A Novel Hand Gesture Tracking Algorithm Fusing Camshift and Particle Filter

M.H.Yao

College of Information Engineering Zhejiang University of Technology China

Q.L.Gu

College of Information Engineering Zhejiang University of Technology China

X.B.Wang College of Information Engineering

Zhejiang University of Technology China

Abstract—This paper proposes an algorithm fusing the Particle Filter and Camshift for improving tracking performance. The algorithm creates a dynamic model which integrates the information of color and motion, uses Camshift to optimize the state of particles and embeds Camshift into Particle Filter. The experimental results show that the new algorithm can effectively handle skin-colored interference and hand gesture transformation, have better robustness when compared with the traditional tracking algorithm.

Keywords-hand gesture tracking; Camshift; Particle Filter; skincolored interference

I. INTRODUCTION

As an interactional way between human and computers, hand gesture is more natural, concise and direct because hand is the most flexible part of the body[1]. The recognition of hand gesture introduces hand gesture to human-computer interface to achieve human-computer interaction(HCI) which is more suitable for human behavior habit. The effectiveness and robustness under complex environment is a challenging task so far[2,3].

There are many algorithms of hand gesture tracking. Definite tracking algorithm can be converted into optimal problem that is finding the optimal match of object. Continuously Adaptive Mean-shift (Camshift) is a kind of definite tracking algorithm which is proposed by Bradski[4,5]. It is based on color probability histogram and tracks the object with certain color effectively[6,7]. But it is interfered by large area of skin color and other part of body under complex background, especially one hand goes through the other hand will result in failure of tracking.

Particle Filter is one of the random tracking algorithms. It implements Bayes Filter based on Monte-Carlo and expresses the uncertainties of tracking through particle set. It W.X.He

College of Information Engineering Zhejiang University of Technology China

Q.Shen

College of Information Engineering Zhejiang University of Technology China

approximates the function of probability density through finding a group of random sample in state space and takes the mean value replacing the integration calculation. The tracking algorithm based on Particle Filter can maintain multiple assumptions and solve some non-linear problems[8,9], especially handle the background interfered and restore from wrong tracking. Usually it can get stable result[10,11].

This paper proposes an algorithm fusing Camshift and Particle Filter to achieve better hand gesture tracking with the combination of color and motion. The experiments are implemented on the OpenCV platform and the experimental results show the effectiveness of the proposed algorithm.

II. TRACKING ALGORITHM

A. CamshiftAlgorithm

Denote the pixel coordinate of the object image P by (x_i, y_i) , where $i = 1, \dots, n$, n is the pixel number of P. Therefore, the color histogram of P can be represented as

$$q(u) = \frac{1}{n} \sum_{(x,y) \in P} \delta(c(f(x,y)) - u)$$
(1)

where f(x, y) is the image function of object image P,

 $c(\bullet)$ is the quantization function of the color space, $\delta(\bullet)$ is the Kroneckerfunction, and u is the total block number of the histogram.

Denote the pixel (probability) value at position (x_i, y_i) in the image by I(x, y). Therefore, I(x, y) is calculated as

$$I(x, y) = \sum_{u=1}^{m} q(u) \delta(c(f(x, y)) - u)$$
(2)

then we get the zero-order moment M_{00} of the probability distribution in the searching window

$$M_{00} = \sum_{x} \sum_{y} I(x, y)$$
(3)

In the same way, the first-order moment f x and y can be represented as

$$M_{10} = \sum_{x} \sum_{y} xI(x, y)$$
(4)

$$M_{01} = \sum_{x} \sum_{y} yI(x, y)$$
(5)

Finally the mean location of searching window (centroid) can be gotten.

$$\begin{cases} x_{c} = \frac{M_{01}}{M_{00}} \\ y_{c} = \frac{M_{10}}{M_{00}} \end{cases}$$
(6)

The center of the searching window is moved to the centroid as soon as the color centroid of the object is found and the size of searching window is adjusted according to the sum of pixel value of object. The moving distance of searching window is compared with default threshold. If the moving distance is larger than the threshold, the centroid of window will be recalculated and the location and size of searching window adjusted until the distance between the center and the centroid is less than the threshold or the cycle time is the maximum. The convergence condition is satisfied and next frame image begins. The threshold is very small compared to searching window.

For current frame, the second-order moment can be calculated as

$$M_{20} = \sum_{x} \sum_{y} x^{2} I(x, y)$$
(7)
$$M_{02} = \sum_{x} \sum_{y} y^{2} I(x, y)$$
(8)

Let
$$a = \frac{M_{20}}{M_{00}} - x_c^2$$
, $b = 2\left(\frac{M_{11}}{M_{00}} - x_c y_c\right)$

 $c = \frac{M_{02}}{M_{00}} - y_c^2$, then length *l* and width *W* of searching

window which are through centroid and perpendicular can be expressed as

$$\begin{cases} l = \sqrt{\frac{(a + c) + \sqrt{b^{2} + (a - c)^{2}}}{2}} \\ w = \sqrt{\frac{(a + c) - \sqrt{b^{2} + (a - c)^{2}}}{2}} \end{cases}$$
(9)

l and w is the long axis and short axis of ellipse respectively if ellipse is used to focus the current object. They can be drawn by computer and l and w is the length and width of next frame respectively.

B. ParticleFilter

Particle Filter, based on Monte Carlo, is a sequential importance sampling (SIS) algorithm by sampling random state particle to express its distribution from posterior probability. The object tracking based on Particle Filter can be described as estimating the system state under a certain group of observing condition.

In object tracking, the state-space model of a dynamic system can be described as

$$\begin{cases} x_{k} = f(x_{k-1}) + u_{k-1} \\ z_{k} = h(x_{k}) + v_{k} \end{cases}$$
(10)

where $f(\bullet)$ is the state transition equation, $h(\bullet)$ is the observation equation, x_k is the system state, u_{k-1} is the system transition noise at the time k-1, z_k is the observation value and v_k is theobservation noise at the time k.

Before getting the observation value atevery moment, we use the state of the formermoment to estimate

$$p(x_{k} | Z_{k-1}) = \int p(x_{k} | x_{k-1}) p(x_{k-1} | Z_{k-1}) dx_{k-1}$$
(11)

where $Z_k = Z_{1:k-1} = \{Z_1, ..., Z_{k-1}\}$ is the observation value at time 1 tok-1.

After getting the observation value, we update the prior probability density by Bayes formula and then get the posteriori probability density

$$p(x_{k} | Z_{k}) = \frac{p(z_{k} | x_{k}, Z_{k-1}) p(x_{k} | Z_{k-1})}{p(z_{k} | Z_{k-1})}$$
(12)

If Z_k only determined by x_k , that is $p(z_k | x_k, Z_{k-1}) = p(z_k | x_k)$, then

$$p(x_{k} | Z_{k}) = \frac{p(z_{k} | x_{k}) p(x_{k} | Z_{k-1})}{p(z_{k} | Z_{k-1})}$$
(13)

where $p(z_k | Z_{k-1})$ is a normalization constant

$$p(z_{k} | Z_{k-1}) = \int p(z_{k} | x_{k}) p(x_{k} | Z_{k-1}) dx_{k}$$
(14)

 $p(z_k | x_k)$ is the similarity between z_k and x_k , $p(x_k | Z_{k-1})$ is the prior probability density of the former step.

III. IMPROVED ALGORITHM

Supposing the state variable of object $X_t = (x_t, y_t)$, where (x_t, y_t) is the center coordinating of the object. The dynamic model describing object moving is

$$X_{t} - X_{t-1} = X_{t-1} - X_{t-2} + w_{t-1}$$
(15)

where w_{t-1} is a Gaussian stochastic component. The particle set is propagated by this simple constant velocity model.

Dynamic model cannot predict effectively sample sequence not included in the training set. This paper adopts the uniform velocity model to achieve a hand gesture tracking, which is weak in learning ability but enough for some common hand movements.

This paper uses non-geometric characteristics to describe hand, such as color feature and motion information. Therefore, we need a reliable observation model. In order to solve the problem of interference caused by skin-colored objects, particle weighting is only done in moving skin-colored regions and motion-color probability distribution in the region can be obtained at the same time.

The zero-order moment of object color probability distribution M_{c00} and the zero-order moment of motion-color probability M_{m00} can be expressed as

$$M_{c00} = \sum_{x} \sum_{y} I_{c}(x, y)$$
(16)

$$M_{m00} = \sum_{x} \sum_{y} I_{m}(x, y)$$
(17)

where $I_c(x, y)$ is the pixel value of color probability distribution at (x, y), and $I_m(x, y)$ is the pixel value of motion-color probability distribution at (x, y). When the hand moves quickly, the motion-color probability distribution contains reliable motion and color moments. On the other hand, when the hand moves slowly, the particles primarily weights by color moments in the color probability distribution. Therefore, the color moment and the motion moment can be combined linearly as

$$M_{00} = (1 - \alpha) \cdot M_{c00} + \alpha \cdot M_{m00}$$

= $\sum_{x} \sum_{y} ((1 - \alpha) I_{c}(x, y) + \alpha I_{m}(x, y))$
(18)

where $\alpha \in [0,1]$ shows the contribution of motion moment to hand tracking, and it is often set as 0.8 to simplify the process.

Given a particle x^i , the corresponding location is $C_0(x_0, y_0)$. We assume that it is the initial location and initialize the iteration number n=0.

Calculating zero-order moment and first-order moment according to Camshift:

$$M_{00} = \sum_{x} \sum_{y} \left((1 - \alpha) I_c(x, y) + \alpha I_m(x, y) \right)$$

$$M_{10} = \sum_{x} \sum_{y} x \left((1 - \alpha) I_c(x, y) + \alpha I_m(x, y) \right)$$

$$M_{01} = \sum_{x} \sum_{y} y \left((1 - \alpha) I_c(x, y) + \alpha I_m(x, y) \right)$$
(19)

So the mean location of searching window (centroid) is

$$\begin{cases} x_{c} = \frac{M_{-01}}{M_{-00}} \\ y_{c} = \frac{M_{-10}}{M_{-00}} \end{cases}$$
(20)

The final location of the window can be found.

Based on color and motion, a new likelihood function can be created, $p(z_k | x_k^i)$, mentioned in the former section. A distance function is as follows to express the similarity between particle and object

$$D = \sqrt{1 - \frac{M}{M_0}} \tag{21}$$

where $M_{0} = \sum_{x} \sum_{y} 1$ is the number of pixels in the searching window.

Then the new likelihood function $p(z_k | x_k^i)$ is

$$p\left(z_{k} \mid x_{k}^{i}\right) = -\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{D^{2}}{2\sigma^{2}}}$$
(22)

where α is an empirical constant obtained by observing the tracking experiment.

Finally, the state of the hand can be estimated by the particle weighting.

IV.EXPERIMENTS

The experiments are implemented on standard PC hardware. The video image size is 640*480 captured by

Logitech QuickCam Pro5000 at 15fps. The software platform is Microsoft Visual Studio 2010 and OpenCV 2.4.2.

At first, Camshift is used for hand gesture tracking. The color histogram of H component is extracted without the pixels in S and V component impacting the H component greatly (the threshold of S and V are as follows, $S_{\min} = 30, V_{\min} = 20, V_{\max} = 235$). The tracking process is shown in Figure 1.



FIGURE I. HAND GESTURE TRACKING BASED ON CAMSHIFT

Figure1 shows that the searching window is disturbed when one hand is passing another one which will result in failure tracking. Next we use Particle Filter to track hand gesture in the same situation and the number of particles is 500. The tracking process is shown in Figure 2.

Figure2 shows that Particle Filter has some advantages to handle skin-colored interference. When dealing with the same situation in Figure1, the Particle Filter still keeps tracking. But the more number of particles also increases the calculation complex which results in Particle Filter not handle the problem of moving with hand gesture. It is worse than Camshift in realtime performance and loses its object when the disturbing hand moving.





GESTURE CHANGING

Considering that both of them have their own advantages and disadvantages, this paper proposes a new approach to track hand gesture based on the fusion of Camshift and Particle Filter. At this time, the number of particles is set to 75. The tracking process is shown in Figure 3.

The proposed algorithm overcomes the skin-colored interference and improves robustness of tracking in Figure 3.You can see that the proposed algorithm still keeps effective tracking and achieves the improving target.

V. CONCLUSION

Camshift is real-time but cannot deal with the problem of large area skin-color interference, while Particle Filter can solve this problem but it is poor in real-time performance. So this paper proposes a novel algorithm combining Camshift and Particle Filter. It takes into account the large area skin-color interference and hand gesture changing when tracking. The experiments show that the proposed algorithm works better in optimizing tracking effect and improving real-time tracking performance. It can track hand gesture of object accurately, handle the problem of interference and make preparations for hand gesture recognition and comprehension next.

ACKNOWLEDGMENTS

The work is supported by Zhejiang Provincial Natural Science Foundation of China(Grant No. LZ14F030001).

REFERENCES

- Qiuyu Zhang, Jianqiang Hu, MoyiZhang, Mean Shift Dynamic Deforming Hand Gesture Tracking Algorithm Based on Region Growth, *Pattern Recognition and Artificial Intelligence*, vol. 23, no. 4, pp.580-585,2010.
- [2] Jialue Fan, Xiaohui Shen, Ying Wu,What Are We Tracking: A Unified Approach of Tracking and Recognition,*Image Processing, IEEE Trans*, vol. 22, no. 2,pp.549-560, 2013.
- [3] C.R. Cameron, L.W.DiValentin, RManaktala, Hand Tracking and Visualization in a Virtual Reality Simulation,*InProceedings of the 2011 IEEE Systems and Information Engineering Design Symposium*, pp.127-132, 2011.
- [4] Gary R.Bradski,Real Time Face and Object Tracking as a Component of a Perceptual User Interface,*InProceedings of the Fourth IEEE Workshop* on Applications of Computer Vision, pp.214-2191998.
- [5] Gary R. Bradski, Computer Vision Face Tracking For Use in a Perceptual User Interface, *Intel Technology Journal*, vol. 2, no.2,pp.12-21, 1998.
- [6] Juanchun Peng, LizhongGu,Jianbo Su, The Hand Tracking for Humanoid Robot Using Camshift Algorithm and KalmanFilter, *Journal* of Shanghai Jiaotong University, vol.40, no.7, pp.1161-1165, 2006.
- [7] Jiaming Xu, LunXie, ZhiliangWang, Targeting and tracking algorithm of humanoid soccer robot, *Journal of Huazhong University of Science and Technology*(*Natural Science Edition*), vol. 39, sup. 2, pp.243-258, 2011.
- [8] M. Bray, E.Koller-Meier, L. Van Gool, Smart particle filtering for 3D hand tracking, *InProceedings of the Sixth IEEE International Conference* on Automatic Face and Gesture Recognition, pp.675-680, 2004.
- [9] Zhiquan Feng, Bo Yang, Yi Li, Research on Hand Gestures Tracking Based on Particle Filtering Aiming at Optimizing Time Cost, *Acta Electronica Sinica*, vol. 37, no. 9, pp.1989-1995, 2009.
- [10] K.Nummiaro, E.Koller-Meier, L Van Gool, Anadaptivecolor-based particle filter, *Image and VisionComputing*, vol. 21, no.1, pp.99-110, 2003.
- [11] Huimin Qian, Yaobin Mao, Jason Geng, Object Tracking with Selfupdatingtracking window, In Proceedings of Pacific AsiaWorkshop on Intelligence and Security Informatics, 2007.