OBJECT DETECTION AND TRACKING USING KALMAN FILTER AND FAST MEAN SHIFT ALGORITHM

A. Ali and K. Terada
The University of Tokushima, 2-1 Minami-Josanjima, Tokushima, 770-8506 Japan

Keywords: Change Detection, Object Tracking, Kalman filter, Mean shift Algorithm.

Abstract: Object detection in videos involves verifying the presence of an object in image sequences and possibly locating it precisely for recognition. Object tracking is to monitor an object's spatial and temporal changes during a video sequence, including its presence, position, size, shape, etc. These two processes are closely related because tracking usually starts with detecting objects, while detecting an object repeatedly in subsequent image sequence is often necessary to help and verify tracking. In this paper, a novel approach is being presented for detecting and tracking object. It includes combination of Kalman filter and fast mean shift algorithm. Kalman prediction is measurement follower. It may be misled by wrong measurement. In order to cater it, fast mean shift algorithm is used. It is used to locate densities extrema, which gives clue that whether Kalman prediction is right or it is misled by wrong measurement. In case of wrong prediction, it is corrected with the help of densities extrema in the scene. The proposed approach has the robust ability to track the moving object in the consecutive frames under some kinds of difficulties such as rapid appearance changes caused by image noise, illumination changes, and cluttered background.

1 INTRODUCTION

Object detecting and tracking has a wide variety of applications in computer vision such as video compression, video surveillance, vision-based control, human-computer interfaces, medical imaging, augmented reality, and robotics. It also plays an important role in video database such as content-based indexing and retrieval. Change detection techniques presented in the literature can be divided in two classes: pixel-based and region-based algorithms(Jain, 1989). Pixel-based algorithms compute the output by analyzing the values assumed by correspondent pixels in the two analyzed images; region-based algorithms compare features extracted from correspondence regions in the two images. Pixel-based methods, (e.g. Change detection based on binary difference) present the advantage of the simplicity that makes possible real-time applications, whereas region-based techniques (e.g. Change detection based on the illumination model(Skifstadt and Jain, 1989)) provide results more robust to false alarms introduced by noise. A further class of CD algorithms detects changed regions by means of edge comparisons(Jain, 1989). The Kalman filter has been extensively used in the vision community for tracking. Broida and Chellappa(Broida and Chellappa, 1986) used the Kalman filter to track points in noisy images. In stereo camera-based object tracking, Beymer and Konolige(Beymer and Konolige, 1999) use the Kalman filter for predicting the objects position and speed in x - z dimensions. Rosales and Sclaroff(Rosales and Sclaroff, 1999) use the extended Kalman filter to estimate 3D trajectory of an object from 2D motion. A common approach to handle complete occlusion during tracking is to model the object motion by linear dynamic models or by nonlinear dynamics and, in the case of occlusion, to keep on predicting the object location until the object reappears. For example, a linear velocity model is used in Beymer and Konolige(Beymer and Konolige, 1999) and a Kalman filter is used for estimating the location and motion of objects. For the image segmentation problem, Mean-Shift Clustering is commonly used. Comaniciu and Meer(Comaniciu and Meer, 2002) propose the mean-shift approach to find clusters in the joint spatial and color space.

In this paper we propose a novel object tracking scheme showing good tracking performance in the consecutive frames under some kinds of difficulties. It combine Kalman filter and fast mean shift algo-

Ahmed A. (2009).
OBJECT DETECTION AND TRACKING USING KALMAN FILTER AND FAST MEAN SHIFT ALGORITHM.
In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 585-589
DOI: 10.5220/0001787705850589

rithm. The paper is organized as follows: section 2 gives a brief overview on the tracking algorithm and also gives details on the proposed tracking algorithm ; The experimental results are demonstrated in Section 3. Finally, the paper is concluded in section 4.

2 TRACKING ALGORITHM

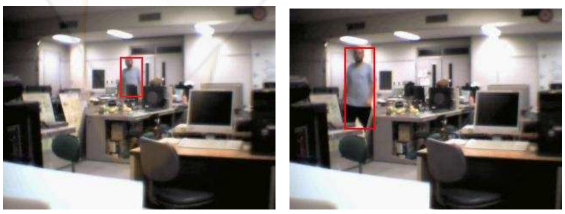
Object tracking is the problem of estimating the positions and other relevant information of moving objects in image sequences. Two-frame tracking can be accomplished using correlation-based matching methods, optical flow techniques, or change-based moving object detection methods. In this paper, the tracking accomplished by using change-based moving object detection method. The tracking algorithm is briefly described in next subsections.

2.1 Change-based Object Tracking Method

Change detection by background subtraction is a common approach to detect moving foreground. The resulting difference image is usually thresholded to obtain objects based on pixel connectedness and resulting blob objects are subsequently tracked(Bovik, 2000). The Change-based tracking algorithm can be described as follows:

- Image subtraction to detection motion blobs:
  - Compute the difference image between two frames
  - Thresholding to find the blobs
  - Locate the blob center as the position of object
- Track the object by track the blob center

For target with not so much cluttered background or little illumination changes and there is no motion in the images being used to generate the stationary background, results obtained by this technique is of very high accuracy as shown in fig.1.



(a) Frame 151 (b) Frame 226

Figure 1: Tracking using change-based detection.

2.2 Difficulties with the Tracking Method

In change detection where the reference frame is a stationary background image that has been generated in an initialization phase using a simple averaging, it is of paramount importance that there is no motion in the images being used to generate the stationary background. However, in practice, it is very difficult or almost impossible to ensure that there is no motion during the initialization phase in real-life applications (in many situations where you do not have full control over the environments – for example, in traffic monitoring and video surveillance applications)(Ong and Spann, 1999). Cheung and Kamath(Cheung and Kamath, 2004) studied the methods for generating a background model, such as frame differencing, median filter, linear predictive filter, non-parametric model, kalman filter and Mixture of Gaussians model. An example of such situation, only change-based tracking technique didn't correctly track the object as shown in fig.2.

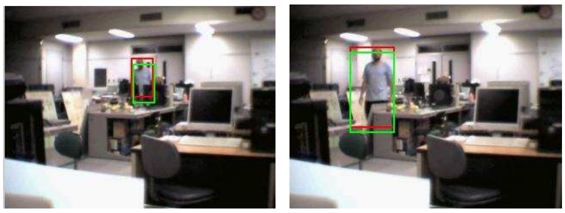
(a) Frame 135 (b) Frame 196

Figure 2: The tracker lost the object.

In order to tackle the above-mentioned problem, Kalman tracker was embedded in the algorithm. Kalman filter continuously predicts the next state of the target(Grewal and Andrews, 2001) based on the measurement of change-based technique. Kalman predictor helps in finding where the object is most likely to be found in the next frame. A good result obtained after embedding Kalman Tracker as shown in fig.3. its clear that Kalman tracker successfully track the object.

2.3 Tracking Using Kalman Filter

A Kalman filter is used to estimate the state of a linear system where the state is assumed to be distributed by a Gaussian. Kalman filtering is composed of two steps, prediction and correction(Bar-Shalom and Foreman, 1988). The Kalman filter a recursive estimator. This means that only the estimated state form the previous time step and the current measurement are needed to compute the estimate for the current state. In contrast to batch estimation techniques, no history of observations and/or estimates is required. Fig.3(a) and (b) show this scenario. Red window shows the location where the target is detected, while green window shows the Kalman prediction window. It was decided that when the measurement differ from the predicted position by more than a certain threshold value, then the measurement will be discarded and the Kalman prediction will be taken as the next position of the target.



(a) Frame 130 (b) Frame 215

Figure 3: The object successfully tracked after embedding Kalman tracker; the red rectangle show the tracking using change-based and the green rectangle show the tracking using Kalman.

2.4 Difficulties Using Kalman Filter

The major problem with the tracking was that of the target detection. Measurements given by the detector are sometimes very noisy due to rapid appearance changes caused by image noise, illumination changes, cluttered background. Kalman prediction is measurement follower. It may be misled by wrong measurement. An example of such situation is shown in fig.4.

(a) (b)

Figure 4: Kalman filter lost the object by wrong measurements.

In order to treat it, fast mean shift algorithm is used. It is used to locate densities extrema, which gives clue that whether Kalman prediction is right or it is misled by wrong measurement. In case of wrong prediction, it is corrected with the help of densities extrema in the scene. Local density maxima in the difference image – usually representing moving objects – are outlined by a fast non-parametric mean shift clustering procedure.

2.5 Fast Mean Shift Algorithm

The mean shift algorithm is a nonparametric technique to locate density extrema or modes of a given distribution by an iterative procedure(C. Beleznai and Bischof., 2005). Starting from a location x the local mean shift vector represents an offset to x', which is a translation towards the nearest mode along the direction of maximum increase in the underlying density function. The local density is estimated within the local neighborhood of a kernel by kernel density estimation where at a data point a kernel weights $K(a)$ are combined with weights associated with the data. .For digital images sample weights are defined by the pixel in intensities at pixel locations a. the new location vector x' obtained after applying the mean shift offset. Fast mean shift algorithm was embedded in the algorithm. If pixel difference between detected measurement obtained by change-based method (the blob center of the object)and Kalman filter predicted position is greater than a certain threshold, then motion region is detected by triggering fast mean shift algorithm and it guides Kalman filter, by finding actual motion region in the scene, for more accurate prediction of next state of the object. Below the proposed tracking algorithm scheme can be described by the following diagram (fig.5).

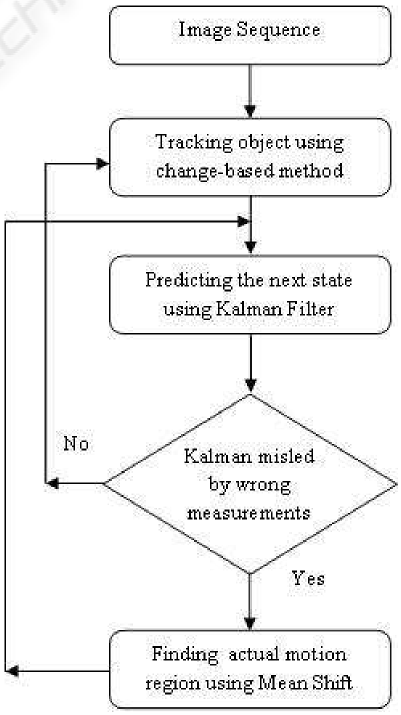
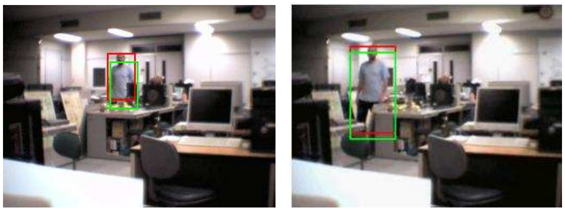


Figure 5: The proposed Tracking Algorithm scheme.

3 EXPERIMENTS AND RESULTS

This section show the results obtained by the proposed algorithm. The system was tested using Intel(R) Core Duo CPU device with 2.10 GHZ and 3GB of RAM and Window XP operation system, and Matlab R2007b are used. In addition, four image sequences(both indoors and outdoors) are used for testing the system and each image sequence consists of video frames with 320x240 resolutions per frame. It is clear from fig.6 (a) that a person walking in a cluttered background under the light sources and lighting changes reflected by the white boards put aside.



(a) Image sequence one

(b) Image sequence two

(c) Image sequence three

Figure 6: Kalman tracker successfully tracked the object guided by mean shift algorithm.

Figure 7: Image sequence four; Kalman tracker successfully tracked pedestrian in outdoor scenes guided by mean shift algorithm.

After emerging mean shift algorithm, the object successfully detected by mean shift algorithm and precisely tracked by Kalman predictor by following the true measurement taken by mean shift detector. Another example to test the system using different noisy images caused by different lighting sources as shown in fig.6(b). the system was chalenged with a very noisy image sequence caused by rapid illumination changes due to different sources of lighting and due to the walking person shadow as shown in fig.6(c). An example of outdoor waking person to test the system by using PETS 2001 Image Sequences. The results show that the system excellently tracked the object as shown in fig.7. The satisfactory results are achieved. By comparing the results taken only using Kalman filter (see fig. 4) and the result taken by the system after embdding Mean Shift Algorithm (see fig. 6), its clear that the system overcomes these difficults by successfully detected the motion area. Finally, the performance of the system was evaluated using the processing time of the test video sequences as shown in Table 1. The average processing time of every frame is between 0.025 and 0.031 second. Therefore, it can be applied to real-time application.

Table 1: The processing time of test video sequences.

Test video seqs	No. Frames	Track. T.(s)
video seq.1	50	1.56
video seq.2	70	1.95
video seq.3	60	1.71
video seq.4	100	2.53

4 CONCLUSIONS

This paper presents a novel approach for object detection and tracking. It includes combination of Kalman filter and fast mean shift algorithm. The system start by detecting the object using change detection technique and track the object by track the blob center of the detected object. To increase the efficiency of the algorithm, Kalman filter works parallel with the tracking algorithm. Kalman filter continuously predicts the next state of the target based on the measurement of change-based technique. Under some kinds of difficulties such as rapid appearance changes caused by image noise, illumination changes, and cluttered background, Kalman prediction may be misled by wrong measurement. If pixel difference between detected measurement obtained by change-based method and kalman filter predicted position is greater than a certain threshold, then fast mean shift algorithm is used. It is used to locate densities extrema, which gives clue that whether Kalman prediction is right or it is misled by wrong measurement. In case of wrong prediction, it is corrected with the help of densities extrema in the scene. The proposed method is an efficient video object tracking algorithm. Furthermore, to consider the situations of tracking multiple objects, every one of multiple objects can be set an Kalman filter to track it. Also since the processing time using the proposed method to track the moving object is short, therefore, the system implemented by the proposed method can afford to track a moving objects in real time.

REFERENCES

Bar-Shalom, Y. and Foreman, T. (1988). *Tracking and Data Association*. Academic Press Inc.

Beymer, D. and Konolige, K. (1999). Real-time tracking of multiple people using continuous detection. In *IEEE International Conference on Computer Vision (ICCV) Frame-Rate Workshop*.

Bovik, A. (2000). *Handbook of Image and Video Processin*. Academic Press.

Broida, T. and Chellappa, R. (1986). Estimation of object motion parameters from noisy images. In *IEEE Trans. Patt. Analy. Mach. Intell.*

C. Beleznai, B. F. and Bischof., H. (2005). Ieee international workshop on visual surveillance and performance evaluation of tracking and surveillance. In *Breckenridge*.

Cheung, S. C. S. and Kamath, C. (2004). Robust techniques for background subtraction in urban traffic video. In *SPIE Electronic Imaging: Video Communications and Image Processing*.

Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. In *IEEE Trans. Patt. Analy. Mach. Intell*.

Grewal, M. S. and Andrews, A. G. (2001). *Kalman Filtering Theory and Practice using MATLAB, second edition*. John Wiley & Sons, Inc.

Jain, A. K. (1989). *Foundamentals of Digital Image Processing*. Prentice-Hall.

Ong, E. P. and Spann, M. (1999). Robust optical flow computation based on least-median of squares regression. In *International Journal of Computer Vision*.

Rosales, R. and Sclaroff, S. (1999). 3d trajectory recovery for tracking multiple objects and trajectory guided recognition of actions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Skifstadt, K. and Jain, R. (1989). Illumination independent change detection for real world sequences. In *Graphics, and Image Processing*.