Moving Target Tracking Based on CamShift Approach and Kalman Filter

Kuo-Hsien Hsia¹, Shao-Fan Lien² and Juhng-Perng Su³

¹ Department of Computer Science and Information Engineering, Far East University, Tainan, Taiwan
² Department of Electrical Engineering, National Yunlin University of Science and Technology, Yunlin, Taiwan
³ Graduate School of Engineering, National Yunlin University of Science and Technology, Yunlin, Taiwan

Corresponding author: Email: khhsia@cc.feu.edu.tw

Abstract: The surveillance system has been developed for decades. It has reduced the crime and protected the lives and properties of people successfully. Some surveillance systems are composed of PTZ cameras. Therefore the moving people or target objects could be tracked by the surveillance system. The surveillance system becomes more and more interment. However, there are two problems occurred, target recognition and target shelter. In this paper, we proposed the method of tracking a moving object for improving the performance of intelligent surveillance system. The method combined CamShift and Kalman filter for tracking the moving object in a complex background or in an occlusion case. The proposed method is very efficient and that is easily implemented in a real-time system. Three major problems of surveillance system design, moving object detecting, moving object tracking and tracking the object in occlusion, have been conquered in the experimental results.

Keywords: CamShift, Kalman filter, HSV color model, Image tracking.

1 Introduction

The surveillance system has been developed for many years. The surveillance systems are usually used for preventing crimes and that also protected the lives and properties of people successfully. In recent years, the fast development of CCD chip and computer science provides the good basis for multimedia video surveillance system. Today, the surveillance system is an important part for security and vigilance tasks. At present, the PTZ camera is very popular for surveillance system due to the Pan-Tilt-Zoom motions [1]. The PTZ camera could be regarded as a servo system with two-axis motion [2]. In previous researches, many algorithms or control schemes have been proposed for making the surveillance system more intelligently and autonomously. Al Haj, Bagdanov, González and Roca proposed recursive Bayesian Iter framework for estimating the camera and world parameters [3]. In the design of intelligent surveillance systems, two important problems, target recognition problem and target shelter problem, have to be considered.

There are usually several moving targets in one frame. Therefore, developing an excellent method for recognizing and tracking the dubious moving target is especially important. In addition, the moving targets are often obstructed by other objects, thus how to produce the object moving trajectory for PTZ camera tracking is another difficult task.

In this paper, three problems for intelligent surveillance system design will be discussed. They are (1) moving object detecting, (2) moving object tracking and (3) tracking the object in occlusion. Image processing technology promoted the development of intelligent surveillance system.

In previous researches, the binary segmentation and background subtraction method is used for detecting moving objects in [4] and [5],
respectively. Such methods are very effective for segmenting the moving object. However, when the background is changed a lot or the objects are in occlusion, the methods in [4] and [5] are not easy to detect moving objects. Therefore, the 3D information from PTZ camera is used for detecting moving objects, and the issue of background changed has been overcome.

But the problem of object in occlusion is still not mentioned. Haritoglu [6] proposed the silhouette projection method for finding out the moving people in occlusion. Wang [7] used the feature matching and kernel tracking for predicting the targets. We will propose in this paper the method which combines TD (Temporal Differencing) [8], CamShift (Continuously Adaptive Mean Shift) [9] and Kalman filter for solving the problems simultaneously.

TD can detect the moving object very effectively, but is unfavorable for tracking for background changing a lot. Therefore, CamShift is utilized for tracking the target. CamShift is the template matching method in a sequence of images. The template of target is established by the color model HSV (hue, saturation, value) [10]. CamShift can easily catch the moving target. Kalman filter provides an optimal estimation for linear systems. Therefore, it is added to estimate the moving target in occlusion. In the following sections, the proposed method will be discussed. In the experimental results, some simulation cases will be used for demonstrating our method.

2 Moving Target Tracking Method

A. Detecting the moving target in image sequence

The proposed method is illustrated as figure 1. There are three major parts in this method. The first is detecting and segmenting the moving target. The second part is template computation. Last, the CamShift and Kalman filter are used for tracking and locking the moving target.

Commonly used methods in moving target detection are BS (Background Subtraction) [8], TD (Temporal Differencing) and OF (Optical Flow). The computation time of OF is larger than BS and TD. Thus OF is not suitable for a real-time system. The background template is not easy to be computed with BS in changing background. Hence TD is used for detecting the moving object in our application.
Kuo-Hsien Hsia et al.: Moving Target Tracking Based on CamShift Approach

1. Wide Range Pan/Tilt: 340 degrees Pan / 120 degrees Tilt.
2. Excellent Low Light Sensitivity: Minimum illumination less than 1 Lux.
3. 37mm mount for optional conversion lens.
4. Fully Remote Controllable via RS-232C (VISCA) or RS-422.

It can be seen that D70 is an excellent device for surveillance system.

B. HSV Color Space

HSV (hue, saturation, value) is a very useful cylindrical-coordinate representation of points in an RGB color model. The HSV color model describes more color features of objects than RGB model. Figure 3 illustrates the HSV color model.

Since the HSV color model is not easy to be implemented in computer graphics, the HSV color model usually transformed into RGB color model, practically. The convolutional relationship between RGB and HSV are

\[ V = \max(R, G, B) \]
\[ S = \frac{(V - \min(R, G, B))}{V} \]
\[ H = \begin{cases} 
(G - B) \times 60 / S, & \text{if } V = R \\
180 + (B - R) \times 60 / S, & \text{if } V = G \\
240 + (R - G) \times 60 / S, & \text{if } V = B 
\end{cases} \]

Since the advantage of HSV is the property of not sensitive to the change of light, the template will be established by HSV.

C. CamShift

The idea of mean shift is first proposed by Fukunaga and Hostetler in 1975 [11]. Mean shift is utilized for estimating the gradient vector of PDF (Probability Density Function) in [12]. In 1995, Cheng proposed some application examples of mean shift [13]. The most representative paper is proposed by Comaniciu, Ramesh and Meer [14]. In [15], mean shift is applied to the analysis of the feature space. In summary, mean shift algorithm is a method for finding the maximum and the most concentrated value of the color probability distribution. In addition, mean shift approach applied to tracking issue has been verified by other researchers. In the following it is a brief description of mean shift tracking.

Assume the target template is \( \{x_i\}_{i=1,2,...,n} \) in \( \mathbb{R}^d \). The multivariate kernel density of point \( X_0 \), \( \hat{q}_u(X_0) \) is

\[ \hat{q}_u(X_0) = \frac{1}{n \cdot h} \sum_{i=1}^{n} k \left( \frac{x_0 - x_i}{h} \right)^2 \]  

(4)

where \( k(x) \) is defined as

\[ k(x) = (2\pi)^{-d/2} \cdot \exp\left(-\frac{1}{2} \|x\|^2\right) \]

(5)

and \( h \) is the window width.

Moreover, the multivariate kernel density of the candidate moving object at position \( Y \) is

\[ \hat{p}_u(Y) = \frac{1}{n \cdot h} \sum_{i=1}^{n} k \left( \frac{Y - x_i}{h} \right)^2 \] 

(6)

Therefore, searching and tracking issue can be regarded as the smallest difference between the values of the templates. The matching cost function is the Bhattacharyya coefficient [16]:

\[ \rho(y) = \rho[p(u, y), q(u)] = \sum_{u=1}^{m} \sqrt{p(u, y) \cdot q(u)} \]

(7)

where \( p(u, y) \) is the \( u \)th color index of the candidate moving object, \( q(u) \) is the \( u \)th color index of the template. Next, the moving vector \( \Delta X \) is

\[ \Delta X = \frac{1}{\sum_{i} K(x_i - x) \cdot w(x) \cdot (x_i - x)} \sum_{i} K(x_i - x) \cdot w(x) \]

(8)

The \( \Delta X \) is converged by iterations. In (5), the \( w(x) \) indicated the correspondence between the template and search block at pixel \( x \).
\[ w(x) = \sqrt{q(I(x)) / p(I(x))} \]  

(9)

where \( q(I(x)) \) and \( p(I(x)) \) are the color distributions of templates and the search block, respectively. The \( K(\cdot) \) is usually a symmetric kernel function.

The steps of mean shift tracking algorithm are

1. Set the region of interest (ROI) of the input image.
2. Calculate the color histogram of the target image as the template.
3. Calculate the target’s color histogram of next image and start to search.
4. Calculate the moving vector \( \Delta X \). (If \( \Delta X \) is divergent, go back to Step 2.)
5. Compute the Bhattacharyya coefficient for finding the best matching block.

CamShift algorithm improves the searching efficiency of mean shift tracking algorithm. From Step 5 of mean shift tracking algorithm, the coordinate of the best matching block to be the initial searching position at next continuous image is recorded. The procedure of CamShift algorithm is illustrated as figure 4.

\[ \Phi I P X I q x w = K(t) \]

D. Kalman filter

Kalman filter [17][18] is a well-known method in engineering applications. It is a time-domain filter in which estimation of states is performed by two procedures, namely, prediction (time update) and correction (measurement update). State estimation is achieved by weighting sum of state prediction and output measurement. In our application, the Kalman filter is utilized for estimating states of the moving target in occlusion. The process of Kalman filter is illustrated in figure 5.

\[ \hat{s}(t) = \Phi(t-1)\hat{s}(t-1) \]  

(10)

\[ C(t) = H(t)s(t) + v(t) \]

(11)

where \( \Phi \) is the state transition matrix, \( H \) is measurement matrix, \( w(t) \) and \( v(t) \) is the noise.

The two major steps of Kalman filter are “Prediction” and “Correction”.

Prediction (Time update):

\[ \hat{s}(t) = \Phi(t-1)\hat{s}(t-1) \]  

(12)

Correction (Measurement update):

\[ \hat{P}(t) = \Phi(t-1)\hat{P}(t-1)\Phi(t-1)^T + Q(t-1) \]

\[ \tilde{s}(t) = \hat{s}(t) + K(t)[C(t) - H(t)\hat{s}(t)] \]

(13)

\[ \hat{P}(t) = [I - K(t)H(t)]\hat{P}(t) \]

(14)

Assume a Newton’s system with measurement is described as

\[ \text{System state: } s(t) = \Phi(t-1)s(t-1) + w(t) \]  

\[ \text{Measurement: } C(t) = H(t)s(t) + v(t) \]

where \( \Phi \) is the state transition matrix, \( H \) is measurement matrix, \( w(t) \) and \( v(t) \) is the noise.

The two major steps of Kalman filter are “Prediction” and “Correction”.

Prediction (Time update):

\[ \hat{s}(t) = \Phi(t-1)\hat{s}(t-1) \]  

(10)

\[ \hat{P}(t) = \Phi(t-1)\hat{P}(t-1)\Phi(t-1)^T + Q(t-1) \]

(11)

Correction (Measurement update):

\[ \hat{s}(t) = \hat{s}(t) + K(t)[C(t) - H(t)\hat{s}(t)] \]

(12)

\[ \hat{P}(t) = [I - K(t)H(t)]\hat{P}(t) \]

(13)
estimate error covariance and posterior estimate error covariance, respectively; \( K \) is the Kalman gain and \( C \) is the measurement vector; \( R \) and \( Q \) are the measurement error covariance and process noise covariance. The Prediction and Correction are cycling alternately.

E. Moving model of target object

Before estimating the moving of object by Kalman filter, the moving model needs to be established. The moving objects, such as moving people or stream of vehicles are usually inertia Newton's systems. Therefore, the states of moving objects in image sequence can be established. The most important states of Newton's system are position, velocity and acceleration. In other words, we will establish the second-order model of target object for Kalman filter calculating.

From CamShift, the position of the target object in the image sequence \( P_t \) is

\[
P_t = P_{t-1} + (P_{t-1} - P_{t-2})
\]

(13)

Therefore, the velocity and acceleration could be obtained by every capturing interval. By using (10) and (13), the state space of Kalman filter is

\[
s(t) = \Phi(t-1)s(t-1) + w(t)
\]

\[
= \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} s(t-1) + \begin{bmatrix} w(t) \\ 0 \end{bmatrix}
\]

(14)

where \( s(t-1) = \begin{bmatrix} P_{t-1} \\ P_{t-2} \end{bmatrix} \) and \( \Phi(t) = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \).

Moreover, the parameters of the measurement in (10) are defined as

\[
s(t) = \begin{bmatrix} P_t \\ P_{t-1} \end{bmatrix}, \quad H(t) = \begin{bmatrix} 1 & 0 \end{bmatrix}
\]

and \( C(t) = [P_t] \).

By this way, the Kalman filter can estimate the moving object continuously.

3 Experimental Results

A. CamShift tracking

The color conversion model used in this paper is HSV. We used a small diskette-case of light green color to be a target and figure 6 demonstrated the tracking results. In figure 6(a), the color histogram of target template is concentrated in green. In figure 6(b-1) and figure 6(c-1), we test the algorithm when target moving in complex environment with light background. Figure 6(b-2) and figure 6(c-2) are the target segmentation results of figure 6(b-1) and figure 6(c-1), respectively. Moreover, for demonstrating the recognition of the algorithm, we added some objects which are of similar color of the target. Figure 6(c-1) and figure 6(c-2) show that the CamShift is very robust. In this case, the image size is 320×240 and the frame rate is 26 fps.

B. Tracking by assistance of Kalman filter

We captured the fragments from the video of tracking diskette-case. The continuous position of target has been computed by CamShift. Figure 7(a) and figure 7(b) show the results estimated by Kalman filter from frame 42 to 107 and frame 206 to 271, respectively. The MSE of the estimations in figure 7(a) and figure 7(b) are 2.025 and 8.234, respectively. The units of x-axis and y-axis in figure 7 are pixels.
The Kalman filter provides an acceptable estimation position of target when the target occluded temporarily. In figure 8, it is the case that the moving person is occluded by another one. From the simulation result, it shows that the tracking result of the method is well. It does not only track the specifically moving person, but also provide the estimation target positions while in occlusion. In the following, descriptions about the tracking and estimating procedures are stated.

![Figure 7: Position estimating results of a moving diskette-case by Kalman filter.](image)

Firstly, the CamShift catches the person wearing a red T-shirt. Next, he is occluded by the person wearing gray jacket. The Kalman filter can provide the estimated positions continuously. Finally the estimation position will be the initial position of CamShift when the target appears again. In this case, the target is occluded in 8 frames, and the Kalman filter provides the possible positions. The square marks in figures 8(4) – (6) are the estimated positions made by Kalman filter. When the target appears again, the CamShift will catch the target immediately based on the estimated positions. The image size is 320×240. The frame rate is 15 fps.

![Figure 8: Moving person tracking results.](image)

**4 Conclusions**

The PTZ camera is a very popular device for surveillance systems. The surveillance systems become more and more intelligent for human life. In this paper, we focused on the problem of moving object tracking with the object in occlusion.

We proposed a simple and efficient method. This method is suitable especially for surveillance system design. The method can detect the moving objects by TD initially and establish the target template in HSV color space for CamShift tracking. Finally, the Kalman filter provides the estimated positions when the target is in occlusion.

In experimental results, we demonstrated that the green CD case could be tracked by CamShift. The
CamShift tracking performance is very good and very robust. In addition, we also showed the case of the moving people passed behind another one. The method can not only track the motion but also estimate the trajectory of the moving people successfully.

Therefore, in this paper, three major problems, moving object detecting, moving object tracking and tracking the object in occlusion, have been conquered successfully by our method.

References


Kuo-Hsien Hsia received the Ph.D. degree in electrical engineering from the National Sun Yat-Sen University, Taiwan, in 1994. He is currently an associate professor of the Department of Computer Science and Information Engineering, Far-East University, Taiwan. His research interests are in the area of differential game theory, fuzzy set theory, image-assistant measurement and control software design.
Shao-Fan Lien was born in Taiwan, R.O.C., on February 13, 1981. He received the B.S. degree from Department of Electronic Engineering of Huafan University in 2004 and received the M.S. degree from Tungnan University in 2006. He is a Ph.D. student of Graduate School of Engineering Science and Technology of National Yunlin University of Science and Technology from 2006 to today. His research interests include computer vision and vision servoing system control.

Juhng-Perng Su was born in Taiwan in 1958. He received the M.S. degree in electrical engineering from the National Taiwan University of Science and Technology, Taipei, Taiwan, in 1986, and the Ph.D. degree in electrical engineering from Sun Yat-Sen University, Kaohsiung, Taiwan, in 1990. Since 1991, he has been with the National Yunlin University of Science and Technology, Yunlin, Taiwan, where he is currently a Professor in the Department of Electrical Engineering. Since 2004, he has been conducting research projects on autonomous unmanned helicopters. His current research interests include the design of autonomous unmanned vehicles, robust estimation and optimal control, intelligent systems, and nonlinear control. Dr. Su was a recipient of the distinguished research award from Yunlin University of Science and Technology in 2005 and awards from various domestic competitions.