Image and Vision Computing 17 (1999) 149–156

Three-dimensional modeling of the human hand with motion constraints ☆

Yoshihiro Yasumuro[a,*], Qian Chen[b], Kunihiro Chihara[a]

[a]*Graduate School of Information Science Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-01, Japan*
[b]*Faculty of Systems Engineering Wakayama University, 930 Sakaedani, Wakayama City 640, Japan*

Received 21 July 1997; received in revised form 21 January 1998; accepted 4 February 1998

Abstract

This paper describes a new approach for modeling the human hand by considering the dynamics and the natural constraints of the motion and the shape of hands. This hand model consists of a dynamic model and a surface model. The dynamic model is used to generate the posture of a hand. We show that the natural hand posture can be generated even when only a few hand parameters are available. The surface model is used to generate the hand shape based on the posture given by the dynamic model. The surface model is built based on the digitized three-dimensional shape of a real hand. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Computer graphics; Hand model; Dynamic model; Range data

1. Introduction

As computer technology advances, we have obtained new environments constructed by computers [1]. To make computers more human-friendly, we need the computers to act like humans in many systems and applications. For example, in virtual reality (VR) applications, we need some human-like agents to do the job with users, or to play the opponent roles for game users.

This paper describes a three-dimensional hand model that can be used to make computer graphics (CG) animation and many other applications such as hand shape recognition. Using CG technology, one can create images of various objects from their three-dimensional models. However, because the motion information is difficult to be obtained from moving object, or to be synthesized without significant artifacts, it is not easy to create CG animation that looks natural. The above is also true in the case of the human hand. The human hand has over 10 joints which are freely movable, and can create numerous positions and generate complex movements.

Recently, motion captors such as data gloves have been widely used to capture the motion sequence of the hand that can be used for hand animation. Although the hand animation generated in this way is smooth and natural, it can only replay the recorded hand motion, thus it will not be able to be used in a computer synthesized hand animation system. Even when there are many degrees of freedom in a human hand, the fingers' motions are not completely independent [2,3]. The angle of each finger joint is influenced by the angles of the neighboring joints and those of the adjacent finger joints. We explored the motion constraints between the joints in the same finger and the relation between the motions in different fingers. We then built a dynamic model to describe these motion constraints and used it to generate the natural hand motions. When we want to generate a particular posture of hand, we give 'force' to the finger to be moved. The postures of one or several of the fingers with a given force are first determined with the model. The postures of the rest of the fingers are then computed with the dynamic model by using the information of the determined posture. The proposed model can also be used to predict the posture of the whole hand when there is only part of the information of the fingers available. The unknown posture parameters of the hand can be determined by the known parameters and the dynamic model by considering the 'inner force' between fingers.

A skin model of the hand was also built for generating the three-dimensional shape of the hand, which can be used to synthesize images of a hand that look real. The skin is modeled as a continuous surface which covers a whole hand. The initial shape of the skin model is constructed based on the three-dimensional shape data of a real hand, and is represented as a polygonal model. The skin shape of a given posture is produced by a deformation of the initial

☆ Based on "3-D modeling of human hand with motion constraints" by Yoshihiro Yasumuro, Qian Chen and Kunihiro Chihara which appeared in Proceedings of 3-D Digital Imaging and Modelling, Ottawa, May 1997; pp. 275–282. © 1999 IEEE.

* Corresponding author.

PII: S0262-8856(98)00118-8

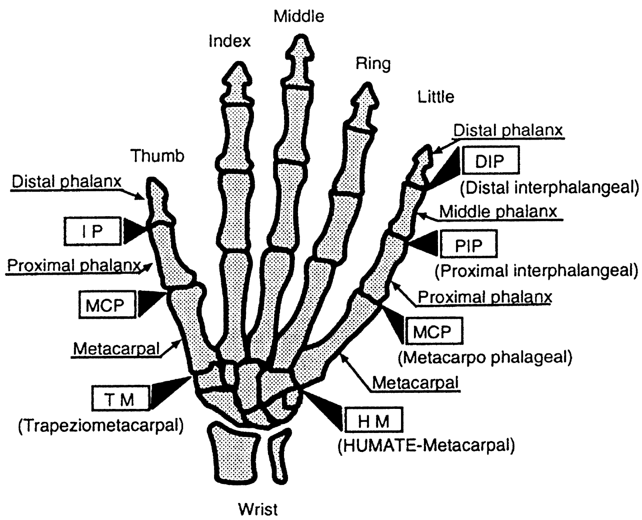

Fig. 1. Diagram for the human hand skeleton.

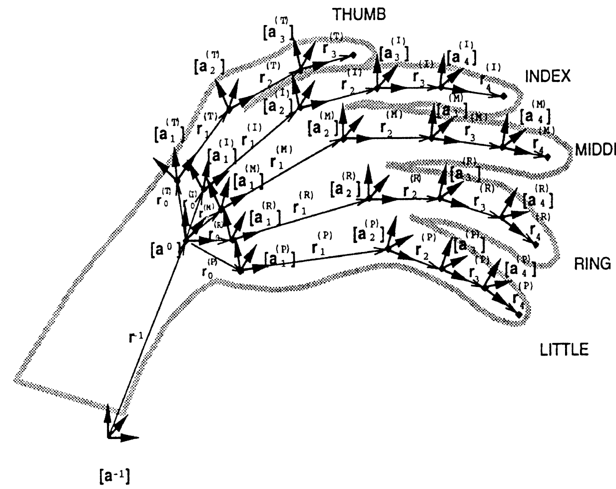

Fig. 2. Coordinate systems.

one. Once the hand posture is determined by the dynamic model, the posture parameters are used to adjust the positions of the control points of the surface model to generate the shape of the surface of the hand.

2. Skeleton model

The proposed hand model is composed of a skeletal model and a skin model. We describe the posture of a hand with the position of the bone and the joint in it. The structure of the bones in a hand is described by a skeletal model based on the anatomical aspects. We modeled the hand skeleton with rigid links and joints.

2.1. The skeleton of a human hand

Fig. 1 illustrates the real hand skeleton [4]. The hand skeleton consists of eight carpal bones, five metacarpal bones and five phalanges which can be divided into proximal, middle and distal phalanges. Only the thumb has two phalanges. The carpus has a delicate deformation since its eight elements articulate subtly with each other. However, since its effect on the surface deformation is small, we neglect it and model the carpus as a rigid link.

2.2. The structure of the skeleton model

We model the skeleton of human hand with a hierarchical structure [2] that consists of rigid links and joints. Each joint has one or two rotational degrees of freedom. This hierarchical structure can be represented by a tree, where the root is the wrist.

We use a three-dimensional coordinates systems shown in Fig. 2 to describe the joint's position. The position of the hand is described by the three-dimensional coordinates of the wrist (root) relative to the world coordinate system. Then the posture of each link, where one end is connected to the root (wrist), is described by a vector, represented under a local coordinate system fixed at the wrist. We put a local coordinate system on each link to describe the position of the next link sequentially. The origin is set at the joint that connects the link and its previous link.

The unit base vector of the world coordinate system is represented as $\boldsymbol{a}_{-11}$, $\boldsymbol{a}_{-12}$, $\boldsymbol{a}_{-13}$.

So we can express the bases vector system as:

$$[\boldsymbol{a}_{-1}]^T = [\boldsymbol{a}_{-11}\ \boldsymbol{a}_{-12}\ \boldsymbol{a}_{-13}] \quad (1)$$

Similarly, we settle a local coordinate system on each link of the fingers. The origin of the coordinate system on the jth joint of finger k (k is one of T (the thumb), I (the index), M (the middle), R (the ring), and L (the little)) is put at the joint $C_j^{(k)}$ and the base vector of the coordinate system is represented as:

$$[\boldsymbol{a}_j^{(k)}]^T = [\boldsymbol{a}_{j1}^{(k)}\ \boldsymbol{a}_{j2}^{(k)}\ \boldsymbol{a}_{j3}^{(k)}] \quad (2)$$

We define the vector on each link as: $\boldsymbol{r}_{-1} =$ the position of the carpus described under $[a_{-1}]$ coordinate system; $\boldsymbol{r}_j^{(k)} =$ the vector from the joint $C_j^{(k)}$ to the joint $C_{j+1}^{(k)}$ described under $[a_j^{(k)}]$ coordinate system.

The position of each joint can be calculated by performing a coordinate transformation with 3-2-1 Euler angular matrices. In this transformation, the local coordinate systems are first rotated around the third axis by angle θ_3, then around the second axis by angle θ_2, and around the first axis by angle θ_1. The transformation from the coordinate system $[a_{j-1}]$ to $[a_j]$ can be expressed by the following equation.

$$[\boldsymbol{a}_j]^T = \mathbf{A}_{jj-1}[\boldsymbol{a}_j - 1]^T \quad (3)$$

where $\mathbf{A}_{jj-1}$ is the transformation matrix:

$$\mathbf{A}_{jj-1} = A_1(\theta_{j1})A_2(\theta_{j2})A_3(\theta_{j3}) \quad (4)$$

The matrix for the transformation from local coordinates to the world coordinates can be obtained with the equation:

$$\mathbf{A}_{j-1}^{(k)} = \prod_{l=0}^{j} \mathbf{A}_{ll-1}^{(k)} \quad (5)$$

Table 1
Degrees of freedom in the model

	The wrist	TM	HM	MCP	PIP, DIP, IP
Number of joints	1	1	2	5	9
Translate	3	—	—	—	—
Rotate	3	2	2	2	1
Sum of DOF	6	2	4	10	9

The three-dimensional position of the joint $C_{(j)}^{(k)}$ described in world coordinates system can be calculated by the following equation:

$$x_n^{(k)} = \boldsymbol{r}_{-1} + \sum_{j=0}^{n} \mathbf{A}_{j-1}^{(k)T} \boldsymbol{r}_j^{(k)}, \ (k = T, I, M, R, L) \tag{6}$$

where the suffix n stands for the index of the joint in each finger, from 1 to 4 for the thumb, 1 to 3 for the other fingers.

2.3. The degrees of freedom in each joint

With the equations described in the previous subsection, the whole posture of the hand skeleton can be determined by giving the position of the wrist and the Euler angles of each joint. The angle of the joint $C_j^{(k)}$ is expressed as $\theta_{(j)}^{(k)}$, which is used as the state variable of the hand skeleton. The direction of the Y-axis is set to be parallel to the axis of jth link, and the direction of X-axis is parallel to the rotational axis of the flexion movement of the joint. We assume that the joint (1) of thumb and the joint (2) of all the other fingers have two degrees of freedom (rotational components around X and Z axes), and the other joints have only one (rotational component around X-axis) (see Table 1 and Fig. 3).

3. Posture determination process

This section describes a dynamic system that models the natural motion constraints in a human hand. Since the fingers in a hand are connected with tendons, muscles, other tissues and covered by skin, the movements of them are not completely independent. There are some rules that determine the interactions of joint motions. We explore these rules, and build a model to describe them. From this model, we can tell whether a given posture is natural, and if the given posture is unnatural, we can estimate the degree of difference between the posture and the natural one. Then the natural posture can be generated by adjusting the given posture in a way that the unnaturalness of the result posture decreases.

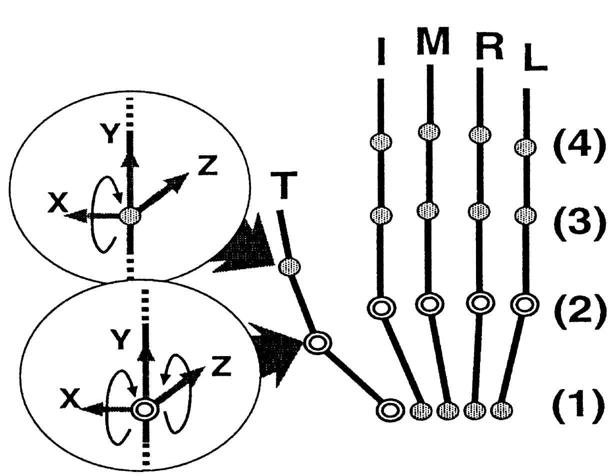

Fig. 3. Hand skeleton model.

In this research, the unnaturalness of a hand posture is modeled by a potential energy function. Thus, the goal of generating natural hand posture can be achieved by producing a hand posture of which the potential energy is zero, or as small as possible. The potential energy can give forces to decrease itself. We apply these forces to the skeletal model to adjust the joint angles. As a result, the hand posture will have less potential energy, and thus become more natural. We use these 'forces' as natural motion constraints and call them 'inner forces'.

The posture determination system is composed of the skeletal model and the dynamic system. We put virtual springs on joints in the skeletal model to express the rigidity of the joints. Once the user input is given to the system, potential energy is changed and inner forces are generated. The output of the process is determined as a equilibrium state of the dynamic system. The computed posture is expressed by the joint angles and is used to generated a hand shape in the skin model.

3.1. System equation

We assign a virtual spring to each rotational degree of freedom of a joint (Fig. 4). State variables (angles of the joints) are determined by equilibrium between the elastic torque of virtual springs and the input torque. This is expressed as the following equation.

$$\mathbf{K}\boldsymbol{\theta} = \boldsymbol{\tau} \tag{7}$$

where $\boldsymbol{\theta}$ = state variable vector, $\boldsymbol{\tau}$ = input torque vector, $\mathbf{K}$ = stiffness matrix. The matrix $\mathbf{K}$ is a diagonal matrix that indicates the stiffness of the virtual springs.

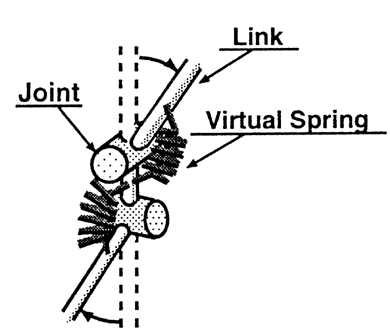

Fig. 4. Virtual spring.

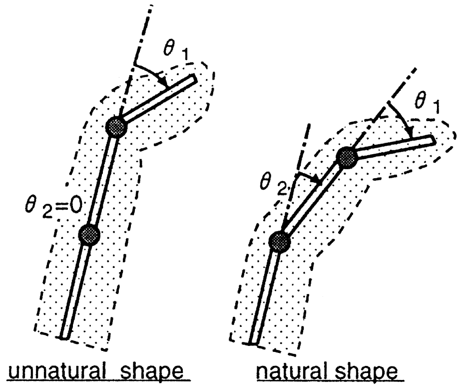

Fig. 5. Natural posture.

3.2. Natural posture

Based on the idea that the joints of the fingers do not move independently (Fig. 5), we represent the rules that describe the natural posture by a linear combination of the state variables of the system.

$$\alpha\theta_1 + \beta\theta_2 + \gamma = 0 \quad \alpha, \beta, \gamma : \ const. \tag{8}$$

We construct a potential energy J_1 to describe the degree of the unnaturalness of a given posture. This potential energy is defined as a quadratic form derived from Eq. (8) as

$$J_1 = \frac{1}{2} w(\alpha\theta_1 + \beta\theta_2 + \gamma)^2 \tag{9}$$

where w is a parameter of elasticity that determines the strength of the inner force. The potential energy J for the whole hand is defined as the linear summation of the potential energies of all the relational joint pairs.

$$J = \frac{1}{2}\Theta^T \mathbf{W}\Theta \tag{10}$$

where $\Theta = \mathbf{A}\boldsymbol{\theta} + \boldsymbol{b}$ is the linear combination of $\boldsymbol{\theta}$, the state variables. The components of $\mathbf{W}$ express the gains of inner forces.

The inner forces to create the natural posture are derived by calculating the partial differentiation respect to the state variables as the following.

$$\tau_{c1} = -\left(\frac{\delta J_1}{\delta\theta_1}\right) = -w\alpha(\alpha\theta_1 + \beta\theta_2 + \gamma) \tag{11}$$

Similarly, the inner force of the whole hand, $\boldsymbol{\tau}_c$, is derived by calculating the partial differentiation of energy with respect to the state variable vector as the following.

$$\boldsymbol{\tau}_c = -\left(\frac{\delta J}{\delta\boldsymbol{\theta}}\right) = -\left\{\left(\frac{\delta J}{\delta\boldsymbol{\theta}^T}\right) + \left(\frac{\delta J}{\delta\boldsymbol{\theta}}\right)^T\right\} - \mathbf{A}^T\mathbf{W}(\mathbf{A}\boldsymbol{\theta} + \boldsymbol{b}) \tag{12}$$

3.3. Determination of hand posture

We call the user input the direct input. Total inputs to the system are the sum of two inputs, the direct inputs and the inner forces:

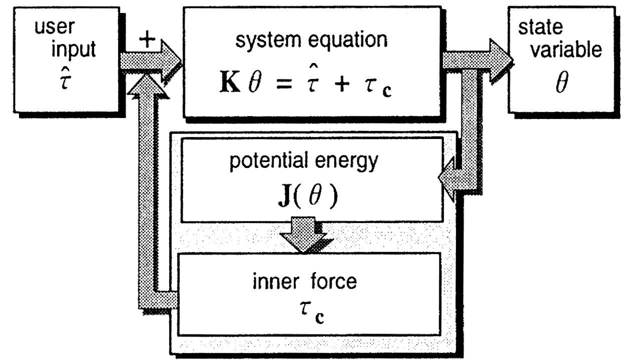

Fig. 6. Diagram of posture determination.

$$\boldsymbol{\tau} = \hat{\boldsymbol{\tau}} + \boldsymbol{\tau}_c \tag{13}$$

where $\hat{\boldsymbol{\tau}}$ = direct input vector, $\boldsymbol{\tau}_c$ = inner force vector.

At first, we assume that the inner forces are zero and estimate the angles of all the joints with Eq. (7). Then the potential energy is calculated with Eq. (10). In order to bring the hand system to a state of equilibrium, we derive the inner forces with Eq. (12) and adjust the input assigned to each joint with Eq. (13) then estimate the angles of all the joints with Eq. (7) again (see Fig. 6). The state variable vector $\boldsymbol{\theta}$ in an equilibrium state can be determined by the following equations.

$$\begin{cases} \mathbf{K}\boldsymbol{\theta} = \boldsymbol{\tau} \\ \boldsymbol{\tau} = \hat{\boldsymbol{\tau}} + \boldsymbol{\tau}_c \end{cases} \tag{14}$$

$$\mathbf{K}\boldsymbol{\theta} = \hat{\boldsymbol{\tau}} + \boldsymbol{\tau}_c \tag{15}$$

$$= \hat{\boldsymbol{\tau}} - \mathbf{A}^T\mathbf{W}(\mathbf{A}\boldsymbol{\theta} + \boldsymbol{b})$$

$$\downarrow$$

$$(\mathbf{K} + \mathbf{A}^T\mathbf{W}\mathbf{A})\boldsymbol{\theta} = \hat{\boldsymbol{\tau}} - \mathbf{A}^T\mathbf{W}\boldsymbol{b}$$

When there are only some of the direct inputs given, we assign the direct inputs not given as zero, and estimate the hand posture similarly as above. By doing this, we can generate natural posture of the hand even when some of the direct inputs are not available.

3.4. Determination of energy parameters

The parameters in the Eq. (8) are determined using the captured motion data from a real hand. We use a video camera to capture the image sequence of the hand motion. The calibration of the camera is done before we take the pictures. We attached tiny color markers to each joint of a hand. The angle of each joint is obtained from the positions of the markers extracted from the image sequence.

We draw line segments connecting the markers in each image and calculate the angles at the joints. We get sequences of joint angles (Fig. 7) and then use them to estimate the parameters in Eq. (8) by fitting a linear function to the

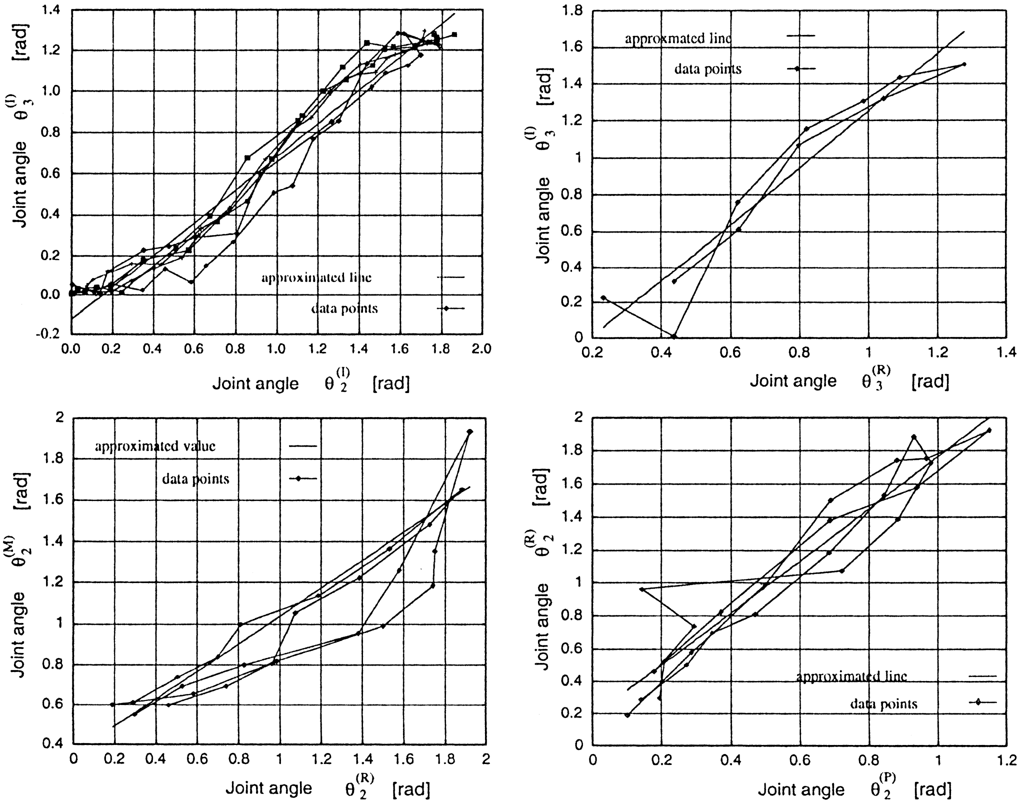

Fig. 7. Some data of joint angles of the finger motion used to determine the parameters in Eq. (8).

dots that indicate the angles of the relational joints when the fingers move. We grab 10–30 images of finger motion for one cycle of flex and extension. The result of the estimation of the linear function give the values of the components of matrix **A** in Eq. (10).

At the same time the dispersion of the linear approximation is also estimated. The weight parameters w_k of constraints in the matrix **W** are calculated as the normalized values of the dispersion of the linear approximation as follows,

$$w_k = \frac{\sum_{i}^{n} l_i^2}{n w_{\max}} \tag{16}$$

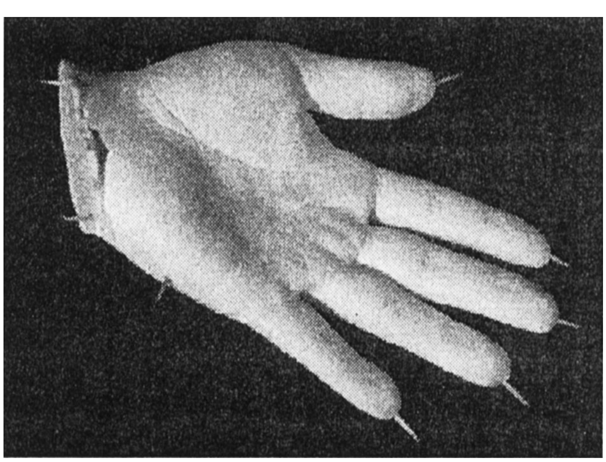

Fig. 8. Plaster replica of a hand.

where, l_i = Euclidean norm between data point and estimated line, n = number of data, $w_{\max}$ = maximum value of w_k.

4. Surface generation

The three-dimensional surface model of the hand is generated based on a digitized three-dimensional shape information of a real human hand. The surface model is a polygon based model. To make it useful, we combine the polygon model with the skeletal model, so we can generate the continuous surface of the hand when we change the posture of the hand skeleton.

4.1. Range data

We measured the three-dimensional shape of a replica of a hand made of plaster. We used a laser range-finder with

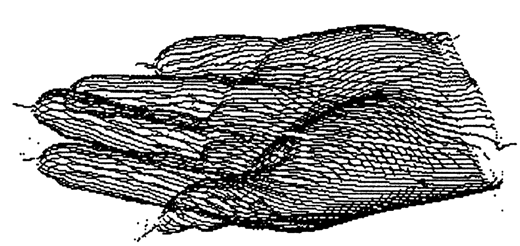

Fig. 9. Range data of a whole hand.

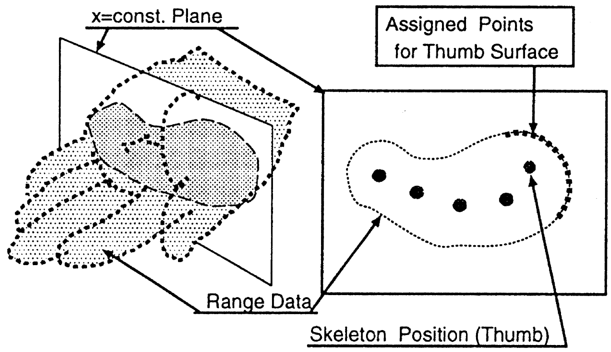

Fig. 10. Procedure to get the polygon vertices.

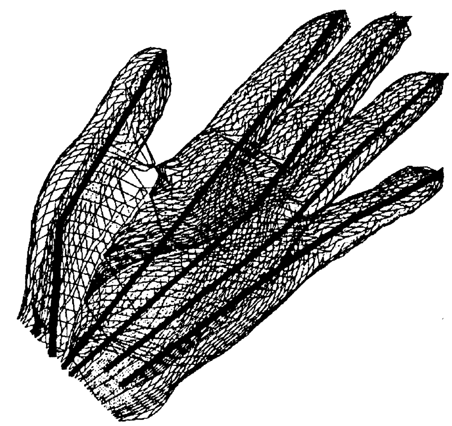

Fig. 12. Wire framed model.

linear motion platform to obtain shape data set of the palmar and the back side separately. Each range data set is composed of 512 × 512 three-dimensional points. The hand shape is extracted by removing the three-dimensional points of motion platform with a proper threshold of height value. Nine markers were attached to the replica (Fig. 8) before measurement. We transformed the dorsal aspect data set to the coordinate system of the palmar side and adjusted the three-dimensional positions of the markers included in both data sets, minimizing the quadratic error. In this way, we could get the range data of a whole hand (Fig. 9).

4.2. Data construction

In order to deform the hand surface properly, according to a given posture of the hand skeleton, we had to establish the relationship between the range data and the skeletal model, indicating which point of the range data belongs to which link of the skeletal model.

First, we assigned a hand posture similar to the one of the digitized hand of the skeletal model, and positioned the skeletal model into the digitized three-dimensional shape of the hand manually. As described in the previous subsection. the data set of the digitized hand shape is composed of palmar side data and dorsal data. We had to transform the dorsal data to the coordinate system of the palmar side in order to obtain the range data of the whole hand. After this process, the scanning lines of palmar and dorsal sides will become unaligned. We construct the polygon model to describe the hand surface as follows (also see Fig. 10).

1. Settle 50 section planes ($x = const.$ plane in the world coordinate).
2. Project all points of range data onto the nearest section plane.
3. Assign the projected points to the nearest link on each section plane.
4. For each group of the projected points belonging to the same link, transform the coordinate of each point to a polar coordinate system, where the origin is set at the link. In order to create the polygon vertices (see Fig. 11), we first select the points with the maximum angle θ_{min} and the one with the minimum angle θ_{min}. Then we create 10 new points of which angles are defined on even intervals, as following.

$$\theta_i^* = \theta_{min} + \frac{(\theta_{max} - \theta_{min})}{10} i \; (i = 0, 1, 2, \ldots, 9) \tag{17}$$

To obtain the radius value of each newly created point, we find the two nearest original points. One has an angle

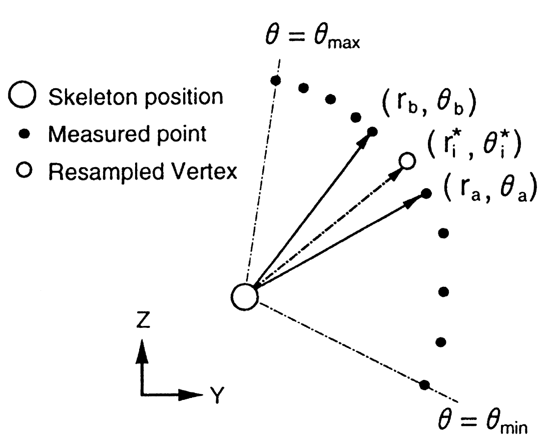

Fig. 11. Resampling to make vertex.

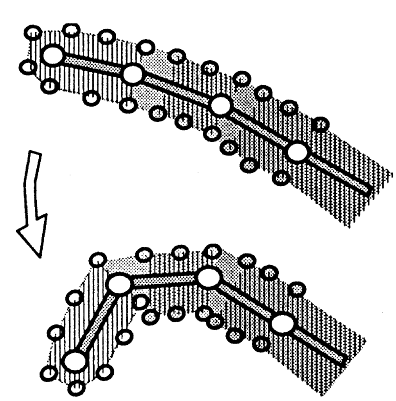

Fig. 13. Deformation of the skin model.

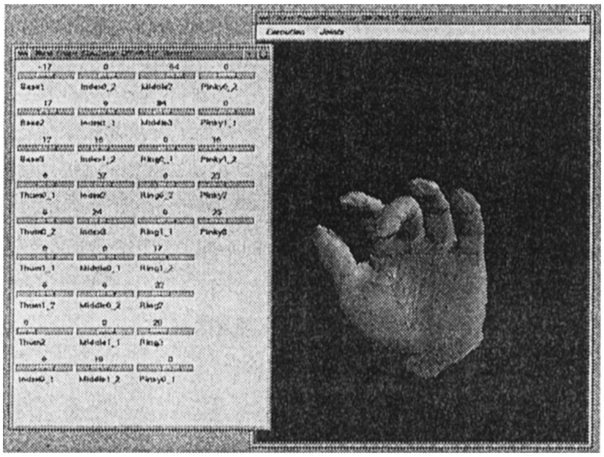

Fig. 14. GUI—scale widgets of inputs.

greater than θ_i^*, (r_b, θ_b) and the other has an angle less than θ_i^*, (r_a, θ_a). Then the radius of the newly created point is calculated by linear interpolation as follows,

$$r_i^* = r_a + \frac{r_b - r_a}{\theta_b - \theta_a}(\theta_i^* - \theta_a) \tag{18}$$

After setting the skeletal model into range data and computing 50 section planes manually, the polygon vertices are automatically produced (see Fig. 12 for a wire framed model).

4.3. Deformation of the surface model

Now we have the polygonal surface that corresponds to the skeletal links. When a skeletal link moves, the surface of the link should deform to follow the moved link simultaneously. Each link surface moves rigidly according to the motion of the skeletal posture, rotating around the joints. We set the edge lines to be shared by the neighboring link surfaces, so that the surfaces around the joints will always stay continuous even when the fingers move. In this way, we can generate the hand surface naturally for various hand postures (Fig. 13).

5. Experiment

We built the skeletal model and the surface model using the method we proposed. In order to evaluate the effectiveness of the hand model, we built a simple GUI (graphic user interface) (Fig. 14) to allow a user to assign the desired angles of the joints and a model rendering engine to generate the hand images. At first, we input the sequence of the angles of the joints (Fig. 7) we want to move. The angles of the joints that we did not give are determined by the dynamic model. After that, the hand surface is generated with the surface model based on pre-determined angle of the joints, which is then sent to model rendering engine to generate the hand animation. The system is implemented on a SGI Indigo^2 workstation.

Fig. 15. Polygon surface.

The hand surface is composed of about 8000 triangle meshes (Fig. 15). Fig. 16 illustrates sample images of the generated hand motion animation. The posture of the hand with zero inputs is shown in the upper frame. The middle and the bottom frame show the deformed hand images in the hand animation, where we only assigned the angles of the

Fig. 16. Motion sequence.

joints of the thumb and the middle finger. The animation is generated in real time corresponding to the user operation of the GUI.

6. Discussion

We obtain sufficient reality for both the shape and the motion (Figs. 15 and 16). The model can be used as an agent in a VR (virtual reality) environment, indicating, touching or grasping something, as well as creating hand animation. We consider that the posture determination process can be applied to biomedical engineering dealing with prosthetics. It will be extremely useful for utilizing the limited inputs that control an object like an artificial hand for disabled people.

As you can see in the figures, the skin surface shape is coarse and noisy in the current implementation. We are thinking of refining it using a curved surface model in stead of polygon model to represent the hand skin.

7. Conclusions

We introduced dynamics to a human hand model by using virtual springs and an 'inner force' energy in a skeletal model. Moreover, we built a surface model of human hand based on the digitized three-dimensional shape information of a real hand and used it to synthesize photo-realistic images. The natural posture of a hand is defined as the equilibrium state of the dynamic model, which can express the complicated effects of the interference between finger motions. The desired natural posture of the hand can be generated by only assigning a few key angles of the joints. The experiment showed that the hand animation can be created easily and the generated hand animation is natural and smooth.

References

[1] K. Singh, J. Ohya, R. Parent, Human figure synthesis and animation for virtual space teleconferencing, in: Proceedings for Virtual Reality Annual international Symposium 1995 (VRAIS'95), 1995, pp. 118–126.

[2] J. Lee, T.L. Kunii, Constraint-based hand animation, in: Models and Techniques in Computer Animation, Springer, Tokyo, 1993, pp. 110–127.

[3] J.J. Kuch, T.S. Huang, Vision based hand modeling and tracking for virtual teleconferencing and tele-collaboration, in: Fifth International Conference on Computer Vision, 1995, pp. 666–671.

[4] W. Kahle, H. Leonhardt, W. Platzer, Taschenatlas der Anatomie, Georg Thieme Verlag, Stuttgart, 1986.