

Available online at

<http://www.TJMCS.com>

The Journal of Mathematics and Computer Science Vol. 4 No.2 (2012) 120 - 128

A new method for video object tracking

saeid bagheri-golzar¹, Fariba Karami-sorkhechaghaei^{*2},
Amir-Masud Eftekhari-Moghadam³

1Dept. of Elec., comp. & IT, Qazvin Islamic Azad University, Qazvin, Iran,
saeid_golzar4439@qiau.ac.ir

2Dept. of Elec., comp. & IT, Qazvin Islamic Azad University, Qazvin, Iran,
fariba_karami@qiau.ac.ir

3Dept. of Elec., comp. & IT, Qazvin Islamic Azad University, Qazvin, Iran,
eftkhari@qiau.ac.ir

Received: January 2012, Revised: April 2012

Online Publication: June 2012

Abstract

In this paper, a new video moving object tracking method is proposed. In initialization, a moving object selected by the user is segmented and the dominant color is extracted from the segmented target. In tracking step, a motion model is constructed to set the system model of adaptive Kalman filter firstly. Then, the dominant color of the moving object in HSI color space will be used as feature to detect the moving object in the consecutive video frames. The detected result is fed back as the measurement of adaptive Kalman filter and the estimate parameters of adaptive Kalman filter are adjusted by occlusion ratio adaptively. The proposed method has the robust ability to track the moving object in the consecutive frames under some kinds of real-world complex situations such as the moving object disappearing totally or partially due to occlusion by other ones, fast moving object, changing lighting, changing the direction and orientation of the moving object, and changing the velocity of moving object suddenly. The proposed method is an efficient video object tracking algorithm.

Keywords: Adaptive Kalman filter, moving object, HSI color space.

1. Introduction

The researches for tracking semantic objects in video have received great attention for the last few years [1-26]. The moving object tracking is an important issue in video system, such as surveillance, sports reporting, video annotation, and traffic

* Corresponding Author

management system. However, if the background and moving objects vary dynamically, such situations will be complicated. In video analysis, we have to know the features of moving objects, such as color, texture, and shape, etc., so that the moving object can be detected and be tracked. There are some situations for video in a real world environment, including camera lens is fixed or not, multiple moving objects, rigid object, or non-rigid object, occlusion by the other objects, one or many cameras, full automatic or semi-automatic semantic object tracking, etc. According to the above discussion, to conclude, we will encounter the following three problems for tracking moving object.

- Initial moving object segmenting problem.
- Detection of moving object.
- Tracking the moving target in occlusion.

In this paper, a semi-automatic method is used. In the environment of still background [3, 8] apply the difference information of two consecutive frames to extract the moving objects.

For the sake of simplicity, most of methods assume that they are in the environment of still background in the initial state. About the detection of moving object, most of techniques base on the features of color, texture, shape, edge, motion, and shape. Generally, the color feature is frequently used [9, 14, 17, 20], since the human perception is sensitive to the color. However, the disadvantage of the single color feature for moving object detection is that it can only be used in the target object with uniformly distributed color.

Jang and Choi [2] propose an active model to detect and track the moving object. However, the method [2] with the Kalman filter technique which predicts the detecting range to reduce the computational complexity can not be applied to solve the target in occlusion.

in this paper, a simple and efficient method for tracking object using adaptive Kalman filter model is developed and its the robustness quality can be applied to a real-world environment.

2. The proposed method

In this paper, the initial moving object is selected by user. Because there may exist more than one moving objects in video, therefore the tracking target should be determined according to the user interested. After the selection, we initially segment the object in frame t by the difference of three consecutive frames, $t-1$, t , $t+1$, and then apply the region growing algorithm to segment the desired object. Afterwards, the dominant color feature of moving target is extracted from the RGB color space by K-means algorithm. After the initial moving object segmentation and feature extraction, to detect the moving object in consecutive frames, the HSI color model is used to build up the resistance to lighting changes, since its three components are relatively independently. In addition, we proposed an adaptive Kalman filtering algorithm to effectively track the moving object. The ratio of the moving object area in frame t to that in frame $t-1$ is used to as the object occlusion ratio .The occlusion ratio will be used to adaptively adjust the estimate parameters of the Kalman filter. Fig. 1 demonstrates the system structure.

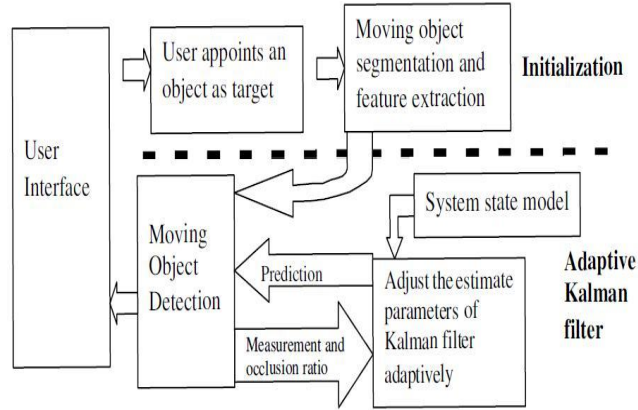


Fig1: The system structure of the proposed method.

2.1. Initial moving object segmentation and feature extraction

The purpose of initial moving object segmentation is to segment the user’s interested target and extract the feature of the selected target. The traditional frame difference and region growing methods are applied to segment the moving object.

2.1.1. Frame difference and region growing

Frame difference is a simple method to segment the moving object in image video. After the user appoints a moving object as target, the target is segmented by the differences of frames in $t-1$, t , and $t+1$. It uses the difference of consecutive frames to detect the change area of frames. The differential frame can be defined as

$$FD(x, y, t) = \begin{cases} 0 & \text{if } |f(x, y, t + 1) - f(x, y, t)| \leq \text{threshold} , \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

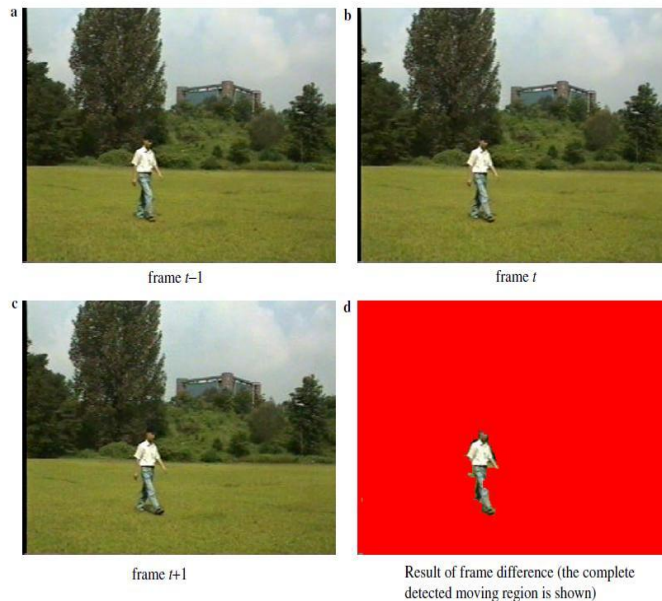


Fig 2: An example of frame difference

To improve the segmentation accuracy, three consecutive frames are used to segment the moving object. The segmented moving object is obtained through the intersect of

the two differences, $FD(x, y, t-1)$ and $FD(x, y, t)$, which are the differences between three consecutive frames. Therefore, the points of segmented moving object is defined as

$$MR(x, y, t) = FD(x, y, t - 1) \cap FD(x, y, t). \quad (2)$$

Fig. 2 demonstrates an example using the frame difference method to segment the moving object. The complete detected moving object is extracted from the original image and is shown in Fig. 2d.

2.1.2. Dominant color feature extraction

To find the dominant color, we process feature extraction in the RGB color space. The RGB color space is represented by a 3-dimensional cube with red, green, and blue additive primaries. When color clustering by K-means method operates in RGB space, it can directly use identical weighting in every dimension to calculate, since the RGB color space forms a cube and the clustering does not need to consider the special property of every dimension. For example, the cyclic property of the hue component in HSI color space needs to be considered. Therefore, the RGB color space is very suitable for dominant color extraction.

First, the moving object colors in RGB components are classified by K-means [27]. The K is experimentally set as 5. The color pixels of moving object are classified into different K groups by color and the average color of the every group in frame t is defined as $MC_k(t)$. If the group d is the highest density (minimum distance summation) one, the $MC_d(t)$ is set as the dominant color feature. The dominant color of the moving object in frame t is defined as

$$DomC(t) = \left\{ MC_d(t) \left| \min \left(\frac{\sum_{(x,y) \in C_k} \sqrt{(x-\bar{x}_{C_k})^2 + (y-\bar{y}_{C_k})^2}}{N_{C_k}}, k = 1, \dots, 5 \right) \right. \right\} \quad (3)$$

2.2. Tracking the moving object by the proposed adaptive Kalman filter

For the tracking problem, if we have to estimate the motion information of object under such a deteriorating condition as occlusion, a simple detecting method will not be able to measure and track the moving object correctly. In the paper, we propose an adaptive Kalman filter to implement. In the proposed adaptive Kalman filter, a moving model is constructed to set the system state model and the result of the moving object detection in consecutive frames is fed back as the measurement for correction. In addition, the estimate parameters of Kalman filter are adjusted adaptively.

2.2.1. The typical Kalman filter

The Kalman filter has been used in engineering application successfully. The Kalman filter has two distinctive features. One is that its mathematical model is described in terms of state-space concepts. The other is that its solution is computed recursively. Usually, the Kalman filter is described by system state model and measurement model. The state-space model is described as

$$s(t) = \ddot{o}(t - 1)s(t - 1) + w(t) \quad (4)$$

The Measurement model is described as

$$Z(t) = H(t) s(t) + v(t) \quad (5)$$

Where $O^-(t-1)$ and $H(t)$ are the state transition matrix and measurement matrix, respectively. The $w(t)$ and $v(t)$ are white Gaussian noise with zero mean and

$$\begin{aligned} E\{W(k)W^T(l)\} &= Q\delta_{kl}, \\ E\{V(k)V^T(l)\} &= R\delta_{kl}, \end{aligned} \tag{6}$$

where δ_{kl} denotes the Kronecker delta function; Q and R are covariance matrices of $w(t)$ and $v(t)$, respectively. The state vector $s(t)$ of the current time t is predicted from the previous estimate and the new measurement $z(t)$.

The tasks of the Kalman filter have two phases: prediction step and correction step. The prediction step is responsible for projecting forward the current state, obtaining a priori estimate of the state $s^-(t)$. The task of correction step is for the feedback. That is to say, it incorporates an actual measurement into the a priori estimate to obtain an improved a posteriori estimate $s^+(t)$.

2.2.2. Motion model construction

In Kalman filter algorithm, the system state model is applied in the prediction step. Before using the Kalman filter, the system state model has to be determined. In our system, the system state model is the moving model. Since the time of the frame interval is very short, it is assumed that the moving object is in uniform velocity within a frame interval. Therefore, the velocity of moving object can be replaced as system parameter by position. In addition, the moving distance in every interval is equal and is set as

$$d(t) = d(t-1) + (d(t-1) - d(t-2)) \tag{7}$$

$$s(t) = \ddot{o}(t-1)s(t-1) + w(t) = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} d(t-1) \\ d(t-2) \end{bmatrix} + \begin{bmatrix} w(t) \\ 0 \end{bmatrix},$$

$$s(t) = \begin{bmatrix} d(t) \\ d(t-1) \end{bmatrix}, s(t-1) = \begin{bmatrix} d(t-1) \\ d(t-2) \end{bmatrix} \text{ and}$$

$$\ddot{o}(t) = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \tag{8}$$

2.2.3. Moving object detection in consecutive frames

In this stage, the position of the moving object in consecutive frame will be detected. Once the dominant color feature in RGB color space is obtained, the color feature is converted from RGB color space to HIS color system, and then a color matching function is used to measure the area of the moving object. The main reason for using HSI color space is to increase tolerance toward light changes in video, The matching function for detection is written as

$HSI_{Matching}(ref, cur)$

$$= \begin{cases} 1 & \begin{aligned} &|ref_H - cur_H| < th_H, \\ &if \ |ref_S - cur_S| < th_S, \\ &|ref_I - cur_I| < th_I \end{aligned} \\ 0 & otherwise \end{cases} \tag{9}$$

where ref is the HSI dominant color in previous frame and cur is image pixels of search range in current frame. The search range depends on the matched object of previous frame. The th_H , th_S , and th_I are the thresholds for matching. In HSI color space, its three components (hue, saturation, and intensity) are relatively independent, therefore, in Eq. (9), the comparisons are made separately. If all of the three components are less than their thresholds, then the pixel is in the target object. Otherwise, it is the background. In Fig. 3, we use an image to evaluate the HSI matching function. The three

components of dominant color in HSI color space are $H = 3.24$, $S = 0.05$, and $I = 0.89$. In addition, the thresholds are set as $th_H = 1.05$, $th_S = 0.5$, and $th_I = 0.5$. Note that the illumination in Figs. 3a and c are different. The results are shown in Figs. 3b and d, respectively. Although the HSI matching function cannot segment the moving object completely, the result is good enough to represent the object area.



Fig3: The experimental results of HSI color matching method. The illumination in (a) and (c) are different.

In addition, to represent the position of the object, we have to find the center of the matched object area. We use a rectangle to represent the moving object's area and the rectangle is the minimum one to cover the moving object completely. Then, the center will be $(cx, cy) = (\frac{left+right}{2}, \frac{top+bottom}{2})$, where (left, top) and (right, bottom) are the rectangle's corner positions at left upper and right lower corners, respectively. The rectangle area will be extended some pixels as the search range of the next frame. The left upper and right lower corner positions of search range, (search_left, search_top) and (search_right, search_bottom), respectively. They are defined as

$$\begin{aligned}
 Search_left &= (pre_left - (d(t-1) - d(t-2))), \\
 Search_top &= (pre_top - (d(t-1) - d(t-2))), \\
 Search_right &= (pre_right + (d(t-1) - d(t-2))), \\
 Search_bottom &= (pre_bottom + (d(t-1) - d(t-2))),
 \end{aligned} \tag{10}$$

Where the prefix pre-represents the previous frame. The $d(t-1)$ and $d(t-2)$ are the object moving distances in frame $t-1$ and $t-2$, respectively. That is to say, the search range considers the object's moving speed. The center of search range is located at the predicted position of adaptive Kalman filter. The Fig. 4 shows the structure of moving object detection using color matching method in HSI color space.

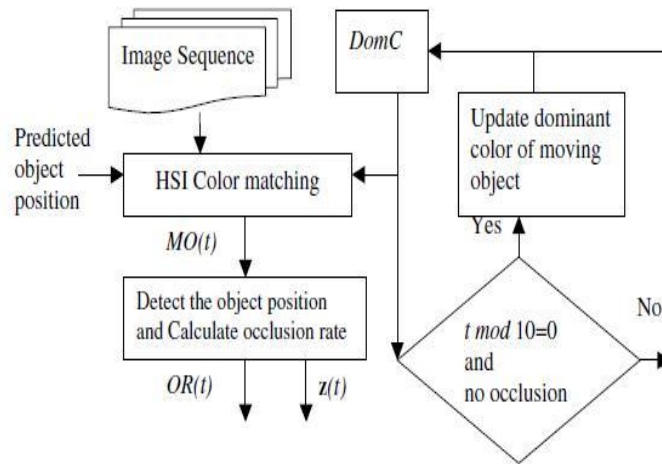


Fig4: The structure of moving object detection.

In Figure 5 an example of the results of the proposed method on a sequence of consecutive frames is provided. At first by the Kalman filter, the coordinates of the moving object center is estimated then dominant color is searched in the around of coordinates is estimated.

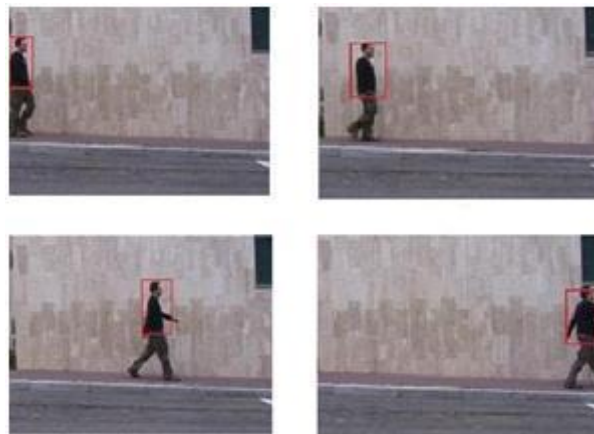


Fig5: The experimental results of moving object

In Figure 6 an example of the results of the proposed method on a sequence of consecutive frames is provided. The example is shown that the proposed method in conditions such as changing velocity and direction works well.

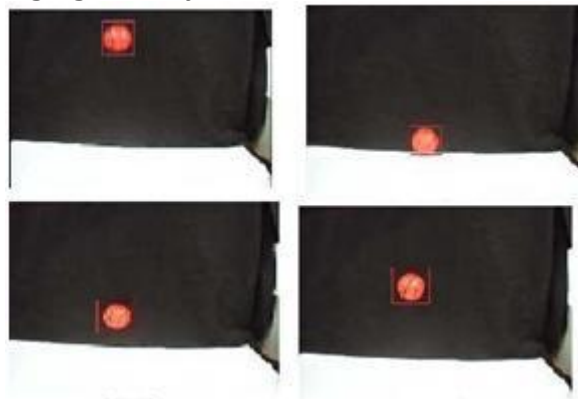


Fig6: The experimental results of moving object with changing direction and velocity.

3. Conclusion

In this paper, an effective adaptive Kalman filter is proposed to track the moving object. In the proposed adaptive Kalman filter, the occlusion rate is used to adjust the error covariance of Kalman filter adaptively. The method can track the moving object in real-time. It successfully estimates the object's position in some kinds of real-world situations such as the fast moving object, partial occlusion, long-lasting occlusion, changing lighting, changing the direction and orientation of the moving object, and changing the velocity of moving object suddenly. Furthermore, to consider the situations of tracking multiple objects, every one of multiple objects can be set an adaptive Kalman filter to track it. Since the processing time using the proposed method to track the single object is short, therefore, the systems implemented by the proposed method can afford to track multiple objects in real time.

References

- [1] Dae-Sik Jang, Seok-Woo Jang, Hyung-Il Choi, 2D human body tracking with structural Kalman filter, *Pattern Recognition* 35 (2002) 2041–2049.
- [2] Dae-Sik Jang, Hyung-Il Choi, Active models for tracking moving objects, *Pattern Recognition* 33 (2000) 1135–1146.
- [3] Shao-Yi Chien, Shyh-Yih Ma, Liang-Gee Chen, Efficient moving object segmentation algorithm using background registration technique, *IEEE Transactions on Circuits and Systems for Video Technology* 12 (7) (2002) 577–586.
- [4] Changick Kim, Jenq-Neng Hwang, Fast and automatic video object segmentation and tracking for content-based applications, *IEEE Transactions on Circuits and Systems for Video Technology* 12 (2) (2002) 122–129.
- [5] Dong Kwon Park, Ho Seok Yoon, Chee Sun Won, Fast object tracking in digital video, *IEEE Transactions on Consumer Electronics* 46 (3) (2000) 785–790.
- [6] Greg Welch, Gary Bishop, An Introduction to the Kalman Filter, UNC-Chapel Hill, TR 95-041, March 11, (2002).
- [7] Yiwei Wang, John F. Doherty, Robert E. Van Dyck, Moving object tracking in video, in: *Proceedings of the 29th workshop of Applied Imagery Pattern Recognition*, 2000, October 16–18, 2000, pp. 95–101.
- [8] Alan J. Lipton, Hironobu Fujiyoshi, Raju S. Patil, Moving target classification and tracking from real-time video, *Proceedings of the Applications of Computer Vision, WACV'98*, October 19–21, 1998, pp. 8–14.
- [9] Johnson I. Agbinya, David Rees, Multi-object tracking in video, *Real-Time Imaging* (1999) 295–304.
- [10] Roberta Piroddi, Theodore Vlachos, Multiple-feature spatiotemporal segmentation of moving sequences using a rule-based approach, *British Machine Vision Conference* (2002) 353–362.
- [11] Volker Rehrmann, Object oriented motion estimation in color image sequences, in: *ECCV '98*, vol. I, *Lecture Notes of Computer Science* 1406, 1998 704–719.
- [12] Michael Mason, Zoran Duric, Using histograms to detect and track objects in color video, *Applied Imagery Pattern Recognition Workshop, AIPR 2001 30th*, October 10–12, 2001, pp. 154–159.
- [13] Georgi D. Borshukov, Gozde Bozdagi, Yucel Altunbasak, A. Murat Tekalp, Motion segmentation by multistage affine classification, *IEEE Transactions on Image Processing* 6 (11) (1997) 1591–1594.

- [14] George V. Paul, Glenn J. Beach, Charles J. Cohen, A Real-time object tracking system using a color camera, Applied Imagery Pattern Recognition Workshop, AIPR 2001 30th, October 10–12, 2001, pp. 137–142.
- [15] Jong Bae Kim, Hye Sun Park, Min Ho Park, Hang Joon Kim, A Real-Time Region-Based Motion Segmentation Using Adaptive Thresholding and K-means Clustering, Kyungpook National University, Computer Engineering, South Korea.
- [16] H. Tao, H.S. Sawhney, R. Kumar, Dynamic layer representation with applications to tracking, Proceedings of the Computer Vision and Pattern Recognition 2 (2000) 134–141.
- [17] Y. Wu, T.S. Huang, A co-inference approach to robust visual tracking, Proceedings of the International Conference and Computer Vision 2 (2001) 26–33.
- [18] A.D. Jepson, D.J. Fleet, T.F. El-Maraghi, Robust online appearance models for visual tracking, Proceedings of the Computer Vision and Pattern Recognition (2001).
- [19] J. Vermaak, P. Perez, M. Gangnet, A. Blake, Towards improved observation models for visual tracking: selective adaptation, Proceedings of the European Conference on Computer Vision 1 (2002) 645–660.
- [20] Hieu T. Nguyen, Arnold W.M. Smeulders, Fast occluded object tracking by a robust appearance filter, IEEE Transaction on Pattern Analysis and Machine Intelligence 26 (8) (2004) 1099–1104.
- [21] T.B. Moeslund, E. Granum, A survey of computer vision-based human motion capture, Computer Vision and Image Understanding 81 (2001) 231–268.
- [22] H. Sidenbladh, M.J. Black, D.J. Fleet, Stochastic Tracking of 3D Human Figures Using 2D Image Motion, Proceedings of the European Conference on Computer Vision (2000) 702–718.
- [23] Y. Song, X. Feng, P. Perona, Towards detection of human motion, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2000) 810–817.
- [24] N.T. Siebel, S. Maybank, Fusion of multiple tracking algorithm for robust people tracking, Proceedings of the European Conference on Computer Vision (2002) 373–387.
- [25] T. Zhao, R. Nevatia, Tracking multiple humans in complex situations, IEEE Transaction on Pattern Analysis and Machine Intelligence 26 (9) (2004) 1208–1221.
- [26] A. Senior, Tracking people with probabilistic appearance models, Proceedings of the International Workshop on Performance Evaluation of Tracking and Surveillance Systems (2002).
- [27] Vance Faber, Clustering and the continuous K-means algorithm, Los Alamos Science (22) (1994) 138–144.