

An Improved Camshift-based Particle Filter Algorithm for Real-time Hand Gesture Tracking

¹Zhang Fang Hu, ²Yun Kai Wang, ¹Yuan Luo, ²Yi Zhang, ¹Bing Xi

¹Key Laboratory of Optical Fiber Communication Technology Chongqing Education Commission, Chongqing University of Post and Telecommunications, Chongqing 400065, P. R. China

²Research Center of Intelligent System and Robot, Chongqing University of Post and Telecommunications, Chongqing 400065, P. R. China

¹Tel.: +86-18725875566

¹E-mail: huzf@cqupt.edu.cn

Received: 24 May 2014 /Accepted: 31 July 2014 /Published: 31 August 2014

Abstract: In the study of dynamic gesture recognition, gesture tracking must be performed reliably in real-time for sufficiently long periods when there is too much background interference. To deal with these problems effectively, an improved particle filter algorithm is proposed in order to track the moving hand quickly and accurately. Firstly, the algorithm improves the traditional hand model and presents a novel hand model, which fuses color and depth cues, to enhance the robustness and accuracy of gesture tracking. Meanwhile, in order to increase the tracking efficiency, the Camshift algorithm is embedded into the particle filter to rearrange the random particles, in which the particles moved toward the maximal posterior probability density of the target state. Experimental results show that compared with the traditional particle filter algorithm or Camshift algorithm, the proposed method achieve fast and robust tracking of the hand with the situations of fast moving hand and strong disturbances in the background. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Particles filter, Camshift, Color and depth model, Gesture tracking, Posterior probability.

1. Introduction

Vision-based gesture recognition is an important field of human-computer interaction research. Using hand gestures instead of an intermediary is interactively natural, intuitive, closer to human communication habits, and its prospects are very broad. Hand gesture analysis and understanding is the key technology to achieve next-generation human-computer interaction, and gesture tracking in complex environments has been a practical gesture analysis bottleneck.

A variety of methods has been proposed to track hand, but most are based on a number of assumptions, such as a single gesture, little of

background interference. When there is too much background interference, tracking accuracy and efficiency drastically decrease. In recent years, the mean shift algorithm has been widely used for its small amount of calculation, but it can't guarantee the global optimum, easy to fall into local optimum [9]. Non-parametric estimation of particle filter in hand tracking fields have developed rapidly, but in order to ensure the stability of the track, the number of particles should be more than enough, which leads to increased amount of computation [3-5]. Therefore, in order to solve the problem that a single algorithm is difficult to achieve real-time, the method combined meanshift and adaptive particles filtering has been proposed to achieve some results, but these

algorithms are based on color information to build gesture model, and there is little prospect. When the color is similar to hand occurs, it will affect the tracking results.

In order to solve the problem of the above algorithm, an improved particle filter algorithm is proposed. Firstly, the algorithm improves the traditional hand model and presents a novel hand model, which fuses color and depth cues, effectively distinguishes between foreground and background, enhances immunity gesture tracking; meanwhile, to increase the tracking efficiency, the Camshift algorithm is embedded into the particle filter to rearrange the random particles, in which the particles moved toward the maximal posterior probability density of the target state. Experimental results show that the proposed method not only overcome the lack of large particles, but also can effectively overcome the disadvantage of camshift into a local maximum, achieve fast and robust tracking with the situations of fast moving hand or strong disturbances in the background.

2. The Traditional Particle Filter Algorithm

Particle filter algorithm is proposed by Monte Carlo. Its core idea is that the discrete states, a random sample from the posterior probability, represent the posterior probability density distribution of the target state variable, and then use the current value and the previous observations to estimate the current state of the target. Simply to say, it is the process that finding a random sample in the state space to represent the probability density function nearly, with the sample means instead of calculus, then get the system state estimation.

Suppose $\{x_k^{(i)}\}_{i=1}^N$ is the particle of moment k , and its normalized weight is $\{w_k^{(i)}\}_{i=1}^N$, the posterior probability can be represented by $\{(x_k^{(i)}, w_k^{(i)})\}_{i=1}^N$.

$$p(x_k / z_{1:k}) \approx \sum_{i=1}^N w_k^{(i)} d(x_k - x_k^{(i)}), \quad (1)$$

where $z_{1:k} = \{z_1, z_2, \dots, z_k\}$ represent all observations until the moment k , and

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(z_k | x_k^{(i)}) p(x_k^{(i)} | x_{k-1}^{(i)})}{q(x_k^{(i)} | x_{k-1}^{(i)}, z_{1:k})}, \quad (2)$$

Particle filter algorithm can be applied to both non-linear system and non-Gaussian distributions system. In line with the actual situation, so it has a good tracking in the complex background system. However, in order to capture changes, a certain number of particles are required to ensure adequate

sampling, and the number grows exponentially with the dimensions of the state, so there are a large number of particles which lead to a high amount of calculation, make the particle filter real poor.

3. Improved Particle Filter Hand Tracking

The improved particle filter gesture tracking, firstly, improves the traditional hand model and presents a novel hand model, which fuses color and depth cues; meanwhile, to increase the tracking efficiency, the Camshift algorithm is embedded into the particle filter to rearrange the random particles.

3.1. Color Information Hand Model

To reduce the computational cost, we use a simple and effective color histogram feature to describe the hand. Color probability distribution is obtained through the hue component of HSV. When interference occurs in the background similar to hand gesture with hue value, tracking performance will drastically decrease. This paper introduces the saturation values will effectively distinguish target from background. Therefore, using the H component as the main color characteristic parameters, and the S component as secondary color characteristic parameters, color model is jointly established.

Firstly, Color histograms of H and S component in hand templates are constructed, then through back projection of their color histogram the probability distribution of H and S component in current frame I_h and I_s is obtain, finally the overall probability distribution I_c of the current frame is obtained,

$$I_c = \alpha I_h + (1 - \alpha) I_s, \quad (3)$$

where $\alpha \in [0.5, 1]$ is the weight which indicates the contribution for color information of hue component, through several experiments, it is appropriate when $\alpha = 0.7$. This method uses the weight hue and saturation obtained corresponds to the color characteristics of the target for any goals versatility.

3.2. Fusion Gesture Model of Color and Depth Information

The traditional model of the target can't reflect the characteristics of the target accurately, while the background is complex or the color distribution between background and foreground are more alike, and then leads to tracking failure. To distinguish between foreground and background, we introduce depth information. Probability distribution of depth information I_d is obtained through extracting depth

information. As a basis for modeling, making target more compact model. Specific methods are as follows.

$$M_c = \sum_x \sum_y I_c(x, y), M_d = \sum_x \sum_y I_d(x, y), \quad (4)$$

where M_c and M_d are the Zero-order moment of the Gesture rectangular area in the color and the depth probability distribution respectively, $I_c(x, y)$ and $I_d(x, y)$ are the pixel values corresponding to the coordinate (x, y) .

Fusion of color and depth information is in the following.

$$M = (1 - \beta) \cdot M_c + \beta \cdot M_d, \quad (5)$$

where M is the joint zero-order moment, and β is the weight which indicates the contribution for tracking of depth information, in the paper $\beta \in [0, 1]$, we use $\beta = 0.8$. The weight of color and depth information corresponding to the two characteristics, based on this model leads the target model reliability and stability. The target model will be applied to the particle filter algorithm.

3.3. Improved Particle Filter Algorithm

In order to increase the tracking efficiency, the Camshift algorithm is embedded into the particle filter to rearrange the random particles, in which the particles move toward the maximal posterior probability density of the target state. Thus the sample sets converges in the area near the true state of the hand. The set of particles obtained by drift can effectively and quickly estimate gestures state. So the algorithm (now called CAMPF) can have a good tracking result by a little number of particles. Steps are as follows.

1) Firstly, we obtain gesture template from the initial frame, and then yield parameters of the target state. While the set of particles $\left\{ \left(x_1^{(i)}, w_1^{(i)} \right) \right\}_{i=1}^N$ of the Gaussian distribution are taken in the vicinity of gesture area. In $\left\{ w_1^{(i)} = 1/N \right\}_{i=1}^N$, and N is the number of particles.

2) Through the following state transition model and the set of particles of the moment $k-1$ $\left\{ \left(x_{k-1}^{(i)}, w_{k-1}^{(i)} \right) \right\}_{i=1}^N$, we can obtain the set of particles of moment k $\left\{ \left(x_k^{(i)}, w_k^{(i)} \right) \right\}_{i=1}^N$.

$$\hat{x}_k - x_{k-1} = x_{k-1} - x_{k-2} + u_{k-1}, \quad (6)$$

where u_{k-1} is in line with Gaussian random numbers of expectation 0 and variance 0.5.

3) We extract color and depth information, each particle drifted by Camshift algorithm.

$$\tilde{x}_k^{(i)} = \text{Camshift}(x_k^{(i)}), \quad (7)$$

Get the set of particles drifted $\left\{ \left(\tilde{x}_k^{(i)}, w_k^{(i)} \right) \right\}_{i=1}^N$. The gesture model we present in the last section is set as the tracking model in the equation (7). And the steps of Camshift algorithm are as follow.

$$x_1 = \frac{M_{10}}{M_{00}}, y_1 = \frac{M_{01}}{M_{00}}, \quad (8)$$

Assuming the initial position of the particle is $C_0(x_0, y_0)$, and then build gesture-sized rectangular area. We can get the center of mass $C_1(x_1, y_1)$ by the equation (8), where M_{00} is the 0th moment, M_{10} is the first moment of x , M_{01} is the first moment of y . If it has the situation $\|C_1 - C_0\| < \varepsilon$, we stop iteration, get the center of mass C_1 , otherwise we put $C_0 = C_1$, return to compute the center of mass. The ε is 2, and we just have 9 iteration. For each particle adjust continuously to determine the centroid position of the particle.

4) We create a new model $p(z_k | x_k^{(i)})$ based on color and depth information. Each particle represents the possibility of the state gesture. The particles similar to the actual have a greater weight; we define one function to measure similarity between particles and gesture template.

$$D = \sqrt{1 - M/M_0}, \quad (9)$$

where M_0 is the 0th moment of gesture template. And we get the observation function of particle filter.

$$p(z_k | x_k^{(i)}) = 1/\sqrt{2\pi} \exp(-D^2/2), \quad (10)$$

Then obtain the weight of particles.

$$w_k^{(i)} = w_{k-1}^{(i)} p(z_k | x_k^{(i)}), i = 1, \dots, N, \quad (11)$$

5) We normalize the weights of particles, and also use resampling techniques to avoid degradation.

6) Calculation of the target state of the current frame.

$$\hat{x}_k \approx E(x_k | z_{1:k}) = \sum_{i=1}^N w_k^{(i)} x_k^{(i)}, \quad (12)$$

The flowchart of improved particle filter algorithm is as shows in Fig. 1.

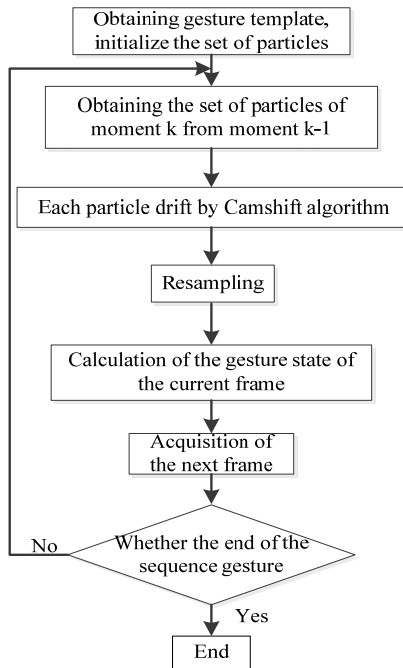


Fig. 1. Flowchart of improved particle filter algorithm.

4. Experimental Results and Analysis

4.1. Verifying the Robustness of Color and Depth Information

In order to verify the robustness of gesture model based on color and depth information, we capture video sequences at 25 fps, with 455×460 pixels, totally 60 frames, and track the palm. Compare the proposed algorithm with algorithm set up in object model only using hue component. The two algorithms use the same number of particles.

Fig. 2 shows the tracking results of frame 22, 34, 46, 58. We can see that palm not only disturbed by color similar to hand in the background, and the hand is blocked. Color information of hand becomes unreliable, but the depth information of the hand performs well. This algorithm can keep tracking of gestures, but the algorithm only using an H component has tracking error, which also leads to follow-up failure.



(a) Tracking results of this algorithm



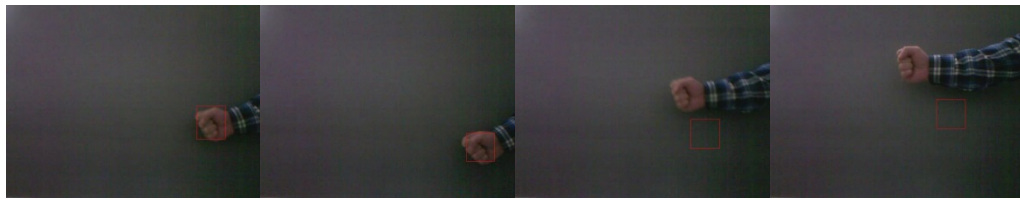
(b) Track results of algorithms only using H component

Fig. 2. Illustration of the robustness for the target model based on color and depth cues.

4.2. The Tracking Performance of the Three Algorithms

In this section, we capture video sequences at 30 fps, with 640×480 pixels, 70 frames, and track the fist. Efficiency and performance of tracking between the algorithm of traditional particle filter and the algorithm proposed in this paper, as well as the Camshift, are compared with each other. The algorithm of traditional particle filter will use the same state transition model and observation model

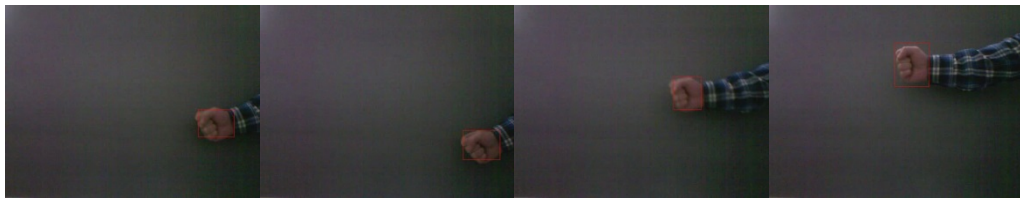
with the algorithm proposed in this paper, algorithm Camshift uses a single color model-based H component. In order to track the gestures better, the algorithm of traditional particle filter requires at least 400 particles, the algorithm that Camshift embedded into the particle filter will be improved, and the set of sample will converge to the true state close to the target, can estimate gestures state effectively and rapidly just by 60 particles. The tracking results of sequences at the frame 23 to 26, 30 to 34 by three algorithms are proposed in Fig. 3.



(a) Tracking results of Camshift



(b) Tracking results of traditional particle filter (N=400)

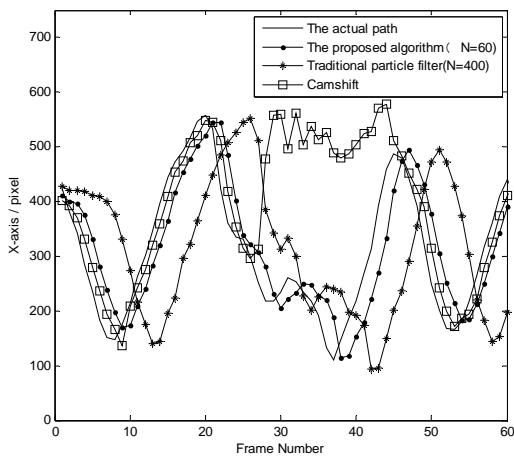


(c) Tracking results of improved particle filter (N=60)

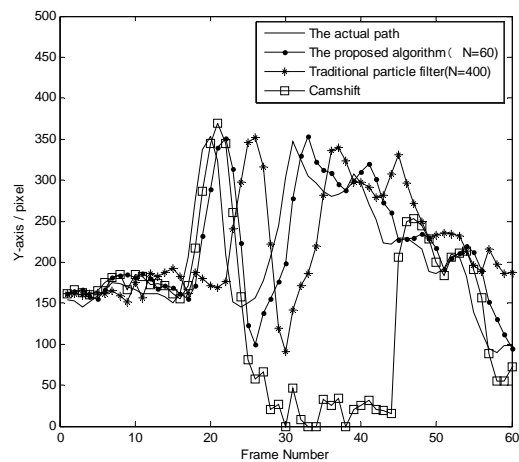
Fig. 3. Comparison of tracking result among three different methods.

As is shown in Fig. 3, the red boxes mark these gestures. The figure shows only the algorithm proposed in this paper accurately track the gesture, as the fist at frame 26 speed to accelerate, leading to deviation from the algorithm of traditional particle filter and Camshift. Until frame 34, the traditional particle filter has recovered, but tracking error by Camshift. It shows that the proposed algorithm is superior to the algorithm of traditional particle filter and Camshift.

Fig. 4 is the comparison of the trajectory coordinates of three algorithms and the actual path. As we can see from Fig. 4 the trajectory of the algorithm proposed in this paper is more accurate than traditional particle filter and Camshift. Thus the algorithm can be more accurately to track the gestures.



(a) Tracking of X-axis



(b) Tracking of Y-axis

Fig. 4. Tracking accuracy comparison between three algorithms.

To illustrate real-time of the proposed algorithm, the average execution time of each frame of the three algorithms is given in Table 1. Comparing with the traditional particle filter, the algorithm proposed in this paper greatly reduces the number of particles; and has small amount of calculation relatively. Although computational overhead of Camshift is smaller than the proposed algorithm, but its tracking performance is far worse than the proposed algorithm.

Table 1. Comparison of tracking efficiency among three algorithms.

	The proposed algorithm (N=60)	Traditional particle filter (N=400)	Camshift
The execution time of each frame /s	0.120	0.231	0.023

5. Conclusion

A novel hand model which fuses color and depth cues and has enhanced the robustness and accuracy of gesture tracking are proposed. Meanwhile, the Camshift algorithm is embedded into the particle filter to rearrange the random particles, and the particles converge to the true state of gesture, which can play a role of particle filter. Reduction in the number of particles can greatly reduce the complexity of the algorithm. And the particles drifted move toward the maximal posterior probability density of the target state. So the tracking is accuracy. Experimental results show that the proposed method not only overcome the lack of large particles, but also can effectively overcome the disadvantage of camshift into a local maximum, achieve fast and robust tracking with the situations of fast moving hand or strong disturbances in the background.

Acknowledgements

This paper is supported by Natural Science Foundation of Chongqing University of Post and

Telecommunications (A2009-49) and the Science and Technology Research Project of Chongqing Education Commission (KJ130512).

References

- [1]. Xiujian Chai, Yili Tang, Shiguang Shan, Xilin Chen, Bingpeng Ma, Image and Depth Fusion Based Hand Tracking and Application, in *Proceedings of the Academic Conference on Harmonious Man-machine Environment (HHME'11)*, Beijing, China, 2011, pp. 68-73.
- [2]. Yuan Luo, Ling Li, Baisheng Zhang, Hong-Mei Yang, Hand Gesture Tracking Method Based on Kalman Filter and Skin Color Feature, *Journal of Northeastern University (Natural Science)*, Vol. 34, Issue 4, 2013, pp. 3547-3550.
- [3]. Del Bimbo A., Dini F., Particle filter-based visual tracking with a first order dynamic model and uncertainty adaptation, *Computer Vision and Image Understanding*, Vol. 115, Issue 6, 2011, pp. 771-786.
- [4]. Okuma K., Taleghani A., De Freitas N., et al., A boosted particle filter: Multitarget detection and tracking, *Computer Vision - ECCV 2004, Lecture Notes in Computer Science*, 3021, 2004, pp. 28-39.
- [5]. Ni Z., Sunderrajan S., Rahimi A., et al., Particle filter tracking with online multiple instance learning, in *Proceedings of the 20th IEEE International Conference on Pattern Recognition (ICPR'10)*, 2010, pp. 2616-2619.
- [6]. Rigatos G. G., Nonlinear Kalman Filters and Particle Filters for integrated navigation of unmanned aerial vehicles, *Robotics and Autonomous Systems*, Vol. 60, Issue 7, 2012, pp. 978-995.
- [7]. Yin M., Zhang J., Sun H., et al., Multi-cue-based Camshift guided particle filter tracking, *Expert Systems with Applications*, Vol. 38, Issue 5, 2011, pp. 6313-6318.
- [8]. Xin Wang, Zhenmin Tang, An improved camshift-based particle filter algorithm for real-time target tracking, *Journal of Image and Graphics*, Vol. 15, Issue 10, 2010, pp. 1507-1514.
- [9]. Hongzhi Wang, Nong Sang, Song Luo, Hand tracking algorithm based on mean shift and region growth, *J. Huazhong Univ. of Sci. &Tech. (Natural Science Edition)*, Vol. 1, Issue 41, 2013, pp. 133-136.
- [10]. Sun Q. Y., Li L. S., Chao Z. L., Target Tracking Based on Particle Filter with Multi-Path Particles, *Applied Mechanics and Materials*, Vol. 130, 2012, pp. 3306-3310.