CONTINUOUS PERSON TRACKING ACROSS MULTIPLE ACTIVE CAMERAS USING SHAPE AND COLOR CUES

N. Bumrungkiat, S. Aramvith*, and T.H. Chalidabhongse†*

*Department of Electrical Engineering
Chulalongkorn University
Bangkok 10330 Thailand
Tel: +66-2218-6909
†Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok 10520 Thailand
Tel: +66-2723-4977

ABSTRACT

This paper proposed a framework for handover method in continuously tracking a person of interest across cooperative pan-tilt-zoom (PTZ) cameras. The algorithm here is based on a robust non-parametric technique for climbing density gradients to find the peak of probability distributions called the mean shift algorithm. Most tracking algorithms use only one cue (such as color). The color features are not always discriminative enough for target localization because illumination or viewpoints tend to change. Moreover the background may be of a color similar to that of the target. In our proposed system, the continuous person tracking across cooperative PTZ cameras by mean shift tracking that using color and shape histogram to be feature distributions. Color and shape distributions of interested person are used to register the target person across cameras. For the first camera, we select interested person for tracking using skin color, cloth color and boundary of body. To handover tracking process between two cameras, the second camera receives color and shape cues of a target person from the first camera and using linear color calibration to help with handover process. Our experimental results demonstrate color and shape feature in mean shift algorithm is capable for continuously and accurately track the target person across cameras.

Keywords: Tracking, Handover, Mean shift algorithm, color calibration.

1. INTRODUCTION

The real time object tracking is used in many applications such as video surveillance, human machine interfaces, robot tracking, and intelligent transportation systems. These applications need good object detection and tracking methods which have been progressed in recent years. However, there are still some difficulties in tracking such as the cases where there are changes in background, view point, or illumination especially when using multiple cameras.

The object tracking in dynamic scene is a typical non-rigid vision tracking problem. The difficulty of the problem with moving cameras is that the environment is changing and the illumination may vary, which will cause the change of the color in the camera view. Appearance-based methods have been investigated for solving this tracking problem. Most appearance-based tracking approaches are based on some kinds of representation of image appearance. Target and localization are robust and more efficient for tracker. In this paper, we focus on the problem of tracking an individual person with cooperative PTZ cameras when viewpoint of camera during a long tracking tends to change.

Some surveillance systems such as the works presented in [1, 2] continuously track person across multiple cameras by handover the tracked person from one camera to another camera. The work in [1] proposed a continuous tracking within and across stationary and PTZ cameras. They use motion and color models for registering the object between overlapping camera view. The works in [2] use a homographic relation mapping between overlapping static cameras to transfer a target object point from one camera to another. This work also can track multiple objects within camera view. Some systems employ two cameras with one providing a fixed wide view and second Pan-Tilt-Zoom to acquire large face image for tracking such as [3]. The work in [4] proposed the process of building a surveillance system with multiple PTZ cameras installed within an indoor laboratory. The system demonstrated multiple PTZ cameras can hand over tracked target to each other.

Most tracking algorithms fall into two categories. The first category is probabilistic methods. These methods view the tracking algorithm as a state solving problem under the Bayesian framework, model uncertainty and propagate the conditional densities through the tracking process. The representative methods are Kalman filter and its derivatives, Condensation [5], particle filter [6], etc. The second category is deterministic methods. These methods compare a model with current frame and find out the most probable region. Mean Shift [7][8] falls into this category.

The mean shift algorithm has success in object tracking due to its simplicity and robustness. It finds local minimum of a similarity measure between the color histogram of model and the candidates in the image. The mean shift algorithm is based on appearance model. Color histogram is the most popular feature that has been used to represent the appearance of the target. But color of an object depends on illumination, viewpoint of camera during a long tracking (important when use PTZ cameras system). Thus the single histogram feature is not always discriminative enough. The work in [9] presents object tracking from multiple stationary and moving cameras using color distribution as the main cue for tracking object across views. Since color information can be easily biased by several factor such as illumination, shadow, and appearance change or difference camera setup, color cue may be not very reliable for tracking moving object. Some other works such as the work in [10] proposed multi-feature tracking approach for adaptive real-time object tracking using a stationary camera. In this paper, we propose a tracking method that employs both color and shape features to improve accuracy of the tracker especially when tracking objects under changes in illumination and viewpoints.

The paper is organized as follows. Section 2 we introduce our system overview. In Section 3, we explain how to extract the color and shape features to represent the target. Section 4 discusses the person tracking method. And then discusses person analysis in cooperative process and control both camera in Section 5 and 6 respectively. We experiment the performance of the proposed method in Section 7. This paper is concluded in Section 8.

2. SYSTEM OVERVIEW

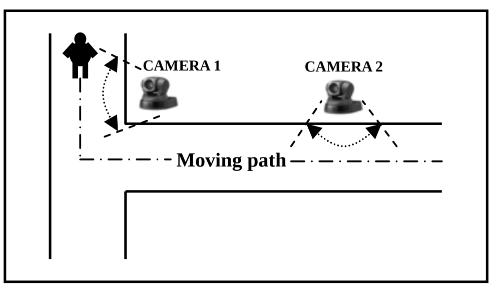


Figure 1 *Cameras set up*

Fig. 1 shows our system setup. We connect two SONY EVI-D100 PTZ cameras to a PC and use them to track the person from different viewpoints. The two cameras are distantly separated and have non-overlapped viewpoint.

Overall architecture is shown in Fig. 2. Both cameras simultaneously run the same algorithm. First, for each video frame, the raw input image is processed to construct a color probability distribution image via color histogram model. Also, the derivatives in x and y directions of the image is calculated to find shape cue that is represented by orientation histogram of the target person being tracked.

In person tracking step, we employ mean shift algorithm using color and shape features of target person. The location of the target person is used for continuously tracking and also used to calculate pan-tilt variables for the camera control in following the moving target. In order to continuously track the same target person across different cameras, we also calibrate color using color calibration matrix between two cameras to solve the color mismatch problem. This will be described in details in Section 5. By using color and shape features for tracking, our system can track and follow the target effectively throughout all reachable cameras' view.

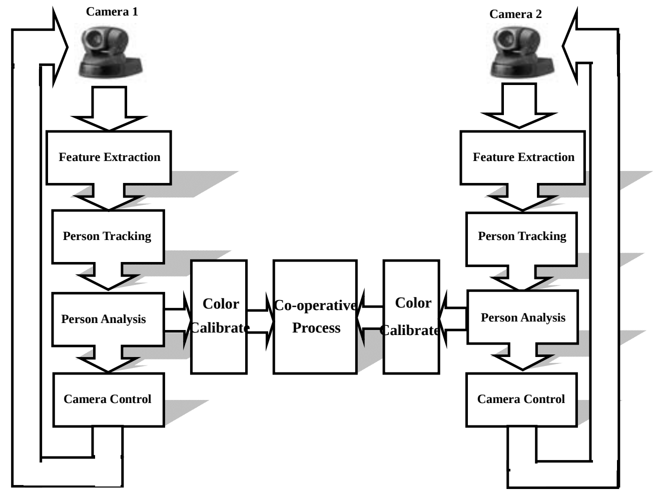


Figure 2 *System diagram*

3. FEATURE EXTRACTION

Fig. 3 show block diagram of feature extraction process which is composed of two sub-processes, that is finding color and shape features for using in mean shift tracking process.

3.1 Color Probability Distribution

In many person tracking processes, color distributions are used as target representation because of its independence from non-rigidity, scaling and partial occlusions. In order to track color objects in a video scene, a probability distribution image of the desired color (cloth and/or skin color) in the video scene must be created. In this paper, we use hue histogram to represent color of the target because of its brightness invariance. The color distributions of the target and candidate are respectively calculated. The evaluating the similarity of two color distributions is done using the Bhattacharyya distance.

3.2 The Orientation Histogram of Shape

In this work, we represent shape feature of the target using an orientation histogram. It is done by calculating image pixel derivatives in x and y directions. We use a standard Sobel masks (S_x and S_y), as shown in Eqs.(1)-(3), because of its simplicity and efficiently computed for real time applications.

$$M_x(x, y) = S_x * I(x, y) \qquad (1)$$

$$M_y(x, y) = S_y * I(x, y) \qquad (2)$$

The Orientation of edge is

$$\theta(x, y) = \arctan\left(\frac{M_x(x, y)}{M_y(x, y)}\right) \qquad (3)$$

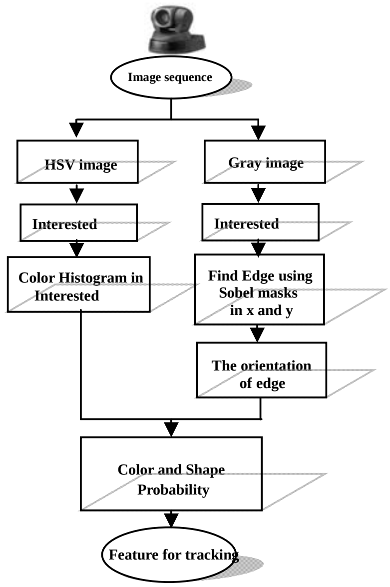


Figure 3 *Block diagram of feature extraction process*

3. PERSON TRACKING

The mean shift algorithm operates on probability distributions. To track interested person in video frame sequences, the color image data has to be represented as a probability distribution; we use color and shape histogram to accomplish feature distributions derived from video image sequences changed over time. Block diagram of mean shift algorithm is shown in Fig. 4.

The mean shift algorithm is a tracking method that finds the mode of the probability distributions of sample data without any assumptions about the priori distribution. This algorithm avoids choosing a distribution of model and estimating its distribution parameters.

There are several researches have been working on object tracking using mean shift algorithm such [8-9]. The basic mean shift tracking algorithm assumes that the target object has to separate sufficiently from background, but this assumption is not always true especially when tracking is carried out in dynamic backgrounds, e.g., surveillance with a moving camera.

The mean shift tracking finds the location corresponding to the target in the current frame based on the appearance of the target. Therefore, a similarity measure is needed between the distributions of an interested region in the current frame and the target model. A popular measure between two distributions is the Bhattacharyya distance considering discrete densities such as two feature (color and shape) histogram $p = \{p^{(u)}\}_{u=1...m}$ and $q = \{q^{(u)}\}_{u=1...m}$ the coefficient is calculated by

$$\rho[p,q] = \sum_{bin=1}^{m} \sqrt{p^{(bin)} q^{(bin)}} \qquad (4)$$

The larger ρ is, the more similar the distributions are. For two identical histograms we obtain $\rho = 1$, indicating a perfect match. As distance between two distributions, the measure can be defined as

$$d = \sqrt{1 - \rho[p,q]} \qquad (5)$$

, where d is the Bhattacharyya distance. The tracking algorithm is to recursively compute of an offset value from the current location y_0 to a new location y_1 according to the mean shift vector y_1 is calculated by using

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\frac{y_0 - x_i}{h}\right)}{\sum_{i=1}^{n_h} w_i g\left(\frac{y_0 - x_i}{h}\right)} \qquad (6)$$

when

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{q^{(u)}}{p^{(u)}(y_0)}} \delta\left[h(x_i) - b_{in}\right] \qquad (7)$$

and

$$g(x) = -k'(x) \qquad (8)$$

The person tracking algorithm is shown in Fig 4.

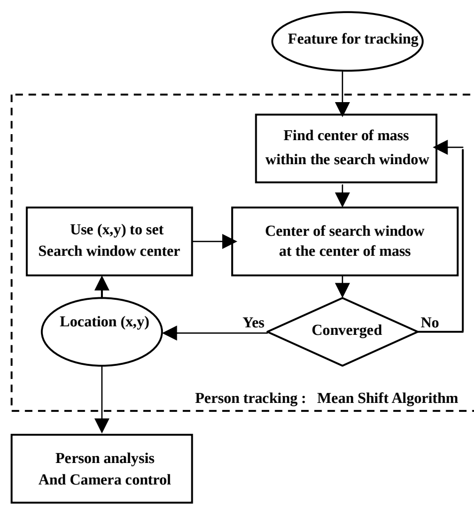


Figure 4 *Block diagram of Person tracking with mean shift algorithm*

5. PERSON MODEL AND COOPERATIVE TRACKING ACROSS CAMERAS

5.1 Person model

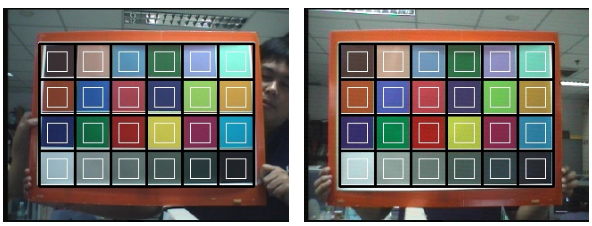
In tracking handover between two cameras, the second camera needs to receive some information of the tracked target from the first camera. As we use color and shape distributions of the target person in tracking, we must handover these distributions in hand-off process.

The color characteristic of tracked target person can be divided into 2 main regions: cloth and skin regions, while the shape characteristic is represented by x and y gradients in tracked region.

5.2 Camera color calibration

To achieve the color matching across the different radiometric characteristic cameras with different viewpoints, we proposed a method that employs linear color calibration. Color correlation transform matrix between a pair of color sets from the two cameras is computed using linear regression method.

To calibrate the cameras, we use a color checker board. A color checker board has 24 reference colors covering the major color palettes in RGB color space. The setup is shown in Fig.5. We adjust the cameras' positions and put the color checker board so that the checker board appears in both cameras' field of views. The images of the checker board are captured by both cameras. In this work, we use the RGB color space due to its being widely accepted as simple and useful probabilistic models. Then the color correlation transformation matrix between the two cameras is computed according to Eqs. (10) – (12).



(a) (b)

Figure 5 *(a) show input image from camera 1 (b) show input image from camera 2*

$$R_N^C = \frac{\sum_{k=1}^{400} I_R(k)}{400} \tag{10}$$

$$G_N^C = \frac{\sum_{k=1}^{400} I_G(k)}{400} \tag{11}$$

$$B_N^C = \frac{\sum_{k=1}^{400} I_B(k)}{400} \tag{12}$$

Let R_N^C (Eq. 4), G_N^C (Eq. 5) and B_N^C (Eq. 5) be average color information from each camera in R, G and B channels respectively. And superscript C be the number of camera in the system, subscript N be number of reference color from checker board and k be pixel of the image.

$$C^1 = \begin{bmatrix} R_1^1 & R_2^1 & \dots & R_{24}^1 \\ G_1^1 & G_2^1 & \dots & G_{24}^1 \\ B_1^1 & B_2^1 & \dots & B_{24}^1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{13}$$

$$C^2 = \begin{bmatrix} R_1^2 & R_2^2 & \dots & R_{24}^2 \\ G_1^2 & G_2^2 & \dots & G_{24}^2 \\ B_1^2 & B_2^2 & \dots & B_{24}^2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{14}$$

Let C^1 (Eq. 7) and C^2 (Eq. 8) be the 24x4 color matrices of the 24 RGB colors taken from checker board images of the camera 1 and camera 2 respectively. Then we can apply linear regression to obtain a color correlation transformation matrix that can be used to convert color of one camera to the color of the other camera. Let *W* is the corresponding regression coefficient of the color of camera 2 to the color of camera 1. We obtain

$$C^1 = WC^2 \tag{15}$$

Then, the *W* can be obtained from inversed matrix analysis. The color correlation transformation matrix, *W*, will be used to transform color from one camera to the color of another camera while we are performing the tracked person handover across the cameras.

6. CAMERA CONTROL

To keep the moving target in the camera's field of view, pan, tilt and zoom variables are computed and used as parameters to control the motion of the camera. To find pan angle (ψ) and tilt angle (ϕ), we calculate those angles using ratio of pixel-based distance change and degree-based distance change, see Eq.(16) and Eq.(17),

$$\psi = \frac{Diff_x}{\beta_x} = \frac{x_f - x_{center}}{\beta_x} \tag{16}$$

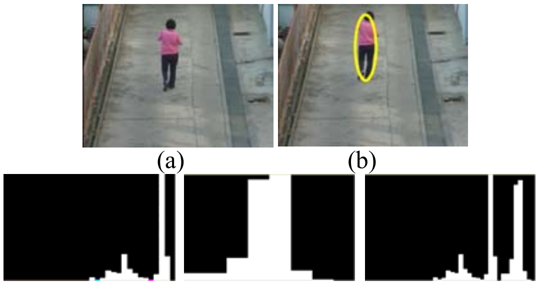
$$\phi = \frac{Diff_y}{\beta_y} = \frac{y_f - y_{center}}{\beta_y} \tag{17}$$

where $Diff_x$ and $Diff_y$ are distance between center of the interested region (x_f, y_f) and the center of image in (x_{center}, y_{center}) horizontal and vertical direction respectively. β_x and β_y are values of pixel distance change when camera moves by 1 degree in horizontal (pan) and vertical (tilt) directions.

7. EXPERIMENTAL RESULTS

We tested our proposed system on a Pentium 4 2.8 GHz PC running Windows XP. The two SONY EVI-D100 PTZ cameras are connected to the system. Distance between both cameras is about 1 meter. The experiment is to show the tracking performance of our proposed handover system when integrating shape cue is feature in mean shift tracking algorithm.

Fig. 6 shows person tracking result when using color and shape mean shift algorithm. Fig.6(a) is an input image, and Fig. 6(b) is the tracked result. Histograms shown in Fig.6(c)-(e) are color histogram, shape histogram, and the concatenation of them respectively.



(a) (b)

(c) (d) (e)

***Figure* 6** *(a) Input image (b) Tracked result (c) 32-bin color histogram (d) 8-bin shape histogram (e) Concatenate histogram*

Figs. **7 -** 9 demonstrate tracking comparisons of mean shift tracking using color only versus using color and shape in various scenarios. Fig.7 shows an outdoor scene where a person walking in a straight trajectory under constant illumination condition, both trackers yielded good results. In Fig.8, a person is walking indoor and there is a brightness change happened in the image sequence. Under illumination change condition, it apparently shows that the color and shape tracker performed better than the color only tracker. For the sequence in Fig. 9, we tried some scenarios where there is change in shape of the target. Fig.9(a) shows a sequence when a person change walking direction, and Fig.9(b) shows a sequence when a person abruptly sits down while walking. We found that both trackers work well. These assure us that by incorporating shape information to the color information helps improving accuracy of tracking.

Fig. 10 shows some snapshots of the tracked videos from the two cameras. It illustrates the handover result using color and shape cues in mean shift tracking algorithm.

8. CONCLUSIONS

We have presented a real-time system for continuously track a person across cooperative pan-tilt-zoom cameras. To handover the tracking process among two cameras, the second camera receives color cue and shape cue of a target person from the first camera. To achieve the mean shift algorithm tracking across the different radiometric characteristic cameras with different view point, our system uses color calibration to help with handover process. The experimental results confirm using color and shape cues helps reduce color mismatch between two cameras and the system is capable for continuously and accurately track the target person across cameras successfully.

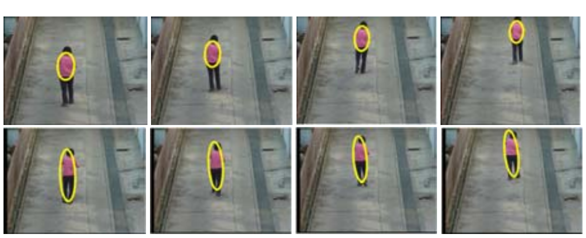


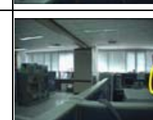
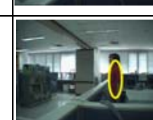
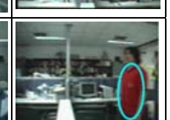
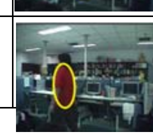
Figure *7 Tracking result using only color cue (top) versus using both color and shape cues (bottom) on an outdoor sequence with constant motion and illumination.*

Figure 8 *Tracking result using only color cue (top) versus using both color and shape cues (bottom) on an indoor sequence with changing illumination.*

(a)

(b)

Figure 9 *Tracking results using only color cue (top) versus using both color and shape cues (bottom) on two indoor sequences with changing motion direction (which makes target's silhouette shape changed) under constant illumination.*

	Camera 1	Camera 2
1. **Normal View**		
2. **The target person enters the first camera's view.**		
3. **The target person go along in first's camera view**		
4. **First camera tracks a target.**		
5. **Same person walks into the second camera's view**		
6. **The target person go along in second's camera view**		

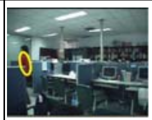
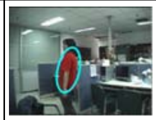
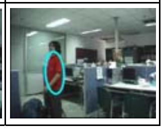
7. **Continuous tracking for the target.**		
8. **Continuous tracking for the target.**		

Figure 11 *Some snapshots of the videos from the two cameras showing the handover moment. Using color and shape cue*

9. REFERENCES

[1] J. Kang, I. Cohen, and G. Medioni, "Cotinuous Tracking Within and Across Camera Streams", in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR'03),* Wisconsin, Jun 2003.

[2] J. Black, T. Ellis, and P. Rosin, "Multi View Image Surveillance and Tracking", in *Proceedings of the IEEE Workshop on Motion and Video Computing(MOTION'02),* Florida, Dec 2002.

[3] L.Machestori, S.Piva and C.Raguzzoni, "Cooperaive Multiple system for Realtime Face Detection and Tracking in Uncontrolled Condition", *SPIE Image and Video communication and processing*,2005

[4] L.Everts, N.Sebe and G.A.John,"Cooperative Object Tracking with Multiple PTZ Cameras", *in International conference On Image Analysis and Processing (ICIAP'07), 2007*

[5] M. Isard and A. Blake, "CONDENSATION - conditional density propagation for visual tracking", *International. Journal of Computer Vision*, pp.5-28, 1998.

[6] K. Nummiaro, E. Koller-Meier and L. Van Gool, "Object tracking with an adaptative color-based Particle Filter", *Image and Vision Computing*, pp.99-I 11, 2002

[7] P. Perez, C. Hue, J. Vermaak and M. Gangnet, "Color-based probabilistic tracking", *European conference on Computer Vision,* pp 661-675, 2002.

[8] G.R. Bradski, "Computer vision face tracking for use in a perceptual user interface", IEEE Workshop on Applications of Computer Vision, pp.214-219, 1998.

[9] Jinman Kang, Issaac Cohen, and Gerard Medioni, "Tracking object from multiple stationary and moving cameras", in *International Symposium on Communication and Information Technologies(ISCIT'04)*, Sapporo, Oct 2005

[10] Junqiu Wang, and Yasushi Yagi, "Intrgrated Shape and Color for Real time Object Tracking ", *in International Conference on Robotics and Biomimetics*, Kunming, China, Dec 2006

[11] P. Amnuaykanchanasin, S. Aramvith, T.H. Chalidabhongse, "Continuous person tracking using two cooperative active cameras," Proceeding of IJWISA 2007, Korea, Feb 2007.

[12] N. Soontranon, S. Aramvith, and T. H. Chalidabhongse, "Face and Hand Localization and Tracking for Sign Language Recognition", in *International Symposium on Communication and Information Technologies(ISCIT'04)*, Sapporo, Oct 2004.