

# Moving Target Tracking based on CamShift Approach and Kalman Filter

Shao-Fan Lien<sup>1</sup>, Kuo-Hsien Hsia<sup>2</sup> and Juhng-Perng Su<sup>3,\*</sup>

<sup>1</sup> Department of Electrical Engineering, National Yunlin University of Science and Technology, Yunlin, Taiwan

<sup>2</sup> Department of Computer Science and Information Engineering, Far East University, Tainan, Taiwan

<sup>3</sup> Department of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan

Received: 15 Apr. 2014, Revised: 16 Jul. 2014, Accepted: 17 Jul. 2014

Published online: 1 Jan. 2015

**Abstract:** The surveillance system has been developed for decades. It reduced the crime and protected the lives and properties of people successfully. Some surveillance system are composed of PTZ cameras. Therefore the moving people or target objects could be tracking by the surveillance system. The surveillance system becomes more and more interment. However, there are two problems occurred, target recognition and target shelter, respectively. 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 filter framework for estimating the camera and world parameters [3]. In the design of the intelligent surveillance system, two important problems need to be concerned, target recognition problem and target shelter problem, respectively.

The several moving targets are usually being in the same 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; how to produce the object moving trajectory for PTZ camera tracking is the other difficult task.

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

In previous researches in moving object detecting method, 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

\* Corresponding author e-mail: [sujp@mail.ndhu.edu.tw](mailto:sujp@mail.ndhu.edu.tw)

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

### 2.1 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.

The camera system for our application is Sony D70 PTZ camera. The capture range of the camera is illustrated as figure 2.

There are many features of D70 PTZ camera, such as:

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, D70 is the excellent device for surveillance system.

### 2.2 HSV Color Space

HSV (hue, saturation, value) is a very useful cylindrical-coordinate representation of points in an RGB

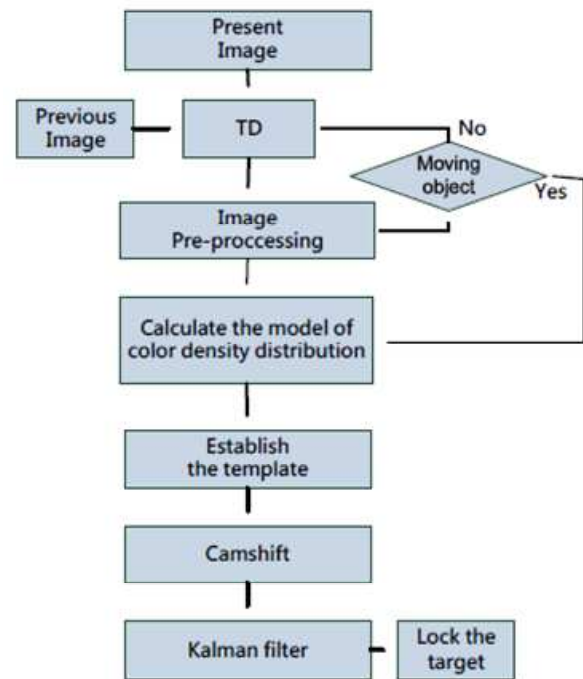


Fig. 1: Flowchart of the proposed methods.

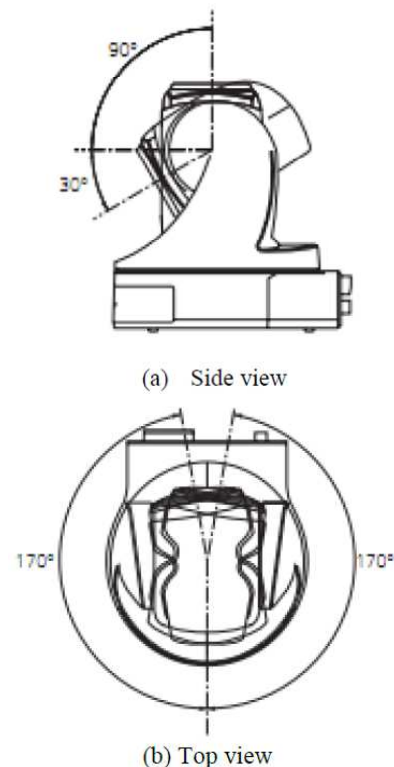


Fig. 2: Capture range of Sony D70 PTZ camera [11].

color model. The HSV color model describe the more

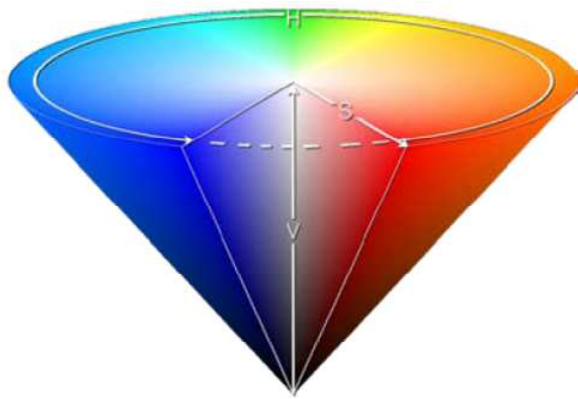


Fig. 3: HSV model [10].

color features of object then RGB model. The figure 3 illustrated the HSV color model.

The HSV color model is not easy to be implemented in computer graphics, therefore, the HSV color model usually transformed into RGB color model, practically.

The conversational relationship between RGB and HSV are

$$V = \max(R, G, B) \tag{1}$$

$$S = \frac{V - \min(R, G, B)}{V} \tag{2}$$

$$H = \begin{cases} (G - B) \times 60/S, & \text{if } V = R, \\ 180 + (B - R) \times 60/S, & \text{if } V = G, \\ 240 + (R - G) \times 60/S, & \text{if } V = B. \end{cases} \tag{3}$$

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

### 2.3 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 the 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 that the target template is  $\{x_i\}_{i=1,2,\dots,n}$  in Rd. The multivariate kernel density of point  $X_0$ ,  $\hat{q}_u(X_0)$  is

$$\hat{q}_u(X_0) = \frac{1}{n \cdot h} \sum_{i=1}^n k(\|\frac{x_0 - x_i}{h}\|^2) \tag{4}$$

where  $k(x)$  is defined as

$$k(x) = (2\pi)^{-d} \cdot \exp(-\frac{1}{2}\|x\|^2) \tag{5}$$

and  $h$  is the window width.

Moreover, the multivariate kernel density of the candidate moving object at position  $Y$  is

$$\hat{p}_u(X_0) = \frac{1}{n \cdot h} \sum_{i=1}^n k(\|\frac{Y - x_i}{h}\|^2) \tag{6}$$

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]:

$$\rho(y) = \rho[p(u, y), q(u)] = \sum_{u=1}^m \sqrt{p(u, y) \cdot q(u)} \tag{7}$$

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

$$\Delta X = \frac{\sum_i K(x_i - x) \cdot w(x) \cdot (x_i - x)}{\sum_i K(x_i - x) \cdot w(x)} \tag{8}$$

The  $\Delta X$  is converged by iterations. In (5), the  $w(x)$  indicated the correspondence between the template and search block at pixel  $x$ .

$$w(x) = \sqrt{q(I(x))/p(I(x))} \tag{9}$$

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

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

CamShift algorithm improves the searching efficiency of mean shift tracking algorithm. From Step 4 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.

### 2.4 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

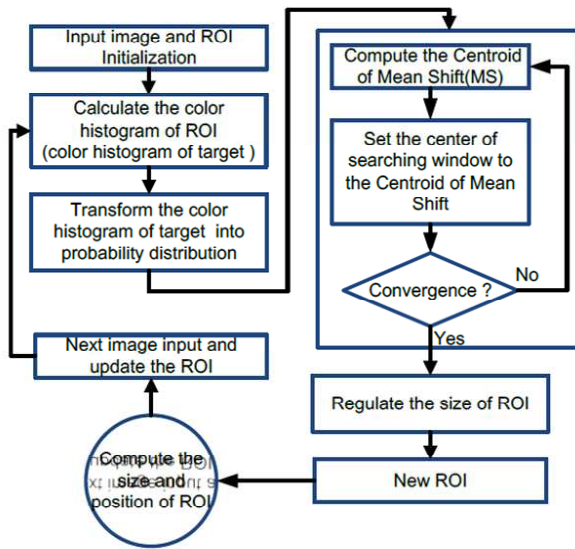


Fig. 4: CamShift algorithm.

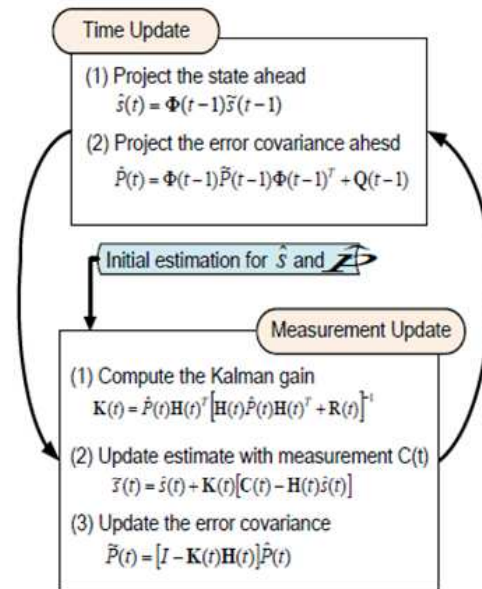


Fig. 5: The process of Kalman filter.

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.

Assume that the Newton's system with measurement is described as

$$\begin{aligned} \text{System state: } s(t) &= \Phi(t-1)s(t-1) + w(t) \\ \text{Measurement: } C(t) &= H(t)s(t) + v(t). \end{aligned} \quad (10)$$

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

The two major steps of Kalman filter are "**Prediction**" and "**Correction**".

Prediction (Time update):

$$\begin{aligned} \hat{s}(t) &= \Phi(t-1)\hat{s}(t-1) \\ \hat{P}(t) &= \Phi(t-1)\hat{P}(t-1)\Phi^T(t-1) + Q(t-1). \end{aligned} \quad (11)$$

Correction (Measurement update):

$$\begin{aligned} K(t) &= \hat{P}(t)H^T(t)[H(t)\hat{P}(t)H^T(t) + R(t)]^{-1} \\ \tilde{s}(t) &= \hat{s}(t) + K(t)[C(t) - H(t)\hat{s}(t)] \\ \tilde{P}(t) &= [I - K(t)H(t)]\hat{P}(t) \end{aligned} \quad (12)$$

where  $\hat{s}$  and  $\tilde{s}$  are a priori estimate and posterior estimate, respectively;  $\hat{P}$  and  $\tilde{P}$  are a priori estimate error covariance and posterior estimate error covariance, respectively.  $K$  is the Kalman gain and  $C$  measurement vector in (12);  $R$  and  $Q$  are the measurement error covariance and process noise covariance. The **Prediction** and **Correction** are cycling alternately. The process of Kalman filter is shown as figure 5.

## 2.5 The moving model of target object

Before estimating the moving of object by Kalman filter, the moving model needs to be established. In this paper, we can produce the linear motion model of target from image sequence. 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

$$P_t = P_{t-1} + (P_{t-1} - P_{t-2}) \quad (13)$$

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

$$\begin{aligned} s(t) &= \Phi(t-1)s(t-1) + w(t) \\ &= \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} P_{t-1} \\ P_{t-2} \end{bmatrix} + \begin{bmatrix} w(t) \\ 0 \end{bmatrix} \end{aligned} \quad (14)$$

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  $s(t) = [P_t \ P_{t-1}]^T$ ,  $H(t) = [1 \ 0]$ , and  $C(t) = [P_t]$ . By this way, the Kalman filter can estimate the moving object continuously.



Fig. 6: The tracking results of CamShift.

### 3 Experimental Results

#### 3.1 CamShift tracking

In this paper, the used color conversion model is HSV. We used a small diskette-case of light green color to be a target and figure 3 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 320240 and the frame rate is 26 fps.

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

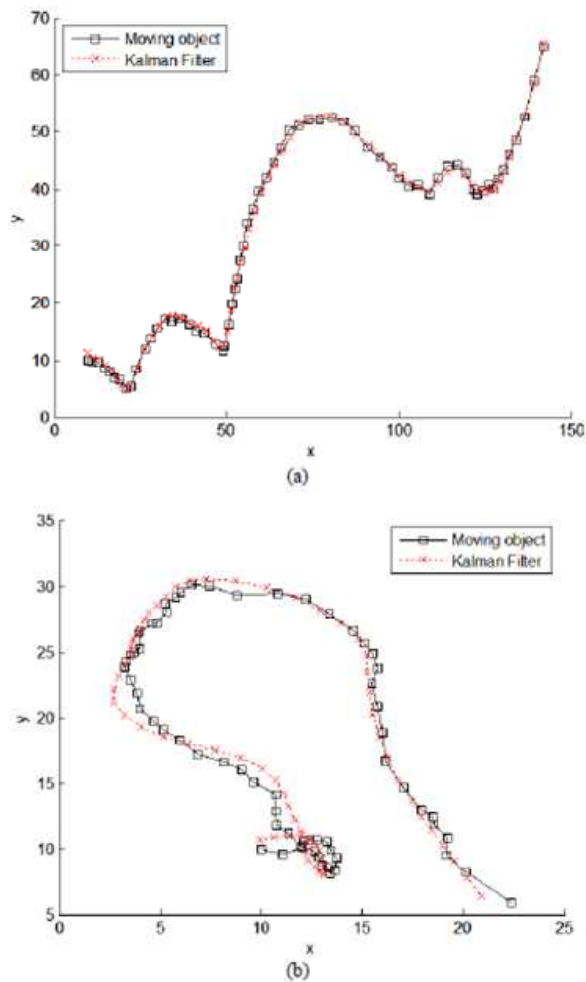


Fig. 7: Position estimating results of a moving diskette-case by Kalman filter.

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.

#### 3.2 Kalman filter

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 results 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.

Firstly, the CamShift catches the person wearing a red T-shirt. Next, he is occluded by the person wearing gray



Fig. 8: Moving person tracking results.

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 figure 8(4) - (6) are the estimated positions made by Kalman filter. When the target appear again, the CamShift will catch the target immediately based on the estimated positions. The image size is 320x240. The frame rate is 15 fps.

## 4 Conclusions

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

We proposed a simple and efficient method for moving object tracking issue. 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 experiment results, we demonstrated the green CD case could be tracking by CamShift. The CamShift tracking performance are very good and very robust. In addition, we also shown the case of to moving people passed by each other. The method are not only tracked the motion but also predicted the trajectory of the moving people.

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

- [1] C. S. Yang, R. H. Chen, C. Y. Lee and S. J. Lin, PTZ camera based position tracking in IP surveillance system, 3rd International Conference on Sensing Technology, 142-146 (2008).
- [2] M. Al Haj, A. D. Bagdanov, J. Gonzalez and F. X. Roca, Reactive object tracking with a single PTZ camera, International Conference on Pattern Recognition (ICPR 20), 1690-1693 (2010).
- [3] T. Wang, G. Liu and W. Xie, Visual servoing control of video tracking system for tracking a flying target, International Conference on Advanced Intelligent Mechatronics, 850-855 (2011).
- [4] D. V. Thombre, J. H. Nirmal and D. Lekha, Human detection and tracking using image segmentation and Kalman filter, International Conference on Intelligent Agent & Multi-Agent Systems, 1-5 (2009).
- [5] Z. Xu and H. R. Wu, Smart Video Surveillance System, IEEE International Conference on Industrial Technology (ICIT 2010), 285-290 (2010).
- [6] I. Haritaoglu, D. Harwood, and L. Davis, Real-Time Surveillance of People and Their Activities, Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, **22**, 809-830 (2000).
- [7] G. L. Wang, D. Q. Liang, Y. C. Wang and Z. H. Hu, Algorithm for Tracking of Fast Motion Objects with Adaptive Mean Shift, Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 359-363 (2007).
- [8] B. Shoushtarian, and H. E. Bez, A practical adaptive approach for dynamic background subtraction using an invariant colour model and object tracking **26**, 5-26 (2005).
- [9] J. G. Allen, Xu R. Y. D., J. S. Jin, Object Tracking Using CamShift Algorithm and Multiple Quantized Feature Spaces, Proceedings of the Pan-Sydney Area Workshop on Visual Information Processing, Darlinghurst, Australia: Australian Computer Society, 3-7 (2004).
- [10] R. Gonzalez, R. E. Woods, Digital Image Processing, 2nd ed. Prentice Hall Press, 295 (2002).
- [11] Sony EVID 70P, <http://pro.sony.com/bbsc/ssr/product-EVID70P/>.
- [12] K. Fukunaga, L. D. Hostetler, The Estimation of the Gradient of a Density Function, with Applications in Pattern

Recognition, IEEE Transactions on Information Theory, **21**, 32-40 (1975).

- [13] Y. Cheng, Mean shift, mode seeking, and clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence, **17**, 790-799 (1995).
- [14] D. Comaniciu, V. Ramesh and P. Meer, Real-time tracking of non-rigid objects using mean shift, IEEE Conference on Computer Vision and Pattern Recognition, **2**, 142-149 (2000).
- [15] M. Y. Ju, C. S. Ouyang and H. S. Chang, Mean shift tracking using fuzzy color histogram, International Conference on Machine Learning and Cybernetics (ICMLC) **6**, 2904-2908 (2010).
- [16] A. Bhattacharyya, On a measure of divergence between two statistical populations defined by their probability distributions, Bulletin of the Calcutta Mathematical Society, **35**, 99-109 (1943).
- [17] R. E. Kalman, A New Approach to Linear Filtering and Prediction Problems, Transactions of the ASME-Journal of Basic Engineering, **82**, 35-45 (1960).
- [18] R. S. Bucy and P.D. Joseph, Filtering for Stochastic Processes with Applications to Guidance, John Wiley and Sons, 1968; 2nd Edition, AMS Chelsea Publ., (2005).



**Juhng-Perng Su** (M'95) 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.



**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 Tunghnan 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 serving system control.



**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 theory, image-assistant measurement and control software design.