TRANSFORMATION OF HUMAN HAND POSITIONS FOR ROBOTIC HAND CONTROL

Lucy Pao
Department of Electrical Engineering
Stanford University
Palo Alto, CA 94305

Thomas H. Speeter
Machine Perception Research Department
AT&T Bell Laboratories
Holmdel, NJ 07733

Abstract

Position control of multifingered robotic hands requires complex position trajectories, defining simultaneous motion of a large number of joints. To teach a multifingered hand manipulation strategies, traditional point to point training is tedious and fails to embody the underlying purpose and constraints driving motion of the hand. In this paper we describe a method for using the human hand as a multidegree of freedom teaching device. The algorithm is based on a functional analysis of the human hand and results in an algebraic transformation of human hand positions to corresponding positions in a target domain. The target domain should be of lower dimensionality (less degrees of freedom) than the human hand but is not constrained in any other way. The target described here is a sixteen degree of freedom robotic hand, with four fingers of for joints each. The target need not, however, be a "hand-like" device but for ease of use, should have a kinematic structure with poses similar in functionality to natural poses of the human hand.

1.0 Introduction

Among the most challenging problems in robotics is to construct a robotic hand having capabilities close to those of the human hand. Several multi-fingered robotic hands have been developed with this aim [3, 4, 8, 9]. Among the difficulties in controlling such devices, however, is teaching them complex manipulation strategies. Control of multifingered robotic hands requires multivariate trajectories, defining simultaneous motion of a large number of joints.

Robotic hand motion can potentially be taught in a point-to-point manner, either by (a) a sequence of joint angles for each joint or (b) by a sequence of fingertip positions for each finger. These methods are time-consuming to perform and do not allow task related modifications or functional input relating to what the hand is trying to do.

A more desirable method for teaching a robotic hand is through direct training by a human hand. This can be done by measuring and recording a sequence of human hand positions and transforming them to corresponding robotic hand positions. The problem then is to transform human hand positions to functionally equivalent robotic hand positions without loss of generality or functional content. The most obvious way is to set up either a joint-joint correspondence or a fingertip-fingertip correspondence between the hands. If there are a different number of joints, however, a joint-joint correspondence requires a set of additional heuristics to map one domain onto the other. This may be feasible after adequate study of the hands' relative structures; however, certain degrees of freedom are likely to be discarded or averaged, resulting in the loss of functionally relevant information. Fingertip-fingertip transformations potentially suffer from the same loss and additionally carry the burden of forward and inverse kinematic calculations to transform finger positions from one hand to the other.

We would like to develop an algorithm that translates human hand poses to corresponding robotic hand positions, without loss of functional information and without the overhead of kinematic calculations. Moreover, we would like to deal at the level of poses rather than at the joint or fingertip levels. In our study we have used the Utah/MIT dextrous hand as the target for training. The Utah/MIT hand is a four fingered robotic hand, each finger having four degrees of freedom. The algorithm applies, however, to all multidegree of freedom devices, the only practical constraint being that they have poses of similar functionality to poses of the human hand.

2.0 Algorithm Development

A major consideration in developing a transformation algorithm is that it be simple and equally useful for all human trainers and robotic targets. Additionally, we are less concerned with appearances than with functionality. That is, we are more interested in having robotic hand positions possess the same functionality rather than the appearance of corresponding human hand positions.

To teach a robotic hand manipulation strategies, we have defined a set of poses that we feel are common and functional and which span the useful range of human hand motion (see Table 1 and Figure 1). There is a wide functional as well as structural variety in this group of poses. Both power and precision poses are included [5], along with several gestures. After the poses are defined, for both the human hand and robotic hand, an interpolation technique is developed to determine corresponding robotic hand positions given arbitrary human hand poses (i.e., those not in the standard set).

A pose, of either the human or robotic hand, is defined by a row vector of joint angles. Let $\mathbf{n_h}$ denote the number of joint angles used to define a human hand pose and $\mathbf{n_r}$ be the number of joint angles used to define a robotic hand pose. Once the poses are defined for both the human hand and robotic hand, the following matrix equation can be written:

$$\mathbf{HT} = \mathbf{R} \tag{1}$$

Here, $\mathbf{H}$ is the human hand matrix. It is an $\mathbf{n_p} \times \mathbf{n_h}$ matrix, where $\mathbf{n_p}$ is the number of poses. Each row corresponds to a human hand pose and each column corresponds to a particular joint in the human hand. Similarly, $\mathbf{R}$ is the robotic hand matrix which is $\mathbf{n_p} \times \mathbf{n_r}$. Each row corresponds to a robotic hand pose and each column to a joint angle of the robotic hand.

$\mathbf{H}$ and $\mathbf{R}$ are known, and the $\mathbf{n_h} \times \mathbf{n_r}$ transformation matrix $\mathbf{T}$ can be solved using the singular value decomposition on $\mathbf{H}$ as follows:

$$\mathbf{T} = \mathbf{H^+R}$$
$$\mathbf{T} = \mathbf{V\Sigma^+U^T R} \qquad (2)$$

where $\mathbf{H} = \mathbf{U\Sigma V^T}$ and $\mathbf{H^+}$ and Σ^+ are the pseudoinverses of $\mathbf{H}$ and Σ, respectively. $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, and Σ is a diagonal matrix of the singular values of $\mathbf{H}$.

Each column of $\mathbf{T}$ is directly related to the corresponding column in $\mathbf{R}$. Equation (1) basically represents $\mathbf{n_r}$ systems of equations, one for each column of $\mathbf{T}$ and $\mathbf{R}$. Let $\mathbf{t_i}$ and $\mathbf{r_i}$ denote the ith columns of $\mathbf{T}$ and $\mathbf{R}$, respectively. Then we can write $\mathbf{n_r}$ equations of the form:

$$\mathbf{H\, t_i = r_i} \qquad (3)$$

Hence, in solving for $\mathbf{T}$, the dependence of each robotic hand joint on each of the human hand joints is determined.

The number of poses defined determines whether the matrix equation (1) represents overdetermined or underdetermined systems of equations. If $\mathbf{n_p} > \mathbf{n_h}$, the systems of equations are overdetermined. If $\mathbf{n_p} < \mathbf{n_h}$, they are underdetermined. If $\mathbf{n_p} = \mathbf{n_h}$, then the transformation is exactly determined by $\mathbf{T} = \mathbf{H^{-1}R}$ (if $\mathbf{H}$ is nonsingular). In this case the transformation is unique since the inverse of $\mathbf{H}$ is unique. Whether $\mathbf{n_p}$ is greater or less than $\mathbf{n_h}$, the solution for $\mathbf{T}$ in (2) minimizes the 2-norm of the error in transforming from the standard poses of the human hand to those in the robotic domain [2].

The condition number of $\mathbf{T}$ indicates how sensitive the transformation is to small changes in human hand positions. The larger the condition number, the more sensitive the transformation. For smooth transitions in the robotic hand motion while using the transformation, the condition number should be reasonably small. What is "reasonably small" depends on the size of the transformation matrix. A larger transformation matrix may have a higher condition number while maintaining smooth robotic hand transitions.

The minimum number of linearly independent poses needed for $\mathbf{T}$ to be full rank is given by

$$\mathbf{n_p} \geq \min(\mathbf{n_h}, \mathbf{n_r}) \qquad (4)$$

If there are fewer than this number of poses, the condition number of the transformation matrix is infinite, and small movements of the human hand can cause large movements in the robotic hand. This is undesirable since we would like the robotic hand transitions to be smooth.

Even if there are $\mathbf{n_p} = \min(\mathbf{n_r}, \mathbf{n_h})$ linearly independent poses, the systems of equations in (1) may still be underdetermined. For instance, suppose $\mathbf{n_r}$ is less than $\mathbf{n_h}$, meaning that the robotic hand has fewer degrees of freedom than the human hand. The minimum number of poses for $\mathbf{T}$ to be full rank is then $\mathbf{n_p} = \min(\mathbf{n_h}, \mathbf{n_r}) = \mathbf{n_r} < \mathbf{n_h}$. Hence, $\mathbf{n_p} < \mathbf{n_h}$ and the systems of equations in (1) are underdetermined. Because the transformation from human hand positions to robotic hand positions is from a higher dimensional space to a lower one, $\mathbf{T}$ is rectangular, being $\mathbf{n_h} \times \mathbf{n_r}$. For $\mathbf{T}$ to be full rank, there need to be only $\min(\mathbf{n_r}, \mathbf{n_h}) = \mathbf{n_r}$ linearly independent columns and rows. This means there needs to be only $\min(\mathbf{n_r}, \mathbf{n_h})$ nonzero singular values of $\mathbf{H}$ for $\mathbf{T}$ to be full rank (see Equation (2)). Therefore, the systems of equations in (1) may be underdetermined and still yield a transformation matrix of full rank.

The difference between $\mathbf{n_h}$ and $\mathbf{n_r}$ does not matter as much in solving for $\mathbf{T}$ as it does in defining poses (see Equation (4)). In defining poses, a large difference between $\mathbf{n_h}$ and $\mathbf{n_r}$ will limit the range of possible poses. If $\mathbf{n_h} > \mathbf{n_r}$, the human hand will have a wider variety of possible poses. If $\mathbf{n_h} < \mathbf{n_r}$, the robotic hand will have a greater diversity of poses. As long as the poses of the robotic hand span the range of useful poses of the human hand, corresponding poses can be defined for both hands and $\mathbf{T}$ can reasonably be solved using Equation (2).

The condition number of the transformation matrix, in general, decreases as more poses are defined. This is shown in Figure 2. Here, 13 joint angles are used to define a human hand pose and 16 joint angles to define a robotic hand pose. By overdetermining the human hand matrix, $\mathbf{H}$, the transformation matrix $\mathbf{T}$ transforms arbitrary human hand poses to robotic hand poses using a least squares fit [2].

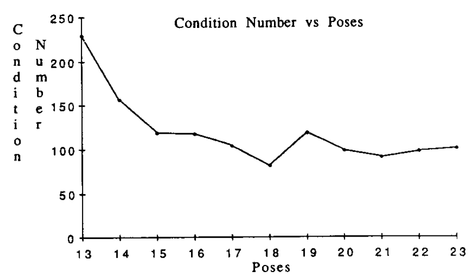

Figure 2

It is important to note that the condition numbers given in Figure 2 have been computed for one human hand. These numbers will vary from human hand to human hand. Further, these numbers will vary with different poses chosen as well as for different robotic hands. Because of all these factors, it is impossible to analytically predict the behavior of the condition number of $\mathbf{T}$ as new poses are added and only heuristic arguments can be made. From our experience, it seems that in order to sustain a decrease in condition number, additional poses, as they are chosen, should be as dissimilar to existing poses as possible.

Once $\mathbf{T}$ is determined, multiplication of a row vector, describing an arbitrary human hand pose, by $\mathbf{T}$, gives a row vector which defines the corresponding robotic hand pose. Since the computations involved in determining $\mathbf{T}$ are simple, a transformation matrix can be learned and stored for each user's hand. For a given set of poses, the $\mathbf{R}$ matrix remains constant for a particular robotic hand. Hence, for each user, an $\mathbf{H}$ matrix is defined, and $\mathbf{T}$ is solved using Equation (2).

The entries in the transformation matrix define correlations between joints of the human hand and those of the robotic hand. Positive entries mean positive correlation and negative entries denote negative correlation. By studying the transformation matrix, a great deal can be learned about the differences in functionality and structure of poses between the human hand and the robotic hand.

3.0 Results and Discussion

Here we describe application of the above algorithm to the Utah/MIT dextrous hand. The dextrous hand has 4 fingers of 4 degrees of freedom each, thus $\mathbf{n_r}$ is 16. Although the Utah/MIT hand was designed to emulate the human hand, there are several fundamental structural dissimilarities [4, 5]. These include (a) a difference in the number of digits, (b) differences in thumb placement on the palm and the resulting angle of opposition of the thumb with the remaining fingers, and (c) differences in the range of motion of the joints. The transformation algorithm must adequately embody these structural differences in order to be useful for training the robotic hand in manipulation strategies.

There are several commercially available devices that can be used to measure human hand joint positions [10, 11]. Nevertheless, the human hand is extremely complex, making accurate measurements of its many degrees of freedom difficult. This difficulty is in fact the fundamental limitation when attempting to use the human hand as a teaching device.

We have experimented with the VPL DataGLove as an initial transducer. The DataGlove is a hand-to-machine interface device which gives position and orientation information about the human hand. It has 13 transducers, 2 on each of the five fingers to give joint flexion and extension data and 3 additional sensors to transduce abduction and adduction between the thumb, index, and middle fingers. Hence, $\mathbf{n_h}$ is 13 in this case.

In our application, $\mathbf{n_h} < \mathbf{n_r}$, so we are transforming from a lower dimensional space to a higher one. Thus, we need at least $\mathbf{n_p} = \min(\mathbf{n_h}, \mathbf{n_r}) = 13$ poses for the transformation to be full rank, as confirmed in Figure 2. The requirement on $\mathbf{n_p}$ as given in (4) will always yield systems of equations in (1) which are determined or overdetermined.

Use of the transformation algorithm on the Utah/MIT dextrous hand results in smooth and functionally related transformations of human hand positions. The Utah/MIT hand moves through poses smoothly and with corresponding functionality. This may suggest at this preliminary stage that the 13 sensor measurements are minimally adequate to define human hand positions.

There is, however, room for improvement, since there are significant limitations in the sensing capabilities of the DataGlove. Since there are 4 degrees of freedom on each of the 5 digits of the human hand, using 20 joint angles to define a human hand pose would be more accurate than the 13 values given by the DataGlove measurements. This incompleteness as well as coupling among sensors and sensitivity of sensors to hand orientation degrade the performance of our algorithm.

The greatest shortcoming of the DataGlove as a measurement device for this algorithm is that it does not completely measure the joint angles associated with the carpometacarpal (CMCP) joint at the base of the thumb. This joint is probably the most important joint of the human hand [6, 7]. A measurement device which completely measures the two joint angles associated with this joint in addition to the 13 joint angles already measured would probably give significantly better results. A measurement device which measures all 20 joint angles would give even better results.

4.0 Conclusions

An algorithm for transforming human hand positions to corresponding robotic hand positions is described. The algorithm allows direct instruction by the human hand for teaching manipulation strategies to a robotic hand. It can be used with any combination of human and robotic hands without loss of applicability. The algorithm results in a linear transformation between human hand poses and robotic hand poses and performs a least squares fit when the transformation between them is either under- or over-specified.

The transformation is sensitive to the specific poses chosen to define the transformation matrix. Nevertheless, given a reasonable variety of poses, the transformation from arbitrary human hand poses to corresponding robotic hand poses is accurate and smooth and embodies the intended functionality as demonstrated by trials with the Utah/MIT dextrous hand and the VPL DataGlove.

There is a great need for a better human hand position measurement device. The results of our algorithm are promising, but there are problems which can be attributed mainly to our human hand measuring device. A second human hand measuring device has recently been introduced by Arthur D. Little, Inc. It is called the Dextrous Hand Master (DHM) and measures four joint angle values on four digits (thumb, index, middle, and ring). The DHM not only gives complete measurements of the important CMCP joint, but also promises to give uncoupled measurements among the sensors.

Development of the described algorithm is one step toward successful training of multifingered robotic hands. The algorithm's main features are its simplicity and generality, making it potentialy useful for a variety of applications. It allows for the control of robotic hands directly by the human hand and for training in the performance of complex manipulation tasks. It provides a natural interface between human and robotic hands without the need for kinematic calculations or point to point trajectory specification.

Table 1. Standard Poses

1. fist
2. index/middle grip (cigarette pose)
3. thumb/index grip
4. thumb/middle grip
5. thumb/index/middle grip
6. thumb/index/middle/ring grip
7. plane
8. lateral pinch
9. open plane
10. neutral
11. plane grip
12. counterclockwise twist
13. cup
14. inward roll
15. rock
16. dynamic tripod (pencil hold)
17. ball grip
18. cylinder grip
19. index point
20. clockwise twist
21. thumb/index pinch
22. thumb/middle pinch
23. thumb/index/middle pinch

Figure 1. Standard Poses

Utah/MIT Hand Human Hand

Utah/MIT Hand Human Hand

fist

lateral pinch

counterclockwise twist

clockwise twist

ball grip

cylinder grip

rock

dynamic tripod (pencil hold)

neutral

open plane

Figure 1. Standard Poses (continued)

Utah/MIT Hand Human Hand

inward roll

cup

plane

index point

plane grip

Utah/MIT Hand Human Hand

thumb/index pinch

thumb/middle pinch

thumb/index/middle pinch

index/middle grip (cigarette pose)

thumb/index/middle/ring grip

Figure 1. Standard Poses (continued)

Utah/MIT Hand Human Hand

thumb/index grip

thumb/middle grip

thumb/index/middle grip

References

[1]Elliott, J. M., and K. J. Connolly. "A Classification of Manipulative Hand Movements." *Developmental Medicine & Child Neurology*, pages 283-296, vol. 26, 1984.

[2]Golub, G. H., and C. F. Van Loan. *Matrix Computations*. Baltimore, Maryland: The Johns Hopkins University Press, 1983.

[3]Hanafusa, H., et. al. "A Robot Hand with Elastic Fingers and its Application to Assembly Process." *Proc. IFAC Symp. Inform. Contr. Prob. Manufact. Tech.*, pages 127-134, 1977.

[4]Jacobsen, S. C., et. al. "The Utah/MIT Dextrous Hand: Work In Progress." *International Journal of Robotics Research*. Vol. 3(4), 1984.

[5]Landsmeer, J. M. F. "Power Grip and Precision Handling." *Annals of the Rheumatic Diseases* , pages 164-169, 1962.

[6]Napier, J. R. "Prehensility and Opposability in the Hands of Primates." *Symp. Zool. Soc. Lond.* Vol. 5, pages 119-132, 1961.

[7]Ocheda, H. "The Emulation of Human Hand Motions by a Dextrous Robot Hand." Bachelor's Thesis. Department of Electrical Engineering, Princeton University, Princeton, NJ. May 1988.

[8]Okada, T. and S. Tsuchiya, "On a versatile finger system." *Proceedings of the 7th International Symposium on Industrial Robots*, pages 345-352, October 1987.

[9]Salisbury, J. K., et. al. "Articulated Hands: Force Control and Kinematic Issues." *International Journal of Robotics Research*. Vol. 1 (1): pages 4-17, Spring 1982.

[10]Zimmerman, T. G., et. al. "DataGlove Model 2: Operating Manual." VPL Research, Inc., Redwood City, CA, 1987.

[11]A.D. Little Inc. DHM Technical Specification, 1988.