

http://dx.doi.org/10.6109/ijice.2011.9.6.633

Automation of an Interactive Interview System by Hand Gesture Recognition Using Particle Filter

Yang-Weon Lee, *Member, KIICE*

Abstract—This paper describes a implementation of virtual interactive interview system. A hand motion recognition algorithm based on the particle filters is applied for this system. The particle filter is well operated for human hand motion recognition than any other recognition algorithm. Through the experiments, we show that the proposed scheme is stable and works well in virtual interview system's environments.

Index Terms—Filtering, Gesture recognition, Interactive interview system, Particle Filter.

I. INTRODUCTION

GESTURES and gesture recognition are terms increasingly encountered in discussions of human-computer interaction. The gestural equivalents of direct manipulation interfaces are those which use gesture alone. These can range from interfaces that recognize a few symbolic gestures to those that implement fully fledged sign language interpretation. Similarly interfaces may recognize static hand poses, or dynamic hand motion, or a combination of both.

In this paper, we focused into the development of hand gesture recognition using particle filter to apply the virtual interview system. As shown in Fig. 1, the virtual interview system is designed for interviewee to make a practice before facing real interview.

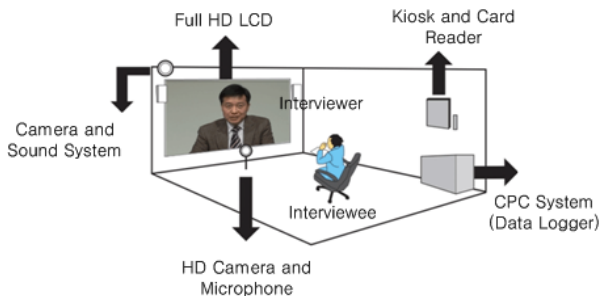


Fig. 1. Overall hand motion recognition operation diagram

But this system is not automatically interactive system between interviewee and its system. So, it is needed to revise the system as a fully automated mode. We applied particle filter for hand motion recognition and tracking. Particle filter [1] is based on the Bayesian conditional probability such as prior distribution and posterior distribution. First of all, we expanded the existing algorithm [2] to derive the CONDENSATION-based particle filter for hand motion recognition. Also, we adopt the nine hand motion model to confirm the algorithm performance such as leftover and paddle. The overall scheme for the hand motion recognition system is shown in Fig. 2.

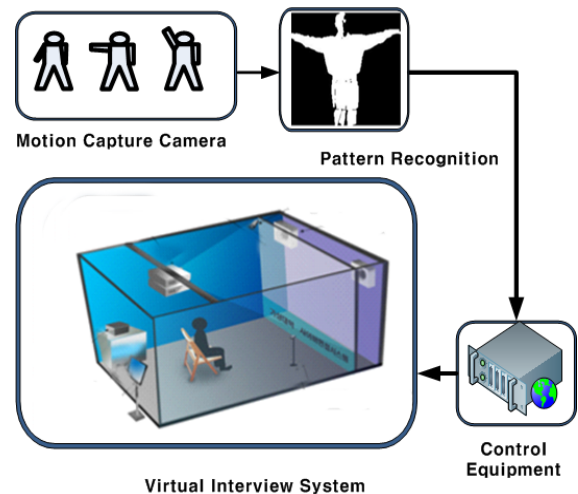


Fig. 2. Overall System Operation Diagram

II. PARTICLE FILTER ALGORITHM

A. Particle Filter

The particle filter approach to track motion, also known as the condensation algorithm [1] and Monte Carlo localization, uses a large number of particles to 'explore' the state space. Each particle represents a hypothesised target location in state space. Initially the particles are uniformly randomly distributed across the state space, and each subsequent frame the algorithm cycles through the steps illustrated in Fig. 3 :

1. Deterministic drift: particles are moved according to a deterministic motion model (a damped

Manuscript received September 18, 2011; revised October 9, 2011; accepted October 15, 2011.

Yang-Weon Lee is with the Department of Information and Communication Engineering, Honam University, Gwangju, 506-714, Korea (Email: ywlee@honam.ac.kr)

- constant velocity motion model was used).
2. Update probability density function (PDF): Determine the probability for every new particle location.
3. Resample particles: 90 percent of the particles are resampled with replacement, such that the probability of choosing a particular sample is equal to the PDF at that point; the remaining 10 percent of particles are distributed randomly throughout the state space.
4. Diffuse particles: particles are moved a small distance in state space under Brownian motion.

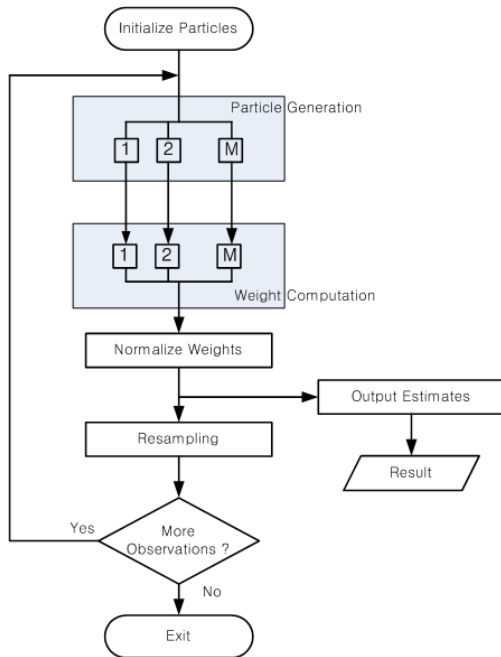


Fig.3. Process of particle filter calculation

This result in particles congregating in regions of high probability and dispersing from other regions, thus the particle density indicates the most likely target states. See[3] for a comprehensive discussion of this method. The key strengths of the particle filter approach to localization and tracking are its scalability (computational requirement varies linearly with the number of particles), and its ability to deal with multiple hypotheses (and thus more readily recover from tracking errors). However, the particle filter was applied here for several additional reasons:

- It provides an efficient means of searching for a target in a multi-dimensional state space.
- Reduces the search problem to a verification problem, i.e. is a given hypothesis face-like according to the sensor information?
- Allows fusion of cues running at different frequencies.

B. Application of Particle Filter for the Hand Motion Recognition

In order to apply the particle filter algorithm to hand motion recognition, we extend the methods described by Black and Jepson[4]. Specifically, a state at time t is described as a parameter vector: $s^t = (\mu, \phi^i, \alpha^i, \rho^i)$ where: μ is the integer index of the predictive model, ϕ^i indicates the current position in the model, α^i refers to

An amplitude scaling factor and ρ^i is a scale factor in the time dimension. Note that i indicates which hand's motion trajectory this $\phi^i, \alpha^i, \text{ or } \rho^i$ refers to left and right hand where $i \in \{l, r\}$. My models contain data about the motion trajectory of both the left hand and the right hand; by allowing two sets of parameters, I allow the motion trajectory of the left hand to be scaled and shifted separately from the motion trajectory of the right hand (so, for example, ϕ^l refers to the current position in the model for the left hand's trajectory, while ϕ^r refers to the position in the model for the right hand's trajectory). In summary, there are 7 parameters that describe each state.

1) Initialization

The sample set is initialized with N samples distributed over possible starting states and each assigned a weight of $1/N$. Specifically, the initial parameters are picked uniformly according to:

$$\left. \begin{aligned} \mu &\in [1, \mu_{max}] \\ \phi^i &= \frac{1-\sqrt{y}}{\sqrt{y}}, y \in [0, 1] \\ \alpha^i &= [\alpha_{min}, \alpha_{max}] \\ \rho^i &\in [\rho^{min}, \rho^{max}] \end{aligned} \right\} \quad (1)$$

2) Prediction

In the prediction step, each parameter of a randomly sampled t is used to $t+1$ determine based on the parameters of that particular s_t . Each old state, s_t , is randomly chosen from the sample set, based on the weight of each sample. That is, the weight of each sample determines the probability of its being chosen. This is done efficiently by creating a cumulative probability table, choosing a uniform random number on $[0, 1]$, and then using binary search to pull out a sample (see Isard and

Blake for details[4]). The following equations are used to choose the new state:

$$\left. \begin{aligned} \mu_{t+1} &= \mu_t \\ \phi_{t+1}^i &= \phi_t^i + \rho_t^i + N(\sigma_\phi) \\ \alpha_{t+1}^i &= \alpha_t^i + N(\sigma_\alpha) \\ \rho_{t+1} &= \rho_t^i + N(\sigma_\rho) \end{aligned} \right\} \quad (2)$$

where $N(\sigma_*)$ refers to a number chosen randomly according to the normal distribution with standard deviation σ_* . This adds an element of uncertainty to each

prediction, which keeps the sample set diffuse enough to deal with noisy data. For a given drawn sample, predictions are generated until all of the parameters are within the accepted range. If, after, a set number of attempts it is still impossible to generate a valid prediction, a new sample is created according to the initialization procedure above.

3) Updating

After the Prediction step above, there exists a new set of N predicted samples which need to be assigned weights. The weight of each sample is a measure of its likelihood given the observed data $Z_t = (z_t, z_{t1}, \dots)$. We define $Z_{t,i} = (z_{t,i}, z_{(t-1),i}, \dots)$ as a sequence of observations for the i th coefficient over time; specifically, let $Z_{(t,1)}, Z_{(t,2)}, Z_{(t,3)}, Z_{(t,4)}$ be the sequence of observations of the horizontal velocity of the left hand, the vertical velocity of the left hand, the horizontal velocity of the right hand, and the vertical velocity of the right hand respectively. Extending Black and Jepsen [5], we then calculate the weight by the following equation:

$$p(z_t | s_t) = \prod_{i=1}^4 p(Z_{t,i} | s_t) \quad (3)$$

$$p(z_{t,i} | s_t) = \frac{1}{\sqrt{2\pi}} \exp \frac{-\sum_{j=0}^{\omega-1} (z_{(t-j),i} - \alpha^* m_{(\phi^* - \rho^* j),i}^\mu)}{2(\omega-1)} \quad (4)$$

where ω is the size of a temporal window that spans back in time. Note that ϕ^*, α^* and ρ^* refer to the appropriate parameters of the model for the blob in question and that $\alpha^* m_{(\phi^* - \rho^* j),i}^\mu$ refers to the value given to the i th coefficient of the model μ interpolated at time $\phi^* - \rho^* j$ and scaled by α^* .

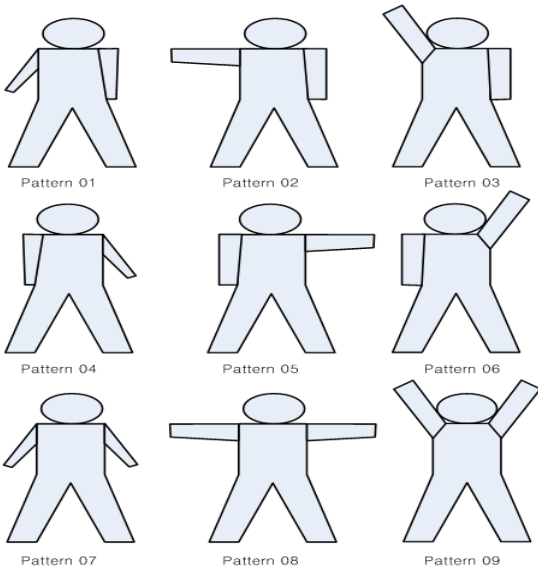


Fig.4. Nine hand motion model

III. EXPERIMENT RESULT

To test the proposed particle filter scheme, we used two gesture models which is shown in Fig. 4. The coefficient of particle filter is $\mu_{max} = 2, \alpha_{min} = 0.5, \alpha_{max} = 1.5, \rho_{min} = 0.5, \rho_{max} = 1.5$ to maintain the 50% scale. Also, the other parameters are $\sigma_\alpha = \sigma_\phi = \sigma_\rho = 0.1$. The value of ω in equation 3 is 10.

We first performed off-line experiments to evaluate the performance of the proposed dynamic gesture recognition method. We had two people perform each gesture eight times. Out of the 80 gesture model sequences, 40 were randomly chosen for training, 16 from each class. The remaining gesture sequences were used to test. The overall recognition rate is 97.6%. Following off-line experiments, we implemented a hand control interface based on dynamic gesture recognition in the real interview system. The control interface worked well in dynamic environments. Their detail decision process are shown in Fig. 5.

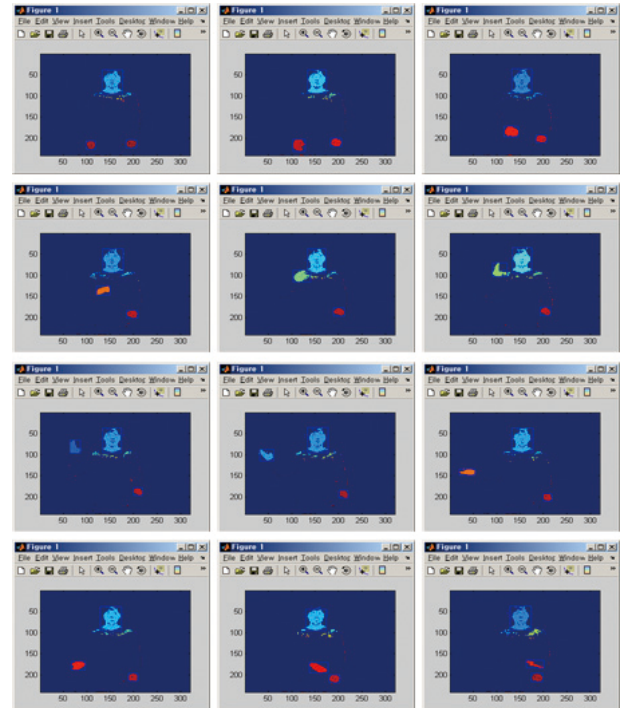


Fig.5. The Tracking process of particle filter for the model 2 (From left to right, top to down)

IV. CONCLUSIONS

Vision-based hand gesture recognition is an important technology for intelligent HCI. In this paper, we have developed the real time hand tracking and gesture recognition in the context of developing a hand control

interface for an interactive interview system. By applying particle filtering, we implemented a real time automatic interview system. Our approach produces reliable tracking while effectively handling rapid motion and distraction with roughly 75% fewer particles. We also present a simple but effective approach for dynamic gesture recognition. The hand control interface based on the proposed algorithms works well in dynamic environments of the virtual interview system. As only color and motion cues are used, our hand tracking algorithm in fact is a general method and could be used in many tracking problems. We are considering including other features, e.g., texture of hands, for more robust hand tracking. Although the proposed algorithm is effective for hand tracking, further investigation should be conducted to verify its effectiveness in other tracking problems, especially the higher dimensional problems such as 3D articulated object tracking, as the number of particles required in high dimensional space is more prohibitive.

REFERENCES

- [1] M.I.Sard and A.Blake: CONDENSATION-conditional density propagation for visual tracking. *International Journal of Computer Vision*,29(1) (1998) 5-28
- [2] Michael J. Black and Allan D. Jepson: A Probabilistic Framework for Matching Temporal Trajectories: Condensation-based Recognition of Gestures and Expressions. In *Proceedings 5th European Conf. Computer Vision*, Vol.1, (1998)909-924
- [3] Michael Isard and Andrew Blake: A mixed-state condensation tracker with automatic model-switching, In *Proceedings 6th International Conference of Computer Vision*, (1998) 107-112
- [4] Yang Weon Lee, Adaptive Data Association for Multi-target Tracking using relaxation, LNCS 35644, pp. 552-561,
- [5] Yang Weon Lee, J. H. Seo and J. G. Lee, A Study on the TWS Tracking Filter for Multi-Target Tracking, *Journal of KIEE*, Vol. 41 No. 4, pp.411-421, Apr. 2004. 2005.



Yang-Weon Lee, He Received B.S degree in the Dept. of Electrical Engineering from Chungang University, Korea, in 1982, and the M.S. degree in the Department of Control and Measurement Engineering from Seoul National University, in 1991. In 1996, he received Ph. D. degrees in the Dept. of Electrical Engineering from POSTECH, Korea. During 1982-1996, he worked for ADD as senior researcher. Since 1996, he has worked

in the Dept. of Information and Communication Engineering at the Honam University, where he now works as a professor. His current research interests include smart computing, target estimation and tracking filter design, communication signal processing, and .NET application.