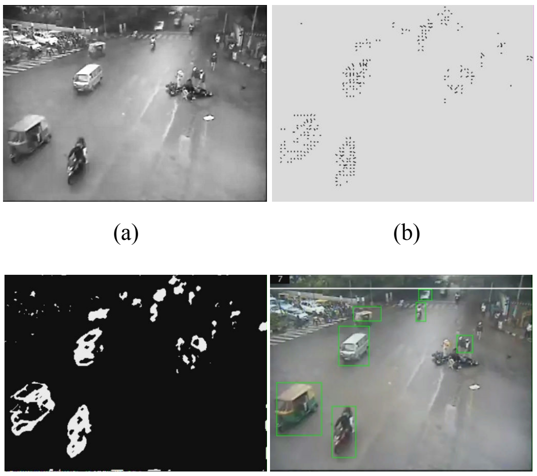
Before applying optical flow estimation on frames, the image format is converted from RGB to gray (Fig. 3 (a)); because Intensity measurements act well on gray-scale frames. Depends on methodology steps, the proper optical flow estimation (Locus-Kanade or Horn-Schunk) has been applied. Then, the median filter is performed to reduce noise corruptions. The optical flow estimation regardless of mobility, results in high contrast spots which are focused on. As it shows in Fig. 2 (a) and Fig. 2 (b), in addition to moving parts, there are some motion vectors pointed regions which are detected as objects, like pedestrian crossing at top left corner of the frame. To overcome this problem morphological close will be operated in algorithm to thin-out the parts of the road and fill holes in the blobs (Fig. 2 (c)).



(a) (b) (c)

Fig. 2 A selected frame (a) original video (b) optical flow motion vectors before and (c) after applying morphological close

The algorithm continues with blob analysis to detect specific objects, and remove unwanted blobs based on size and features like motion (Fig. 3 (c)). The blob constraints, depend on their utilizations, are fairly varied. The most important one which we have set is the density, i.e., the amount of pixels that exist in one blob. Another is connectivity coefficient, which illustrate the number of pixels that are connected to each other. The larger the number, the higher the accuracy and the more time consuming to compute. The blobs after this stage are surrounded by boxes and are ready for the next step.



(c) (d)

Fig. 3 An urban video frame (a) Gray-scale image, (b) optical flow motion vectors, (c) binary image with blobs and (d) tracking motion vehicles with boundary boxes

The final point in the algorithm is the tracking of moving objects and counting the number of detected vehicles that is shown above the images (Fig. 3 (d)). Before tracking objects, i.e., vehicles (cars, trucks and motorcycles …), it is needed to ensure whether all of the bounding boxes contain vehicles or not. In our experiments, when the ratio of the sides of the bounding box is above 0.4 (40%), it is classified as a vehicle.

Table I depicts the average processing times (in ms) of proposed framework when applied on a 120-frame-video and RGB (120×160) image format using MATLAB.

TABLE I
AVERAGE PROCESSING TIMES (IN MILLISECONDS) OF VARIOUS STAGES WITHIN OUR FRAMEWORK

Stages	**Methods**	
	Lucas-Kanade	***Horn-Schunk***
Optical Flow Estimation	5.2	6
Filtering+Segmentation	1.6	1.6
Motion Vectors Computing	0.44	0.58
Tracking+ Count Detected Number	11.3	11.3
Initializing and Finalizing	5,534.4	6,201.1
Total	5,552.9	6,220.5

As mentioned in section III, due to relatively large number of iterations, Horn-Schunck method is slower but more accurate. This computing cost also appears in motion vectors calculations. Other stages are identical and have the same time consuming. The most processing time belongs to initializing and finalizing steps, which are related to preprocessing methods, filters, their constraints, motion vectors embedded in original frames and computing the results; that demonstrate the moving vehicles bounded by boxes and the number of them in each frame.

VII. CONCLUSION

This paper has presented a system for motion object detection and tracking in image sequences from aerial or stationary camera images. Traffic surveillance in the urban environment is one of the most applicable usages of this system. The proposed system employs several methods to detect, filtering, segmentation and tracking objects. We used Horn-Schunk algorithm, as the most suitable method of optical flow estimation, to detect moving objects by the intensity changes of frames. The median filter performance significantly surpasses that of the other filters under study. The morphological close extracted significant features of region shapes from binary images and then blob analysis introduced these shapes to the next step as foregrounds. A great advantage of blob analysis is the low computation cost. Finally, as shown in the experimental result, the system is able to remove unwanted motion object which are not vehicles in the image sequence, using constraints on blob areas.

For the future work, ego motion estimation can be employed to compensate the camera shakings or to use this method on aerial videos. An adaptive filter helps this system to improve noise reduction. Also the algorithm may be developed to detect and identify overlapping objects and occlusion or transparencies during object tracking.

REFERENCES

[1] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Systems and experiment performance of optical flow techniques," Int. J. Comput. Vis., vol.12, no. 1, pp. 43–77, Feb. 1994.

[2] J. Xiao, H. Cheng, H. Sawhney, C. Rao, and M. Isnardi, “Bilateral filtering- based optical flow estimation with occlusion detection,” ECCV, 2006, pp. 211–224.
[3] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, “High accuracy optical flow estimation based on a theory for warping,” ECCV,2004, pp. 25–36.
[4] C. L. Zitnick, N. Jojic, and S. B. Kang, “Consistent segmentation foroptical flow estimation,” IEEE ICCV, 2005, pp. 1308–1315.
[5] M. J. Black, “The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields,” Comput. Vis. Image Underst., vol.63, no. 1, pp. 75–104, Jan. 1996.
[6] Rakshit, S.; Anderson, C.H., "Computation of optical flow using basis functions," Image Processing, IEEE Transactions on , vol.6, no.9, pp.1246,1254, Sep 1997.
[7] Zhan, Wei, and Xiaolong Ji. "Algorithm Research on Moving Vehicles Detection," Procedia Engineering, vol. 15, 2011,pp. 5483-5487.
[8] Ballard, Dana H. "Generalizing the Hough transform to detect arbitrary shapes," Pattern recognition, vol. 13, no. 2, 1981, pp. 111-122.
[9] Illingworth, John, and Josef Kittler. "A survey of the Hough transform," Computer vision, graphics, and image processing, vol. 44, no. 1, 1988, pp. 87-116.
[10] Shi, Xuejie. "Research on Moving Object Detection Based on Optical Flow Mechanism." University of Science and Technology of China, 2010, pp. 6-14.
[11] Qiaona Pei. “Moving Objects Detection and Tracking Technology based Optical Flow”. North China University of Technology, 2009, pp. 11-14.
[12] Lucas B. D., Kanade T. “An Iterative Image Registration Technique with an Application to Stereo Vision,” DARPA Image Understanding Workshop, 1981, pp. 121–130.
[13] Silar, Z., and M. Dobrovolny. "Comparison of two optical flow estimation methods using Matlab." Applied Electronics (AE), 2011 International Conference on. IEEE, 2011.
[14] Horn B. K. P., Schunck B. G. “Determining Optical Flow,” Artificial Intelligence, vol. 17, 1981, pp. 185–204.
[15] Hageman, L., Young, D. “Applied Iterative Methods,” Academic Press, New York, 1981, pp. 18-22.
[16] Eng, How-Lung, and Kai-Kuang Ma. "Noise adaptive soft-switching median filter," IEEE Transactions on Image Processing, vol. 10, no. 2, 2001, pp. 242-251.
[17] Tsekeridou, Sofia, Constantine Kotropoulos, and Ioannis Pitas. "Morphological signal adaptive median filter for still image and image sequence filtering", IEEE International Symposium on Circuits and Systems, vol. 4, 1998, pp. 21-24.
[18] Gonzalez, Rafael C., and E. Richard. Woods, “digital image processing," Prentice Hall Press, ISBN 0-201-18075-8. 2002.
[19] Tang, Sze Ling, Zulaikha Kadim, Kim Meng Liang, and Mei Kuan Lim. "Hybrid blob and particle filter tracking approach for robust object tracking," Procedia Computer Science, vol. 1, no. 1, 2010, pp. 2549-2557.
[20] Telagarapu, Prabhakar, MV Nageswara Rao, and Gulivindala Suresh. "A novel traffic-tracking system using morphological and Blob analysis." International Conference on In Computing, Communication and Applications (ICCCA), 2012, pp. 1-4.
[21] Gangodkar, D.; Kumar, P.; Mittal, A., "Robust Segmentation of Moving Vehicles Under Complex Outdoor Conditions," Intelligent Transportation Systems, IEEE Transactions on , vol.13, no.4, pp.1738,1752, Dec. 2012.