Initialization of 3D Human Hand Model and Its Applications in Human Hand Tracking

Zhiquan Feng
School of Information Science and Engineering,
Provincial Key Laboratory for Network Based Intelligent Computing,
University of Jinan,
Jinan, 250022, P.R.China
e-mail: fzqwww@263.net

Bo Yang[1], Yanwei Zheng[1], Haokui Tang[1], Yi Li[2]
[1]School of Information Science and Engineering,
[2]School of Control Science and Engineering,
Provincial Key Laboratory for Network Based Intelligent Computing,
University of Jinan,
Jinan, 250022, P.R.China
e-mail: ise_fengzq@ujn.edu.cn

***Abstract*—Initialization of 3D human hand gesture is one of the fundamental and key steps in the study of 3D hand tracking, and a novel approach to initialize 3D human hand gesture is put forward in this paper. This paper will cover the following points. First, a new approach to selecting a human hand gesture from the hand postures database is presented. Second, both techniques of visualization and human-computer interaction are used into the initialization process, through which the 3D human hand model is fine-tuned time after time until the required accuracy is satisfied. Lastly, the proposed initialization method is applied to 3D human hand tracking system based on PF (Particle Filtering), with real video data under complex background. In order to address high dimensional problem of 3D hand structures, a new concept, which is called key factor in this paper, is introduced to guide sampling; to improve robustness to changing light conditions, a new skin model is proposed. The main contributions of this paper include: (1) combine the techniques of the interaction between operator and computer and the visualization to achieve initialization of 3D human hand model. (2) use the key factors to address the issue of high dimensionality of 3D hand model, and (3) a new hand skin model is presented, as well as self-occlusion problem is effective addressed. Our experimental results show that the proposed approach is not only fast, accurate and robust, but also direct, natural and convenient for operators to handle.**

***Index Terms*—3D Freehand pose model, features extraction, initialization, visualization, cognitive behavioral models**

I. INTRODUCTION

3D nature human hand tracking based on computer vision will provide a new human-computer interaction model (HCI) for virtual reality (VR), which makes the HCI more direct, natural and harmonious, and so research on human hand tracking has attracted a great deal of attentions from all over the world [1-4]. Motivated by the application of human-computer interface in AR and VE, bare hand tracking requires a reliable initial pose in the first frame.

Of all the recursive and model-based human hand tracking approaches, initialization of 3D human hand gesture is a necessary step, but it is also a troublesome problem. Of all the automatic methods, single frame pose estimation approach is a basic way to fulfill initialization of 3D human hand gesture.

As early as the year of 2001, Triesch and Malsurg [5] proposed an EGM (Elastic Graph Matching) method to find out the best hand gesture of candidate model. In this method, Gabor jets, color Gabor jets and skin color are fused into EGM. N. Shimada et al [6] proposed a method to estimate arbitrary 3D human hand postures in real-time. It can accept not only pre-determined hand signs but also arbitrary postures and it works in a monocular camera environment. The estimation is based on 2D image retrieval. More than 16,000 possible hand appearances are originated from a given 3D shape model by rotating model joints and stored in an appearance database. Every appearance is tagged with its own joint angles which are used when the appearance was generated. By retrieving the appearance in the database well-matching to the input image contour, the joint angles of the input shape can be rapidly obtained. The search area was reduced by using an adjacency map in the database. To prevent tracking failures, a fixed number of the well-matching appearances are saved at every frame. Rómer Rosales [7] proposed Specialized Mappings Architecture (SMA) approach to map image features to likely 3D hand poses. It was based on the idea of learning a mapping from a feature space to the parameter space. Rotation and scale invariant moments of the hand silhouette were utilized to implement the mapping by employing a machine learning architecture. The SMA's fundamental components are a set of specialized forward mapping functions, and a single feedback matching function. The joint angle data in the

Project number: 60973093, 60873089.

doi:10.4304/jcp.7.2.419-426

training set is obtained via a CyberGlove, a glove with 22 sensors that monitor the angular motions of the palm and fingers. In training, the visual features are generated by using a computer graphics module that renders the hand from arbitrary viewpoints given the 22 joint angles. The viewpoint is encoded by two real values, therefore 24 real values represent a hand pose. The system can automatically detect and track both hands of the user, calculate the appropriate features, and estimates the 3D hand joint angles and viewpoint from those features. Generally speaking, SMA provides continuous pose estimates through regression.

Athitsos and Sclaroff [8] used chamfer distance, edge orientation histogram and moment to retrieve in segmented human hand images database. The basic idea is to estimate 3D hand shape and orientation by retrieving appearance-based matches from a large database of synthetic views. The hand shape in the input image is assumed to be close to one of 26 predefined shapes. The data base views are computer generated renderings of the 26 hand shape prototypes from viewpoints that was sampled uniformly along the surface of the viewing sphere.

In the year of 2003, Stenger [9] proposed a tree-based representation of the distribution, where the leaves define a partition of the state space with piecewise constant density. Their experimental results demonstrated the effectiveness of the technique by using it for tracking 3D articulated and non-rigid motion.

The single frame pose estimation approach is based on a local search and keeps track of only the best estimation at each frame. This type of tracker is expected to work well at the initialization phase, because no more history information can be used. One of the distinct features of single frame pose estimation approach is to retrieve hand gestures from hand images database. The advantage of using appearance-based matching for 3D parameter estimation is that the estimation is done indirectly, by looking up the ground truth labels of the retrieved synthetic views. This way avoids the ill-posed problem of recovering depth information directly from the input image.

Martinde LaGorce, et al [10] assumed that the hand is parallel to the image plane at initialization and linear constraints are defined on relative length of the parts within each finger. Furthermore, they supposed the hand color to be constant across the surface.

However, it is clear that in most of the studies, initialization state of a hand gesture is manually complemented [11-12].

To summarize, the hand pose tracker requires a reliable initial gesture in the first frame. Estimating the hand pose in a single frame without strong prior information of the hand pose is challenging. Unfortunately, initialization has not been paid enough attention by now and there are few related or specialized papers available [1].

Recently, we proposed a 3D-freehand-pose initialization method based on operator's cognitive behavior models [13]. This approach is composed of three phases, hand pose recognition, coarse-tune and fine-tune. The objective of hand pose recognition is to obtain an approximate 3D hand model. The operator moves his/her hand onto the projection of the 3D hand model in the period of coarse-tune time. Then, the computer repeatedly fine-tunes the 3D hand model until the projection of the 3D hand model is completely superposed onto the operator's hand image. we focused on exploring and modeling cognitive behavior of operator's hand. Our research shows that cognitive behavior models are not only beneficial to reducing cognitive loads for operators because it make the computers cater for the changes of the operators' hand poses, but also helpful to address high dimensionality of articulated 3D hand model. However, this study exists several issues. (1) dimensionality reduction of 3D hand model depends on the extracted cognitive behavioral model. (2) self-occlusion problem is not taken into consideration.

II. OVERVIEW

There are many reasons for the difficulties of 3D initialization of hand gesture. Firstly, to recover 3D structure from a single hand image is an open problem in the field of computer vision. Secondly, human hand is a typical articulated object with high dimensionality, and to find out the real 3D hand model from nearly unlimited hand gestures is almost impossible unless approximate methods are used, just as most of the single frame pose estimation approaches do. Thirdly, how to acquire measurements or image features automatically and robustly from frame images sequence is still a problem to be further researched. In order to overcome the above difficulties, we use indexing techniques to retrieve the nearest neighbors of a given input, which include the following key points. (1) the correspondence between 2D hand images and pre-determined 3D hand models, and (2) the invariable features of hand images for scale and rotation transformation. On the another hand, it should be noticed that in the process of initialization, the hand gestures of the operator are little possible to be the same as the models in the database, because the hands of the operator, the distance to the camera, the positions and gestures, as well as the lighting conditions, are not likely to be in the same situation when human images were sampled. In order to address this problem, we regard the 3D models in the database just as the start state from which search process will set out to look for the so-called real solutions, which is regarded as the same as the real positions and gestures of the operator.

At the same time, both techniques of human-computer interaction and visualization are introduced into initialization process. The operator continuously adjusts his/her hand gestures and positions according to the cues which are visualized by graphics and images; then the computer modifies the 3D hand models according to the identified gestures and positions. Fig.1 shows the main idea of our initialization method.

III. HUMAN HAND SKIN MODEL

Skin color provides a relatively simple method to determine the probability whether the pixel belongs to an image of a hand, and skin color model impacts the accuracy of features extraction of hand images.

In order to improve the robustness to light change, a novel skin model based on skin luminance distribution is put forward in this paper. First of all, a great deal of skin samples under different light conditions are drawn,

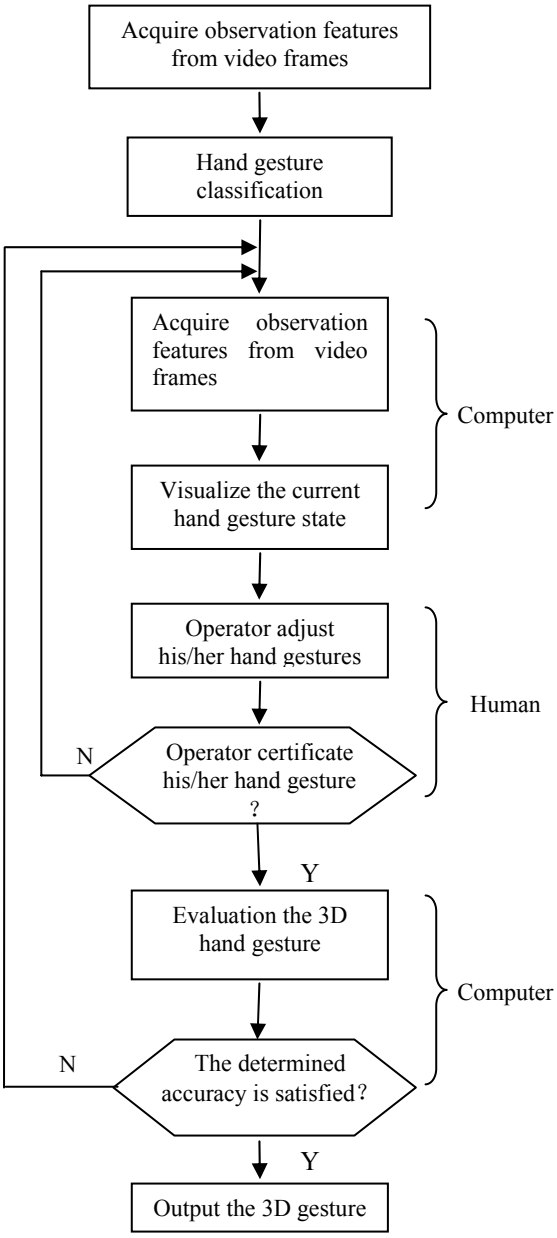


Figure 1. The global frame of our 3D model initialization approach.

forming a skin samples database. Then, compute skin luminance distribution. Thirdly, skin samples are grouped according to luminance, building index for each group. At last, compute features in each group.

By tries and errors, we found the following hand skin features can obtain the best results to robustly segment human hand from complex background. The proposed skin hand model is shown in Fig. 2.

$$f_1 = (R+G+B)/3 \tag{1}$$

$$f_2 = R - B \tag{2}$$

$$f_3 = (2G - R - B)/2 \tag{3}$$

$$f_4 = \alpha_1 (B - Y) \tag{4}$$

$$f_5 = \alpha_2 (R - Y) \tag{5}$$

the vector $\boldsymbol{F}$=(f_1, f_2, f_3, f_4, f_5) is the human hand skin model used in this paper. In the formulae (1)-(4), α_1=0.492, α_2=0.877, and Y is the luminance value of the color (R, G, B).

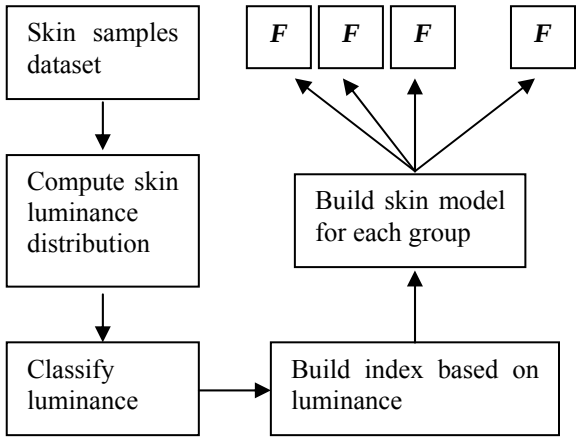


Figure 2. The hand skin model proposed in this paper.

IV. Human Hand Gesture Classification Based on Invariable Features

Invariable features are widely used in order to keep recognition correct and stable under the conditions of rotation, translation and scale of the operator's hands.

Our hand gesture classification approach based on invariable features is presented as follows.

a) Acquire features of human hand images based on the human *hand skin model.*

b) Acquire the centroid of the frame image.

c) Compute the distances from each feature to the centroid.

d) For all features find out the maximum distance Dmax to the centroid.

e) The circumcircle, with the centroid as the centre of the circle and Dmax as the radius, is divided into M concentric cirque with equal distance between the adjecent concentric circle, producing M subarea.

f) Compute density

$$r_k = \frac{S_k}{S_{\max}} \tag{6}$$

with the probability p_k where k is 1, [M/2] and M respectively.

g) Normalize all r_k, and form a density distribution features vector.

In the formula (6), S_k is the number of pixels of hand skin in the kth subarea, and $S_{\max}$ is the max number of pixels of hand skin in the all subareas.

The proposed density distribution features are invariant to translation, scale and rotation, and is faster than the methods proposed by Huang Chunmu [14].

We employ a fast feature extraction method [15] to acquire hand image features. First of all, the hand contour is approximately described by a polygon with concave and protruding, and the relationship between hand gesture polygon and its bounding box is studied. Secondly, a hand gesture contour algorithm proposed to obtain the hand

gesture polygon. Then, based on the hand gesture contour algorithm, the approaches to gain the several main feature points are presented, including fingertips, roots of fingers, joints and the intersection of knuckle on different fingers.

V. Evaluation and Sampling Based on Key Factors

Suppose the Hausdorff distance [16] between the set A and the set B is H, and H is also the distance between $i \in A$ and $j \in B$, then the point i and the point j are called key factors. In fact, according to the definition of Hausdorff distance, within some scopes, some of the elements in A and B do not affect the Hausdorff distance, the Hausdorff distance is determined by the two elements, one is in set A and another is in set *B*.

A. Sampling for Particle Filtering (PF)

The key factors can be easily acquired. Of the hand gesture vector, the variables impacting the key factors can be identified, upon which sampling is performed.

For example, in Fig. 3, suppose the hand gesture vector ***X*** has 5 variables which have something to do with Hausdorff distance, then these variables belong to key factors, and they need to be sampled while the others are not need sampled.

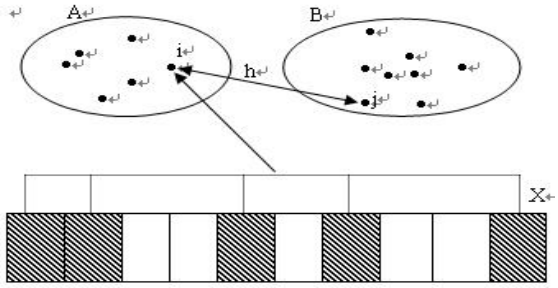


Figure 3. The relationship between key factors and variables needed be sampled in our algorithm. The variables with shadow in the vector X are corresponding to the key factor i.

Suppose at time *t*, the frame image is *I(t)*, the 3D hand gesture is ***M(t)***, the projection of ***M(t)*** onto *I(t)* is ***P(t)***, and Hausdorff distance between *I(t)* and *P(t)* is *h(i, j)*. The sampling algorithm is described as follows.

a) Determine the key factors by

$$h(i, j)=\text{Hausdorff}(I(t), P(t)) \tag{7}$$

b) Determine the selected variables in the hand gesture vector by means of camera projection model, acquiring those variables dependent on the key factors.

c) Sampling is taken for the selected variables.

For simplification, the state distribution at time t is replaced by that at time t-1, and take the assumption that Gauessian distribution is satisfied, or

$$(\mu_t^x, \delta_t^x) \rightarrow (\mu_{t-1}^x, \delta_{t-1}^x) \tag{8}$$

B. Address Self-occlusion Problem

Self-occlusion is an ubiquitous property, as well as an intrinsic visual property of human hand. Essentially, the problem of self-occlusion judge is boiled down to finding intersections among sphere, column and cuboid. The occluded joints will not be projected onto the frame image, therefore, they are not concerned with Hausdorff distance and effectively avoid the influence of self-occlusion on 3D model.

Fig. 4 shows us the method to judge whether or not a joint A on a knuckle is occluded by other knuckle, and the original point O is the viewpoint. If exist the intersection B or C, then the joint A is occluded.

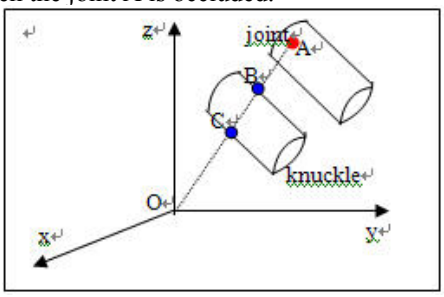


Figure 4. The way to judge if the joint A is occluded by the other knuckle, O is the viewpoint, OA is the line of sight and B and C are the possible intersections of the line of sight and the knuckle on a finger.

The values of occluded parts in 3D hand gesture are reasonably assigned in this paper.

VI. Integration Techniques of HCI and Visualization

In the age of ubiquitous computing, how to make computers adapt human or how to make human-centered rather than computer-centered is one of the key problems in human computer intelligent interaction or perceptual user interface design, which is also one of the objectives in this paper.

On one hand, in the process of initialization, the operators adjust their hand's position or gestures to make the real-time human hand images be superposed while keeping their hand 3D models fixed.

On the other hand, after the operators identify their hand states, computer begins to fine-tune the 3D hand gesture until the error between the human image and the projection of the 3D hand gesture is in the required scope while the operators' hands keep fixed.

The above two phases go by turns between human and computer until the operators feel satisfied at their hands status.

This human-computer interaction to some extent depends on the way of feedback and effectiveness of state visualization. In our study, several approaches are used for visualization. For example, the 3D hand model is rendered and outputted by means of OpenGL, the newly hand images are displayed in real time with virtual style, and the key factors are revealed with an intelligible way.

The next remaining problem is the way to determine whether or not the operators certificate their hands' states. A protocol based on context is proposed in our study. If the new hand images keep unchanged within fixed time, it is believed that the operators identify their hands' situations.

VII. Our Algorithm

A. Our Algorithm

Our algorithm is described as follows:

For computer:

a) recognize hand posture and obtain a initial 3D hand model.

b) extract hand image features from the current frame image.

c) compute the Hausdorff distance between the frame image features and the projection onto the current image plane of the 3D hand model.

d) compute the key factors.

e) feed the key factors back to the operator with information visualization.

For operator:

a) cognize the feedbacks on screen.

b) adjust the position and posture of hand.

The above algorithm is an interactive process between the operator and the computer; the algorithm ends with a special feedback on screen. When the Hausdorff distance is small than some designated threshold, computer will feed a special mark back to the operator.

B. Characteristics of Our Method

Our algorithm is featured with following aspects.

a) Astringency

Based on evaluation of the key factors, a Hausdorff distance series $\{H^k|, k=1,2,\ldots\ldots\}$ is obtained and it satisfies

$$0 \le H^{k+1} \le H^k \quad (9)$$

and so, after several iterative process, the real solution can be reached.

b) Interaction

In our algorithm, the computer and human are equivalent in position. The operators do not wait for the computer to search the real 3D model, but actively adjust their hands gestures and positions, making the initialization process interesting and funny.

c) Visualization

In our algorithm, temporary data and feedback information are visualized by means of graphics or images, and the protocol based on context is used as the turning point between human and computer, making HCI feasible, amusing, natural and convenient.

d) Pertinency

According to N. Vaswani [17], most state changes occur in a small number of dimensions, while the change in the rest of the state space is small. As a result, only those states that are extremely needed to be modified will be sampled with Gauessian distribution based on key factors analysis, providing a novel and effective approach to address high dimensionality problem.

VIII. Experimental Results

A calibrated gun-style camera is connected with our computer via a color data image collection card, DH-CG410. Frame images are inputted into the computer frame after frame, and initialized and rendered by OpenGL and outputted onto the screen. The basic configuration of our PC computer is pentium 4 with 3.00 GHz CPU and 512M memory. We employ our previous 3D hand model [13]. The experimental platform is illustrated in Fig. 5 and the on-line experimental scene is shown in Fig. 6.

The experiments of our method and the referenced way, PF are performed for different human hand gestures.

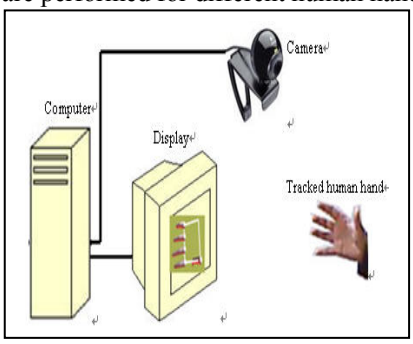


Figure 5. The experimental hardware platform. Only one camera is used in our system.

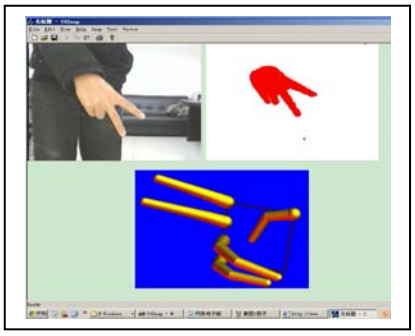


Figure 6. The on-line experimental scene. The top-left is the frame image, the top-right is the segmented human hand and the bottom is the tracked 3D model. This is gesture classification process.

A. Experimental Process and Results

a) For cloth-wrapper hand gesture

First of all, the segmentation based on our skin color model is conducted and is shown in Fig. 7. The procedures of initialization by our method of cloth-wrapper hand gesture are demonstrated in detail in Fig. 8.

In Fig. 8, the operator can see that his hand is above the 3D model, so he actively adjusted his hand towards the model, fitting the 3D model. At the moment when his hand kept unchanged within fixed time, the computer identified the operator's suggestion and started to evaluate the difference between the situation of the operator's hand and the 3D hand model. Subsequently, turn the computer to modulate the 3D hand model. In order to enhance effectiveness of visualization, the color of visualized graphics is changed at the turning point between human and computer.

b) Scissors hand gesture

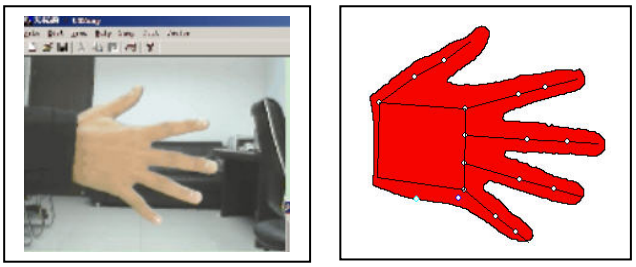
Another example is about the initialization of scissors hand gesture, which is demonstrated in Fig. 9. The thumb is occluded in this example, and the state of the thumb and other occluded fingers in the 3D hand model are properly assigned the value that is the same as that in the last frame.

B. Accuracy and Time Cost

We compare the proposed method in this paper with our previous approach [13]. For different methods, we experimented with cloth-wrapper hand gesture, scissors hand gesture and fist gesture, each of the gestures are

repeated 10 times. The average of the three postures are shown in Fig. 10.

It is demonstrated from Fig. 10 that the performance of our method is improved compared with our previous method. The smaller the Hausdorff distance is, the higher the accuracy is.



(a) (b)

Figure 7. The skin segmentation. (a) The raw frame image. (b) The segmented human hand from chaotic background and the projection onto the image of 3D model by hand.

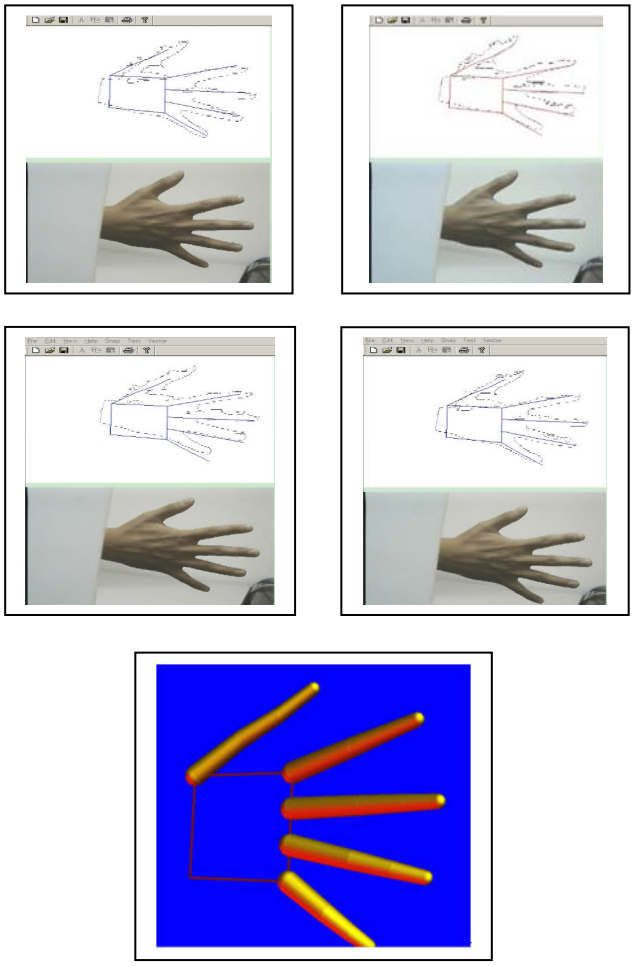


Figure 8. The process of initialization for the cloth-wrapper hand gesture, in which the dashed lines stand for the new positions of human hands, the real lines stand for the temporarily identified positions, and the blue and real lines stand for the projections of adjusted 3D hand gesture. The last row is the initialized 3D hand gesture.

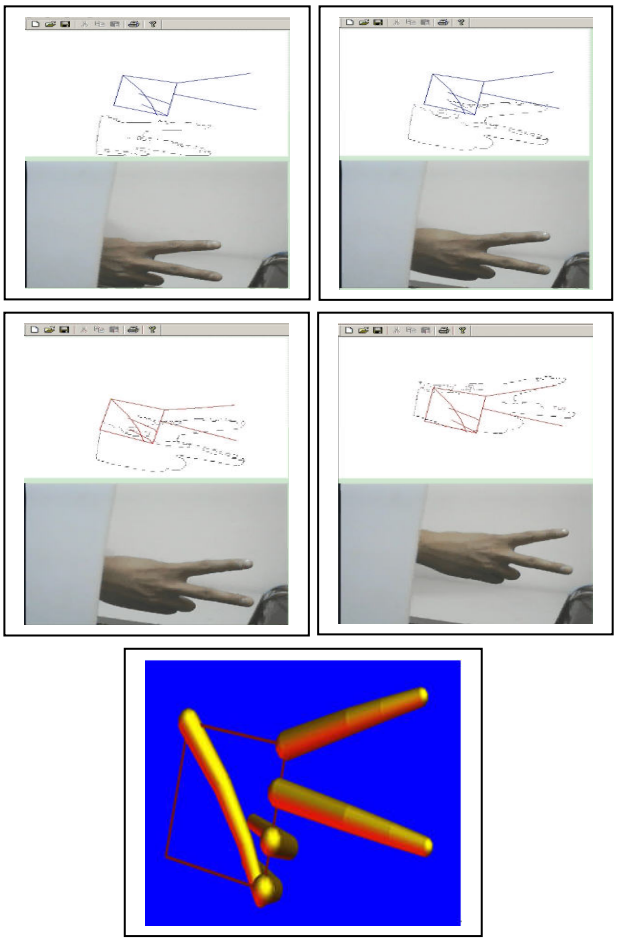


Figure 9. The process of initialization for the scissors hand gesture. The last row is the initialized 3D human hand model. Self-occlusion problem can be effectively addressed.

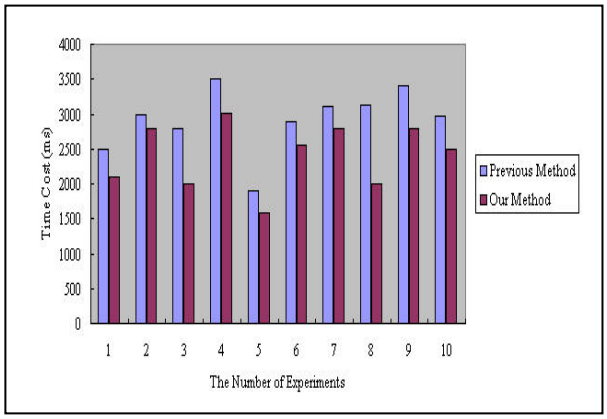
C. Analysis to Experimental Results

Our initialization system benefits much from visualization. Because of the use of visualization, it is possible that human–computer interaction is carried through harmoniously, and that the operators can observe intuitionistically the directions and ranges their hands should be adjusted, no matter how far the distance between the objective position and the current position is. Experimental results show that PF may not acquire the accuracy that our algorithm obtains, or that PF maybe far from the real solution.

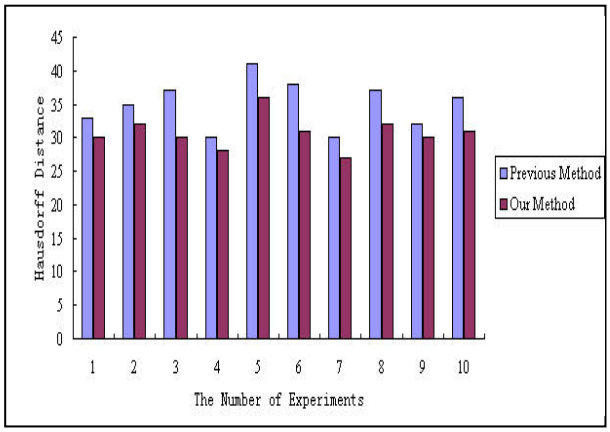
The time cost of our algorithm depends mainly on both the speed of hand gesture segmentation and of being adjusted by the operator, while the time cost of PF is mostly impacted on the speed of hand gesture segmentation and the number of the particles used. We notice the fact that with our method the time cost on adjusting human hand image can be further improved by training and learning, but with PF method to increase the number of used particles is limited by required time because the time cost increases exponentially with the particle number.

D. Applications in Human Hand Tracking

At last, we apply the proposed method into our 3D hand tracking system, some of the tracked scenes are shown in Fig. 11. The average time of processing a frame is 180ms on our PC computer.



(a)



(b)

Figure 10. Comparison of our method with previous method [13]. (a) average time cost. (b) average Hausdorff distance.

IX. Conclusion and Future Work

In the first frame, a prediction is not available, therefore, a separate initialization procedure is needed, which is evidently lagged in the research of 3D human hand tracking. A novel algorithm for initializing human 3D human hand model is presented in this paper. The typical techniques and characteristics, such as astringency, interaction, visualization and pertinency, are fused into our methods. Our visualization system has been implemented with VC++6.0, and many experimental results demonstrate that our system is more feasible, amusing, natural and convenient and robust than PF approach.

Compared with our previous Feng [13], the main contributions of this paper: (1) combine the techniques of the interaction between operator and computer and the visualization to achieve initialization of 3D human hand model. (2) use the key factors to address the issue of high dimensionality of 3D hand model, and (3) a new hand skin model is presented, as well as self-occlusion problem is effective addressed.

There are still some problems to be further researched. First of all, to build a huge hand gestures database is a toilsome work and it takes a great deal of time. Second, human hand skin recognition is impacted when lighting condition greatly changed. These problems should be intensively studied in our future work.

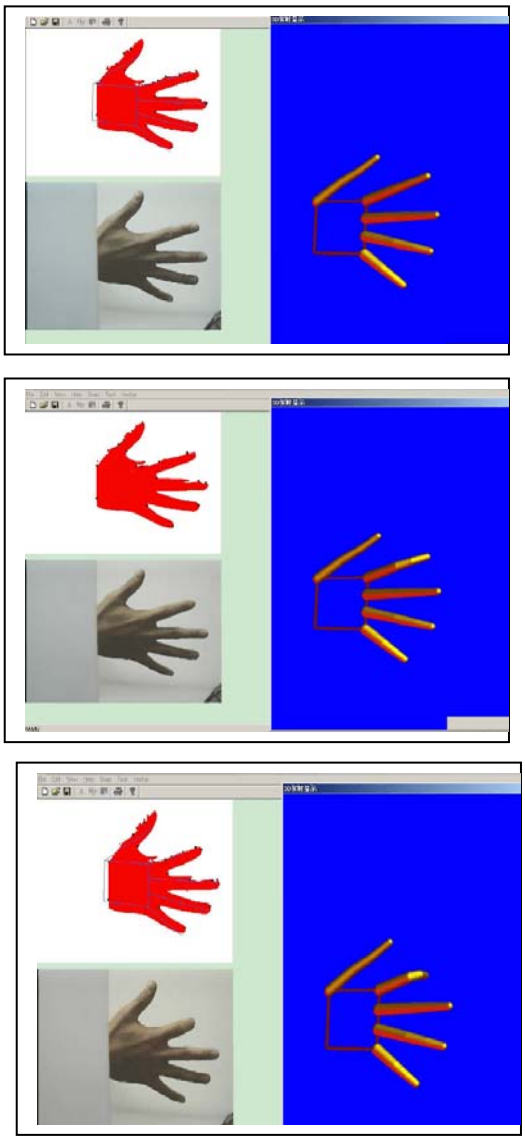


Figure 11. The experimental results of 3D human hand tracking using real video data. In this experiment, the PF method is used to the tracking system, in which the initialization of 3D human hand model is generated by our method. The tracking with 26 DOFs are conducted in this experiment.

Acknowledgment

This paper is supported by NSFC (No.60973093), Natural Science Foundation of Shandong Province (Y2007G39), Natural Science Foundation for Distinguished Youth Scholar of Shandong Province (No.JQ200820), Key Project of Natural Science Foundation of Shandong Province (2006G03), Science and Technology Plan of Shandong Province Education Department (J07YJ18).

References

[1] Ali Erol et al,. Vision-based hand pose estimation: A review. Computer Vision and Image Understanding. 2007, 108: pp. 52–73.

[2] Julier SJ,Uhlmann JK.A new extension of the Kalman filter to nonlinear systems[A].The Proco AeroSense:11th Int. Symposium Aerospace/Defense Sensing, Simulation and Controls[C].Orlando, 1997, pp. 54-65.

[3] Cui JinShi. Studies on Three-Dimensional Model Based Posture Estimation and Tracking of Articulated Objects[D]. Ph.D. Thesis, Beijing: Tsinghua University, China, 2004
[4] Ali Erol, George Bebis, Mircea Nicolescu, Richard Boyle, Xander Twombly. A Review on Vision-Based Full DOF Hand Motion Estimation, Proceedings of the IEEE Workshop on Vision for Human-Computer Interaction (V4HCI), San Diego, California, 2005, 3: pp. 75.
[5] J. Triesch and C. von der Malsurg, "A system for person-independent hand posture recognition against complex background," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 23, No. 12, Dec. 2001.
[6] N. Shimada, K. Kimura, and Y. Shirai, "Real-time 3-D hand posture estimation based on 2-D appearance retrieval using monocular camera," Proc. Int. Workshop RATFG-RTS, 2001, pp. 23-30.
[7] Rómer Rosales.The Specialized Mappings Architecture with Applications to Vision-Based Estimation of Articulated Body Pose. Ph.D. Thesis. BOSTON University graduate school of arts and sciences, 2002.
[8] V. Athitsos and S. Sclaroff, "Estimating 3D Hand Pose from a Cluttered Image," Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, 2003, 2: pp. 432-439, ,.
[9] B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla, "Filtering using a tree-based estimator," Proc. IEEE Int. Conf. Computer Vision, 2003, 2: pp. 1063-1070.
[10] Martinde LaGorce, NikosParagios, DavidJ.Fleet. Model Based Hand Tracking with Texture, Shading and Self-occlusions. IEEE Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp.1-8.
[11] Zhiquan Feng, Bo Yang, Yuehui Chen , Yanwei Zheng , Yi Li , and Zhonghua Wang. "Research on Sampling Methods in Particle Filtering Based upon Microstructure of State Variable". Lecture Notes in Computer Science, 2008, 5226: pp. 251-257,
[12] Zhiquan Feng, Bo Yang, Yanwei Zheng, Zhonghua Wang, Yi Li. "Research on 3D Hand Tracking Using Particle Filtering. The 4th International Conference on Natural Computation. New York: IEEE Computer Society Press,2008, 7: pp. 367-371.
[13] Zhiquan Feng, Minming Zhang, Zhigen Pan, Bo Yang, Tao Xu, Haokui Tang, Yi Li. 3D-Freehand-Pose Initialization Based on Operator's Cognitive Behavior Models. The Visual Computer. 2010, 26: pp. 607- 617.
[14] Huang Chunmu, Zhou lili. "Density D istribution Feature and its Application in Binary Image Retrieval". Journalof Image and Graphics, 2008, 13(2) : pp. 307-311
[15] Zhiquan Feng, Peilin Du, Xiaona Song, Ling Zhang, Zhengxiang Chen. "Research on Features Extraction From Frame Image Sequences". 2008 International Symposium on Computer Science and Computational Technology (ISCSCT2008). 2008: 762-766, New York: IEEE Computer Society Press.
[16] Huttenlocher D. P., Klanderman G. A., Rucklidge W. J.. Comparing Images Using the Hausdorff Distance. IEEE Trans. Pattern Analysis and Machine Intelligence, 1993, 15(9): pp. 850-863.
[17] N. Vaswani. "Particle Filtering For Large Dimensional State Spaces with Multimodal Observation Likelihoods". IEEE Transactions on Signal Processing, 2008, 56(10): pp. 4583-4597.

Zhiquan Feng is a professor of School of Information Science and Engineering, Jinan University. He got the Master degree from northwestern polytechnical university, china in 1995, and PhD degree from Computer Science and Engineering Department, shandong university in 2006. He has published more than 50 papers on international journals, national journals, and conferences in recent years. His research interests include: human hand tracking/recognition/interaction, virtual reality, human-computer interaction and image processing.

Bo Yang is a professor and vice-president of University of Jinan, Jinan, China. He is the Director of the Provincial Key Laboratory for Network-based Intelligent Computing and also acts as the Associate Director of Shandong Computer Federation, and Member of the Technical Committee of Intelligent Control of Chinese Association of Automation. His main research interests include computer networks, artificial intelligence, machine learning, knowledge discovery, and data mining. He has published numerous papers and gotten some of important scientific awards in this area.

Yanwei Zheng, born in 1977, is a Lecturer of University of Jinan, Jinan, China. His main research interests include image processing.

Haokui Tang is a Lecturer of University of Jinan, Jinan, China. His main research interests include image processing.

Yi Li, born in 1961, Associate professor and master supervisor. Her main research interests include information control.