Mashadi, Syamsudhuha, MDH Gamal dan M. Imran, (Eds)
Proceedings of the International Seminar on Mathematics and Its Usage in Other Areas
November 11 – 12, 2010
ISBN. 978-979-1222-95-2

TRACKING PERSON IN OUTDOOR AND DYNAMIC ENVIRONMENT

Elfizar[1], Liang Shing Ng[2]

[1] FMIPA University of Riau - Indonesia
[2] AI Dept, Faculty of Computer Science and Information Technology
University of Malaya – Malaysia

Abstract

This research is intended to track person in the video sequences. The videos used in this research consist of persons walking in the outside of building with dynamic environment. There are two cameras used in this research with different field of view but with overlapping views. Each video has specifications i.e. frame size of 576x768 pixels, rate of 30 frame per second, and avi format. After reading the input, we do background estimator. The background is taken from the initial frame where there is no people walking in the observed area. After the segmentation process, then we use Kalman Filter to predict the tracked person walking in the view. The filtered position yielded by the above algorithm is used to mark the rectangles which state the foreground. When there is occlusion in one view, the homography from this view to other views is estimated from previous tracking results and used to infer the correct transformation for the occluded view. The experiment results show that the tracking process in outdoor and dynamic environment can be done correctly, although some problems still occur to be a challenge for the next research.

Keywords: *Outdoor environment, person tracking.*

1. Introduction

Despite the legitimacy of a number of privacy issues, many systems have been deployed for surveillance applications. The important of video surveillance techniques [1], [2], [3] has increased considerably since the latest terrorist incidents. Security has become critical in many public areas, and there is a specific need to enable human operators to remotely monitor activity across large environments such as airports, banks, shopping malls, military bases, etc. They generate large amounts of data that needs to be filtered out either for online detection of dangerous situations, or for offline information retrieval.

Up to now those tasks are performed by human operators, of which a huge number is required if online analysis is required. Most of the data is stored in video archives without even being analyzed and is currently only used after the fact as forensic tool, losing its primary benefit as an active real time media.

Therefore the demand for automated surveillance systems, providing a decision support interface to enhance the performance of a human operator seems reasonable. These unlawful acts could be detected in time or even prevented. This is not easy task, as the system has to deal with large crowds resulting in severe occlusion, difficult and fast changing lighting situations, and views that are very narrow or too wide.

In order to track the person in the views, the method we propose in this paper uses Kalman filter to accomplish the single view tracking. Occlusion is detected using certain threshold. When occlusion is detected in one view, the homography between two views is estimated from previous tracking results.

Kalman filter is used as basis in order to predict an object's position x_k in the next frame x_{k+1} using a motion model [4]:

$$x_{k+1} = F_{k+1,k} * x_k + w_k$$

In a subsequent step a measurement is performed to find the object fitting best to the prediction:

$$y_k = H_k * x_k + v_k$$

ISBN. 978-979-1222-95-2

The random variables w_k and v_k represent the process and measurement noise respectively. The n x n matrix F relates the state at the previous time step k to the state at the current step $k+1$, in the absence of process noise. Note that in practice A might change with each time step, but here we assume it is constant. The m x n matrix H relates the state to measurement y_k. In practice H might change with each step or measurement, but here we assume it is constant.

Suppose P is a scene point lying on a plane π. Let p and p' be the projections of P in view 1 and view 2, respectively. Then there exists a 3x3 matrix H_π such that $p' \cong H_\pi p$ where H_π is called the homography matrix of the plane π [5,6].

Suppose at frame j occlusion is detected for view 2, but not for view 1. Denote $T_1^{\,j}$ and $T_2^{\,j}$ is the transformations from frame j-1 to frame j for view 1 and view 2 respectively. We need to derive $T_2^{\,j}$ from H and $T_1^{\,j}$. Let x^{j-1} and x'^{j-1} be a pair of corresponding points at frame $j-1$ for view 1 and view 2 respectively. Then we have

$$x^j = T_1^{\,j} x^{j-1} \quad ; \quad x'^j = T_2^{\,j} x'^{j-1} \tag{1}$$

and

$$x'^{j-1} = Hx^{j-1} \quad ; \quad x'^j = Hx^j \tag{2}$$

Knowing H and $T_1^{\,j}$, it is easy to derive from (1) and (2) that

$$T_2^{\,j} = HT_1^{\,j} H^{-1} \tag{3}$$

Although (3) gives a theoretically correct solution for $T_2^{\,j}$, practically $T_2^{\,j}$ can be obtained from x'^{j-1} and x'^j.

2. Methodology

Each video input we do background segmentation to get the foreground. Then we find blobs for the detected person. Using the position of the blob we use Kalman filter to determine the next position of each blob. When occlusion is detected in one view, the homography between two views is estimated from previous tracking results to infer the target motion in the occluded view.

3. Experimental Result

Experiments were conducted using two synchronized videos shot by two cameras with overlapping views. These videos were taken from PETS research result (http://www.cvg.cs.rdg.ac.uk).

Figure 1 shows the snapshots of tracking result. Fig 1.a and Fig 1.b are two snapshots of the tracking result for the un-occluded view in camera 1. Their corresponding views in camera 2 can be seen in Fig 1.c and Fig 1.d respectively. Furthermore tracking occlusion handling result is depicted in Fig 1.f where the person can be followed by the tracker. It is clearly different to Fig 1.e where the person is loss from the tracker.

4. Conclusion

We have described a two view person tracking approach which uses Kalman filter to predict the tracking position and uses the homography to handle the occlusion. Experiment results show that this method is robustly to track person in outdoor and dynamic environment whether there is occlusion or not.

5. References

[1] Chen, H.M., Lee, S., Rao, R.M., Slaman, M.A., and Varshney, P.K., "Imaging for concealed weapon detection". IEEE signal processing Mag. Vol.22, No.2, pp.52-61. Mar.2005

ISBN. 978-979-1222-95-2

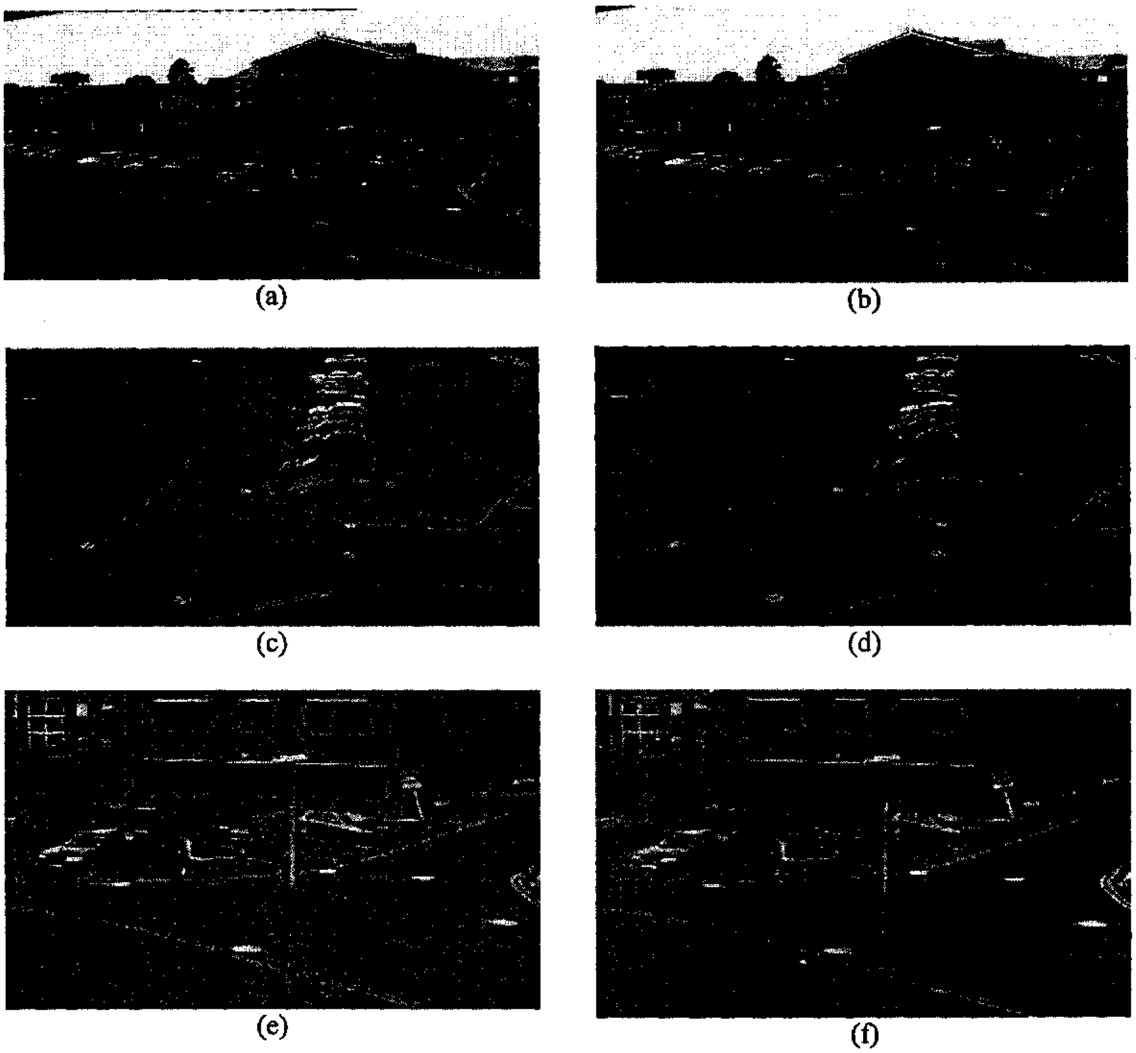

Figure.1. tracking results: (a)-(b) tracking result for the un-occluded view in view 1. (c)-(d) the corresponding tracking result for the un-occluded view in view 2.. (e) tracking result for occluded view without occlusion handling. (f) tracking result for occluded vie with occlusion handling

[2] Foresti, G.L., Regazzoni, C.S., and Visvanathan, R. "Scanning the issue/technology-special issue on video communications, processing and understanding for third generation surveillance systems. Proc. IEEE, vol 89, no 10, pp.1355-1367. oct 2001.

[3] James, B.,Tim, E. & Rosin, P., "A novel method for video tracking performance evaluation" Proc. Joint Int. IEEE workshop: visual surveillance and performance evaluation of tr... and surveillance . Nice-france, pp. 125-132, Oct 2003.

[4] Arsic, D., Hristov, E.,Lehment,N.,Hornler,B., Schuller,B., & Rigoll, G. Applying multi layer homography for multi camera person tracking. In Interational conference on distributed smart camera. 2008.

[5] Hartley, R., and Zisserman, A. Multiple view geometry in Computer Vision. Cambridge University Press. 2000

[6] Eshel, R & Moses, Y. "Tracking in a dense crowd using multiple cameras". Int J Computer Vision Vol 88 pp.129-143. 2010.