Hand Gesture Target Model Updating and Result Forecasting Algorithm based on Mean Shift

Xiao Zou
Corresponding author, School of Computer and Communication, Lanzhou University of Technology, Lanzhou, China
Email: zouxiao@lut.cn

Heng Wang, Qiuyu Zhang
School of Computer and Communication, Lanzhou University of Technology, Lanzhou, China
Email: sjwdyh@126.com, zhangqylz@163.com

***Abstract*—To propose a gesture model updating and results forecasting algorithm based on Mean Shift, and to solve the problem of target model changing and influenced tracking results in gesture target tracking process. Firstly, the background difference and skin color detection methods are used to detect and get gesture model, and the Mean Shift algorithm is used to track gesture and update the gesture model, and finally to use the Kalman algorithm to predict the gesture tracking results. The experimental results show that this algorithm reduces the influence of surrounding environment in gesture tracking process, and get better tracking result.**

***Index Terms*—Mean shift algorithm, Hand tracking, Gesture model, Result forecasting**

I. INTRODUCTION

As a preprocessing step for gesture recognition, gesture tracking has a great influence on the gesture recognition. The exact gesture tracking results can plays a positive role in the gesture recognition effect [1]. So the gesture tracking research has attracted the attention of many scholars.

Currently the most popular method is Mean Shift algorithm which was put forward by Fukunage and Hostetler in 1975 [2, 3]. The rapid development of Mean Shift algorithm is implemented by Comaniciu etc. The algorithm is applied to tracking field by them [4], and the variable window width Mean Shift algorithm is also studied [5]. Peng et al [6] proposed the forecast Camshift tracking results by using Kalman filter, in order to reduce the interference of large area skin color. Wang et al [7] used K-means clustering algorithm and particle filter method to solve the problem of mutual interference of finger tracking, and achieved good tracking effect. Feng et al [8] proposed a particle filter gestures tracking algorithm based on microscopic structure of state variables, which could describe the microstructure of the hand state variables from the continuous deformation time model, the period of time for deformation model, and mutations strength and self occlusion characteristics. Gan et al [9] proposed a kind of target tracking algorithm based on the Mean Shift algorithm and normalized moment of inertia characteristics, and it could reach real-time and stable tracking when the moving air target was occluded and in large deformation.

However, when the gesture was affected by the changing surrounding environment, such as the light suddenly becomes strong, or become weak or the non target interference became more serious, tracking may be make a fail because of the excursion of the target model. This paper proposes a gesture tracking algorithm based on the updating and forecasting of the target model. First establish the model of gesture, and then use Mean Shift algorithm to track the gesture. In the tracking process, judge the target model according to the Bhattacharyya coefficient, and update the target model according to the judgment result. Finally predict the tracking results by using Kalman filter. Experiments show that the algorithm proposed in the paper is effective when the gesture in a complex environment.

II. HAND EXTRACTING

This article will first need to extract the target hand to be tracked. Therefore, the use of combination of color information, motion information and the surrounding scene information gain the gesture center to obtain good tracking performance.

A. Skin Color Detection

There are many skin color detection methods, and each algorithm has its characteristics and use occasions [10]. To ensure the computing efficiency, and real-time, the method of literature [11] is used to detect the skin color.

In color space, skin color has its special distribution characteristics. These distribution characteristics may be more effective to gestures segmentation from the surroundings. But, skin color in different color space shows different clustering features. Even though the same skin color in different color space still have differences in the expression of clustering features [12]. So there is a relative important influence to the gestures division to choose the right color space. In the study of gestures tracking and recognition, the main color application space are YC_bC_r, HSV, HIS, YUV, etc.

YC_bC_r color space is a kind of color space with brightness separation. It is a color space developed for

doi:10.4304/jmm.8.1.1-7

digital *TV* system [13]. The color signal C_bC_r is independent with brightness signal Y . Based on these characteristics, in gestures tracking and recognition process, we can reduce the influence of considering brightness chromaticity signal, and then we can get relative ideal test results of color skin.

In YC_bC_r color system, skin color distributes relatively concentration in C_b , C_r , mainly in the $R_{Cb} = [77,127]$, $R_{Cr} = [133,173]$ [10-11]. Therefore, judge each pixel of the image as follows:

$$O(x,y)=\begin{cases}1,[C_r(x,y)\in R_{Cr}]\bullet[Cb(x,y)\in R_{Cb}]\\0,else\end{cases} \quad (1)$$

The judgment result is the skin color detection result.

B Background Subtraction Algorithm

It is unable to complete division of the hand only by using skin color detection method, as the existence of the same or similar color around (such as facial skin, the furniture with skin color etc.). Therefore, by using the appropriate background subtraction algorithm based on motion information, and combining with skin color detection results to segment the gesture region, in order to obtain the better effect [12].

$$D_t(x,y)=\min|I_t(x,y)-B_t(x,y)| \quad (2)$$

where $I_t(x,y)$ is the input value of pixels (x,y) in t time; $B_t(x,y)$ is the constructing background image of pixels (x,y) in t time, and then, on segment movement part:

$$M_t(x,y)=\begin{cases}1,D_t(x,y)>\sigma_b\\0,else\end{cases} \quad (3)$$

where σ_b is threshold.

Thus, gain the gestural information from skin color information and motion information. This can ensure the gesture segmentation and tracking.

C Elliptic Skin Model

To set a skin color model to gestures detection and segmentation, and distinguish the color area by the modeling results. The traditional test method includes threshold value method and Gaussian model method. Threshold value method is simple, but the test results are acceptable, and the tracking results have a distance with the ideal value. Gaussian model although has high accuracy, but it is more difficult to establish the model, and the application is relatively complex, not good at application in real time. Therefore, in this paper, we consider the elliptical model to set the skin color model to get the efficient simple test results.

Elliptic skin model was firstly proposed by Rein and applied in the color image face detection [15]. The model is based on YC_bC_r color space, and suppose that chromaticity component C_b and C_r is weight function of brightness Y . Skin model composed of the center and color range. According to the literature definition, ($\dot{C}_b(Y)$, $\dot{C}_r(Y)$) expressed the center, $Wc_r(Y)$, $Wc_b(Y)$ expressed the range of color. The color of skin area is expressed as following:

$$\dot{C}_b(Y)=\begin{cases}108+\dfrac{(K_l-Y)(118-108)}{K_l-Y_{\min}} & ifY<K_l\\[2ex]108+\dfrac{(Y-K_h)(118-108)}{Y_{\max}-K_h} & else\end{cases} \quad (4)$$

$$\dot{C}_r(Y)=\begin{cases}154+\dfrac{(K_l-Y)(154-144)}{K_l-Y_{\min}} & ifY<K_l\\[2ex]154+\dfrac{(Y-K_h)(154-132)}{Y_{\max}-K_h} & else\end{cases} \quad (5)$$

$$Wc_i(Y)=\begin{cases}WLc_i+\dfrac{(Y-Y_{\min})(Wc_i-WLc_i)}{K_l-Y_{\min}} & ifY<K_l\\[2ex]WHc_i+\dfrac{(Y_{\max}-Y)(Wc_i-WHc_i)}{Y_{\max}-K_h} & else\end{cases} \quad (6)$$

And give C_i value C_b or C_r .The skin samples are trained to get parameter values $WLc_b = 23$, $WHc_b = 14$, $Wc_r = 38.76$, $WLc_r = 20$, $WHc_r = 10$, $K_l = 125$, $K_h = 188$ .Threshold of brightness Y : $Y_{\max} = 235$ and $Y_{\min} = 16$. Based on the formula, transform the skin model to $C_r'C_b'$ coordinate system, the coordinate system C_r' and C_b' both are function of the brightness Y . Conversion formula is shown as in (7).

And $i \in \{b,r\}$, Wc_i is the skin color range of original model, $Wc_i(Y)$ is the color value of the converted brightness function. Therefore, in the $C_r'C_b'$ color space, elliptic skin model expression is as follows:

$$C_i'(Y)=\begin{cases}(C_i(Y)-\dot{C}_i(Y))\dfrac{Wc_i}{Wc_i(Y)}+\dot{C}_i(K_h) & ifY<K_l orK_h<Y\\[2ex]C_i(Y) & else\end{cases} \quad (7)$$

$$\frac{(x-ec_x)^2}{a^2}+\frac{(y-ec_y)^2}{b^2}=1 \tag{8}$$

$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} C_b'-c_x \\ C_r'-c_y \end{bmatrix} \tag{9}$$

θ is the angle, ec_x , ec_y are the two focus of the elliptic, a is the long axis of the elliptic, b is the short axis. (cx,cy) is the center of skin model in $C_r' C_b'$ coordinate system, the value is (109.38, 152 03).

III. Gesture Segmentation and Tracking

Based on visual gestures tracking first need to gestures segmentation, to get a complete gesture characteristics set. A single color features or sports information is not fully to isolate the gestures. Thus, in this paper the characteristics fusion method for gestures division is introduced. The YC_bC_r color space is used in skin detection, the brightness component and the chroma of the separation characteristics in this color space can reduce light influence of the skin. In order to ensure detection results, based on the elliptic skin model for skin detection. Although the judgment method based on threshold is simple, convenient and feasible and can ensure the real-time tracking test division, but at the cost of performance of sacrifice. The background model established by Gaussian mixture model can avoid the small influence of surrounding environment, and the change of the real-time updating background. Therefore we use Gaussian mixture model to build the background model, and to get the motion gestures by using background difference method, and at last to complete gestures division based on the color of skin test results. Gestures tracking used the classical Mean Shift algorithm. The gestures tracking algorithm based on a nuclear density theory is simple and efficient, also can guarantee the real-time tracking effect.

A. Mean Shift Algorithm

Mean Shift algorithm is a common algorithm in the gesture tracking methods based on vision. It's simple and efficient characteristics have attracted the attention of many researchers.

The mean shift iterative formula which comes from kernel density estimation theory expressed as [16]:

$$y_{j+1}=\frac{\sum_{i=1}^{n} x_i\omega_i g\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)} \tag{10}$$

where $\{y_j \mid j=1,2...\}$ is a series of target position, and y_1 was initialized for nuclear density center.

Assuming the object model:

$$\hat{q}=\{\hat{q}_i \mid i=1...n,\sum_{i=1}^{n}\hat{q}_i=1\} \tag{11}$$

Candidate targets:

$$\hat{p}(y)=\{\hat{p}_i(y) \mid i=1...n,\sum_{i=1}^{n}\hat{p}_i=1\} \tag{12}$$

The target model and the candidate model can be expressed as:

$$\hat{q}_i=C\sum_{u=1}^{n}k\left(\|x_{ui}\|^2\right)\delta\left[b(x_u)-i\right] \tag{13}$$

$$\hat{p}(y)=C_h\sum_{u}^{n_h}k\left(\left\|\frac{y-x_u}{h}\right\|^2\right)\delta\left[b(x_u)-i\right] \tag{14}$$

where y is a candidate target position. Accordingly, the evaluation coefficient Bhattacharyya of the target is defined as follows:

$$\hat{\rho}(y)\equiv\rho[\hat{p}(y),\hat{q}]=\sum_{i=1}^{n}\sqrt{\hat{p}_i(y)\hat{q}_i} \tag{15}$$

The distance between the two distributions is:

$$d(y)=\sqrt{1-\rho[\hat{p}(y),\hat{q}]} \tag{16}$$

Target tracking is to find the location y with maximum similarity ρ in each frame. If the target position of the first frame is y_0 , Taylor expansion the $\rho(y)$ in y_0 , and the similar functions are expressed as:

$$\rho(y)\approx\frac{1}{2}\sum_{1}^{m}\sqrt{\hat{p}_i(y_0)\hat{q}_i}+\frac{C_h}{2}\sum_{u=1}^{n_h}\omega_u k\left(\left\|\frac{y-x_u}{h}\right\|^2\right) \tag{17}$$

where

$$\omega_u=\sum_{i=1}^{m}\sqrt{\frac{q_u}{p_u(y_0)}}\delta[b(x_u)-i] \tag{18}$$

B. Target Model Updating Method

In the target tracking process, through the target detection or selection, build the model according to target characteristics such as texture, color, edge traits, so that as the target tracking prototype to be compared. By comparison with the target model, the area that the most similar to it is the next frame image of the target location. However after the effects of the illumination change, shade appearing etc from surrounding environment, the original target characteristics and the aim features appear not consistent, so it will lead to the results deviation. Therefore, in the process of the target tracking, model updating is particularly important.

Usually, the target object is described by using the following normalization formula [17]:

$$q_{i,m} = C\sum_{k=1}^{n} k\left(\left\|\frac{x_0 - x_k}{h}\right\|^2\right)\delta[b(x_k) - u] \quad \text{u=1, 2}\cdots\text{m} \tag{19}$$

All the candidate goals are: $\{q_i, u\} i = 1, 2\ldots N; u = 1, 2\ldots m$., Get the normalized combination for nuclear histogram:

$$\hat{Q}_u = \frac{1}{N}\sum_{i=1}^{N} q_{i,u} \tag{20}$$

According to the reference [18] definition and construct auxiliary template:

$$B_u = \begin{cases} 1 & Q_u \geq Q \\ 0 & Q_u < Q \end{cases} \quad \text{u=1, 2}\cdots\text{m} \tag{21}$$

where Q is a threshold, when a color features component larger than the threshold, then update it; Otherwise not update. Updating formula is as follows:

$$q_t = (1-\alpha)q_{t-T} + \alpha q_t \tag{22}$$

where q_t is the current state new target model, q_{t-T} is the target model of before frame; α is the updating rate.

Tong et al proposed a feedback loop tracking algorithm, introducing of tracking state decision as updating basis of a follow-up template, choose appropriate sample and the right time to update through the decision feedback information. This algorithm can effectively avoid the appearance of the target model change [19]. Shen et al proposed the updating strategy of choice sub-model, calculated the contribution of each component and according to the circumstances of the necessary choice update the model, and got a more robust tracking algorithm than whole model updating strategy [20]. Gan et al use the normalized inertia as the identification of standards of target model updating, and establish tracking strategy combining of false-alarm probability minimum principle and similarity level 2 decision threshold [9].

From the renewal of the target model, the paper looked for method improving the accuracy of gestures tracking.

IV. Target Model Updating

At present the main updating methods of tracking target model are: the updating strategy based on the whole [21-22] and the updating strategy based on selecting factor [20, 23]. From the perspective of real-time hand tracking, tracking process needs to reduce the calculating amount, and reduce the computational cost. So, this paper take the target gesture model as a whole, make a weighted summation to the current target gesture model and target historical gesture model, and then get the object model in the next frame. The definitions as follow:

$$q^t = (1-\alpha)q^{t-1} + \alpha p^t \tag{23}$$

where q^t is the updated target gesture model; p^t is the tracked target gesture model; q^{t-1} is the previous frame target gesture model, α is the tracking result weights of the current frame, also known as the speed factor of model updating.

Use coefficient Bhattacharyya on gesture tracking effect evaluation in the updating process. Set update model threshold as ρ_{Th}, and the current tracking results is ρ_t, and the target gesture model is updated as follows:

$$q^t = \begin{cases} q^{t-1}, & \rho_t \geq \rho_{Th} \\ (1-\alpha)q^{t-1} + \alpha p^t, & \rho_t < \rho_{Th} \end{cases} \tag{24}$$

V. Hand Gesture Tracking and Prediction

In a gesture tracking process, the gesture tracking results may appear deviation due to the disturbance of the ambient environment. So the gesture tracking prediction and correction is necessary. At short notice, a gesture of motion can be viewed as uniform motion in straight line. This paper uses the Kalman filter to predict the gesture tracking results. Kalman filter is a linear recursive filter, its prediction is unbiased, stability and optimality [24]. It makes the next state optimal estimation based on the state sequence of the former system.

According to the definition of literature [24], assumed that x_k, , $x_k^{'}$, $y_k^{'}$ are respectively the position of the center of the hand at K moment in X axis and Y axis as well as the position of the center of the hand that Kalman filter predicts. u_k, v_k, $u_k^{'}$, $v_k^{'}$ are expressed as movement velocity of the center of the hand at K moment in X axis, and Y axis and the moving speed Kalman filter predicts. Therefore, define the observation vector as: $Z_k = (x_k, y_k)$, state vector: $X_k = (x_k, y_k, u_k, v_k)$, then the predicted state vector is: $X_k^{'} = (x_k^{'}, y_k^{'}, u_k^{'}, v_k^{'})$. State transition matrix:

$$A = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{25}$$

where T is time difference of k moment and $k-1$ moment. The observation matrix:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{26}$$

Get the system state equation:

$$X_{k+1} = AX_k + W_k \tag{27}$$

System of observation equation:

$$Z_k = HX_k + V_k \tag{28}$$

where the measurement noise v_k obey P (V)~N (0, R), dynamic noise w_k obey P (W)~N (0, Q), and the two noise are uncorrelated white noise. R is the correlation matrix of measurement noise, Q is the correlation matrix of motion noise.

The prediction equations of the system is:

$$x_k' = A \cdot x_{k-1} + B \cdot u_k \tag{29}$$

The prediction equations of the priori error matrix is:

$$P_k' = A \cdot P_{k-1} \cdot A^T + Q \tag{30}$$

Kalman gain equation:

$$K_k = P_k' \cdot H^T \cdot (H \cdot P_k' \cdot H^T + R)^{-1} \tag{31}$$

State change equation:

$$x_k = x_k' + K_k \cdot (z_k - H \cdot x_k') \tag{32}$$

Covariance change equation:

$$P_k = (I - K_k \cdot H) \cdot P_k' \tag{33}$$

VI. ALGORITHM DESCRIPTION

According to the former description, the gesture model updating and results forecasting algorithm based on mean shift is described as follows:

Step 1: Judge skin color regions by using thresholding method, and obtain the gesture tracking target preliminarily;

Step 2: Use the background subtraction algorithm which based on motion information, combine with the skin color detection in step 1, get the tracking target accurately;

Step 3: Calculate the candidate target model $p(y_0)$ and the target model of Bhattacharyya coefficient $\rho(y_0)$;

Step 4: Calculate ω_i by using formula (12), calculate the target new position y_1 according to equation (4)and calculate $p(y_1)$ and $\rho(y_1)$;

Step 5: If $\rho(y_1) < \rho(y_0)$, then $y_1 < -(y_0 + y_1)/2$ and go to step 6;

Step 6: If $\|y_1 - y_0\| < \varepsilon$, then iterative end; otherwise make $y_0 = y_1$, and turn to step 3;

Step 7: After the final tracking, calculate the similarity coefficient ρ_t, if $\rho_t < \rho_{th}$, update the target model, otherwise, turn to the next tracking.

Step 8: Predict the tracking results using the last tracking results, and modify tracking results according to the predicted result.

VII. SIMULATION EXPERIMENT AND RESULTS ANALYSIS

In order to verify the accuracy, real-time and continuous of the algorithm, use a 30 frames per second, with dimensions of 640×480 sign language video as the contrast material. Take VC++ 6.0 as the development platform at the Windows XP SP3 environment, with AMD Athlon 7750Dual-Core Processor 2.70GHz, RAM 2GB.

Compare with the CamShift algorithm, the experiment results is shown in Table I. The experimental results chart is divided into three layers, where the first layer is a video extracted from the original image, the second layer is tracking results of CamShift algorithm, the third layer is the tracking results of the algorithm proposed in the paper. Seen in the table, in tracking process CamShift tracking algorithm appeared tracking error because of the other color background. Especially when the gesture tracking through the same or similar color region, such as the seventy-seventh frame, eighty-third frame, eighty-fifth frame and ninety-first frame, CamShift algorithm tracked the wrong target. While our approach got better tracking results.

In addition, we can clearly see the two tracking results from the trajectory diagram and the X and Y axes coordinate diagram. When the gesture tracked into the skin or similar skin color region with the CamShift algorithm, the target will appear deviation. The algorithm shows good robustness, tracking results had small difference with the actual position.

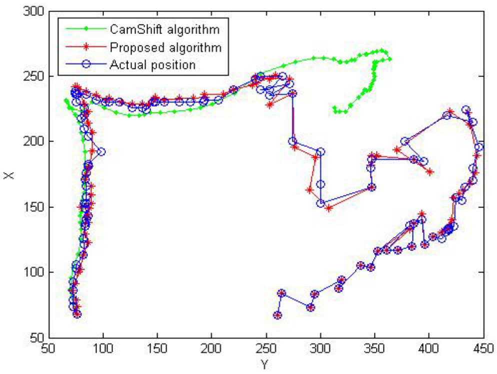

Figure 1. Comparison of tracking results position.

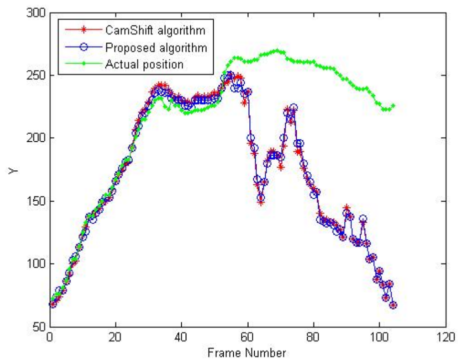

Figure 2. The X coordinates.

TABLE I.
COMPARISON OF TRACKING RESULTS

Frame	The Original Video Image	CamShift Algorithm Tracking Results	This Algorithm Tracking Results
The third frame			$img_1 47.jpg$
The seventeenth frame			
The thirty-sixth frame			
The forty-eighth frame			
The fifty-seventh frame			
The seventy-third frame			
The seventy-seventh frame			
The eighty-third frame			
The eighty-fifth frame			
The ninety-first frame			

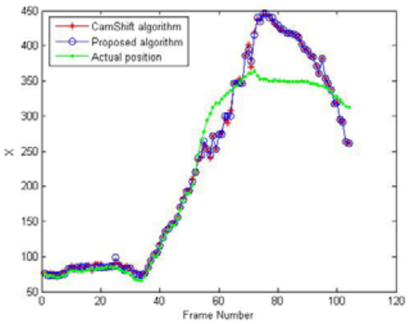

Figure 3. The Y coordinates.

VIII. CONCLUSIONS

To propose a gesture model updating and results forecasting algorithm based on Mean Shift, and to solve the problem of target model changing and influenced tracking results in gesture target tracking process. Firstly, the background difference and skin color detection methods are used to detect and get gesture model, and the Mean Shift algorithm is used to track gesture and update the gesture model, and finally to use the Kalman algorithm to predict the gesture tracking results. The experimental results show that this algorithm reduces the influence of surrounding environment in gesture tracking process, and get better tracking result.

However, the algorithm also has its flaws. For example in the shade and shelter, the hand gesture tracking is inaccurate. And also the contour description problem is still not solved. We will study it in future.

REFERENCES

[1] Ren H.B., Zhu Y.X., Xu G.Y etc. Vision-Based Recognition of Hand Gestures: A Survey. Acta Electronica Sinica, Vol. 2, No. 28, pp. 118-122, 2000.
[2] Dorin Comaniciu, Peter Meer. Mean Shift Analysis and Applications. In: Proceedings of IEEE International Conference on Computer Vision, Kerkira, Greece, pp. 1197-1203, 1999.
[3] Hou Z.Q., Han C.Z., A Survey of Visual Tracking. Acta Automatica Sinica, Vol. 32, No. 48, pp.603-617, 2006.
[4] Dorin Comaniciu, Visvanathan Ramesh, Peter Meer. Kernel-Based Object Tracking, IEEE transactions on pattern analysis and machine intelligence, Vol.25, No.5, pp.564-577, 2003.
[5] Dorin Comaniciu, Visvanathan Ramesh, Peter Meer, The Variable Bandwidth Mean Shift and Data-driven Scale Selection. In: Proceedings of IEEE International Conference on Computer Vision, Vancouver, BC, pp. 438-445, 2001.
[6] Peng J.C., Gu L.Z., Su J.B., The Hand Tracking for Humanoid Robot Using Camshift Algorithm and Kalman Filter. Journal of Shanghai Jiaotong University, Vol. 40, No. 7, pp. 1161-1165, 2006.
[7] Wang X.Y., Zhang X.W., Dai G.Z., An Approach to Tracking Deformable Hand Gesture for Real-Time Interaction. Journal of Software, 2007, Vol. 18, No. 10, pp. 2423-2433.
[8] Feng Z.Q., Yang B., Li Y etc., Research on Hand Gestures Tracking Based on Particle Filtering Aiming at Optimizing Time Cost. Acta Electronica Sinica, 2009, Vol.37, No.9, pp. 1989-1995.
[9] Gan M.G., Chen J., Wang Y.L etc. A Target Tracking Algorithm Based on Mean Shift and Normalized Moment of Inertia Feature. Acta Automatica Sinica, 2010, Vol.36, No.9, pp.1332-1336.
[10] Chen D.S., Liu Z.K. A Survey of Skin Color Detection. Chinese Journal of Computers, 2006, Vol.29, No.2, pp.194-207.
[11] Douglas Chai, King N. Ngan. Locating Facial Region of a Head-and-Shoulders Color Image, In: Proceedings of the 3th International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998, pp. 124-129.
[12] Kovac J, Peer P, Solina F. Human Skin Color Clustering for Face Detection, The IEEE REGION 8 EUROCON 2003, Computer as a Tool, 2003, Vol.8, No.2, pp.144-148.
[13] Xu Z.W, Zhu M.L. Optimum Colorspace for Skin-tone Detection. Journal Of Computer-Aided Design & Computer Graphics, 2006, Vol. 18, No. 9, pp.1350-1356.
[14] Xiao M., Han C.Z., Zhang L. Moving Object Detection Algorithm Based on Space-Time Background Difference. Journal of Computer-Aided Design & Computer Graphics, Vol. 7, No. 18, pp. 1044-1048, 2006.
[15] Rein-Lien Hus, Mohamed Abdel-Mottaleb, Anil K. Jain. Face Detection in Color Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24 No. 5, pp. 696-706, 2002.
[16] Dorin Comaniciu, visvanathan Ramesh, Peter Meer, Real-Time Tracking of Non-Rigid Objects using Mean Shift, In: Proceeding Of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 142-149, 2000.
[17] Wen Z.Q, Cai Z.X., Convergence Analysis of Mean Shift Algorithm, Journal of Software, Vol.2, No.18, pp. 205-212, 2007.
[18] An G.C., Mean Shift Algorithm Based on Fusion Model for Hhead Tracking. Journal of Southeast University (English Edition), Vol. 25, No. 3, pp. 299- 302, 2009.
[19] Tong X.M, Zhang Y.N, Yang T., Robust Object Tracking Based on Adaptive and Incremental Subspace Learning. Acta Automatica Sinica, Vol. 37, No. 12, pp.1483-1494, 2011.
[20] Shen Z.X., Yang X., Hang X.Y., Study on Target Model Update Method in Mean Shift Algorithm. Acta Automatica Sinica, vol. 35, No. 5, pp. 478-483, 2009.
[21] Katja Nummiaro, Esther Koller-Meier, Luc Van Gool. An Adaptive Color-based Particle Filter, Image and Vision Computing, vol. 21, No. 1, pp.99-110, 2002.
[22] Katja Nummiaro, Esther Koller-Meier, Luc Van Gool, Object Tracking with an Adaptive Color-Based Particle Filter, In: Proceedings of the 24th DAGM Symposium on Pattern Recognition. London, UK: Springer, pp. 353-360, 2002.
[23] R. Venkatesh Babu, Patrick Perez, Patrick Bouthemy, Robust Tracking with Motion Estimation and Local Kernel-based Color Modeling. Image Vision Computing, Vol. 25, No.8, pp.1205-1215, 2007.
[24] Luo Y., Li L., Zhang B.S., Yang H.M. Video Hand Tracking Algorithm Based on Hybrid CamShift and Kalman Filter. Application Research of Computers, Vol.26, No.3, pp.1163-1165, 2009.

Xiao Zou received the degree of master in School of Computer and Communication from Lanzhou University of Technology in 2006.Now, she server as an associate professor of School of Computer and Communication at Lanzhou University of Technology. Her current research interests include computer vision, pattern recognition, image processing, computer networks and information security.

Heng Wang received the BS and MS degrees both in School of Computer and Communication from Lanzhou University of Technology in 2009 and 2012, respectively. His current research interests include image processing, pattern recognition and computer vision.

Qiuyu Zhang is a researcher and doctoral tutor, vice dean of Technology & Research Center of Gansu Manufacturing Information Engineering, director of “software engineering” characteristic research direction and academic group of Lanzhou University of Technology. His research interests include image processing, pattern recognition, multimedia information processing, information security and software engineering.