13—29

The Scalable Model-Based Hand Posture Analysis System

Cheng-Chang Lien[1]
Department of Computer Science and Information Engineering
Chung Hua University, Taiwan, R.O.C.

Abstract

Recently, the model-based hand posture analysis researches have become increasingly popular for many applications, such as hand gesture recognition, intelligent human-machine interface, virtual reality, and computer animation. However, the main problems that make the conventional marker-based hand posture analysis system inapplicable are inefficient calculation of the inverse kinematics, unable to scale the size of the hand model, and unable to analyze the hand posture when some markers are occluded. In this paper, a scalable hand posture analysis system is proposed to overcome these problems.

1. Introduction

The model-based methods can be categorized into hand without markers [1-3] and hand with markers [4-6] methods. Most system without markers require the hand to be placed in certain pose and position, and the projection of 3-D model will lose the depth information so that it cannot guarantee the accuracy of the estimated joint angles. With the 3-D positions of the markers, the methods in second category [5-7] track the hand motion with 27 DOF precisely. The marker-based hand posture analysis systems in [5,6] have not considered the scale calibration and the marker occlusion problems. The scale inconsistency between the 3-D hand model and the real hand will cause the wrong identification of the hand posture, i.e., the inaccuracy estimation of the joint angles for each finger. In order to improve the scale inconsistency problem and generalize the application of the marker-based hand posture analysis system, a scale calibration method that applies the scalable closed form solution functions to efficiently calibrate the scale of the hand model is proposed. For the marker occlusion problem, the path coherence function is applied to predict the 3-D positions of the occluded markers.

2. The Generation of the Scalable Inverse Kinematics Solution Functions

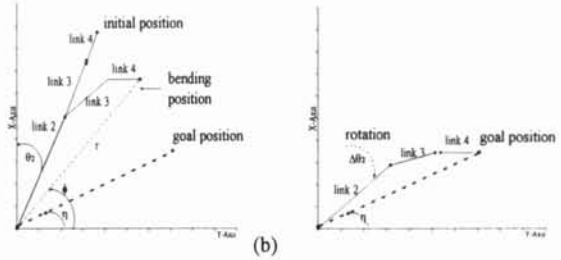
2.1 The Kinematics of the Hand Model

There exist 27 DOF in the hand, including three DOF for wrist translation and three DOF for wrist rotation. According to the D-H rules [7], the local coordinate system [6] is defined for each finger.

2.2 The Scalable Solution Function for Finger Inverse Kinematics

In the finger motion analysis process [6], the following two steps are applied to make the fingertip reach the goal position (r, η).

1. Bend the finger along a certain ϕ angle to the desired r distance.(see Fig.1 (a))
2. Rotate the link 2(MP joint) of the finger from angle ϕ to the angle η. (see Fig.1 (b))



(a) (b)

Fig. 1. The two steps that are applied to make the fingertip reach the goal position (r, η).

Based on the separable finger movements, the closed form solutions [6] for the inverse kinematics are generated. However, these closed form solutions are not scalable. Here, the scaling factor s is considered to generate the scalable closed form solutions for the inverse kinematics. Therefore, the first step may be modified to describe the relationship between the distance r, scaling variable s, and the joint angles (θ_2, θ_3) when the finger is bending along a certain ϕ angle to the desired r distance. Since the bending movements of the joint angles θ_2 and θ_3 are independent, two solution functions: $r=D_1(s,\theta_3)$ (see Fig, 2(a)) and $r=D_2(s,\theta_2)$ (see Fig, 2(b)) are defined to describe the relationship between the distance r and (s,θ_2,θ_3). Because $\theta_4=2/3\,\theta_3$ [5], the DOF of each finger can be reduced to 2. The second step describes the rotation movement of the link 2 (MP joint) from angle ϕ to the angle η and here, a linear function $\Delta\theta_2=f(\eta)$ is applied to describe the rotation movement of the MP joint. The first function that can be used to analyze the relationship among r, s, and θ_3 for a certain elevation angle ϕ is

$$r = D_1(s,\theta_3) = a_1 s^2 + b_1 s\theta_3 + c_1\theta_3^2 + d_1 s + e_1\theta_3 + f_1. \qquad (1)$$

The second function that can be used to analyze the

[1] Address: 30 Tung Shiang, Hsin Chu, Taiwan, 30012, R.O.C. Email: cclien@chu.edu.tw

relationship among r, s, and θ_2 for a certain elevation angle ϕ is

$$r = D_2(s,\theta_2) = a_2 s^2 + b_2 s\theta_2 + c_2\theta_2^2 + d_2 s + e_2\theta_2 + f_2 . \quad (2)$$

The value of θ_2 obtained from equation (2) is an interim value of MP joint angle and the correct value is found from the third function. The third function $\Delta\theta_2 = f(\eta)$ describes the rotation movement of the MP joint (see Fig. 1(b)) for the slant angle rotating from angle ϕ to the angle η, and it is defined as

$$\Delta\theta_2 = f(\eta) = \phi - \eta . \quad (3)$$

Finally, the correct value of the MP joint angle is $\theta_2 + \Delta\theta_2$. Here, the total least-square regression method is used to find the coefficients of Eq. (1) and Eq. (2).

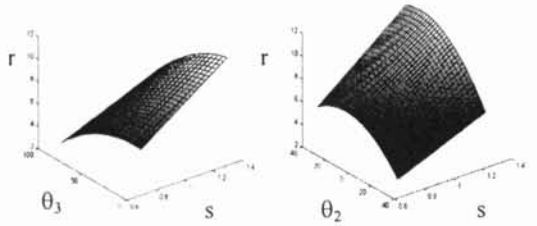


Fig. 2. Two closed form solution functions: (a) $r=D_1(s,\theta_3)$ and (b) $r=D_2(s,\theta_2)$.

2.3 The Bounding Surface for the Scalable Closed-Form Solution Functions

The variation range of r and η for various scales of the hand are specified by the second order conical surfaces $\eta = B_1(s, r)$ and $\eta = B_2(s, r)$.

3. The Hand Model Fitting

The hand model fitting process uses the alignment measure [8] for the wrist fitting process and the closed form solutions for the finger inverse-kinematics [8]. However, the scale calibration is not considered in the hand model fitting method developed in [8]. In order to improve the scale inconsistency problem and generalize the application of the marker-based hand posture analysis system, a new scale calibration method is proposed to efficiently calibrate the scale of the hand model.

3.1 The Scale Calibration of the 3-D Hand Model

In the scale calibration system, the 3-D positions of the seven markers are used to calibrate the scale of the 3-D hand model. In the calibration process, an image sequence (Fig. 2) of hand grasping movements is applied to calibrate the scale of the hand model. In the image sequence, the finger movements is carefully observed and then some measurements is applied to calibrate the length of each link in the 3-D hand model accurately. The measurements are (1) the fitting error, (2) the penalty for violating the constraints of finger motion, and (3) the penalty for violating the constraints of the length ratio between the fingers.

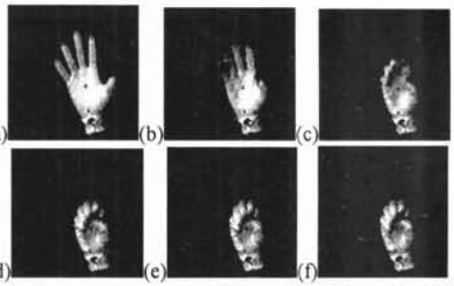


Fig. 2. The image sequence for calibrating the 3-D hand model.

To reduce the computation complexity for the scale calibration process, the links of all fingers are separated into two segments (l_{i1}, l_{i2}, i is the finger index), and the links in each segment are scaled concurrently with the same scaling factor. The first segment of each finger comprises the links below the MP joint and the second segment consists of the links above the MP joint. Two scaling factors s_{i1} and s_{i2} are used to scale the length of two segments for each finger. Based on the these measurements, a scale calibration measure function $\boldsymbol{C}(\cdot)$ which consists of three functions measuring the calibration error is defined as

$$C(s_{i1}, s_{i2}) = w_1 E_{fit}(s_{i1}, s_{i2}) + w_2 E_{mov}(s_{i1}, s_{i2}) + w_3 E_{rat}(s_{i1}, s_{i2}). \quad (4)$$

a) The fitting error measure function E_{fit}. The fitting error between the scale fitted hand model and the real hand is analyzed by the following formula:

$$E_{fit} = \sum_{i=1}^{5} d_i \Big/ \sum_{i=1}^{5} l_i ,$$

where d_i is the distance from the tip position of finger i of the fitted hand model to the tip position of the same finger i of the real hand, and l_i is the length of the finger i.

b) The movement measure function E_{mov}. The function E_{mov} measures the bending movements of a finger exceeding its movement constraints. Inconsistent scale of the hand model, especially when the scale of the hand model is larger than the scale of real hand, may have the value of slant angle η exceeding the bound surfaces that are mentioned in section 2.3. The function E_{mov} is defined as:

$$E_{mov} = \begin{cases} w(\eta' - B_1(s, r')), & \text{if } \eta' > B_1(s, r'), \\ w(B_2(s, r') - \eta'), & \text{if } B_2(s, r') > \eta', \\ 0, & \text{if } B_1(s, r') < \eta' < B_2(s, r'). \end{cases} \quad (5)$$

where, s', r' and η' are determined by different scaled hand model, and w is the weighting coefficient.

c) The length and ratio measure function E_{rat}. The ratio constraint applied on the 3-D hand model fitting describes the general finger length range and the reasonable length ratios between segments. Based on the above two constraints, the function $\boldsymbol{E}_{rat}$ is defined as

$$E_{rat} = \lambda \left\| R_i - \overline{R_i} \right\|^2 + \kappa P_i ,$$

where R_i and $\overline{R_i}$ are the current value and the mean value of R_i respectively, λ and κ are the weighting coefficients, and P_i is the finger length penalty function.

3.2 The Calibration Process

By minimizing the calibration measure function of each finger, the fitness of various scale of 3-D hand model to the hand postures is evaluated in the calibration sequence. Therefore, each finger is calibrated by minimizing the calibration measure function, $C_i(s_{i1}, s_{i2})$ for i=1~5.

4. The Dynamic Analysis of the Hand Postures

The hand grasping movements or the hand rotating movements often introduces the occlusion problem of the markers. In this section, the temporal information of the hand motion is utilized to estimate the hand posture when some markers pasted on the hand are occluded.

4.1 The Dynamic Tracking of the Hand Postures

The rough searching process of the wrist fitting is implemented by using the modGA optimization method, which can find the approximate rotation angle of the wrist. After the wrist fitting process, the finger fitting process is performed by using the inverse kinematics solutions. The detailed description can be seen in [6].

4.2 The Hand Posture Analysis for Some Occluded Markers

During the hand movements, two kinds of marker occlusion problems may happen. The first one is the occlusion of the markers on the palm or the tip of the middle finger, and it will make the wrist fitting process fail. The second one is the occlusion of the markers on the five fingertips, and it will jeopardize the finger fitting process. The differential values between the two hand models are applied to predict the occluded hand posture. Here, the path coherence function [8] is applied to identify the occluded marker. The mathematical formulation of the path coherence for the 3-D positions of the corresponding markers located at frame k-1, k, and k+1 (p_i^{k-1}, p_i^k, p_i^{k+1}) is defined in the following:

$$\Psi(p_i^{k-1}, p_i^k, p_i^{k+1}) = w_1\varphi_d + w_2\varphi_s \qquad (6)$$

where

$$\varphi_d = 1 - \frac{\vec{l}\cdot\vec{m}}{|\vec{l}\,||\vec{m}|}, \quad \varphi_s = 1 - 2\frac{\sqrt{|\vec{l}\,||\vec{m}|}}{|\vec{l}|+|\vec{m}|},$$

$\vec{l} = \vec{p}_i^{\,k} - \vec{p}_i^{\,k-1}$, $\vec{m} = \vec{p}_i^{\,k+1} - \vec{p}_i^{\,k}$, w_1 and w_2 are weights.

By applying the path coherence function, some occluded markers can be identified. If there are m markers detected in frame k, and n markers detected in frame k+1 (assume $m > n$), then, there are m-n occluded markers are found in frame k+1. The method of identifying which markers are occluded is described in the following:

1. Let the 3-D positions of m markers detected in frame k be described as a set $A=\{p_1^k, p_2^k, \ldots, p_m^k\}$, and the 3-D positions of n markers detected in frame k+1 be in set $B=\{p_1^{k+1}, p_2^{k+1}, \ldots, p_n^{k+1}\}$ (assume m>n). Here, a point mapping function $M(\cdot)$ is defined to find the marker p_i^k in set A which generates the minimum path coherence measure to a specified marker p_j^{k+1} in set B as:
$$M(p_i^k \rightarrow p_j^{k+1}) = \min_{1\le i\le m} \Psi(p_i^{k-1}, p_i^k, p_j^{k+1}).$$
2. By applying the point mapping function $M(\cdot)$ for all points in set B, there are m-n un-mapped points in set A (see Fig. 3) may be found. These un-mapped points in set A are the occluded markers.

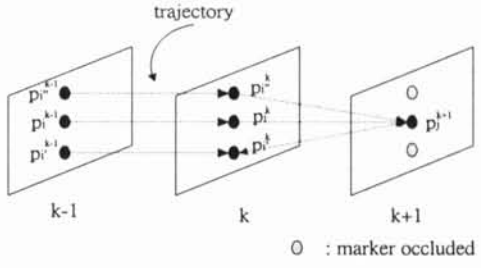


Fig. 3. The method of identifying the occluded markers.

On the other hand, if there are n markers detected in frame k, and m markers detected in frame k+1 (m > n), then, m-n markers will appear in frame k+1. The method of identifying which markers appear is similar to the above algorithm.

5. Experimental Results

The system flowchart of the scalable hand posture analysis system is illustrated in Fig. 4. The experiments of the scalable hand model fitting process consist of two parts. The first one is to analyze the error and efficiency of the scale calibration process for fitting the hand model to the various size of hand. The second one is to analyze the hand postures when some markers are occluded.

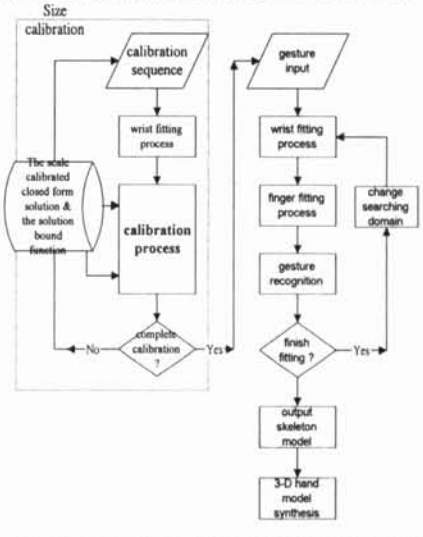


Fig 4. The flowchart of the scalable hand posture analysis system.

a) The accuracy and efficiency analysis of the scale calibration process. The initial 3-D hand model is constructed by using the average size of twenty people's hand. Fig. 5 shows the accuracy of the scale calibration process for index finger and the accuracy is analyzed by performing the calibration process ten times and

calculating the normalized mean absolute error defined as

$$E(i) = \sum_{j=1}^{2} \frac{|l'_{ij} - l_{ij}|}{l_{ij}} \Big/ 2 \,, \; I = 1, 2, \ldots, 5.$$

L_{ij}': the length of the segment j of finger I for the calibrated hand model.
L_{ij}: the length of the segment j of finger I for the real hand.
The computation time of the calibration process running on a Pentium-II 300 MHz PC is listed in Table 1

Table 1. The calibration time for the five fingers.

Finger	Thumb	Index	Middle	Ring	Little
Calibration time	58	6	4	6	7

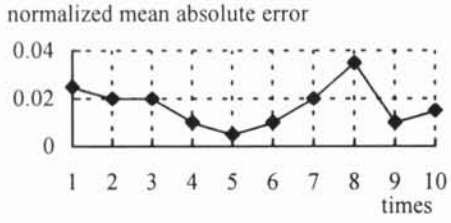


Fig. 5. The accuracy of the scale calibration process for index finger.

b) The dynamic hand posture analysis with occluded markers. Fig. 6 shows the image sequence of the hand posture that some markers are occluded for a short time interval. The image sequence is sampled at 10Hz. Figure 7 shows the 3-D hand model of the recognized postures in Fig. 6. Figure 8 shows the fitting error for 3-D model fitting to the postures shown in Fig. 6. The fitting error is defined as

$$e(j) = \sum_{i=1}^{5} d_i(j) \Big/ \sum_{i=1}^{5} L_i \,,$$

where $e(j)$ is the error rate for jth adjustment, d_i is the distance from the tip position of finger i of the fitted hand model to the tip position of the same finger i of the real hand, and L_i is the length of the finger i.

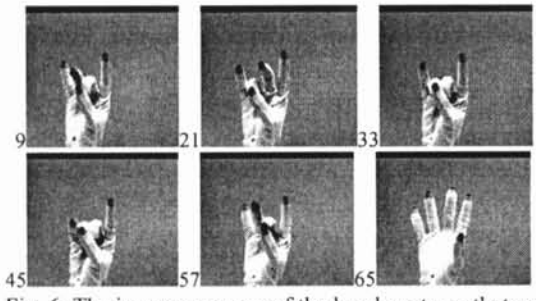


Fig. 6. The image sequence of the hand postures that some markers are occluded in some frames.

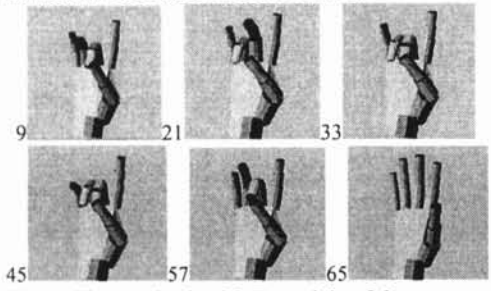


Fig. 7. The synthesized 3-D models of the postures in Fig. 6.

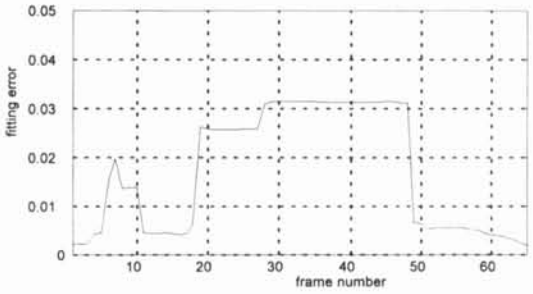


Fig. 8.The error for the model fitting to the postures in Fig.6

6. Conclusion

There are three main contributions in this paper. The first one is the generation of the scalable inverse kinematics solution functions. The second one is to propose a scale calibration method that can efficiently calibrate the scale of the hand model. The average calibration error is below 0.05. The third one is the dynamic hand posture analysis system that consists of the estimation process for solving the marker occlusion problem. However, the complex hand movements are not considered in this paper. In the future work, the analysis of the complex hand movements will be considered.

References

[1] J.M. Rehg, T. Kanade, "Digiteyes: vision-based hand tracking for human-computer interaction," IEEE Workshop on Motion of Non-rigid and Articulated Objects, pp. 16-22, Nov. 1994.
[2] J.J. Kuch, T.S. Huang, "Vision based hand modeling and tracking for virtual teleconference and telecollaboration," ICCV Proceeding, pp. 666-671, June 1995.
[3] T. Heap, D. Hogg, "Towards 3D tracking using deformable model," IEEE Proc. of 2nd International Conference on Automatic Face and Gesture Recognition, pp. 140-145, Oct. 1996.
[4] I.J. Mulligan, A.K. Mackworth, P.D. Lawrence, "A model-based vision system for manipulator position sensing," IEEE Workshop on Interpretation of 3D Scenes, pp. 186-193, Nov. 1989.
[5] J. Lee, T.L. Kunii, "Model-based analysis of hand posture," IEEE computer graphics & application, Vol. 15, No. 5, pp. 77-86, Sep. 1995.
[6] C.C. Lien, C.L. Huang, "The Model-based Dynamic Hand Posture Identification Using Genetic Algorithm," Journal of Machine Vision and Application, No. 11, pp. 107-121, June 1999.
[7] A.J. Koivo, Fundamentals for control of robotics manipulators. John Wiley & Sons, Chichester.
[8] C.L. Huang, C.H. Wu, "Dynamic scene analysis using path and shape coherence constraint, Pattern Recognition," Vol. 25, No. 5, pp. 445-461, 1992.