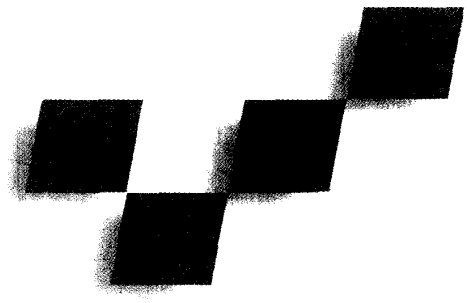


Model-Based Analysis of Hand Posture

Jintae Lee and Tosiyasu L. Kunii
University of Aizu

Automatic input and analysis of hand motion have attracted attention from computer animation and virtual reality researchers. Computer animation applications directed at the human hand must handle about 30 interacting parameters to specify a particular motion. Consequently, researchers have investigated several methods to derive these parameters from real hand motions. Virtual reality research has addressed the removal of the distinctions between the computer system and user environment; that is, the computer system presents a virtual space to the user's sense organs such as eyes, ears, and skin, and the user reacts using gestures and speech. The hands obviously play a major role in this scenario.

Although mechanical gloves can capture hand motion in real time, they are expensive, and the associated measuring equipment inhibits free movement.[1] More advantageous, a camera-based approach requires no mechanical gloves or motion-constraining monitoring equipment. The virtual reality and computer vision communities have targeted capturing hand motion via cameras. Researchers have successfully recognized specific fingers of the hand by silhouette images[2] and distinguished a small set of hand signs by contour features of images.[3,4] However, the silhouette or contour features recovered from the images do not provide sufficient information to generate 3D hand posture with the fingers positioned properly.

This failure led to our study, which developed a new method employing a hand model that can automatically analyze 3D hand posture using stereo images. Our method differs from previous methods in that it captures 27 interacting hand parameters, including finger joint angles, enabling suitable reconstruction of 3D hand posture.

Human hand modeling

A model-based approach for image analysis is common in computer vision applications. In fact, comparing an internal model with an external image is considered the basis of natural vision. However, computer analysis of hand posture from actual images involves complex problems even when employing conventional, simple models. In addition to the three common difficulties arising in practical computer vision—occlusion, noise, and spurious data—the analysis process for the human hand is further complicated by the fact that

1. The human hand is an articulated structure with about 30 degrees of freedom and changes shape in various ways by its joint movements. Contributing to this, hand images change by both finger movements and hand movement as a whole.
2. Because four fingers are adjacent to each other, occlusion of the fingers frequently occurs. The many tiny wrinkles existing on the skin also make it difficult to detect meaningful edges of the image. Thus, the conventional strategy of segmenting an image into parts by edge extraction and sequentially tracing the parts along the central axes does not apply.
3. The fingers are comparatively short, and the surfaces around the joints are significantly deformed by joint movements, making it very difficult to accurately estimate joint positions in images.

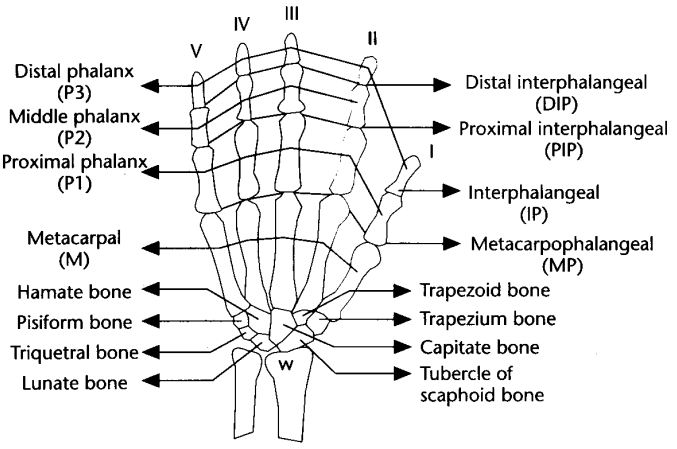
This new method employs a hand model and static images to analyze hand posture. Guided by internal constraints and external forces, the model is automatically fitted to the hand image.

To solve such problems associated with the recognition of human hand images, we used a hand model that incorporates constraints based on the natural movements of a human hand.

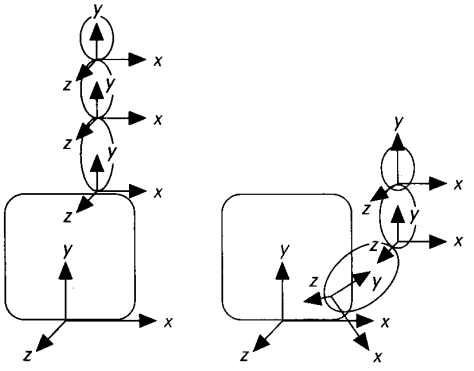
Hand movements

The human hand is a complex mechanical structure comprising bones, ligaments loosely connecting bones, muscles serving as tension motors, tendons acting as cables connecting muscles to bone, and a covering of

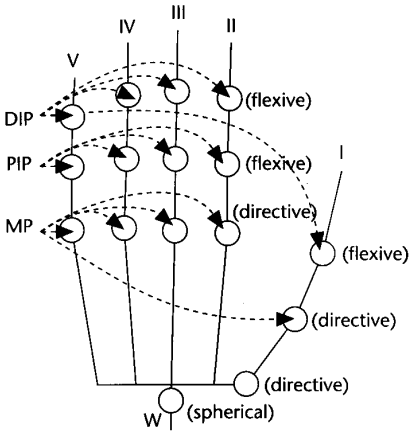
0272-17-16/95/$4.00 © 1995 IEEE



1 A hand skeleton observed from the palmar side (adapted from *Pernkopf's Anatomy*[5]). We use the indicated nomenclature hereafter. For example, M(III) represents the metacarpal bone of the middle finger.



2 Local coordinate systems of the fingers and thumb defined with conventional right-handed coordinate systems. Each joint's rotation is expressed by a sequence of rotations occurring around the *x*, *y*, z axes of the local coordinate system.



3 Joints of the hand and their movement types.

protective soft tissue and skin. The bones are linked at the joints and do not change in size. Muscles produce torque and/or joint movements through tension, and for every muscle there exist one or more muscles that serve to oppose it through counter-torque and/or opposing motion. Figure 1 illustrates the skeleton of the right hand observed from the palmar side. We use the terminology and abbreviations from that figure hereafter.

In modeling the human hand, analysis of constraints is essential to achieve the following:[6] (1) avoid unrealistic motions during hand animation and (2) reduce the search space in model-based analysis of hand images. Inevitably, a trade-off arises between the degree of constraints contained in a model and its resultant performance; that is, a lack of constraints leads to a useless model, whereas too many of them requires complex procedures necessitating expensive computation time. Our hand model, therefore, attempts to effectively balance these considerations.

Description of joint movement. The movement of a particular segment of the hand is produced by rotations of its proximal joint and can be specified by the joint's rotation angles. Here we define local coordinate systems on every joint position (Figure 2) and represent any joint rotation by a sequence of rotations occurring around the three axes of the local coordinate system. Joint rotation around an axis is expressed by $\theta_\beta^\alpha(\gamma)$, where α represents the rotation axis, β the joint, and γ the finger.

Human hand joints can be classified as *flexion*, *twist*, *directive*, or *spherical* according to the type of movement or possible rotation axes.[7] A flexion movement joint with a single degree of freedom (DOF) is the knee or elbow, while a twist movement joint having one DOF is the pronation of the forearm. Directive movement with two DOF permits flexion movement in more than two directions. Spherical movement, as in the shoulder or hip, has three DOF and permits simultaneous directive and twist movements.

Figure 3 shows the joints employed in our hand model, along with their associated allowable movement types. Each finger (II – V) has four DOF (two at the metacarpophalangeal or MP, one at the proximal interphalangeal or PIP, and one at the distal interphalangeal or DIP), while the thumb (I) has five DOF (two at the MP, two at the PIP, and one at the DIP). The wrist's twist movement is included because the hand must be considered separately from the lower arm. According to this joint movement classification, 27 DOF exist for the hand, including six DOF to position and orient it.

Constraints on joint movements

Normally, movements of the finger joints are coordinated by constraints that make some configurations impossible. We therefore analyzed and incorporated some prominent constraints into the hand model,

broadly classified as follows (based on movement types and the involved joints):

1. Joint angle limits and movement types
2. Flexion of the interphalangeal joints
3. Flexion of the metacarpophalangeal joints
4. Adduction and abduction of the metacarpophalangeal joints

Constraints on joint angle limits and movement types. In Figure 3, note that possible movement of the MP joint of fingers II – V is only flexion/extension or side movements, and that of the PIP and DIP joints is only flexion/extension in the same direction. Hence, the P1, P2, and P3 segments occupy the same plane.

Constraint 1. The four fingers are planar manipulators with the exception of the MP joint.

Although the allowable ranges of joint angles vary slightly from person to person, they do fall into general ranges.[8] Here, we distinguish between *passive* and *active* movements. The former movement is externally forced, whereas the latter is activated by tendons and muscles of the hand without external interaction. Joints generally have a greater range for passive movement. We only consider active hand motions, since inclusion of external forces exceeds the scope of our study.

Constraints on flexion of the interphalangeal joints. In the human finger, it is nearly impossible to move the DIP without moving the adjacent PIP joint—and vice versa—without forcing one of them to move in an unnatural manner. Namely, active movement involves a dependency between the DIP and PIP joints.[9] Anatomical studies have been directed at this phenomena, and an empirical study[10] revealed that an almost linear relationship exists between these two joints.

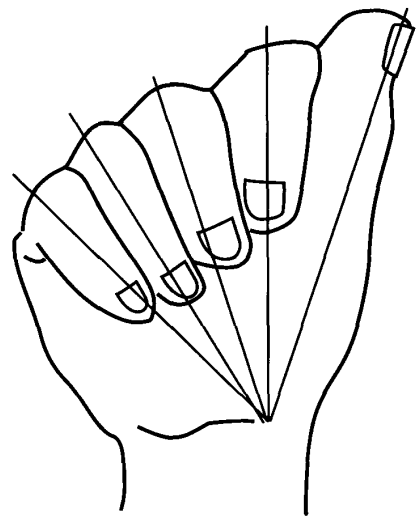
Constraint 2. The joint angles of the DIP and PIP joints have a dependency represented by

$$\theta_{DIP} = (2/3)\,\theta_{PIP}$$

Constraints on flexion of the metacarpophalangeal joints. The MP joint has a flexion range of 90 degrees, slightly less for the index finger (II) and progressively increasing for fingers III – V.

However, since isolated flexion of a finger is restricted by accompanying tension in the palmar interdigital ligament, such flexion might cause flexion of the adjacent fingers. In the same way, a finger's extension is hindered by the flexion of others. After measuring several people, we found this behavior could be reasonably approximated through inequalities. We use the *static joint angle limit* and *dynamic joint angle limit* to distinguish respectively between the joint angle limit calculated for all possible hand configurations and that calculated with the other joints in some specific configuration.

Constraint 3. The joint angle limits of the MP joints depend on those of the neighboring fingers



4 Convergence of fingers when the fist is clenched (based on Kapandji's *The Physiology of the Joints*[11]).

according to the following inequalities:

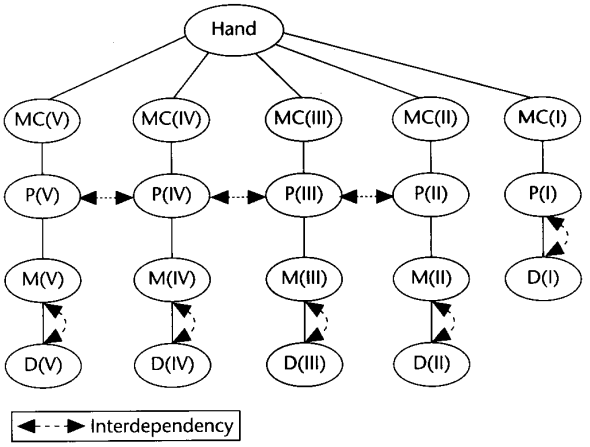
$$dmax\,(\theta^X_{MP(I)}) = min(\theta^X_{MP(M)} + 25, smax(\theta^X_{MP(I)}))$$
$$dmin(\theta^X_{MP(I)}) = max(\theta^X_{MP(M)} - 54, smin(\theta^X_{MP(I)}))$$
$$dmax(\theta^X_{MP(M)}) = min(\theta^X_{MP(I)} + 54, \theta^X_{MP(R)} + 20, smax(\theta^X_{MP(M)}))$$
$$dmin(\theta^X_{MP(M)}) = max(\theta^X_{MP(I)} - 25, \theta^X_{MP(R)} - 45, smin(\theta^X_{MP(M)}))$$
$$dmax(\theta^X_{MP(R)}) = min(\theta^X_{MP(M)} + 45, \theta^X_{MP(L)} + 48, smax(\theta^X_{MP(R)}))$$
$$dmin(\theta^X_{MP(R)}) = max(\theta^X_{MP(M)} - 20, \theta^X_{MP(L)} - 44, smin(\theta^X_{MP(R)}))$$
$$dmax(\theta^X_{MP(L)}) = min(\theta^X_{MP(R)} + 44, smax(\theta^X_{MP(L)}))$$
$$dmin(\theta^X_{MP(L)}) = max(\theta^X_{MP(R)} - 48, smin(\theta^X_{MP(L)}))$$

where *dmax*, *dmin*, *smax*, and *smin* are the variables for the dynamic and static joint angle limits defined in Constraint 5, and *max* and *min* are the maximum and minimum functions.

Constraints on adduction and abduction of the metacarpophalangeal joints. Adduction (approximation) and abduction (separation) movements of the fingers are referenced by the axis of the hand running through the third metacarpal bone M(III) and finger III. In such movement, finger III does not move appreciably without intentional forcing.

Constraint 4. The MP joint of finger III displays limited adduction and abduction.

In their naturally open position, the fingers can freely carry out adduction or abduction. However, when clenched into a fist with the DIP joints extended as shown in Figure 4, the axes of the two distal phalanges P2 and P3 of fingers II – V and the axis of the thumb (I) converge to a point. That is, the abduction and adduction angles continuously decrease as the flexion angles of the MP joints increase. Such behavior might seem initially to prevent collisions between the fingers, but clos-



5 Hand tree. Each node represents a segment of the hand. Inter dependency in movement between segments is represented by the dashed lines drawn between the corresponding nodes.

er observation reveals that this convergence takes place independently of collisions with other fingers.

Constraint 5. Define

θ^Z_{MP} = abduction or adduction angle of an MP joint
θ^X_{MP} = flexion angle of a MP joint
smax: maximum joint angle of static constraint
dmax: maximum joint angle of dynamic constraint
smin: minimum joint angle of static constraint
dmin: minimum joint angle of dynamic constraint.

Then

$$dmax(\theta^Z_{MP}) = k \times smax(\theta^Z_{MP})$$

where

$$k = \left(1 - \frac{1}{smax\left(\theta^X_{MP}\right)}\right)\theta^X_{MP}$$

Obviously, other miscellaneous constraints exist. We did not include them in our model because we felt they increase model complexity without providing meaningful gain.

Hand tree

To represent the articulate data structure of the hand model, imagine a hierarchical tree in which each node represents a segment (Figure 5). Understand that some of these segments might have interdependent movement.

Model driving forces

Our hand model is subject to two types of driving forces: the model's internal constraints and the external driving forces derived from images. This means that the model moves according to the external driving forces derived from images as long as the movement does not violate the internal constraints. We will describe these driving forces in detail to explain how their systematic combination within the hand model leads to target hand shapes.

Internal constraints

To obtain the finger positions needed to reach given points, the hand model must calculate corresponding joint vectors—an *inverse kinematics problem*. We simplify this problem by applying reasonable postulates based on the internal constraints. Namely, we use the previously described internal constraints of the hand model to either reduce the hand model's DOF (Constraints 1, 2, and 4) or reject infeasible inverse kinematics solutions (Constraint 3). Among the two kinds of constraints, Constraints 1, 2, and 4 are crucial for successful system operation, since removal of any one of them requires using spatial positions of additional hand joints that lie inside the fingers or the palm, which drastically increases the analysis complexity.

Since a dependency exists between the PIP and DIP joints, joint angle θ_{DIP} can be calculated from θ_{PIP} using the equation in Constraint 2 and vice versa.

Postulate 1. For each finger there is a dependency between θ_{DIP} and θ_{PIP}, represented by the equation in Constraint 2.

The joint vector for the inverse kinematics of a finger takes the form

$$\mathbf{Q}_f = (\theta^X_{MP}, \theta^Z_{MP}, \theta_{PIP}, (2/3)\theta_{PIP}) \text{ for } f = \text{II, III, IV, V}$$

Another way to simplify the problem relies on the fact that in active movement, finger III does not have much adduction or abduction movement. Because the other fingers are planar manipulators with the exception of the MP joint, it follows that θ^Z_{MP} can be calculated from the displacement of the fingertip in the object coordination frame of finger III (see Figure 6).

Postulate 2. Finger III does not naturally make side movements (Constraint 4):

$$\theta^Z_{MP}(\text{III}) = 0$$

Thus, the joint vector for finger III becomes

$$\mathbf{Q}_{III} = (\theta^X_{MP}(\text{III}), 0, \theta_{PIP}(\text{III}), (2/3)\theta_{PIP}(\text{III}))$$

Finger III is now a nonredundant manipulator, since the joint's DOF (2) equals the number of freedoms associated with the end-effector task. For the fingertip to reach an arbitrary point, a unique solution must therefore exist.

However, when the palm is included in a finger's movement chain, it becomes a redundant manipulator. That is, the finger has more DOF (on the palm) than the minimum number required to reach a specific position. The joint vector for the whole hand is

$$\mathbf{Q}_h = (\mathbf{Q}_p, \mathbf{Q}_f)$$

where $\boldsymbol{Q}_p$ is the joint vector of the palm. To further simplify the problem, we impose a final hypothesis on palm movement by assuming that the metacarpals are firmly connected together inside the palm such that the palm minimizes its "hollowness." Hollowness indicates that the shape of the palm is rounded by the flexion of metacarpal bones M(II) – M(V) inside it. In actuality, anatomical studies have shown that the palm has minor hollowness when not grasping an object.[11]

> Postulate 3. Assume that the palm has zero hollowness in nonprehensile motion.

In general, three points on a plane are necessary to align it in space. However, if the palm has zero hollowness and the fingertip of finger III is fixed, the wrist and another MP joint in the palm determine the palm's posture. This can be explained as shown in Figure 7 by the fact that finger III including M(III) is a planar manipulator, where palm rotation around the axis through W and X is limited to ensure planarity of finger III. The configuration of other fingers besides finger III are uniquely decided when the palm is fixed by the position of their fingertips, since their θ^Z_{MP} value is calculated from finger III's value (that is, 0).

The thumb (I) is a much more complex manipulator because much of it is part of the palm, and the joints move along nontrivial axes. However, its inverse kinematics can be calculated almost uniquely by experimental observations,[10] that is,

$$\theta^X_{MP}(\mathrm{I}) = 2\,(\theta^X_M(\mathrm{I}) - (1/6)\pi)$$
$$\theta^Z_{MP}(\mathrm{I}) = \theta^Z_M(\mathrm{I})(7/5)$$

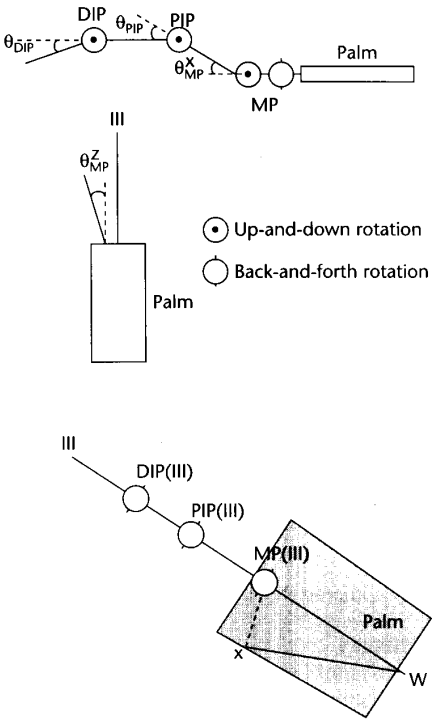
In summary, we can use Postulates 1 through 3 to approximately determine the posture of the human hand.

> Postulate 4. Under Postulates 1 through 3, the hand's posture can be identified by determining the positions of all five fingertips, the wrist, and an extra point on the palm.

Because the MP joint of the little finger (V) is more visible than other MP joints from both the front and side views, we specifically chose it as the extra point on the palm.

External image forces

Other forces acting on the hand model are derived from the image content. Called "image force," this type of force enables us to transform the model into the shape of a real hand. Although in principle we can exploit a variety of image-based cues, including profiles (also known as occluding contours), edges, shading, and texture, we use information exclusively from seven selected points of the hand (Postulate 4), called "characteristic points," to determine hand posture. In image analysis, the use of these characteristic points, including the fingertips, has an advantage over other joint positions or edges of the hand: the fingertips are more visible and do not change position or get "squeezed" out of shape as the surface of the skin slides on the finger joint. Moreover, the characteristic points remain stable as the viewpoint varies, whereas edges do not.

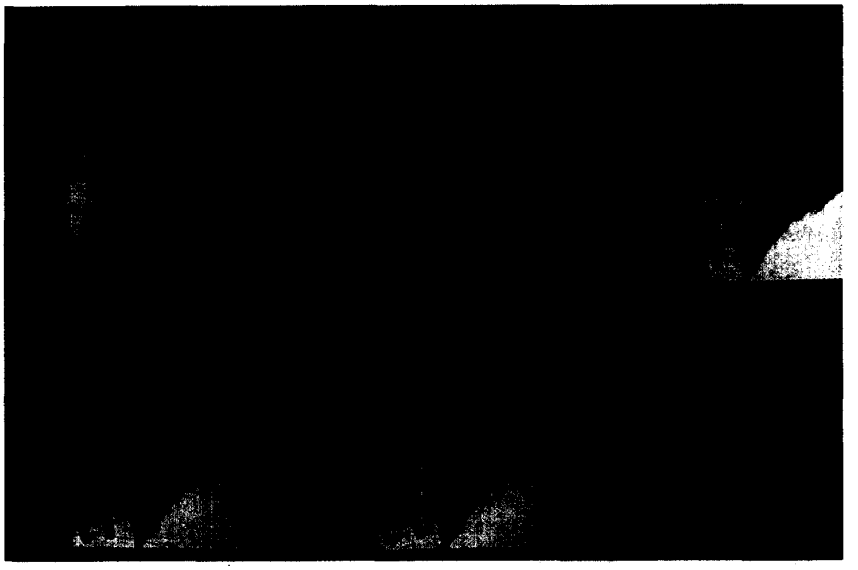


6 Hand mechanisms represented by joint symbols.

7 Diagram showing how the planarity of finger III restricts the palm's rotation.

Color-coded gloves helped us identify the seven characteristic points in images of a real hand. We chose red, blue, green, magenta, cyan, and yellow from the six points of the HSV (hue, saturation, value) single hexcone model and painted them on white gloves at the tips of five fingers, and at the MP joint of finger V. The same green marks the tip of finger III and the wrist, since the width of the area discriminates between the points. The image's RGB coordinates converted into HSV vectors segment it according to hue level. The discrimination ratio is inversely proportional to the number of different colors used when the color vectors are uniformly distributed over the hue range [0, 360]. Two cameras positioned at different viewpoints obtained the spatial positions of the characteristic points.

Imagine that for each characteristic point of the real hand, a "virtual spring" exists that pulls the corresponding characteristic point of the hand model, with each such spring possibly having a different spring constant, for example its weight value. Since the connectivity and joint constraints must be conserved, the forces induced by the springs will subsequently be reflected by translating, rotating, or twisting the model. The model-fitting problem can then be solved if we let these springs act on the model and change its posture so that its characteristic points reach the corresponding characteristic points of the hand calculated from the images. The characteristic points of the real hand are simply termed the "goal posi-



8 Movement of the hand model during model fitting (front view). From left to right, top to bottom: original video image, initial state of the model (superimposed), wrist-positioning phase, intermediate stage in palm/fingers positioning phase, final state of palm/fingers positioning, and final result.

tions" (of the characteristic points of the hand model).

Model fitting

By "model fitting" we mean that the parameters of the hand model are automatically determined such that postulated positions of the hand model match the images of the real hand. Based on our conclusion that the posture of the whole hand can be determined by the position of the hand's seven characteristic points, the model fitting problem reduces to how to reach the characteristic points of the hand model for a given set of goal positions.

In our hand model algorithm incorporating internal constraints, the model-fitting process divides into two distinct, successive phases: (1) a wrist-positioning phase and (2) a palm/fingers positioning phase (see Figure 8). The wrist-positioning phase consists of a sequence of actions that move the hand model from an arbitrary position and orientation to one in the vicinity of the hand image. During this phase, the wrist of the hand model is positioned at the location of the wrist image such that its orientation roughly matches the central axis of the hand image. The more complex palm/fingers positioning phase fits the palm and fingers by iteratively twisting the palm and moving the fingers as explained next.

```
HTnode {
   HTnode *parent, *son[];
   int numsons; /* number of son-nodes */
   Vector origin; /* position of the segment in the
      object coor. sys. */
   Orientation orientation; /* orientation of the
      segment */
   Goal goal; /* goal associated with the node */
   Vector touchpoint; /* position of the segment in the
      world coor. sys. */
      :
      :
}
Goal {
   Vector location; /* location of the goal */
   Float weight; /* weight of the goal */
}
```

9 Variables used in the hand tree for model fitting.

Model-fitting principles

Several considerations must be emphasized in our model-fitting process:

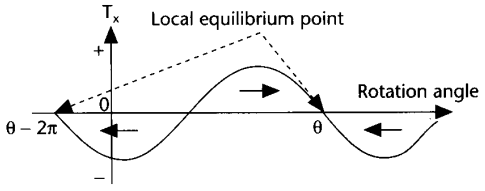
1. The goal positions of the seven characteristic points obtained from real hand images might contain errors. A small inconsistency could also arise given different sizes of the model and the object. Using an analytic calculation to align the segments one by one would therefore not lead to optimal positioning.
2. The best solution would be to simultaneously satisfy multiple goal positions. For the case when the hand model cannot simultaneously satisfy more than two of them, each goal position may have a weighted value representing its "importance."
3. The position of the palm (the parent of the hand tree) should be decided based on the displacements of the fingers (sons of the hand tree). Likewise, the fingers should move in synchronization whenever the palm moves. Consequently, the algorithm must be able to traverse the hand tree bottom-up and top-down in turns.

Based on these considerations, we developed an iterative algorithm for hand-model fitting. Higher-weighted values are assigned to the wrist (10) and to the fingertip of finger III (3), since their alignments have a stronger influence on the posture of the whole hand model.

Hand tree nodes

Each node of the hand tree provides information on the associated segment of the hand model, as well as containing the variables used for model fitting (see Figure 9). The "touchpoint" represents the current position of

the segment in the world coordinate system for which a goal position can be specified. The goal associated with each node has two fields: location and weight. The location represents the goal position of the corresponding segment's touchpoint if a characteristic point has been defined on the segment, while the weight represents the importance of this goal relative to those of other segments. Goals with zero weight are considered inactive, that is, the corresponding segment has no characteristic point.



10 Curve showing the change of the torque vector component in the direction of the rotation axis T_x as the hand model rotates. Arrows represent rotation direction for the hand model to reach its local equilibrium point.

When the hand model with joint angle limits is pulled by virtual springs from goal positions and the wrist position is fixed, the natural movement is a rotation around the axes passing through the wrist. By calculating the direction and size of the net torque to the hand model, we can guide it to an orientation that minimizes the torques caused by virtual springs while maintaining equilibrium. For example, if the hand model rotates around axis x passing through the wrist, the component of the torque vector in the direction of rotation axis T_x will change like a sine function (Figure 10). Therefore, the hand model reaches a local equilibrium point by turning to the left when the sign of T_x is negative and by turning to the right when its sign is positive (represented by arrows in Figure 10).

To calculate torques, we added some variables to the nodes of the hand tree: The local weighted displacement `lwd` is the displacement between the characteristic point of the current segment and its goal position, scaled by the goal weight. The virtual spring is assumed to pull the segment toward its goal position with a force proportional to `lwd`:

$$\boldsymbol{f} = w\,(\boldsymbol{v}_{goal} - \boldsymbol{v}_{cur})$$

where w is the goal weight. The subtree-weighted displacement or `stwd` associated with a node is the sum of the subtree-weighted displacements of all subtree nodes rooted to the current node, including that of the current node itself.

In a similar manner, the local weighted torque `lwt` is the torque at the wrist produced by all the virtual springs connected to the current segment. The subtree-weighted torque `stwt` is the sum of the local weighted torques of all the subtree nodes rooted to the current node. We define torque as

$$\tau = \frac{\boldsymbol{d} \times \boldsymbol{f}}{|\boldsymbol{d}|}$$

where $\boldsymbol{f}$ is the force and $\boldsymbol{d}$ the distance vector between the force and a given axis. Note that τ is not proportional to the distance between the force and rotation axis. That is, in model fitting the characteristic points further away from the wrist are not necessarily more important than those closer to the wrist.

Model-fitting algorithm

Our model-fitting algorithm uses the following iterative algorithms to perform an optimizing process that minimizes the displacement of characteristic points of the hand model from their goal positions. To reduce solution complexity, the algorithm does not consider collision constraints, though we realize that doing so would improve solution accuracy.

Algorithm `Fitting-hand`

1. Determine the subtree-weighted torque `stwt` for the hand tree root.
2. Select the rotation axis passing through the wrist as that axis having the greatest torque component.
3. Incrementally rotate the hand model around the selected rotation axis such that the direction causes the torque on the axis to decrease to zero. At each increment, relocate all fingers to optimal positions (using Algorithm `Fitting-fingers`).
4. If torque of the whole hand has a smaller magnitude than the threshold value or if the iterative limit is reached, then quit; else repeat.

Algorithm `Fitting-fingers`

1. Call the inverse kinematic subroutine to calculate the joint angles of each finger. These angles must be within their respective joint angle limits such that the characteristic point of each finger approaches its goal position as closely as possible.
2. While maintaining the current orientation of the whole hand, extend each finger using the calculated joint angle values.
3. Return.

Experiment

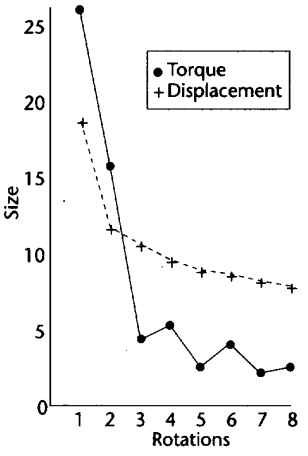
We carried out experiments on an Iris 70GT/40 system to automatically input and analyze hand shapes of American Sign Language letters (ASLL). First, we constructed the hand skeleton model (Figure 3) after precisely measuring the anthropometric parameters of a specific human subject. We used the following anthropometric parameters: size of the segments (M, P1, P2, and P3), joint angle limit of each finger joint, and interdependency of MP joints (Constraint 3). We then fitted the hand model to the stereo hand images, extracted necessary data from the model, and subsequently analyzed the data (Figure 11, next page).

Convergence

Our model-fitting algorithm builds on the basic assumption that the displacement between the model

```
****** CURRENT GEO. DATA ******
hand[0]
  position 0.000000 -80.000000 0.000000
  orientation -15.625000 26.875000 9.534454
  angle(thumb, W) 0.000000 0.000000 -24.285713
  angle(thumb, W1) 78.571434 0.000000 -65.714287
  angle(thumb, MP) 38.571426 -45.000000 0.000000
  angle(thumb, DIP) 8.571426 0.000000 0.000000
  angle(index, MP) 0.000000 0.000000 20.000000
  angle(index, PIP) 72.000000 0.000000 0.000000
  angle(index, DIP) 54.000000 0.000000 0.000000
  angle(middle, MP) 0.000000 0.000000 0.000000
  angle(middle, PIP) 81.000000 0.000000 0.000000
  angle(middle, DIP) 60.750000 0.000000 0.000000
  angle(ring, MP) 27.000000 0.000000 -7.000000
  angle(ring, PIP) 90.000000 0.000000 0.000000
  angle(ring, DIP) 67.500000 0.000000 0.000000
  angle(little, MP) 36.000000 0.000000 6.000000
  angle(little, PIP) 72.000000 0.000000 0.000000
  angle(little, DIP) 54.000000 0.000000 0.000000
  angle(wrist) 180.000000 0.000000 0.000000
  loc(thumb,tip) 13.288818 -67.524338 -19.453569
  loc(index,tip) 53.241577 21.177095 -36.654503
  loc(middle,tip) 37.136475 16.749365 -32.839592
  loc(ring,tip) 19.905762 -21.597826 -12.802109
  loc(small,tip) -3.130005 -74.665726 -34.014271
```

11 Sample data of hand parameters extracted from the hand model after fitting to images (for ASLL "N"). Each joint angle is represented by a triple of rotation angles around the *x*, *y*, and *z* axes of the local coordinate system.



12 Typical results showing that the torque and displacement of the whole hand tree decrease as the hand model's rotation amount increases (ASLL "N").

and image will be at a minimum when the torque of the whole hand tree produced by the displacements becomes zero. This is a nonlinear optimization problem, where given the continuous torque function $\tau: R^3 \rightarrow R^3$, we wish to find $min\{|\tau(v)| : v \in R^3\}$ and $v \in R^3$ of the orientation of the whole hand model where the minimum is achieved. Naturally, the algorithm must converge to obtain a solution. Experiments subsequently showed that the torque and displacement of the whole hand tree continuously decreased as the whole hand model's rotation amount increased (Figure 12).

Efficiency

We define the error rate after i rotations as

$$ER(i) = D(i)/L(i) = \sum_s displacement / \sum_s length$$

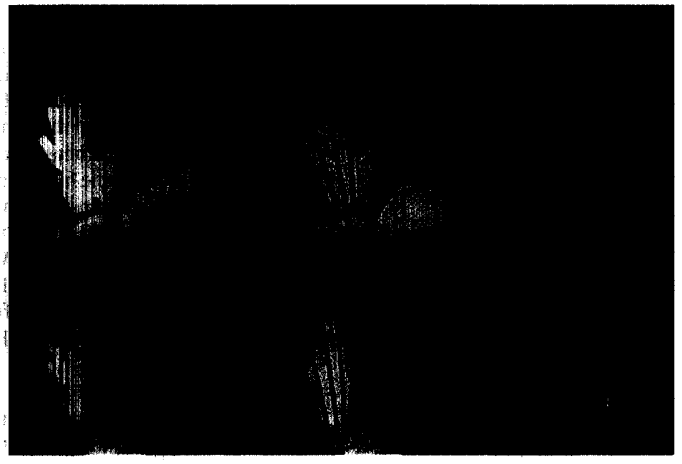
where $\sum_s displacement$ is the summed displacement between a segment s and its goal position, and $\sum_s length$ is the summed length of the chain from root node to the segment s. $ER(i)$ represents how closely the characteristic points of the hand model fit their goal positions after i rotations. Most of the experimental results shown here were obtained in five to eight rotations, with $ER(7)$ for random samples having an average value of 0.023. The values for the best and worst case were 0.0048 (for the open hand) and 0.032 (for ASLL "S").

The iterative search method employed to solve the inverse kinematics of the fingers is not very efficient and ultimately becomes the bottleneck of our model-fitting algorithm. Obtaining a precise result ($ER \leq 0.01$) requires about 45 minutes of runtime on an Iris workstation, with about 99.2 percent of the processing time spent solving the inverse kinematics of the fingers.

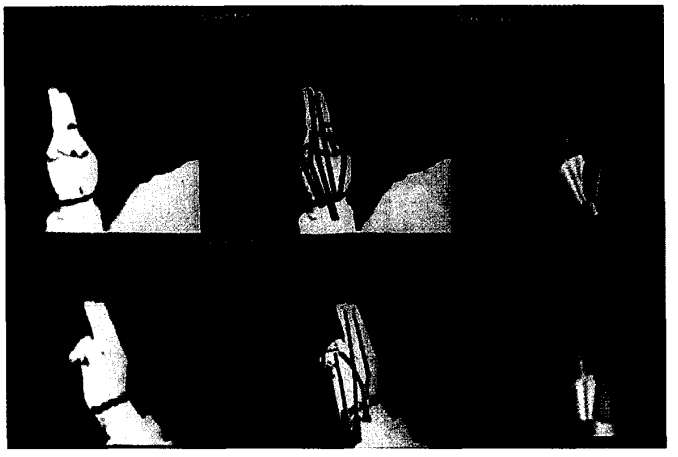
Sources of error

Although the error rate is quite low, various sources of error influence experimental results.

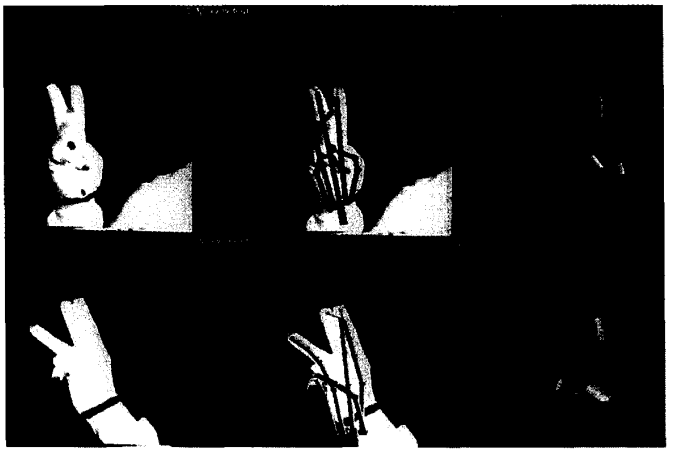
■ *Hollowing of the palm.* We assumed zero hollowness



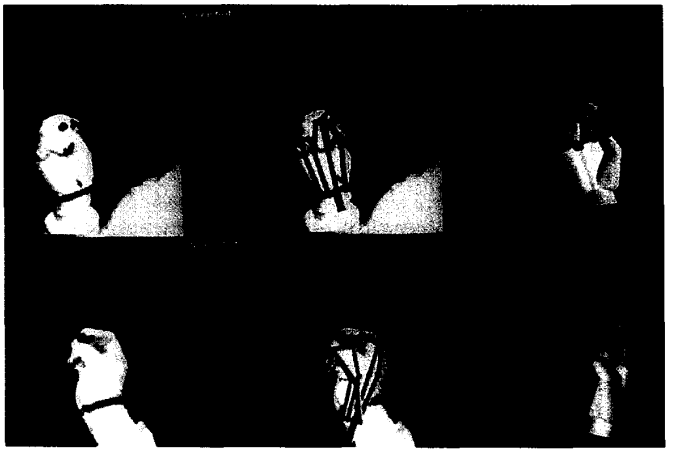
13 Open hand ($D = 4.8$). Upper three frames are front view images, respectively, of the hand, the hand model skeleton fitted to the hand image, and the final hand model. Lower three frames are the corresponding right view images. The hand model corresponds almost perfectly with the real hand.



14 ASLL "H" ($D = 9.8$). **Better results are obtained when the fingers are stretched and the palm is not hollowed so much.**



15 ASLL "W" ($D = 16.1$). **Displacement is greater than that of ASLL "H" because the palm is hollowed more.**



16 ASLL "N" (D = 12.3). **Displacement increases because of the bent fingers, although the hand's overall configuration is correctly captured.**

in the palm (Postulate 3), an assumption that results in errors if the hand is cupped. By comparing Figures 13 through 15 it is apparent that cupping the hand increases displacement. Side movement of the middle finger (Postulate 2) also introduces error in some special hand shapes (for example, ASLL "V"), though it is trivial compared with that due to palm hollowness.

■ *Size inconsistency between the model and hand.* Another error that influences the experimental results comes from the inconsistency between the size of the hand model and that of the real hand. This inconsistency increases the displacement when the fingers bend. As shown in Figure 16, the workspace of a finger is reduced and the finger has little freedom to

adjust to the displacement.

- *Numerical iterations.* The inverse kinematics of the fingers and the rotation of the palm are solved by numerical iterations, therefore, more iterations obtain a better solution.

Future work

Future research will target reducing the sources of error and performing calculations in real time. Possible system enhancements include

- Incorporation of a more accurate model for simulating palm movement.
- Collision prevention of finger segments using a knowledge base of skin surface topology.
- Development of special indicators to identify the seven characteristic points on the real hand in an ordinary indoor environment.
- Use of parallel processing and recently developed techniques[12] such as calculating the inverse kinematics in real time.

Once the speed and accuracy of our technique improve, interesting and diverse applications will naturally follow. ■

Acknowledgments

We would like to express our sincere gratitude to professors of the University of Tokyo, Hiroshi Imai, Kei Hiraki, and Yoshihisa Shinagawa, for their valuable discussions. We also thank Kyung-Il and Jung-Sook Han for their encouragement, and the anonymous reviewers for their helpful comments.

References

1. D. Sturman and D. Zeltzer, "A Survey of Glove-Based Input," *IEEE CG&A*, Vol. 14, No. 1, Jan. 1994, pp. 30-39.
2. M. Krueger, *Artificial Reality*, 2nd edition, Addison Wesley, Reading, Mass., 1990.
3. S. Tamura and S. Kawasaki, "Recognition of Sign Language Motion Images," *Pattern Recognition*, Vol. 21, No. 4, 1988, pp. 343-353.
4. W. Freeman and M. Roth, "Orientation Histograms from Hand Gesture Recognition," Tech. Note MN 94-03, Mitsubishi Electric Research Lab., Cambridge, Mass., 1994.
5. E. Pernkopf, *Pernkopf's Anatomy*, Vol. 2, 2nd ed., Urban & Schwarzenberg, 1989.
6. J. Lee and T. Kunii, "Constraint-Based Hand Animation," in *Models and Techniques in Computer Animation World* (Proc. Computer Animation 93), N. Magnenat Thalmann and D. Thalmann, eds., Springer-Verlag, Berlin, 1993, pp. 110-127.
7. J. Korein, *A Geometric Investigation of Reaching*, MIT Press, Cambridge, Mass., 1985.
8. J.O. Society, "A Method of Display and Measuring Joint Angle Limits," *Rehabilitation Medical Science*, Vol. 11, No. 2, 1974, pp. 127-132. In Japanese.
9. J. Landsmeer, "The Coordination of Finger-Joint Motions," *J. of Bone and Joint Surgery*, Vol. 45-A, Dec. 1963, pp. 1,654-1,662.
10. H. Rijpkema and M. Girard, "Computer Animation of Knowledge-Based Human Grasping," *Computer Graphics* (Proc. Siggraph), Vol. 25, No. 4, July 1991, pp. 339-347.
11. I. Kapandji, *The Physiology of the Joints*, Vol. 1, Churchill Livingstone, New York, 1982.
12. C. Phillips, J. Zhao, and N. Badler, "Interactive Real-Time Articulated Figure Manipulation Using Multiple Kinematic Constraints," *Computer Graphics*, Vol. 24, No. 2, 1990, pp. 245-250.

Jintae Lee *is assistant professor at the University of Aizu. His research interests include virtual reality, multimedia, human interfaces, and computer-aided translation. Lee received a BS in computer science and statistics from Seoul National University in 1981, an MS in computer science from Korea Advanced Institute of Science and Technology in 1983, and a PhD in information science from the University of Tokyo in 1993. He is a member of IEEE, Information Processing Society of Japan, and Korea Information Science Society.*

Tosiyasu L. Kunii *is president and professor at the University of Aizu. He is founder of the Computer Graphics Society, founding editor and editor-in-chief of* The Visual Computer *and* International Journal of Shape Modeling, *and associate editor-in-chief of* The Journal of Visualization and Computer Animation. *Kunii received his BS, MS, and DS degrees in chemistry, all from the University of Tokyo. He is a fellow of IEEE and a member of ACM, BCS, Information Processing Society of Japan, and Institute of Electronics, Information, and Communication Engineers.*

Readers may contact Lee at Multimedia Device Lab, University of Aizu, Aizu-Wakamatsu City, Fukushima 965-80, Japan, e-mail j-lee@u-aizu.ac.jp