IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 1, Ver. I (Jan – Feb. 2016), PP 24-29
www.iosrjournals.org

Competent Tracking of Moving Object Using Affine & Illumination Insensitive Template Matching

Moiz Abdul Hussain[1], Dr. Govind Ukhandrao[2]
[1]*(Asst. Prof. ENTC Deptt. Anuradha Engineering College, Chikhli, India)*
[2]*(Principal Sharadchandra Pawar College of Engineering, Otur, India)*

Abstract : *Moving object detection & tracking in real world scene is becoming significant problem in today's era. The extensive study in this area is motivated by potential number of applications of object tracking. In this paper, we analyze a method for motion segmentation & tracking of nonstationary objects that uses the complete image information using the affine flow equations. These affine flow parameters are then combined with illumination insensitive template matching tracker to efficiently segment & track the moving objects. Experimental results performed on various database videos shows the method effectively tracks the object without miss detection of the target compared to other methods.*
Keywords: *Affine Flow, Video Surveillance, Object Tracking, Template matching.*

I. Introduction

Motion segmentation & tracking has been a crucial component for many computer vision applications such as structure from motion, objects tracking/recognition, advanced video editing and behaviour analysis. A variety of computational models, *e.g*. optical flow fields [7, 8], layers [1] and boundaries [9], have been developed to approximate the motion from a video sequence. It is significant to assess the performance of various motion analysis algorithms to gain insight and design better models. However, little has been done for evaluating motion analysis algorithms compared to the tremendous effort put into developing these algorithms.

The extraction of the affine parameters has fundamentally two components: the recognition of a suitable set of spatial patches to characterize every surface in a scene; and the tracking of the patches during the frame sequence. The tracking of the patches (2-D) and assessment of their connected affine motion parameters is realized using weighted linear regression over an approximated optical flow field. The weights are provided by truncated Gaussian windows, each defining the spatial extent of one of the patches being tracked. The preference of a gaussian window function is deliberate: the group of such functions are closed under the action of affine transformation and hence naturally represent the evolution of the patches as they warp according to the affine motion approximation.

The main contribution of this paper is the analysis of fast version of method for obtaining affine flow from the image sequence. These affine flow parameters are then combined with the illumination insensitive tracker to appropriately track the nonstationary objects. The motion evaluation is based on three observations. First, humans are experts at segmenting layers in a video sequence because human being can easily recognize the moving objects and their relative depth relationships. Second, humans are sensitive to any differences between two images when these two images are displayed back and forth. Third, humans have knowledge of the smoothness and discontinuities of the motion that a moving object undergoes.

We thoroughly tested our method by conducting experiments on several videos collected for our experimentation purpose in both indoor and outdoor surroundings. The experimental results demonstrate that the method performs favorably well in real world surroundings. One of the advantages of our approach compared to the other state-of-the-art methods is that it reduces the number of false detections, as pixel-level comparison can be per-formed in regions with significant motion only. Another advantage is that the illumination insensitive template matching eliminates the effects of illumination changes which are a major problem in visual surveillance system.

The paper is structured as follows. Section 2 explores some of the related methods. Section 3 presents our system methodology in detail. Section 4 describes the database collection for experimentation. Section 5 presents some experimental results & Finally, Section 6 concludes the paper.

II. Literature Review

A large number of methods are proposed for video surveillance system using human motion detection & tracking. Bramble [10] detects people in indoor environments and uses explicit 3-D cylindrical modelling of target objects. Haritaoglu et al. proposed a system which uses projection histograms of the single detected object as features to classify different actions [2]. Hager and Belhumeur [11] perform photometric stereo using at least three images of a scene under linearly independent lighting conditions.

In [3] Hidden Markov Models are used. However, these methods are usually much slower and adaptation to changes of the scene is difficult. Hedvig [12] uses optical flow features for training directly. Support vector machine is used for training. This method can get acceptable detection result in outdoor scene. McKenna used a combination of motion, skin color and face detection to track people and their faces [6]. Karmann and Brandt [13] and Kilger [4] respectively proposed an adaptive background model based on Kalman filtering to adapt temporal changes of weather and lighting.

Chen et al. [5] use a human attention model to create different cropped and scaled versions of images according to the screen size. An algorithm was presented in [14] to track moving objects in the scene by considering dynamic occlusions among moving objects; this is especially useful to track pedestrians moving in indoor environments. Robust detection of illumination changes in outdoor video surveillance images was effectively achieved in [15] by using an adaptive algorithm. This approach dynamically learns and updates the background by using a segmentation algorithm, especially for sudden changes.

Since successful tracking depends a lot on precise object detection, the evaluation of object detection algorithms within a surveillance system plays a significant part in overall performance study of the whole system.

III. system framework

Proposed analysis is directed to build a system that is able to detect nonstationary/mobile objects in the scene and track those using affine parameters & illumination insensitive tracker. Figure 1 below represents a block diagram of the system.

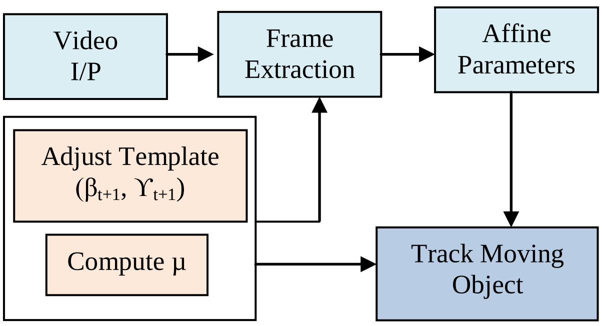


Fig. 1: Framework for object detection & tracking

Affine flow model combined with illumination insensitive tracker overcomes the problem of traditionally used optical flow technique like motion changes due to camera movement, dynamic background objects (e.g. tree leave or branches), and changes in the background geometry (e.g. parked cars). For the surveillance of outdoor scenes, sturdiness against noise and adaptivity to illumination changes are also essential. We coupled the method of affine flow extraction with template matching algorithm with illumination intensity minimization model which will take care of illumination changes in real images caused due to automatic exposure alterations of the camera, adjustment of light cause irradiance, appearance of shadows or movement of the tracked objects. Our objective is to provide a robust and improved method to find the moving objects in the video frame as well as to track them. The proposed method is effective in reducing the number of false alarms that may be triggered by a number of reasons. Fig.1shows the different stages involved in the proposed system. First, image frames are taken out from video, and then, affine flow parameters are extracted from an image. Simultaneously, illumination insensitivity hyperplane method for template matching is also applied to perfectly track the object.

3.1 Affine Flow

The starting point is the assumption of brightness constancy, which assumes that the brightness structures of local-time varying image regions are unchanging under motion for a short period of time. Formally, this is defined as,

$$I(x,y,t) = I(x-u, y-v, t-1) \qquad \ldots (1)$$

Where $(x,y)^t$ represents image position in pixel coordinates, t represents the temporal coordinate, $(u,v)^t$ represents the motion at image position $(x,y)^t$ over the time $t+1$ and $I(x,y,t)$ represents the image brightness function. Using the least-squares criteria, we seek the motion that minimizes the error ϵ over a region R of the image, formally,

$$\in = \sum_{x,t,\in R}[(x,y,t) - I(x-u, y-v, t-1)]^2 \qquad \ldots (2)$$

Assumint that $I(x,y,t)$ is approximately locally linear (2) is simplified by taking a Taylor series expansion and omitting terms higher than first order.

$$\in = \sum_{x,y,\in R}[I(x,y,t) - (I(x,y,t) - I_x(x,y,t)u - I_y(x,y,t)v - I_t(x,y,t))]^2 \quad \ldots (3)$$

$$\in = \sum_{x,y,\in R}[I_x(x,y,t)u + I_y(x,y,t)v + I_t(x,y,t))]^2 \quad \ldots (4)$$

Where, I_x I_y and I_t represent the partial derivatives of the intensity function with respect to the spatial parameters x, y and temporal parameter t, respectively. To obtain the estimate of velocity in the pure translation case, the partial derivatives of the error in (4) are taken with respect to the translational velocity $(u,v)^T$ and set to zero. This yields the following set of linear equations,

$$\begin{pmatrix} \sum {I_x}^2 & \sum I_x I_y \\ \sum I_x I_y & \sum {I_y}^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = -\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix} \quad \ldots (5)$$

A minimum of two points are required for the full solution. The under constrained problem given a single point is commonly referred to in the literature as the aperture problem. In the case where the motion is modelled locally by an affine transformation, corresponding to a composition of rotation, dilation, shear and translation,

$(u,v)^T$ are defined as follows,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_{41} & a_5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_0 \\ a_3 \end{pmatrix} \quad \ldots (6)$$

As in the case of translational motion, to obtain the estimates of the affine parameter (4) is differential w.r.t the six unknown affine parameters (a_i $where$ $i = 0, \ldots., 5$) and set to zero.

3.2 Template Matching

For eradicating the effects of illumination changes, a technique is to estimate illumination compensation parameters for the current template, in order to adjust its gray-level values with respect to the reference template. Using the linear model

$$F_{new}(x) = \beta f(x) + \gamma \, \forall \times \varepsilon \, r \quad \ldots (7)$$

Contrast and brightness variations are represented by illumination compensation parameters β and γ. Least-squares minimization problem is solved when corresponding points of the initial template $f(r, t0)$ and the current template $f(g(r, \mu(t)))$ are known.

$$(\beta_t, \gamma_t) = argmin \sum_{x \in r}[\beta f g(x, \mu(t), t) + \gamma - f(x, t0)] \quad \ldots (8)$$

We replace $g(x, \mu(t))$ with $x(t)$. Differentiating equation (8) with respect to the motion compensation parameters yields the linear system

$$\left[\sum_{x \in r} \begin{pmatrix} f^2(\check{x}(t),t) & f(\check{x}(t),t) \\ f(\check{x}(t),t) & 1 \end{pmatrix}\right] \begin{pmatrix} \beta_t \\ \gamma_t \end{pmatrix} = \sum_{x \in r} \begin{pmatrix} f^2(x,t_0) \; f(\check{x}(t),t) \\ f(x(t),t_0) \end{pmatrix} \quad \ldots (9)$$

To avoid time consuming operation of computing the matrix on the left, we exchange the reference template with the current template and hence obtain the motion compensation parameters and for adapting the reference template to the current template. Consequently, the matrix in equation (9) has to be computed only once. As we still want to adapt the current template, we revert to the original illumination compensation parameters.

3.3 Centroid Estimation

A set of two integers, $C\ (Cx, Cy)$ which determines the position of the moving object in the given scene. The centroid coordinates are calculated by:

$$Cx = Cx + x \quad \ldots (10)$$

$$Cy = Cy + y \quad \ldots (11)$$

$$T = T + 1 \quad \ldots (12)$$

For each pixel, where (x, y) is the current pixel location. The resulting centroid is then divided by the total value:

$$Cx = Cx \,/\, T \quad \ldots (13)$$

$$Cy = Cy \,/\, T \quad \ldots (14)$$

This centroid positional information is then transferred to the object tracking module which is coupled with the illumination insensitive template matching model which gives the good result for detection & tracking.

IV. Dataset

For the evaluation of our experiments we used two different data bases. First, Weizmann database which is publically available and another is our own created database, this database contains 86 sequences

having five & four of classes of actions (Jump, Run, Walk, Side, Handwaving, Jogging) respectively performed by 19 different subjects in two different conditions d1-d2.
d1: Indoor Environment;
d2: Outdoor (Playground) + illumination variations
Few samples of Indoor environment
Outdoor (Playground) + illumination variations Environment

‘Running’ ‘Jogging’ ‘Walking’

‘Jumping’ ‘Handwaving’

“Self Database” Images

‘Bend’ ‘Jack’ ‘Pjump’

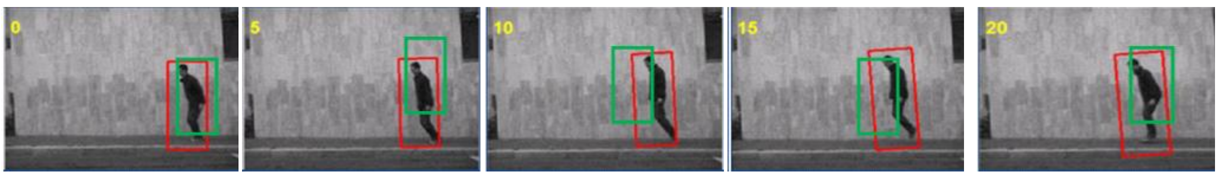
‘Run’ ‘Side’ ‘Skip’

“Weizmann Database” Images

Fig. 2: Sample images of Self & Weizmann database

V. Experimental result

This section summarizes the experimental results for motion detection & tracking system. The experiments were performed on MatLab 7.10.0 software. The PC was equipped with Pentium corei3, 2.13 GHz processor & 4 GB RAM Memory. The algorithm was implemented on two different datasets i.e. Weizmann publically available & self made database. Weizmann database [180*144, Frame rate 25frames/sec] has different actions Jump, Run, Walk, Bend, Skip and Side. All the actions were captured with steady camera. Our own data-base [1280*720, Frame rate 29frames/sec] was captured when both camera & target are moving in dynamic environment.



Proposed Model Without Illumination insensitive tracker

Fig. 3: Experimental Results on database videos

Figure 3 shows the comparison result of the proposed tracker system with the motion detection & tracking system of affine flow without illumination insensitive template tracker. As it can be observed from the figure 3 that, proposed model tracks the nonstationary object precisely as compared to the other state of the art method.

The algorithm implemented has achieved highly robust & clean results in cluttered backgrounds and illumination change. This shows the effectiveness of algorithm. Table 1 below shows the variation in the pyramid level of affine flow calculations for the 3 number of iterations. From the value represented in the table it can be seen that as the pyramid level increases the value of affine flow in between the pixel value decreases.

Table 1: Norm of optical flow values for 3 numbers of iterations

	Iteration 1	Iteration 2	Iteration 3
Pyramid 1	318.47	311.17	309.05
Pyramid 2	191.28	190.77	189.08
Pyramid 3	104.53	101.16	99.82
Pyramid 4	47.29	43.43	41.96
Pyramid 5	23.48	21.43	21.10
Pyramid 6	10.86	9.87	9.65

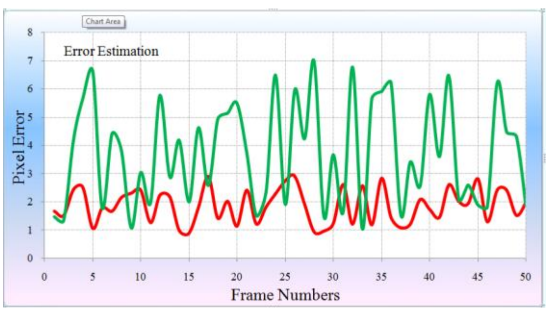


Fig. 4: Error Estimation of targets tracked by proposed method & affine flow without template tracker method.

Figure 4 displays the results in terms of average tracking errors. Proposed method shows the error maximum upto 3 pixels where as the method of the affine flow without illumination insensitive template matching tracker have error upto 7 pixels. This error comparison shows that the proposed system is very much precise in the motion segmentation and tracking.

VI. Conclusion

In this paper, we have presented a new proposal for object detection & tracking based using the combination of affine flow extraction & illumination insensitive template matching model. The system uses efficient technique of affine flow extraction which removes the drawback of traditionally used optical flow technique like detecting non-stationary background objects and shadows cast by moving objects. Further module of template matching eliminates the problem of illumination & dynamic environment conditions.

The integrated system was successfully tested on a number of sample videos containing various moving entities. Analysis was made on its performance. Experimental results shows that this algorithm has better tracking precision and can resolve the robust tracking problem in illumination changes & dynamic background, so it has a very wide range of applications.

Further work will concentrate on occlusion handling & multiple object tracking problems that will work in more complex cluttered environments. We will also focus on increasing the performance of the system by decreasing the time duration from frame to frame processing.

References

Journal Papers:

[1] J. Y. A. Wang and E. H. Adelson, Representing moving images with layers, IEEE Trans. Image Processing, , 1994, 3(5):625 – 638.

[2] I.Haritaoglu, D.Harwood and L.S.Davis, W4: Realtime surveillance of people and their actions, IEEE Trans. on Pattern Anal. and Mach. Intell. , August 2000, 22(8),pp.809-830.

[3] J. Kato, S. Joga, J. Rittscher, and A. Blake, An HMM-based segmentation method for traffic monitoring movies, IEEE Trans. on PAMI, , 2002, vol. 24, no. 9, pp. 1291–1296.

[4] M. Kilger, A shadow handler in a video-based real-time traffic monitoring system, Proc. of IEEE Workshop on Applications of Computer Vision. 1992, pp. 1060-1066.

[5] L.-Q. Chen, X. Xie, X. Fan, W.-Y. Ma, H.-J. Zhang, and H.-Q. Zhou, A visual attention model for adapting images on small displays, Multimedia Syst., Oct. 2003. vol. 9, pp. 353–364.

Chapters in Books:

[6] McKenna, Stephen J., Tracking groups of people Computer Vision and Image Understanding 80.1 (2000): 42-56.

Proceedings Papers:

[7] B. K. P. Horn and B. G. Schunck, Determing optical flow, Artificial Intelligence, 17:185–203, 1981.

[8] B. Lucas and T. Kanade, An iterative image registration technique with an application to stereo vision, In Proc. of the Intl. Joint Conf. on Artificial Intelligence, pages 674–679, 1981.

[9] C. Liu, W. T. Freeman, and E. H. Adelson, Analysis of contour motions, In NIPS, 2006.

[10] Isard, Michael, and John MacCormick, BraMBLe: A Bayesian multiple-blob tracker, Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol. 2. IEEE, 2001.

[11] G.D. Hager and P.N. Belhumeur, Real-time tracking of image regions with changes in geometry and illumination, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition: 403 – 410, 1996.

[12] Hedvig Sidenbladh, Detecting Human Motion with Support Vector Machines, ICPR 2004.

[13] K.P. Karmann, A. Brandt, Moving object recognition using an adaptive background memory, In Cappellini, Time-varying Image Processing and Moving Object Recognition, Elsevier, Amsterdam, The Netherlands, 1990.

[14] Marchesotti, L. Marcenaro, L. Regazzoni C. S, Tracking and counting multiple interacting pedestrian in indoor scenes, In Proceedings of third IEEE international workshop on performance evaluation of tracking and surveillance, Copenhagen Denmark, June 1, 2002.

[15] Regazzoni, C, Adaptive change detection approach for object detection in outdoor scenes under variable speed illumination changes, In Proceedings of Eusipco 2000.