Jurnal Teknologi, 39(D) Dis 2003: 57–68

INTEGRATION OF PROJECTION HISTOGRAMS AND LINEAR PREDICTION FOR OBJECT TRACKING

YEOH PHAIK YONG[1] & SYED ABDUL RAHMAN ABU BAKAR[2]

Abstract. In this paper, a real-time, efficient and robust visual tracking system for a single moving object using image sequences captured by a stationary camera is presented. The proposed tracking algorithm integrates the Projection Histograms technique with the Linear Prediction method in order to achieve a faster tracking speed. The Projection Histograms technique is applied to obtain the actual location of the tracked target, whereas the Linear Prediction method is incorporated in the proposed tracking algorithm to predict the location of the moving object in the next frame based on its several past centroid measurements. The Projection Histograms technique coupled with a second order Linear Prediction method has enabled the proposed algorithm to accurately track a single moving object. The potential applicability and efficiency of the proposed tracking algorithm has been validated by good tracking results on several experimental image sequences.

Keywords: Motion tracking, projection histograms, linear prediction, frame differencing

Abstrak. Kertas kerja ini mengemukakan satu sistem visual untuk menjejak objek bergerak yang dapat beroperasi secara masa-nyata, efisien dan tahan lasak, dengan menggunakan turutan imej yang diperolehi daripada satu kamera statik. Algoritma menjejak yang dikemukakan menggabungkan teknik Unjuran Histogram dan kaedah Ramalan Linar untuk mencapai kelajuan menjejak yang lebih tinggi. Unjuran Histogram dilakukan untuk mendapatkan lokasi sebenar bagi objek yang dijejak, manakala kaedah Ramalan Linar disertakan dalam algoritma menjejak yang dikemukakan untuk meramal lokasi bagi objek bergerak dalam imej seterusnya, berdasarkan beberapa ukuran centroid sebelumnya. Hasil gabungan teknik Unjuran Histogram dan tertib kedua Ramalan Linar telah membolehkan algoritma yang dikemukakan untuk menjejak objek bergerak dengan tepat. Potensi and kecekapan bagi algoritma menjejak yang dikemukakan telah dibuktikan oleh keputusan menjejak yang baik pada beberapa turutan imej eksperimen.

Kata kunci: Menjejak pergerakan, unjuran histogram, ramalan linar, perbezaan bingkai

1.0 INTRODUCTION

Motion tracking is an iterative process of determining the trajectory of a moving object during a video sequence, by monitoring the object's spatial and temporal changes, including its presence, position, size, shape, etc [1]. It has become one of the most popular areas of real-time video processing. The applications that can benefit from

[1, 2] Department of Microelectronics and Computer Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia. 07-5535274. Email: [1]yeohpy@hotmail.com & [2]syed@fke.utm.my

motion tracking include security surveillance systems, traffic monitoring, intelligence robots control and medical technology.

As motion tracking gains more practical usefulness, much research efforts have been devoted to developing and exploring new techniques for the motion tracking analysis. Although this area has been studied for the past couple of decades, a fast, reliable and robust object tracking algorithm remains a great challenge.

This paper is focused on presenting an efficient object tracking algorithm. The structure of this paper is as follows: Section 2 gives an overview of the existing tracking approaches as proposed by many researches. Section 3 briefly illustrates the overall proposed tracking system. Section 4 discusses in detail the fundamental techniques used in the proposed motion detection and tracking algorithm (including the Projection Histograms technique, the Linear Prediction and Maximum Entropy Method). Section 5 demonstrates the performance of the proposed method applied on some real-world image sequences. Finally, conclusions are drawn in Section 6.

2.0 BACKGROUND

In general, existing approaches formulated to deal with the tracking of a single moving object or multiple objects can be subjectively classified into a few categories, i.e. feature-based approach, template-based method, gradient-based method, statistical model and prediction approach.

The feature-based approach utilizes certain characteristics [2,3] of the target image (such as line segments, corners, contours and curvilinear) for establishing identification in the following frame. An efficient representation model for the target is crucial for this method and sometimes the exact dynamic models are either difficult to obtain or need complex mathematical descriptions. On the other hand, the template-based method takes the template as a whole [4,5]. Both of these methods are less robust to changing shape of the tracked target and temporary occlusions.

The gradient-based method performs tracking based on spatio-temporal gradient of the image intensity [6] or the depth and motion information of the moving object relative to the camera [7]. The statistical model [8] involves obtaining some prior information about the dynamics or certain feature characteristics of the target and subsequently analyzing the corresponding posterior distribution. In general, this method is computationally demanding and presently hardly useful in real-time systems with high frame-rate.

The prediction approach involves forecasting the subsequent trajectory of the tracked object based on previous measurements. The existing prediction techniques are mainly Kalman Filter (KF) and the Extended Kalman Filter (EKF), where a state model that well represents the dynamics of the target is developed, and the state parameters are estimated based on the process and measurement equations by taking into account the additive stochastic noise [9,10,11]. However, KF can be error-prone if the assumption of local linearity is violated or when the tracked target actively maneuvers. The predictive tracking approach proposed in the work [12] deals with the motion estimation from regression analysis, where the motion parameters are fitted by the optimal orders of the time-depending polynomials models. Many conventional position based tracking strategies usually define the dynamics models for position prediction on target. Consequently, the implemented system becomes inflexible and less robust.

The tracking algorithm proposed in this paper is based on the prediction approach as it is more robust to temporary occlusions and it enables faster tracking speed to be achieved.

3.0 THE PROPOSED ALGORITHM

The proposed motion tracking algorithm integrates the Projection Histograms technique [13] and the Linear Prediction method [14] with the aim of enabling the implemented tracking system to achieve real-time performance. Although the Linear Prediction method is not new, its efficiency and potential applicability in vision-based predictive object tracking system has not been explored.

The Projection Histograms technique is used to identify the actual location of the tracked target in the current frame. The Linear Prediction method is incorporated into the tracking algorithm so as to predict the centroid location of the moving object in the subsequent frame based on its several past centroid measurements. Using a proper order of the Linear Prediction method enables the predictor to accurately track the moving object. The Levinson Recursion in the Linear Prediction method is solved by using the Maximum Entropy Method (MEM) due to its significant accuracy results [14].

The proposed approach enables accurate and robust tracking without constraining the system to know the characteristics of the object being tracked. It can be used for tracking a large variety of objects with different shape, size and movements. Apart from that, based on the predicted centroid location of the target, the problem of exhaustive search over the entire frame (during motion detection and tracking) can be substantially alleviated by constraining the search window to certain confidence region. The prediction accuracy is always guaranteed by recursively updating its database at certain interval or when the prediction value begins to deviate.

The implemented tracking system is summarized in the block diagram shown in Figure 1. An input live video (captured by a stationary camera at a pre-fixed rate) is fed to the tracking system that consists of the motion detection and the motion tracking modules.

At the initial stage, each frame of the image sequence is directed to the motion detection module, where the Hexagonal Edge Detector and the frame differencing operation are applied. Details for motion detection will be discussed in Section 4.1. The next stage is the tracking process that integrates the Projection Histograms technique and Linear Prediction. The former is performed to obtain the actual location of the moving object, followed by centroid computation. The latter deals with predicting the next centroid location of the target, as well as constraining the search region of the image to certain confidence zone where the tracked target is estimated to exist. This region corresponds to the area where the detection and tracking modules will be applied at the next time interval.

Subsequently, centroid error is computed to determine the accuracy and reliability of the predicted centroid. If the error exceeds the pre-fixed threshold value, the predicted centroid in the prediction database will be updated based on the actual centroid value to prevent further deviation.

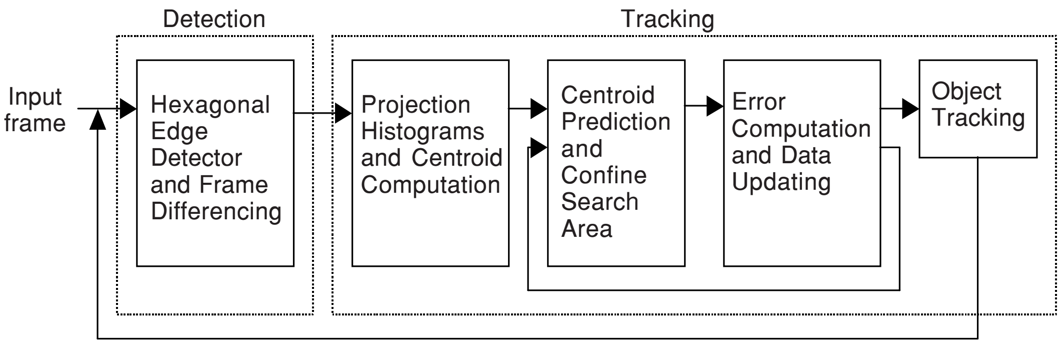


Figure 1 The proposed detection and tracking algorithm

4.0 MOTION DETECTION AND TRACKING

The tracking system consists of the motion detection and motion tracking modules. The implementation methods for both modules will be discussed in the following sub-topics.

4.1 Motion Detection

The motion detection scheme employed is based on the frame differencing technique for extracting edges of the moving object from the static-background image sequences. In this stage, the Hexagonal Edge Detector proposed in [15] is applied to the raw image pixel values in order to obtain the edge map for the current frame. It has been selected, instead of other traditional gradient edge operators (such as Sobel, Prewitt, Roberts) due to its significant performance in computation time, as well as accuracy in generating edge pixels as verified in [15].

Comparison of edge maps (frame differencing) is carried out by subtracting the edge pixels of the previous frame from the edge pixels at the corresponding spatial location of the current frame. If the pixel difference exceeds the threshold value for binarization, its value will be marked as '255' to indicate availability of motion. Otherwise, the pixel will be assigned with value '0' which implies static background.

The resulting object mask contains information about the location, size as well as the shape of the moving object. The moving edges extraction process is shown in Figure 2.

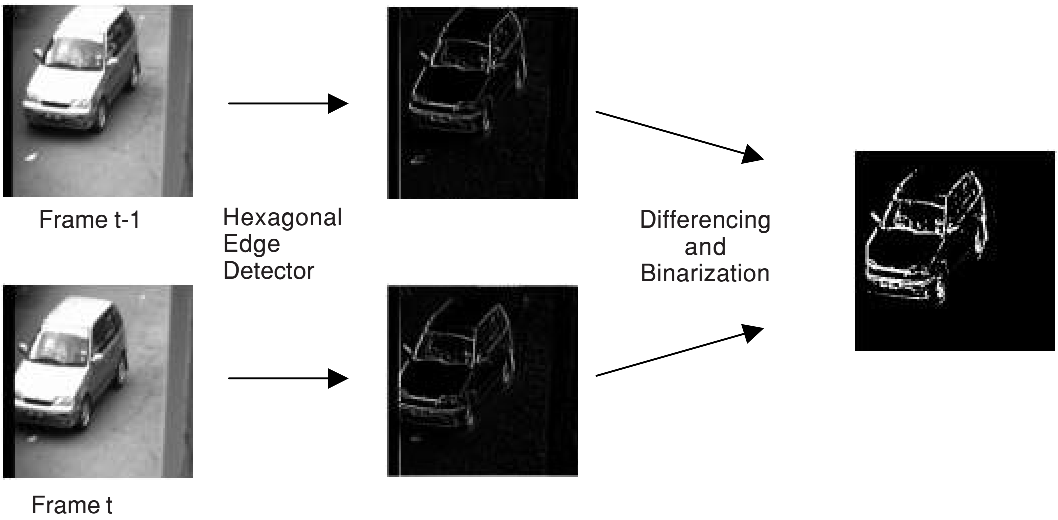


Figure 2 Moving edges extraction process

4.2 Tracking Algorithm

The proposed tracking algorithm incorporates the Projection Histograms technique with Linear Prediction so that the Linear Prediction method could enable the implemented tracking system to achieve real-time performance.

4.2.1 Projection Histograms

The Projection Histograms technique is applied on the binarized edge map to determine the actual location (centroid) of the moving object in the current frame, based on the projections in the horizontal and the vertical directions.

In order to track the location of the detected moving object, initially, the values of both the horizontal and vertical projections are obtained. Projection of the horizontal axis is obtained by summing up all the pixels column-wise, i.e.:

$$P_h(i) = \sum_i f(i, j)$$

On the other hand, projection of the vertical axis is obtained by adding all the pixels row-wise:

$$P_v(j) = \sum_i f(i, j)$$

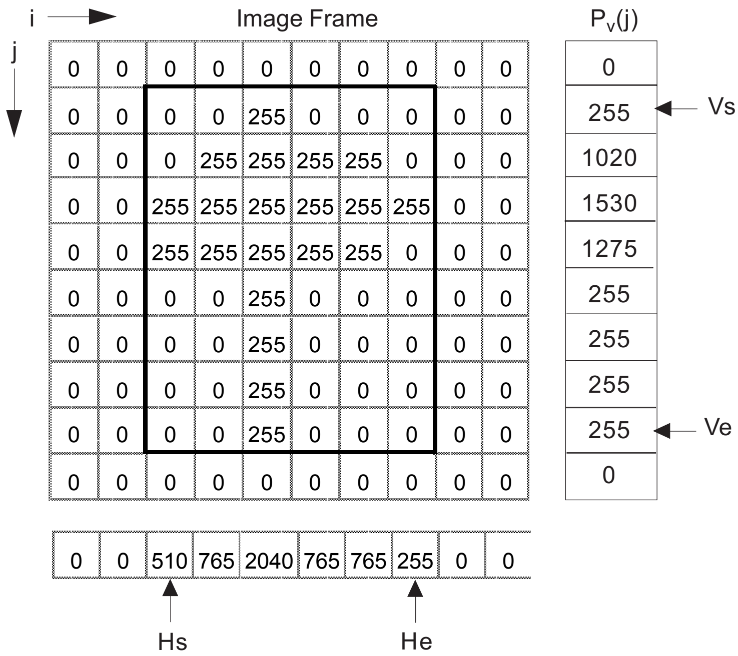


Figure 3 Summations of rows and columns to obtain vertical and horizontal projections

where: i is column, $i = 0, 1, 2, 3, ..., 255$

j is row, $j = 0, 1, 2, 3, ..., 255$

$f(i, j)$ is the pixel value at column i and row j

Next, based on the projection values, a bounding box that encompasses all the binarized edges of the tracked object will be drawn (as shown in Figure 3).

The center of the bounding box (C_x and C_y) will then be identified as the centroid of the moving object. The centroid of the object is computed as follows:

x coordinate of the centroid, $C_x = (\text{Hs} + \text{He}) / 2$

y coordinate of the centroid, $C_y = (\text{Vs} + \text{Ve}) / 2$

where:

Hs = x coordinate at the starting of the horizontal perimeter

He = x coordinate at the end of the horizontal perimeter

Vs = y coordinate at the starting of the vertical perimeter

Ve = y coordinate at the end of the vertical perimeter

Each centroid location is stored in the prediction database. Once a sufficient number of data is available, prediction (based on Linear Prediction and Maximum Entropy Method) is executed to determine the centroid of the object in the subsequent frame. The number of data required corresponds to the order of Linear Prediction being employed. Each prediction of $\hat{C}_x$ and $\hat{C}_y$ is performed independently and the corresponding prediction values are stored in the separate prediction database.

4.2.2 Linear Prediction

Linear Prediction is used to predict the next centroid location of the tracked target denoted as $\hat{C}_n$, based on its finite past centroid measurements [14].

In the developed tracking system, the second order Linear Prediction is adopted. This second order Linear Prediction has been identified as the optimal order for the employed Linear Prediction, by analyzing the trade-off between the computational time and accuracy based on the acquired image sequences. From our empirical experiments as shown in [16], it has been found that the second order Linear Prediction enables the resultant predictor to achieve minimum mean error as well as fast computation time.

For the second order predictor being employed, prediction is obtained from the past 3 centroid coordinates, namely $\{C_{n\text{-}i}; i = 1,2,3\}$. The linear predictor of order 2 can be written in the form as shown in (1):

$$\hat{C}_n = -\sum_{i=1}^{3} a_i C_{n-i} = -\left[a_i C_{n-1} + a_2 C_{n-2} + a_3 C_{n-3}\right] \quad (1)$$

The prediction coefficients a_1, a_2 and a_3 are chosen by minimizing the mean-squared prediction error:

$$\varepsilon = E[e_n^2] = \min \quad (2)$$

where e_n is the prediction error (shown in (3)):

$$e_n = C_n - \hat{C}_n = C_n + a_1 C_{n-1} + a_2 C_{n-2} + a_3 C_{n-3} \quad (3)$$

The best linear predictor of order 2, $\hat{C}_n$ can be obtained efficiently by using Levinson's algorithm, where the lattice realizations of best linear prediction filters for orders p = 0, 1 and 2 are determined. In the proposed algorithm, the Maximum Entropy Method (MEM) is employed to solve the prediction coefficients, due to its superior performance as compared to the autocorrelation and the covariance method [14]. It is capable to ensure that the predictor does not run off the block of data, and always results in a minimal-phase filter. Moreover, its minimization criterion is able to produce more precise prediction (compared to the autocorrelation and covariance method), by minimizing the sum-squared of both the forward and backward prediction errors, i.e.:

$$\varepsilon = \sum_{n=p}^{2} [e_p^+(n)^2 + e_p^-(n)^2] = \min \tag{4}$$

The iterative procedure of Levinson recursion is applied to determine the prediction-error filter of order p:

$$\begin{bmatrix} 1 \\ a_{p,2} \\ a_{p,2} \\ \vdots \\ a_{p,p-1} \\ a_{p,p} \end{bmatrix} = \begin{bmatrix} 1 \\ a_{p-1,1} \\ a_{p-1,2} \\ \vdots \\ a_{p-1,p-1} \\ 0 \end{bmatrix} - \gamma_p \begin{bmatrix} 0 \\ a_{p-1,p-1} \\ a_{p-1,p-2} \\ \vdots \\ a_{p-1,1} \\ 1 \end{bmatrix} \tag{5}$$

The following lattice relationships are valid for n that falls in the range of $p \le n \le 2$ to ensure that the filter does not run off the data:

$$\begin{aligned} e_p^+(n) &= e_{p-1}^+(n) - \gamma_p e_{p-1}^-(n-1) \\ e_p^-(n) &= e_{p-1}^-(n-1) - \gamma_p e_{p-1}^+(n) \end{aligned} \tag{6}$$

The reflection coefficient, γ_p can be computed with the following equation, i.e.:

$$\gamma_b = \frac{\sum_{n=p}^{2} [e_{p-1}^+(n) e_{p-1}^-(n-1)]}{\sum_{n=p}^{2} [e_{p-1}^+(n)^2 + e_{p-1}^- n - 1)^2} \tag{7}$$

Elaborate explanation on the Linear Prediction method and the Maximum Entropy Method used in the proposed tracking algorithm can also be referred in [17].

Once the predicted centroid is obtained from the second order Linear Prediction, the search region for the subsequent frame will be constrained to certain confined region based on the predicted centroid location, as shown in Figure 4.

This inevitably helps to speed up the tracking process, as both the motion detection and tracking modules can be applied to a smaller area for the next time interval, rather than covering the whole frame.

Subsequently, the centroid error C_e is computed. Due to simplicity, we have chosen to determine the error based on the Euclidean distance between the actual centroid and the predicted centroid, as follows:

$$C_e = \sqrt{(C_x - \hat{C}_x)^2 + (C_y - \hat{C}_y)^2}$$

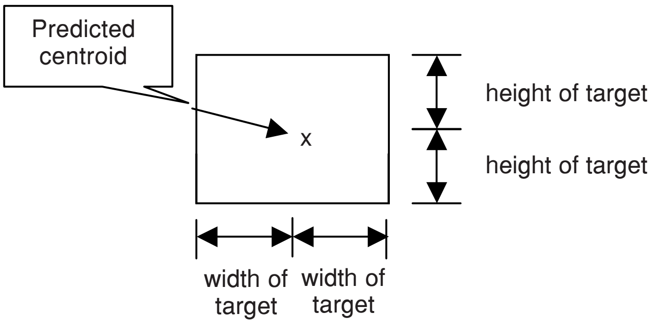


Figure 4 Search window region and its size

where: (C_x, C_y) = actual centroid coordinates
(C_x, C_y) = predicted centroid coordinates

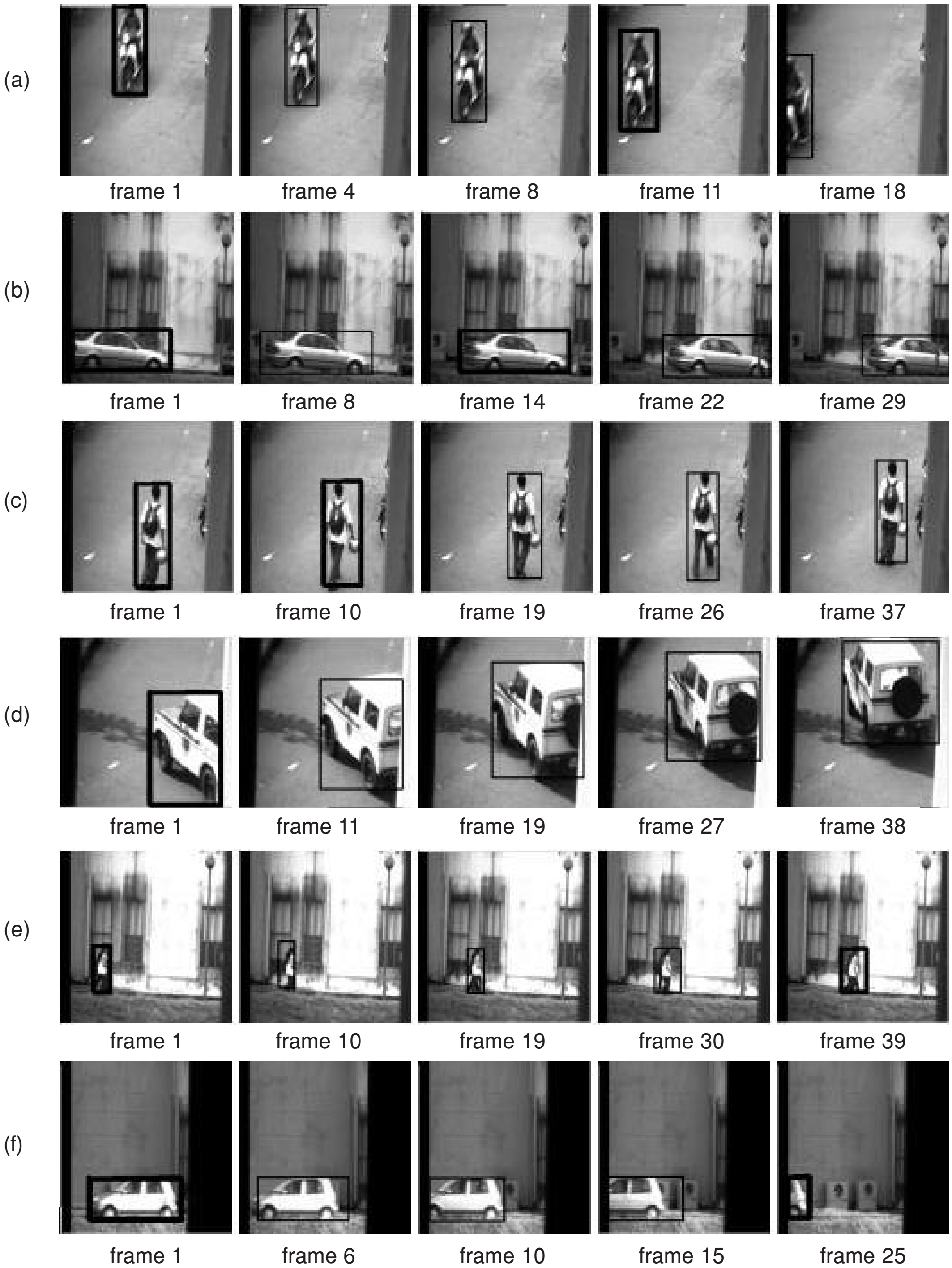
If the error exceeds the pre-fixed threshold value, the centroid value in the prediction database will be updated with the actual value obtained using the Projection Histograms technique. Otherwise, the predicted centroid value is considered reliable and the location of the moving object for the next time instant will be determined merely by this value.

5.0 EXPERIMENTAL RESULTS

The proposed object tracking algorithm has been implemented on a Pentium IV 1.7 GHz PC. It has been evaluated with many real-world image sequences containing vehicles and people in different landscapes. The image sequences were acquired using the stationary UNIQ-UP610 CCD camera and a Picolo frame grabber. The frame is being processed in the grayscale format and the size of each frame is fixed to a size of 256×256 pixels.

Figure 5 shows some results of the applied tracking algorithm on the tested real-world image sequences. The detected moving object in each frame is marked by a black bounding box. The average tracking speed achieved is approximately 5 to 20 fps.

From the experimental image sequences shown in Figure 5, it is apparent that the proposed tracking algorithm is able to track the detected moving object accurately at a real-time speed. The average tracking speed achieved is approximately 5 to 20 fps, over a 256×256 pixel image.



(a) MOTORCYCLE sequence (b) CAR_1 sequence (c) WALKING STUDENT sequence (d) JEEP sequence (e) STUDENT sequence (f) CAR_2 sequence

Figure 5 Tracking results by the proposed tracking algorithm

6.0 CONCLUSIONS

This paper has proposed a single object tracking algorithm by integrating the Projection Histograms technique and the Linear Prediction method. It is used for performing reliable real-time object tracking from the dynamic video sequence captured by a stationary camera. The incorporated Linear Prediction method predicts the next centroid location of the moving object, which is based on the Projection Histograms technique. Prior to the Projection Histograms technique, motion detection based on the Hexagonal Edge Detector is employed. The experimental results have demonstrated the possibility of integrating the Projection Histograms technique and Linear Prediction for motion tracking. As shown in the experimentation, our proposed system is able to achieve real-time, efficient and robust performance.

REFERENCES

[1] [Online] Available at http://trident.mcs.kent.edu/~zguo/courses/survey.htm#thresholding

[2] Lars-Peter, B., K. Talmi and J. Liu. 1997. "Automatic Detection and Tracking of Faces and Facial Features in Video Sequences". Proceeding of The Picture Coding Symposium 1997.

[3] Wai-Pak, C., K. M. Lam and W. C. Siu. 2001. "An Adaptive Active Contour Model for Highly Irregular Boundaries". *Pattern Recognition*. 34(2): 323-331.

[4] Mehmet, C. and H. Reza. 1998. "Moving Object Tracking Using Local Windows". Proceeding of the IEEE International Symposium on Intelligent Control. 180-185.

[5] Frederic, J. and M. Dhome. 2002. "Hyperplane Approximation for Template Matching". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 24(7): 996-1000.

[6] Wilson, M. B. and S. Dickson. 1999. "Poppet: A Robust Road Boundary Detection and Tracking Algorithm". Proceeding of the British Machine Vision Conference (BMVC99). 163-172.

[7] Jun-ya, T., M. Mada, S. Ohyama and A. Kobayashi. 2002. "Detection and Analysis of Moving Objects Using Time Series Stereo Difference Images". Proceeding of the 2002 41st SICE Annual Conference (SICE 2002). 228-233.

[8] Hedvig, S. and M. J. Black. 2001. "Learning Image Statistics for Bayesian Tracking". Proceeding of the International Conference on Computer Vision (ICCV 2001). Vol. II 709-716.

[9] Wira, P. and J. P. Urban. 2000. "A New Adaptive Kalman Filter Applied to Visual Servoing Tasks". Proceeding of the Fourth International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies (KES '2000).

[10] Kyu-Bum, H. and Y. S. Baek. 1998. "Visual Servo Tracking Strategy Using Time-Varying Kalman Filter Estimation". Proceeding of The Fourth Conference on Motion and Vibration Control (MOVIC '98).

[11] Sou-Chen, L., Y. C. Huang and C. Y. Liu. 1999. "Trajectory Estimation for Tactical Ballistic Missiles in Terminal Phase Using On-line Input Estimator". Proc. Natl. Sci. Counc. ROC(A). 23(5): 644-653.

[12] Frau, J. and V. Llario. 1991. "Predictive Tracking of Targets Using Image Sequences". Proceeding of IEEE/RSJ International Workshop on Intelligent Robots and Systems (IROS '91). 848-854.

[13] Wan-Ayub, W. A. 2002. "*Menjejak Objek Yang Bergerak Dalam Satu Jujukan Imej*". M.Sc. Thesis: Universiti Teknologi Malaysia, Malaysia.

[14] Sophocles, J. O. "*Optimum Signal Processing: An Introduction*". 2nd Edition. McGraw-Hill Publishing Company.

[15] Syed, A. B., and R. J. Green. 1996. "*Detection of Edges Based on Hexagonal Pixel Formats*". Proceeding of the 3rd. International Conference on Signal Processing (ICSP). Vol. II 1114 –1117.

[16] Yeoh, P. Y. and S.A.R. Abu-Bakar. 2003. "Performance Study on Linear Prediction For Tracking of A Moving Object". To be appear in the Proceeding of the Malaysian Science and Technology Congress (MSTC).

[17] Yeoh, P. Y. and S.A.R. Abu-Bakar. 2003. "The Application of Linear Prediction To Object Tracking". Proceeding of The First Malaysia-France Regional Workshop on Image Processing In Vision Systems and Multimedia Communication (IMAGE 2003).