
Coordinate Transformation Learning of Hand Position Feedback Controller by Using Change of Position Error Norm

Eimei Oyama*
Mechanical Eng. Lab.
Namiki 1-2, Tsukuba Science City
Ibaraki 305-8564 Japan

Susumu Tachi
The University of Tokyo
Hongo 7-3-1, Bunkyo-ku
Tokyo 113-0033 Japan

Abstract

In order to grasp an object, we need to solve the inverse kinematics problem, i.e., the coordinate transformation from the visual coordinates to the joint angle vector coordinates of the arm. Although several models of coordinate transformation learning have been proposed, they suffer from a number of drawbacks. In human motion control, the learning of the hand position error feedback controller in the inverse kinematics solver is important. This paper proposes a novel model of the coordinate transformation learning of the human visual feedback controller that uses the change of the joint angle vector and the corresponding change of the square of the hand position error norm. The feasibility of the proposed model is illustrated using numerical simulations.

1 INTRODUCTION

The task of calculating every joint angle that would result in a specific hand position is called the inverse kinematics problem. An important topic in neuroscience is the study of the learning mechanisms involved in the human inverse kinematics solver.

We questioned five pediatricians about the motor function of infants suffering from serious upper limb disabilities. The doctors stated that the infants still were able to touch and stroke an object without hindrance. In one case, an infant without a thumb had a major kinematically influential surgical operation, transplanting an index finger as a thumb. After the operation, the child was able to learn how to use the index finger like a thumb [1]. In order to explain the human motor learning

*Phone:+81-298-58-7298, Fax:+81-298-58-7201, e-mail:eimei@mel.go.jp
Advances in Neural Information Processing Systems 11, MIT Press, pp.1038-1044, 1999

capability, we believe that the coordinate transformation learning of the feedback controller is a necessary component.

Although a number of learning models of the inverse kinematics solver have been proposed, a definitive learning model has not yet been obtained. This is from the point of view of the structural complexity of the learning model and the biological plausibility of employed hypothesis. The *Direct Inverse Modeling* employed by many researchers [2] requires the complex switching of the input signal of the inverse model. When the hand position control is performed, the input of the inverse model is the desired hand position, velocity, or acceleration. When the inverse model learning is performed, the input is the observed hand position, velocity, or acceleration. Although the desired signal and the observed signal could coincide, the characteristics of the two signals are very different. Currently, no research has successfully modeled the switching system. Furthermore, that learning model is not “goal-directed”; i.e., there is no direct way to find an action that corresponds to a particular desired result. The *Forward and Inverse Modeling* proposed by Jordan [3] requires the back-propagation signal, a technique does not have a biological basis. That model also requires the complex switching of the desired output signal for the forward model. When the forward model learning is performed, the desired output is the observed hand position. When the inverse kinematics solver learning is performed, the desired output is the desired hand position. The *Feedback Error Learning* proposed by Kawato [4] requires a pre-existing accurate feedback controller.

It is necessary to obtain a learning model that possesses a number of characteristics: (1) it can explain the human learning function; (2) it has a simple structure; and (3) it is biologically plausible. This paper presents a learning model of coordinate transformation function of the hand position feedback controller. This model uses the joint angle vector change and the corresponding change of square of the hand position error norm.

2 BACKGROUND

2.1 Discrete Time First Order Model of Hand Position Controller

Let $\boldsymbol{\theta} \in \mathbf{R}^m$ be the joint angle vector and $\mathbf{x} \in \mathbf{R}^n$ be the hand position/orientation vector given by the vision system. The relationship between \mathbf{x} and $\boldsymbol{\theta}$ is expressed as $\mathbf{x} = \mathbf{f}(\boldsymbol{\theta})$ where \mathbf{f} is a C^1 class function. The Jacobian of the hand position vector is expressed as $\mathbf{J}(\boldsymbol{\theta}) = \partial \mathbf{f}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$. Let \mathbf{x}_d be the desired hand position and $\mathbf{e} = \mathbf{x}_d - \mathbf{x} = \mathbf{x}_d - \mathbf{f}(\boldsymbol{\theta})$ be the hand position error vector. In this paper, an inverse kinematics problem is assumed to be a least squares minimization problem that calculates $\boldsymbol{\theta}$ in order to minimize the square of the hand position error norm $S(\mathbf{x}_d, \boldsymbol{\theta}) = |\mathbf{e}|^2 / 2 = |\mathbf{x}_d - \mathbf{f}(\boldsymbol{\theta})|^2 / 2$.

First, the feed-forward controller in the human inverse kinematics solver is disregarded and the following first order control system, consisting of a learning feedback controller, is considered:

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \Delta \boldsymbol{\theta}(k) \quad (1)$$

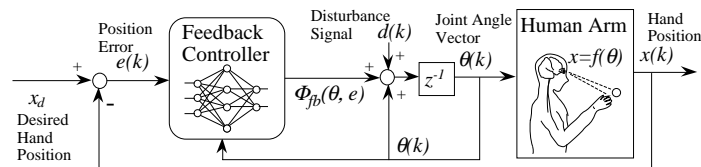


Figure 1: Configuration of 1-st Order Model of Hand Position Controller

$$\Delta\boldsymbol{\theta}(k) = \Phi_{fb}(\boldsymbol{\theta}(k), \mathbf{e}(k)) + \mathbf{d}(k) \quad (2)$$

$$\mathbf{e}(k) = \mathbf{x}_d - \mathbf{f}(\boldsymbol{\theta}(k)) \quad (3)$$

where $\mathbf{d}(k)$ is assumed to be a disturbance noise from all components except the hand position control system. Figure 1 shows the configuration of the control system. In this figure, z^{-1} is the operator that indicates a delay in the discrete time signal by a sampling interval of Δt . Although the human hand position control system includes higher order complex dynamics terms which are ignored in Equation (2), McRuer's experimental model of human compensation control suggests that the term that converts the hand position error to the hand velocity is a major term in the human control system [5]. We consider Equation (2) to be a good approximate model for the analysis of human coordinates transformation learning.

The learner $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}) \in \mathbf{R}^m$, which provides the hand position error feedback, is modeled using the artificial neural network. In this paper, the hand position error feedback controller learning by observing output $\mathbf{x}(k)$ is considered without any prior knowledge of the function $\mathbf{f}(\boldsymbol{\theta})$.

2.2 Learning Model of the Neural Network

Let $\Phi'_{fb}(\boldsymbol{\theta}, \mathbf{e})$ be the desired output of the learner $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$. $\Phi'_{fb}(\boldsymbol{\theta}, \mathbf{e})$ functions as a teacher for $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$. Let $\Phi_{fb}^+(\boldsymbol{\theta}, \mathbf{e})$ be the updated output of $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$ by the learning. Let $E[t(\boldsymbol{\theta}, \mathbf{e})|\boldsymbol{\theta}, \mathbf{e}]$ be the expected value of a scalar, a vector, or a matrix function $t(\boldsymbol{\theta}, \mathbf{e})$ when the input vector $(\boldsymbol{\theta}, \mathbf{e})$ is given. We assume that $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$ is an idealistic learner which is capable of realizing the desired output signal, completely. $\Phi_{fb}^+(\boldsymbol{\theta}, \mathbf{e})$ can be expressed as follows:

$$\Phi_{fb}^+(\boldsymbol{\theta}, \mathbf{e}) \approx E[\Phi'_{fb}(\boldsymbol{\theta}, \mathbf{e})|\boldsymbol{\theta}, \mathbf{e}] = \Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}) + E[\Delta\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})|\boldsymbol{\theta}, \mathbf{e}] \quad (4)$$

$$\Delta\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}) = \Phi'_{fb}(\boldsymbol{\theta}, \mathbf{e}) - \Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}) \quad (5)$$

When the expected value of $\Delta\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$ is expressed as:

$$E[\Delta\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})|\boldsymbol{\theta}, \mathbf{e}] \approx \mathbf{G}_{fb}\mathbf{e} - \mathbf{R}_{fb}\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}), \quad (6)$$

$\mathbf{R}_{fb} \in \mathbf{R}^{m \times m}$ is a positive definite matrix, and the inequality

$$\left| \frac{\partial\Phi_{fb}^+(\boldsymbol{\theta}, \mathbf{e})}{\partial\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})} \right| = \left| \frac{\partial(\mathbf{G}_{fb}\mathbf{e} - (\mathbf{R}_{fb} - \mathbf{I})\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}))}{\partial\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})} \right| < 1 \quad (7)$$

is satisfied, the final learning result can be expressed as:

$$\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}) \approx \mathbf{R}_{fb}^{-1}\mathbf{G}_{fb}\mathbf{e} \quad (8)$$

by the iteration of the update of $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$ expressed in Equation (4).

3 USE OF CHANGE OF POSITION ERROR NORM

3.1 A Novel Learning Model of Feedback Controller

The change of the square of the hand position error norm $\Delta S = S(\mathbf{x}_d, \boldsymbol{\theta} + \Delta\boldsymbol{\theta}) - S(\mathbf{x}_d, \boldsymbol{\theta})$ reflects whether or not the change of the joint angle vector $\Delta\boldsymbol{\theta}$ is in proper direction. The propose novel learning model can be expressed as follows:

$$\Phi'_{fb}(\boldsymbol{\theta}, \mathbf{e}) = -\alpha\Delta S\Delta\boldsymbol{\theta} \quad (9)$$

where α is a small positive real number. We now consider a large number of trials of Equation (2) with a large variety of initial status $\boldsymbol{\theta}(0)$ with learnings conducted at the point of the input space of the feedback controller $(\boldsymbol{\theta}, \mathbf{e}) = (\boldsymbol{\theta}(k-1), \mathbf{e}(k-1))$ at time k . ΔS and $\Delta\boldsymbol{\theta}$ can be calculated as follows.

$$\Delta S = S(k) - S(k-1) = \frac{1}{2}(|\mathbf{e}(k)|^2 - |\mathbf{e}(k-1)|^2) \quad (10)$$

$$\Delta\boldsymbol{\theta} = \Delta\boldsymbol{\theta}(k-1) \quad (11)$$

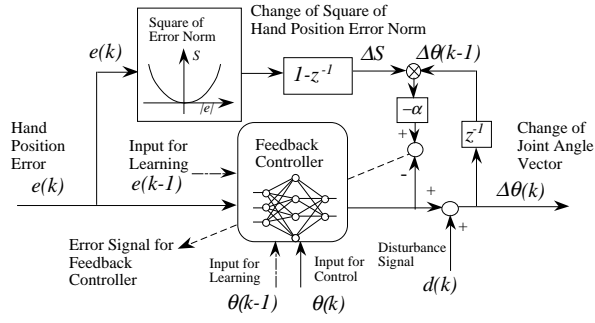


Figure 2: Configuration of Learning Model of Feedback Controller

Figure 2 shows the conceptual diagram of the proposed learning model.

Let $p(\mathbf{q}|\boldsymbol{\theta}, \mathbf{e})$ be the probability density function of a vector \mathbf{q} at the point $(\boldsymbol{\theta}, \mathbf{e})$ in the input space of $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$. In order to simplify the analysis of the proposed learning method, $\mathbf{d}(k)$ is assumed to satisfy the following equation:

$$p(\mathbf{d}|\boldsymbol{\theta}, \mathbf{e}) = p(-\mathbf{d}|\boldsymbol{\theta}, \mathbf{e}) \quad (12)$$

When $\Delta\boldsymbol{\theta}$ is small enough, the result of the learning using Equation (9) can be expressed as:

$$\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}) \approx \alpha \left(\frac{\alpha}{2} \mathbf{R}_\theta \mathbf{J}^T(\boldsymbol{\theta}) \mathbf{J}(\boldsymbol{\theta}) + \mathbf{I} \right)^{-1} \mathbf{R}_\theta \mathbf{J}^T(\boldsymbol{\theta}) \mathbf{e} \quad (13)$$

$$\mathbf{R}_\theta = E[\Delta\boldsymbol{\theta} \Delta\boldsymbol{\theta}^T | \boldsymbol{\theta}, \mathbf{e}] \quad (14)$$

where $\mathbf{J}^T(\boldsymbol{\theta})\mathbf{e}$ is a vector in the steepest descent direction of $S(\mathbf{x}_d, \boldsymbol{\theta})$. When $\mathbf{d}(k)$ is a non-zero vector, \mathbf{R}_θ is a positive definite symmetric matrix and $(\frac{\alpha}{2} \mathbf{R}_\theta \mathbf{J}^T \mathbf{J} + \mathbf{I})^{-1}$ is a positive definite matrix. When α is appropriate, $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e})$ as expressed in Equation (13) can provide appropriate output error feedback control. The derivation of the above result will be illustrated in Section 3.2. A partially modified steepest descent direction can be obtained without using the forward model or the back-propagation signal, as Jordan's forward modeling [3].

Let \mathbf{R}_d be the covariance matrix of the disturbance noise $\mathbf{d}(k)$. When α is infinitesimal, $\mathbf{R}_\theta \approx \mathbf{R}_d$ is established and an approximate solution $\Phi_{fb}(\boldsymbol{\theta}, \mathbf{e}) \approx \alpha \mathbf{R}_d \mathbf{J}^T(\boldsymbol{\theta}) \mathbf{e}$ is obtained.

3.2 Derivation of Learning Result

The change of the square of the hand position error norm $\Delta S(\mathbf{x}_d, \boldsymbol{\theta})$ by $\Delta\boldsymbol{\theta}$ can be determined as:

$$\begin{aligned} \Delta S(\mathbf{x}_d, \boldsymbol{\theta}) &= \frac{\partial S(\mathbf{x}_d, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Delta\boldsymbol{\theta} + \frac{1}{2} \Delta\boldsymbol{\theta}^T \mathbf{H}(\mathbf{x}_d, \boldsymbol{\theta}) \Delta\boldsymbol{\theta} + O(\Delta\boldsymbol{\theta}^3) \\ &= -\mathbf{e}^T \left(\mathbf{J}(\boldsymbol{\theta}) + \frac{1}{2} \frac{\partial \mathbf{J}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \otimes \Delta\boldsymbol{\theta} \right) \Delta\boldsymbol{\theta} + \frac{1}{2} \Delta\boldsymbol{\theta}^T \mathbf{J}^T(\boldsymbol{\theta}) \mathbf{J}(\boldsymbol{\theta}) \Delta\boldsymbol{\theta} + O(\Delta\boldsymbol{\theta}^3) \end{aligned} \quad (15)$$

where \otimes is a 2-operand operator that indicates the Croneker's product. $\mathbf{H}(\mathbf{x}_d, \boldsymbol{\theta}) \in \mathbf{R}^{m \times m}$ is the Hessian of $S(\mathbf{x}_d, \boldsymbol{\theta})$. $O(\Delta\boldsymbol{\theta}^3)$ is the sum of third and higher order terms of $\Delta\boldsymbol{\theta}$ in each equation. When $\Delta\boldsymbol{\theta}$ is small enough, the following approximate equations are obtained:

$$\Delta \mathbf{x} \approx \mathbf{J}(\boldsymbol{\theta}) \Delta\boldsymbol{\theta} \approx \mathbf{J}(\boldsymbol{\theta} + \frac{1}{2} \Delta\boldsymbol{\theta}) \Delta\boldsymbol{\theta} \approx \left(\mathbf{J}(\boldsymbol{\theta}) + \frac{1}{2} \frac{\partial \mathbf{J}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \otimes \Delta\boldsymbol{\theta} \right) \Delta\boldsymbol{\theta} \quad (16)$$

Therefore, ΔS can be approximated as follows:

$$\Delta S \approx -\mathbf{e}^T \mathbf{J}(\boldsymbol{\theta}) \Delta\boldsymbol{\theta} + \frac{1}{2} |\Delta \mathbf{x}|^2 \quad (17)$$

Since $e^T J \Delta \theta \Delta \theta = \Delta \theta \Delta \theta^T J^T e$ and $|\Delta x|^2 \Delta \theta = \Delta \theta \Delta \theta^T J^T J \Delta \theta$ are determined, $\Delta S \Delta \theta$ can be approximated as:

$$\Delta S \Delta \theta \approx -\Delta \theta \Delta \theta^T J^T(\theta) e + \frac{1}{2} \Delta \theta \Delta \theta^T J^T(\theta) J(\theta) \Delta \theta \quad (18)$$

Considering $\Delta \theta_{n,fb}$ defined as $\Delta \theta_{n,fb} = \Delta \theta - \Phi_{fb}(\theta, e)$, the expected value of the product of $\Delta \theta$ and ΔS at the point (θ, e) in the input space of $\Phi_{fb}(\theta, e)$ can be approximated as follows:

$$\begin{aligned} E[\Delta S \Delta \theta | \theta, e] &\approx -R_\theta J^T e + \frac{1}{2} R_\theta J^T J \Phi_{fb}(\theta, e) \\ &+ \frac{1}{2} E[\Delta \theta \Delta \theta^T J^T J \Delta \theta_{n,fb} | \theta, e] \end{aligned} \quad (19)$$

When the arm is controlled according to Equation (2), $\Delta \theta_{n,fb}$ is the disturbance noise $d(k)$. Since $d(k)$ satisfies Equation (12), the following equation is established.

$$E[\Delta \theta \Delta \theta^T J^T J \Delta \theta_{n,fb} | \theta, e] = \mathbf{0} \quad (20)$$

Therefore, the expected value of $\Delta \Phi_{fb}(\theta, e)$ can be expressed as;

$$E[\Delta \Phi_{fb}(\theta, e) | \theta, e] \approx \alpha R_\theta J^T e - \left(\frac{\alpha}{2} R_\theta J^T J + I\right) \Phi_{fb}(\theta, e) \quad (21)$$

When α is small enough, the condition described in Equation (7) is established. The learning result expressed as Equation (13) is obtained as described in Section 2.2.

It should be noted that the learning algorithm expressed in Equation (9) is applicable not only to $S(\mathbf{x}_d, \theta)$, but also to general penalty functions of hand position error norm $|e|$. The proposed learning model synthesizes a direction that decreases $S(\mathbf{x}_d, \theta)$ by summing after weighting $\Delta \theta$ based on the increase or decrease of $S(\mathbf{x}_d, \theta)$.

The feedback controller defined in Equation (13) requires a number of iterations to find a correct inverse kinematics solution, as the coordinates transformation function of the controller is incomplete. However, by using Kawato's feedback error learning [4], the second feedback controller; the feed-forward controller; or the inverse kinematics model that has a complete coordinate transformation function can be obtained as shown in Section 4.

4 TRACKING CONTROL SYSTEM LEARNING

In this section, we will consider the case where \mathbf{x}_d changes as $\mathbf{x}_d(k)$ ($k = 1, 2, \dots$). The hybrid controller that includes the learning feed-forward controller $\Phi_{ff}(\theta(k), \Delta \mathbf{x}_d(k)) \in \mathbf{R}^m$ that transforms the change of the desired hand position $\Delta \mathbf{x}_d(k) = \mathbf{x}_d(k+1) - \mathbf{x}_d(k)$ to the joint angle vector space is considered:

$$\Delta \theta(k) = \Phi_{ff}(\theta(k), \Delta \mathbf{x}_d(k)) + \Phi_{fb}(\theta(k), e(k)) + d(k) \quad (22)$$

$$e(k) = \mathbf{x}_d(k) - \mathbf{x}(k) \quad (23)$$

The configuration of the hybrid controller is illustrated in Figure 3.

By using the modified change of the square of the error norm expressed as:

$$\Delta S = \frac{1}{2} (|\mathbf{x}_d(k-1) - \mathbf{x}(k)|^2 - |e(k-1)|^2) \quad (24)$$

and $\Delta \theta(k)$ as defined in Equation (22), the feedback controller learning rule defined in Equation (9) is useful for the tracking control system. A sample holder for memorizing $\mathbf{x}_d(k-1)$ is necessary for the calculation of ΔS . When the distribution

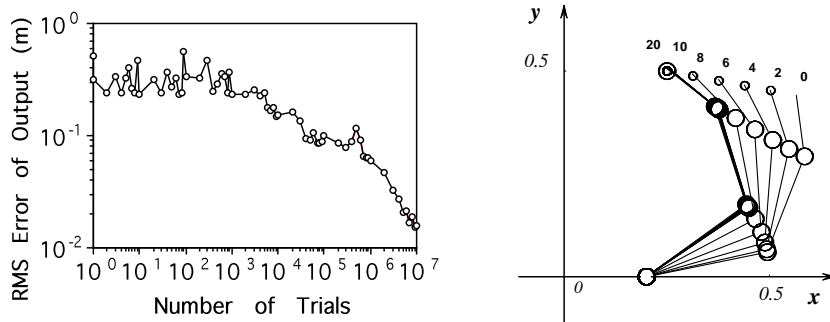


Figure 4: Learning Process of Controller Figure 5: One Example of Tracking Control

Figure 4 shows the progress of the proposed learning model. It can be seen that the RMS error decreases and the precision of the solver becomes higher as the number of trials increases. The RMS error became $9.31 \times 10^{-3}m$ after 2×10^7 learning trials. Figure 5 illustrates the hand position control by the inverse kinematics solver after 2×10^7 learning trials. The number near the end point of the arm indicates the value of k . The center of the small circle in Figure 5 indicates the desired hand position. The center of the large circle indicates the final desired hand position. Through learning, a precise inverse kinematics solver can be obtained. However, for RMS error to fall below 0.02, trials must be repeated more than 10^6 times. In such cases, more efficient learner or a learning rule is necessary.

6 CONCLUSION

A learning model of coordinate transformation of the hand position feedback controller was proposed in this paper. Although the proposed learning model may take a long time to learn, it is capable of learning a correct inverse kinematics solver without using a forward model, a back-propagation signal, or a pre-existing feedback controller.

We believe that the slow learning speed can be improved by using neural networks that have a structure suitable for the coordinate transformation. A major limitation of the proposed model is the structure of the learning rule, since the learning rule requires the calculation of the product of the change of the error penalty function and the change of the joint angle vector. However, the existence of such structure in the nervous system is unknown. An advanced learning model which can be directly compared with the physiological and psychological experimental results is necessary.

References

- [1] T. Ogino and S. Ishii, "Long-term Results after Pollicization for Congenital Hand Deformities," *Hand Surgery*, 2, 2, pp.79-85, 1997
- [2] F. H. Guenther and D. M. Barreca, "Neural models for flexible control of redundant systems," in P. Morasso and V. Sanguineti (Eds.), *Self-organization, Computational Maps, and Motor Control*. Amsterdam: Elsevier, pp.383-421, 1997
- [3] M. I. Jordan, "Supervised Learning and Systems with Excess Degrees of Freedom," *COINS Technical Report*, 88-27, pp.1-41, 1988
- [4] M. Kawato, K. Furukawa and R. Suzuki, "A Hierarchical Neural-network Model for Control and Learning of Voluntary Movement," *Biological Cybernetics*, 57, pp.169-185, 1987
- [5] D.T. McRuer and H. R. Jex, "A Review of Quasi-Linear Pilot Models," *IEEE Trans. on Human Factors in Electronics, HFE-8*, 3, pp.38-51, 1963