

TOPS (Task Optimization in the Presence of Signal-Dependent Noise) Model

Hiroyuki Miyamoto,^{1,*} Eri Nakano,² Daniel M. Wolpert,³ and Mitsuo Kawato²

¹Japan Science and Technology Corporation, Kawato Dynamic Brain Project, Kyoto, 619-0288 Japan

²ATR International, Information Science Division, Kyoto, 619-0288 Japan

³University College London, Sobell Department of Neurophysiology, Institute of Neurology, London WC1N, 3BG United Kingdom

SUMMARY

This paper proposes a movement trajectory planning model, which is a maximum task achievement model in which signal-dependent noise is added to the movement command. In the proposed model, two optimization criteria are combined, maximum task achievement and minimum energy consumption. The proposed model has the feature that the end-point boundary conditions for position, velocity, and acceleration need not be prespecified. Consequently, the method can be applied not only to the simple point-to-point movement, but to any task. In the method in this paper, the hand trajectory is derived by a psychophysical experiment and a numerical experiment for the case in which the target is not stationary, but is a moving region. It is shown that the trajectory predicted from the minimum jerk model or the minimum torque change model differs considerably from the results of the psychophysical experiment. But the trajectory predicted from the maximum task achievement model shows good qualitative agreement with the hand trajectory obtained from the psychophysical experiment. © 2004 Wiley Periodicals, Inc. *Syst Comp Jpn*, 35(11): 48–58, 2004; Published online in Wiley Inter-

Science (www.interscience.wiley.com). DOI 10.1002/scj.10377

Key words: signal-dependent noise; variance; unscented filter; optimization; reaching movement.

1. Introduction

When a human moves his hand quickly from a point to another point, the hand follows a trajectory which exhibits common and invariant characteristics in time and space [1]. The movement is as follows. The hand (finger) follows a trajectory which is almost straight but slightly curved. The velocity in the tangential direction exhibits a bell-shaped waveform in time. The acceleration exhibits a smooth waveform. Various optimization models have been proposed as movement trajectory generation models for this human movement. Examples are the minimum jerk model [2], the minimum torque change model [3], the minimum motor command change model [4], and the minimum commanded torque change model [5].

Among these, the minimum jerk trajectory has the feature that the analytical solution can be derived relatively easily. But it is very difficult to derive strict solutions for other models based on other dynamic optimal criteria. It is required to solve a constrained nonlinear optimization

*Now affiliated with the Graduate School of Biological Engineering, Kyushu Institute of Technology.

problem in which the solution minimizing the evaluation function is derived under a constraint (nonlinear dynamics) and subject to boundary conditions (start and goal points). It is generally very difficult to derive the solution.

When the hand is moved repeatedly, almost the same trajectories are observed, but it is impossible to reproduce exactly the same trajectory. Such a variation of movement has been ignored in past movement control models as an unnecessary disturbance. Is it meaningless, however, to consider that the variation of movement has an essential role in achieving the movement? Harris and Wolpert proposed a new trajectory planning model in which the trajectory is generated so that the variance of the hand position in the period specified after the movement is minimized when noise dependent on the magnitude of the movement command is added to the movement command [6]. This variance minimization model can effectively reproduce saccadic eye movement and the reaching movement trajectory of the human arm without using the minimum jerk criterion or the dynamic optimization criterion. The model can also account well for Fitt's law (speed-accuracy trade-off) as well as the 2/3-power law (relation between trajectory curvature and hand velocity).

Past models have focused on generation of a trajectory which maximizes the smoothness (weak constraint) while strictly satisfying the boundary conditions (strong constraint). In point-to-point movement, the purpose is to generate a trajectory which gives strictly zero error for the prespecified position, velocity, and acceleration of the hand at the start and stop points of the movement, while maximizing smoothness, that is, minimizing jerks (changes of acceleration) or minimizing movement command changes. It is questionable, however, whether it makes sense to satisfy strictly the boundary condition so long as there is noise added to the movement command and the final position fluctuates. The purpose is to achieve a satisfactory task, and it is not always necessary for the boundary conditions to be strictly satisfied.

The authors are trying to extend the minimum variance model in the presence of movement command-dependent noise to the maximum task achievement model. The model can handle in a unified way trajectory generation under various task conditions, such as catching, tracking control, and impedance control. The criterion is to maximize task achievement, and it is not necessary for the boundary condition to be satisfied strictly or for the variance to be minimized. In this study, the maximum task achievement model is proposed, and the results of numerical and psychophysical experiments are compared. It is shown that the results of the numerical and psychophysical experiments agree well qualitatively when an optimization criterion combining maximum task achievement and minimum energy consumption (squared sum of joint torque in this study) is applied.

2. Optimization Model for Trajectory Planning

Representative past models of trajectory planning and the model proposed in this paper are briefly described below. Table 1 summarizes the major features of the models.

2.1. Past optimal trajectory generation models

2.1.1. Minimum jerk model

In the minimum jerk model, the trajectory is planned so that the end-point boundary conditions are strictly satisfied (in the point-to-point movement, zero error is realized for prespecified position, velocity, and acceleration of the hand at the start and stop points of movement), and the smoothness (jerk index) of the hand trajectory is minimized over the movement period t_f :

$$C_J = \frac{1}{2} \int_0^{t_f} \left\{ \left(\frac{d^3x}{dt^3} \right)^2 + \left(\frac{d^3y}{dt^3} \right)^2 \right\} dt \quad (1)$$

The jerk index is represented by the third-order time derivative of the hand position (x, y) in an orthogonal coordinate system. The parameters are defined in kinematic space, and

Table 1. Difference between models

Model	Signal-dependent noise	End-point boundary condition	Optimization criterion
Minimum jerk	not considered	strong constraint	smoothness (minimum jerk)
Minimum motor command change	not considered	strong constraint	smoothness (minimum motor command change)
Minimum variance	considered	strong constraint	minimum end-point position variance
Maximum task achievement	considered	no constraint	maximum task achievement and minimum energy consumption

the arm dynamics and movement time are not related to trajectory planning. Straight-line movement is predicted in the hand orthogonal coordinate system, and the tangential velocity along the trajectory has a bell-shaped waveform in time. It is easy to solve Eq. (1) analytically so that a fifth-order polynomial in the time t is obtained for the trajectory prediction.

2.1.2. Minimum torque change model

In the minimum torque change model, the end-point boundary conditions are strictly satisfied and the trajectory is planned so that the time change of the torque generated at each joint is minimized over the movement time t_f :

$$C_T = \frac{1}{2} \int_0^{t_f} \sum_{i=1}^n \left(\frac{d\tau_i}{dt} \right)^2 dt \quad (2)$$

where τ_i is the torque generated at the i -th joint.

In this model, in addition to the external coordinates of the start and the end points, dynamical quantities such as the arm length and the moment of inertia are related to trajectory planning, and the trajectory changes depending on the arm posture and the positions of the start and end points. The constrained nonlinear optimization problem must be solved, and it is not easy to derive a solution, but the solution once obtained well reproduces a gradually curved hand trajectory.

2.1.3. Minimum variance model

In the minimum variance model, it is assumed that the trajectory is planned so that the variance of the finger position in a specified time after the movement (the post-movement time from t_f to t_e) is minimized when the movement is repeated a certain number of times in the presence of noise dependent on the magnitude of the movement command):

$$C_V = \sum_{i=1}^n \int_{t_f}^{t_e} \{ (\bar{x} - x)^2 + (\bar{y} - y)^2 \} dt \quad (3)$$

where \bar{x} and \bar{y} are the mean finger position coordinates at the end of movement. There is no criterion for determining the postmovement time [6].

This model does not explicitly specify the criterion for smoothness, and a smooth trajectory is realized by minimizing the variance in the presence of noise dependent on the magnitude of the movement command. Only a kinematic variable—the variance of the end position—is explicitly used, but the dynamics has an effect on the trajectory planning. In the minimum variance model of Harris and Wolpert, the trajectory is approximated by a spline curve and the end-point node is fixed at the stop position.

2.2. Maximum task achievement model

In the trajectory generation model proposed in this paper, the trajectory is planned so that task achievement is maximized and energy consumption is minimized in the presence of the noise depending on the magnitude of the movement command. The end-point boundary condition is not specified beforehand. The situation is the same as in the minimum variance model in which noise dependent on the magnitude of the movement command plays the essential role, resulting in implicit consideration of the smoothness as a criterion.

In past trajectory planning models, the boundary conditions (position, velocity, and acceleration at the start and the end points) must be prespecified. But the trajectory is planned in the proposed model so that task achievement is maximized. In other words, the task is specified and the optimization problem is solved. Thus, as a result, the position, velocity, and acceleration at the end points of the trajectory are determined. It is not necessary, as in past models, for the end-point boundary condition to be prespecified. Consequently, the method can easily be applied to any task.

3. Psychophysical Experiment

In most past studies, only stationary targets have been used in which the end-point boundary conditions for the velocity and the acceleration are zero. In contrast, the target is not a point in this experiment: it is a region which moves with constant acceleration.

3.1. Experimental setup and procedure

The subjects were six males and two females, all right-handed, aged 22 to 46 years. They used the right hand in all trials. Figure 1 outlines the experimental setup. The subject sits in an otolaryngology chair with the shoulders fixed to the back of the chair by a belt. The arm is placed on a working table, and the height of the table is aligned to the height of the shoulder. The table surface is covered by a frictionless Teflon sheet.

An infrared LED marker is set on the finger, and its position is measured with a sampling frequency of 200 Hz, using a three-dimensional position measuring device (OPTOTRAK3020). The measured position of the finger is displayed on a CRT monitor as a cursor. The background of the CRT monitor screen is white. The subject watches the CRT monitor screen and performs the task. The measured position data are smoothed by a sixth-order Butterworth filter with a cutoff frequency of 5 Hz.

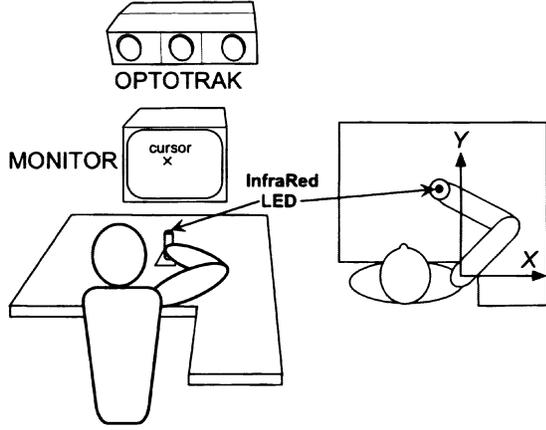


Fig. 1. Experimental setup.

3.2. Definition of task

The following task is used in the experiment.

(1) (Fig. 2a) The subject aligns the cursor representing the finger position to the inside of a circle with a diameter of 3 cm, indicating the position of the start point. The start point is at the position $(-0.25, 0.35)$ [m], with the origin at the right shoulder of the subject.

(2) (Fig. 2b) When the instruction is given by the examiner, the subject moves his finger to the right. When the cursor X leaves the start circle, the circle disappears.

(3) (Fig. 2c) At a certain time after the cursor leaves the start circle (ballistic movement time, $t_f = 0.5$ s), the target (a circle of diameter $D_T = 0.05$ m) appears 0.3 m to

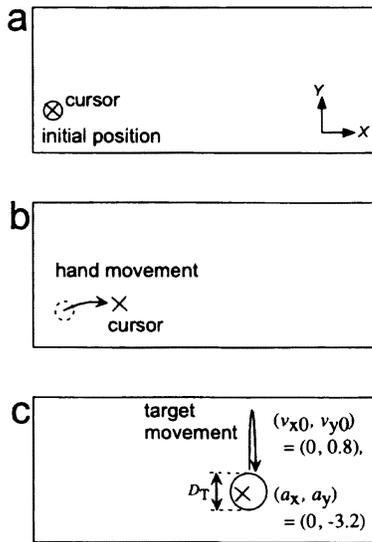


Fig. 2. A schematic diagram of the task. (a) Initial state; (b) immediately after the start of the movement; (c) after ballistic movement time (time t_f).

the right of the start point. The target moves with an initial velocity $(v_{x0}, v_{y0}) = (0, 0.8)$, and a constant acceleration $(a_x, a_y) = (0, -3.2)$. After a certain time (tracking movement time, $t_e = 1$ s), the target returns to the first position (0.3 m to the right of the start), and then disappears.

(4) The subject is not instructed particularly regarding the shape of the trajectory, etc., but is instructed to set the cursor within the target circle so long as the target is displayed (from time t_f to t_e). Utilizing the change of color of the cursor and the target and a beep tone, the subject is informed when the cursor stays within the circle. If the subject succeeded in keeping the cursor within the target circle, the trial is judged a success. Before acquiring the trajectory data, the subject practices the procedure 100 to 200 times in order to master the performance of the task. After starting the measurement of the trajectory data, the task is terminated when the number of successful trials reaches the specified value.

4. Numerical Experiment

4.1. Optimization

In this experiment, the hand trajectory is generated by a fifth-order spline function with nodes set at constant time intervals. The nodes are initialized so that the generated trajectory is the minimum jerk trajectory. The nodes are updated as follows by the simplex method.

(1) The trajectory of the finger position (x, y) is generated by a fifth-order spline function.

(2) By the inverse kinematics of the arm, the trajectory of the finger position (x, y) is converted to the trajectory of the joint angle (θ_1, θ_2) . Then, the result is converted by the inverse dynamics function to the trajectory of the movement command (u_1, u_2) .

(3) It is assumed that noise dependent on the magnitude of the movement command is added to the movement command, and the trajectories of the mean x_{mean} and the standard deviation x_{std} of the finger position are estimated by using an unscented filter [8].

(4) The evaluation function C_T for optimization is calculated.

(5) The nodes are updated so that the evaluation function C_T for optimization is decreased.

4.2. Evaluation function for optimization

In this experiment, the evaluation function C_T for optimization is defined as follows.

(1) The mean $\mathbf{x}_{\text{mean}} = (x_{\text{mean}}, y_{\text{mean}})$ and the standard deviation $\mathbf{x}_{\text{sd}} = (x_{\text{sd}}, y_{\text{sd}})$ of the finger position are estimated

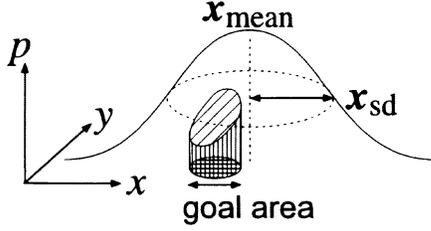


Fig. 3. A measure of task achievement.

at each sampling time by the unscented filter. Using the result, the probability $p(x, t)$ that the finger is at the position $\mathbf{x} = (x, y)$ is approximated by a normal distribution.

(2) The probability $g(t)$ that the finger is in the target region G (shaded region in Fig. 3) is calculated each time:

$$g(t) = \int_{\mathbf{x} \in G} p(\mathbf{x}, t) d\mathbf{x} \quad (4)$$

(3) The minimum value of $g(t)$ in the tracking movement time (from t_f to t_e) is defined as the task achievement T_A :

$$T_A = \min\{g(t)\} (t_f \leq t \leq t_e) \quad (5)$$

(4) The weighted sum of the task achievement T_A and the supplementary term C_{sup} is defined as the evaluation function, which is then minimized as the optimization:

$$C_T = (1 - T_A) + \alpha C_{\text{sup}} \quad (6)$$

As the supplementary term C_{sup} in this experiment, we use the temporal mean of the square-sum torque in the ballistic movement time and the tracking movement time combined (1 s in all):

$$C_{\text{sup}} = \frac{\sum_{i=1}^2 \int_0^{t_e} \tau_i(t)^2 dt}{t_e} \quad (7)$$

5. Comparison of Psychological Experiment and Prediction by Numerical Experiment

Figures 4 and 5 show the trajectories obtained in the psychophysical experiment. In Fig. 4, all measurements for eight subjects are stacked. 1185 trajectories were measured: 169 (40), 186 (38), 153 (40), 200 (15), 115 (39), 106 (47), 126 (47) and 130 (44). The numbers in parentheses are the number of trials judged to be successful. In this study, all

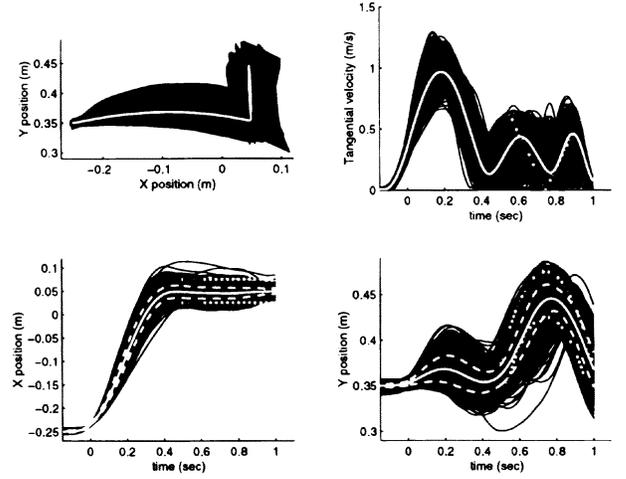


Fig. 4. Hand trajectories of all subjects.

trajectories were used in the analysis, regardless of whether the trial was a success or a failure. The white solid line is the mean of the hand trajectories, the dashed line is the trajectory of the mean standard deviation, and the dotted line is the trajectory of the target region.

Figure 5 shows the mean hand trajectory of subject AS. Figures 6 and 7 show the trajectories after optimization computation by the maximum task achievement trajectory generation model proposed in this paper. In Figs. 4 to 7, the upper left is the trajectory on the x - y plane, the upper right is the time waveform of the tangential velocity of the finger, the lower left is the time waveform of the x component of the finger position, and the lower right is the time waveform

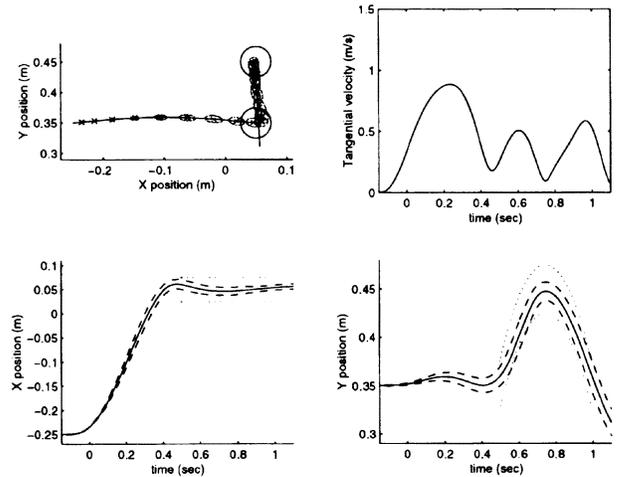


Fig. 5. Mean trajectory of hand trajectory of subject AS.

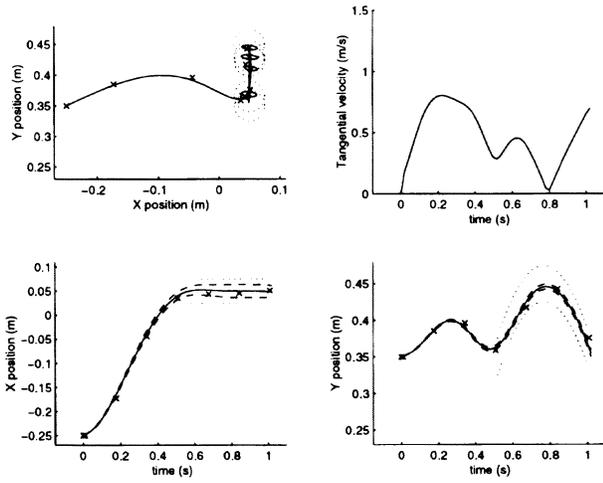


Fig. 6. Hand trajectory predicted by the max task achievement criterion.

of the y component of the finger position. The solid line is the trajectory of the mean, the dashed line is the trajectory of the mean \pm standard deviation, and the dotted line is the trajectory of the target region. X indicates the nodes of the spline function.

Figure 8 shows the hand trajectory obtained from the psychophysical experiment and the trajectory predicted by various models within the ballistic movement time, that is, from time 0 to t_f . The eight dotted lines are the mean trajectories of the eight subjects. TOPS ($\alpha=0$), TOPS

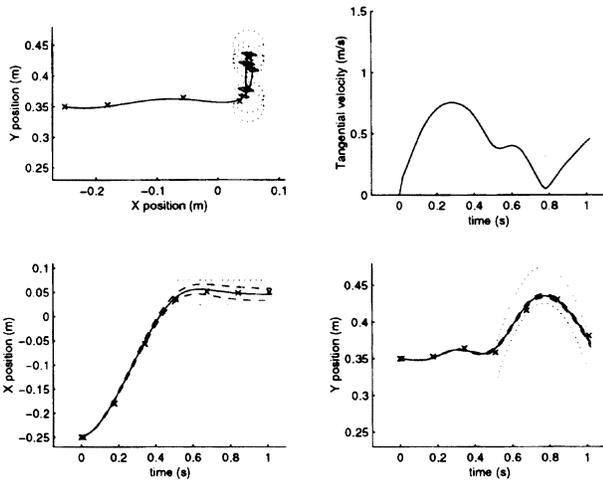


Fig. 7. Hand trajectory prediction by the max task achievement and minimum squared torque sum criterion.

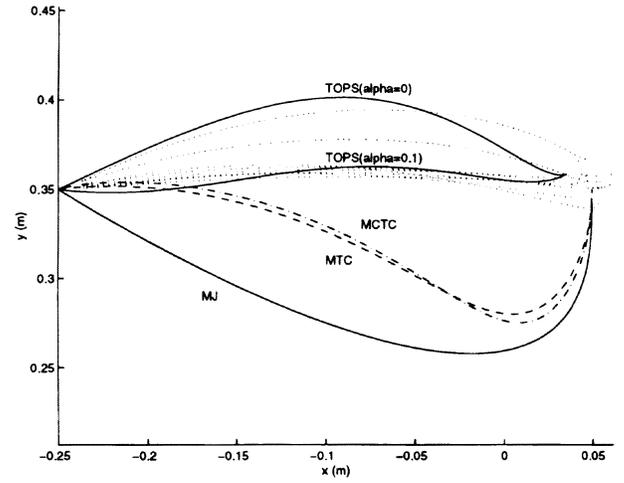


Fig. 8. Hand trajectories of eight subjects and each model. Eight dotted lines show mean trajectories of the eight subjects.

($\alpha=0.1$), MCTC, MJ and MTC are the predicted trajectories based on the optimization criteria of maximum task achievement only, maximum task achievement plus minimum energy consumption, minimum commanded torque change, minimum jerk, and minimum torque change, respectively.

We see from Fig. 5 that the position, velocity, and acceleration of the finger of the subject at time $t_f = 0.5$ s differ considerably from the values at the center of the target. In other words, it is estimated that the subject does not consider the movement of the center of the target at time $t_f = 0.5$ s as the end-point boundary condition. In the case of the conventional minimum commanded torque change, minimum jerk, and minimum torque change models, however, the end-point boundary conditions must be precisely specified. For the conventional model considered in this paper, the movement of the target center at time $t_f = 0.5$ s [the initial velocity $(v_{x0}, v_{y0}) = (0, 0.8)$ and constant acceleration $(a_x, a_y) = (0, -3.2)$] is used as the end-point boundary condition, and the optimal trajectory is determined.

Figure 9 shows the total deviation [7] of the hand trajectory estimated by each model from the hand trajectory obtained from the psychophysical experiment. The vertical axis is the frequency and the horizontal axis is the total deviation. Deviation in the positive y direction is defined as positive, and deviation in the negative y direction is defined as negative. The area between the straight line connecting the start point and the point at time t_f and the trajectory from time 0 to t_f is defined as the total deviation. For the trajectory measured for the subject, the point at which the curvature

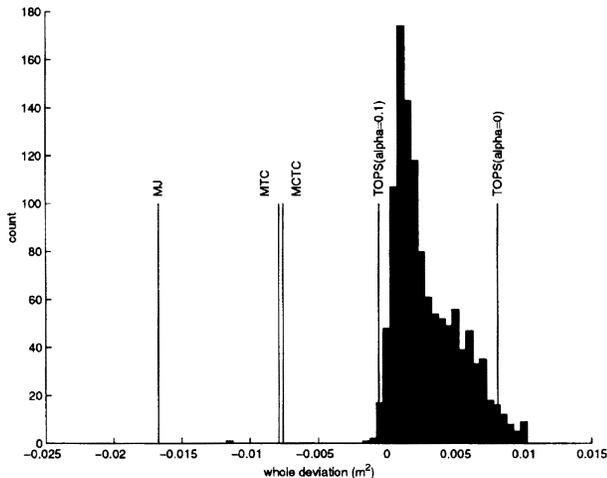


Fig. 9. Whole deviation from eight subjects and each model.

of the trajectory is the maximum near time t_f on the x - y plane is used instead of the point at time t_f .

As shown in Figs. 8 and 9, the trajectories predicted by the conventional models—the minimum commanded torque change, the minimum jerk, and the minimum torque change—are markedly curved in the negative y direction. As is evident from Fig. 9, the total deviation of the hand trajectory obtained from the psychophysical experiment is mostly distributed in the range of small positive values. Thus, in contrast to the predictions of the conventional models, the trajectory obtained by the psychophysical experiment is slightly curved in the positive y direction.

Figure 6 is the trajectory predicted by the optimization criterion based only on maximum task achievement [$\alpha = 0$ in Eq. (6)]. In contrast to the trajectory predicted by a conventional model such as the minimum jerk model, the trajectory is curved in the positive y direction. This trajectory is closer to the trajectory obtained by the psychophysical experiment. When only maximum task achievement is used as the optimization criterion, however, the trajectory is excessively curved in the positive y direction compared to the hand trajectory obtained from the psychophysical experiment.

Figure 7 is the trajectory predicted by the optimization criterion combining maximum task achievement and minimum energy consumption [$\alpha = 0.1$ in Eq. (6)]. We see from Figs. 4, 8, and 9 that the trajectory predicted by the optimization criterion combining maximum task achievement and minimum energy consumption shows good qualitative agreement with the trajectory obtained from the psychophysical experiment. We see that by varying α in Eq. (6) from 0 to 0.1, the model can account for the characteristics of the hand trajectories of almost all trials of the subjects.

6. Conclusions

This paper has proposed the maximum task achievement model, a new model that can handle trajectory generation in a unified way under a wide range of task conditions. It has also been shown that the hand movement trajectory which is predicted by the optimization criterion combining maximum task achievement and minimum energy consumption agrees well with the movement trajectory data obtained from the psychophysical experiment.

Recently, Nakano and colleagues (personal communication) examined the relation between stiffness and performance in single-joint reaching movement, and found the following tendency. When the simultaneous activity of the arm muscle is increased, the noise generated in the muscle increases, but the variation of the final position is decreased with increasing stiffness and noise. Todorov found that when the task requires that the trajectory pass through an intermediate point of a certain size, the generated trajectory varies depending on the size of the intermediate region to be passed through [9]. This suggests that humans do not always use the criterion that the variation of the final position is minimized (minimum variance).

There exists a supplementary term C_{sup} to be combined with the maximum task achievement term in the optimization criterion [Eq. (6)]. Muscle fatigue [10], reduction of energy consumption [11], the commanded torque change [5], and the smoothness of the movement command [4] may be used for this term. In this paper, the temporal mean of the square-sum torque over the ballistic movement time is used as the supplementary term C_{sup} .

As shown in Fig. 8, the trajectory is curved differently depending on the subject. By adjusting the weight [α in Eq. (6)] in the optimization criterion combining maximum task achievement and minimum energy consumption, the curved shape of the trajectory depending on the subject can be predicted. When a larger weight is assigned to task achievement, the trajectory is more curved. When a larger weight is assigned to energy consumption minimization, the trajectory becomes straighter. It is suggested that humans adjust the weight of the supplementary term according to the complexity of the task and the difficulty of its accomplishment.

The rough shape of the trajectory obtained by the numerical experiment in this paper is quite close to the actual data for the subjects, but the velocity waveform differs considerably. Various factors can be considered as responsible for this deviation, such as incompleteness of the maximum task achievement model or the numerical model, or other factors. In order to compare the theoretical prediction and the experimental data more quantitatively, the numerical model should be further refined and behavioral and numerical experiments should be continued.

REFERENCES

1. Abend W, Bizzi E, Morasso P. Human arm trajectory formation. *Brain* 1982;105:331–348.
2. Flash T, Hogan N. The coordination of arm movements: An experimentally confirmed mathematical model. *J Neurosci* 1985;5:1688–1703.
3. Uno Y, Kawato M, Suzuki R. Formation and control of optimal trajectory in human multijoint arm movement—Minimum torque-change model. *Biol Cybern* 1989;61:89–101.
4. Kawato M. Optimization and learning in neural networks for formation and control of coordinated movement. In Meyer DE, Kornblum S (editors). *Attention and performance XIV*. MIT Press; 1992.
5. Nakano E, Imamizu H, Osu R, Uno Y, Gomi H, Yoshioka T, Kawato M. Quantitative examination of internal representations for arm trajectory planning: Minimum command torque change model. *J Neurophysiol* 1999;81:2140–2155.
6. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature* 1998;394:780–784.
7. Osu R, Uno Y, Koike Y, Kawato M. Possible examination of trajectory curvature in multi-joint arm movement. *J Exp Psychol Human Perception and Performance* 1997;23:890–913.
8. Julier SJ, Uhlmann JK, Durrant-Whyte HF. A new approach for filtering nonlinear systems. *Proc 1995 American Control Conference*, p 1628–1632, Seattle.
9. Todorov EV. *Studies of goal directed movements*. PhD thesis, Massachusetts Institute of Technology, 1998.
10. Dul J, Johnson GE, Shiavi R, Townsend MA. Muscular synergism II. A minimum-fatigue criterion for load sharing between synergistic muscles. *J Biomech* 1984;17:675–684.
11. Alexander RM. A minimum energy cost hypothesis for human arm trajectories. *Biol Cybern* 1997;76: 97–105.

APPENDIX

1. Numerical Model for Arm

The equation of motion for the planar two-link two-joint arm used in the numerical experiment (Fig. A.1) is as follows:

$$\begin{aligned} \tau_1 = & (I_1 + I_2 + 2M_2L_1S_2 \cos \theta_2 + M_2L_1^2)\ddot{\theta}_1 \\ & + (I_2 + M_2L_1S_2 \cos \theta_2)\ddot{\theta}_2 \\ & - M_2L_1S_2(2\dot{\theta}_1 + \dot{\theta}_2)\dot{\theta}_2 + b_{11}\dot{\theta}_1 + b_{12}\dot{\theta}_2 \end{aligned}$$

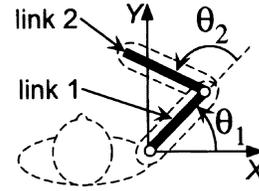


Fig. A.1. Arm configuration.

$$\begin{aligned} \tau_2 = & (I_2 + M_2L_1S_2 \cos \theta_2)\ddot{\theta}_1 + I_2\ddot{\theta}_2 \\ & + M_2L_1S_2 \sin \theta_2 \dot{\theta}_1^2 + b_{21}\dot{\theta}_1 + b_{22}\dot{\theta}_2 \end{aligned}$$

Here, τ_1 , θ_1 and τ_2 , θ_2 are the torque commands and the joint angles of the shoulder and the elbow joints, respectively. I_i , M_i , L_i and S_i are the moments of inertia with the joint as the axis, the mass and length of the joint, and its distance to the center of gravity of link i ($i = 1, 2$), respectively. Link 1 is the upper arm and link 2 is the forearm. B_{ij} represents the viscosity, which expresses the effect of the joint angular velocity of link j ($j = 1, 2$) on link i ($i = 1, 2$).

It is assumed that r_1 and r_2 follow the following second-order differential equation of the second-order muscle model:

$$u_i = (t_1 * t_2)\ddot{\tau}_i + (t_1 + t_2)\dot{\tau}_i + \tau_i$$

Here, u_i and τ_i represent the movement command and the torque command of joint i ($i = 1, 2$). $t_1 = 0.04$ and $t_2 = 0.03$ are the time constants of the muscle.

2. Unscented Filter

In the estimation of the mean and the variance of a system, the usual approach is to apply the Monte Carlo method. In this study, however, the unscented filter [8] is used to estimate the state (mean and variance) of the planar two-joint arm (Appendix 1), which is a nonlinear dynamical system. The unscented filter is a new method of estimating the state of a nonlinear dynamical system. It is easier to apply than the extended Kalman filter, and estimation with higher accuracy can be achieved.

The idea is to estimate the mean and variance from a small number of sigma points, under the assumption of a Gaussian distribution propagating in the nonlinear system dynamics. In the Monte Carlo method, in which a large number of sample points are selected at random to estimate the mean and variance, a very large sample is required in order to achieve highly accurate estimation. Fluctuations are produced in the Monte Carlo method due to the use of a random variable each time the mean and variance of the

trajectory are derived, even if the same target trajectory is used. This makes it difficult to achieve stable convergence of learning. In parameter updating in the Monte Carlo method, small ups and downs occur in the learning gradient, making it difficult for the procedure to converge to the correct solution.

In contrast to the Monte Carlo method, the unscented filter is a deterministic sample extraction procedure without using a random variable. Thus, the learning gradient is smoothed and stable convergence is realized. The unscented filter has the advantages over the Monte Carlo method that the procedure is simple, the accuracy of estimation is high, the computation cost is extremely low, and learning converges stably, since it is a deterministic sample extraction procedure without using a random variable.

Figure A.2 shows examples of the mean and variance estimated by the Monte Carlo method and the unscented filter. The horizontal and vertical axes are the x and y coordinates of the finger position, respectively, when the finger is moved by 0.3 m in the x direction in 0.5 s. The ellipse in the dashed line shows the variance for 1000 sample points (X) in the Monte Carlo method. The ellipse in the solid line shows the variance for 13 sigma points (O) in the unscented filter. We see that the unscented filter can estimate the mean and variance with high accuracy with a much smaller number of samples than the Monte Carlo method. The unscented filter is briefly described below.

2.1. The filtering problem

The filtering problem considered in this section is as follows. Consider the discrete-time nonlinear equation

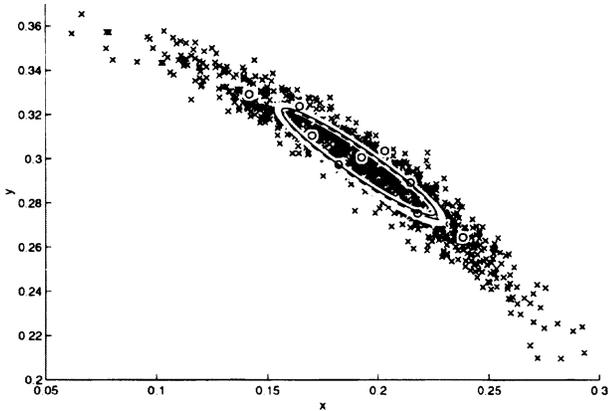


Fig. A.2. Comparison with Monte Carlo method and unscented filter method.

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{f}(\mathbf{x}(k), \mathbf{u}(k), \mathbf{v}(k), k), \\ \mathbf{z}(k) &= \mathbf{h}(\mathbf{x}(k), \mathbf{u}(k), k)\end{aligned}$$

The problem is to derive the optimal linear estimates for the state vector $\mathbf{x}(k)$ and the observation vector $\mathbf{z}(k)$ with the transition following the above equation.

Here, $\mathbf{f}(\cdot)$ is the process model, $\mathbf{x}(k)$ is the state variable at discrete time k , $\mathbf{u}(k)$ is the input vector, $\mathbf{v}(k)$ is the process noise, $\mathbf{z}(k)$ is the observation vector, $\mathbf{h}(\cdot)$ is the observation model, and $\mathbf{v}(k)$ is white noise satisfying $E[\mathbf{v}(k)] = 0$, $E[\mathbf{v}(i)\mathbf{v}^T(j)] = \delta_{ij}\mathbf{Q}(i)$. In this paper, it is assumed that the process model is the forward dynamics of the musculoskeletal system, the observation model is the forward kinematics, the process noise is noise dependent on the movement command, and the observation noise is 0.

2.2. Unscented transform

Suppose that \mathbf{x} has mean $\bar{\mathbf{x}}$ and variance \mathbf{P}_{xx} , and that $\mathbf{y} = \mathbf{f}(\mathbf{x})$ is a nonlinear function. The problem is to calculate the mean $\bar{\mathbf{y}}$ and variance \mathbf{P}_{yy} of \mathbf{y} . The unscented transform procedure to be used in the unscented filter is described briefly below. Let the mean and variance of the n -dimensional variable \mathbf{x} be $\bar{\mathbf{x}}$ and \mathbf{P}_{xx} , respectively. Then, $2n + 1$ sigma points \mathcal{X}_i are selected so that the sample mean and sample variance agree with $\bar{\mathbf{x}}$ and \mathbf{P}_{xx} , respectively.

The nonlinear transformation $\mathcal{Y}_i = \mathbf{f}(\mathcal{X}_i)$ is applied to each sample point. The mean $\bar{\mathbf{y}}$ and variance \mathbf{P}_{yy} after transformation are obtained as follows, as the weighted sum and the weighted outer product of the sample points:

$$\begin{aligned}\bar{\mathbf{y}} &= \sum_{i=0}^{2n} W_i \mathcal{Y}_i, \\ \mathbf{P}_{yy} &= \sum_{i=0}^{2n} W_i (\mathcal{Y}_i - \bar{\mathbf{y}})(\mathcal{Y}_i - \bar{\mathbf{y}})^T\end{aligned}$$

The $2n + 1$ sigma points are as follows:

$$\begin{aligned}\mathcal{X}_0 &= \bar{\mathbf{x}} \\ \mathcal{X}_i &= \bar{\mathbf{x}} + \left(\sqrt{(n + \kappa)\mathbf{P}_{xx}} \right)_i \\ \mathcal{X}_{i+n} &= \bar{\mathbf{x}} - \left(\sqrt{(n + \kappa)\mathbf{P}_{xx}} \right)_i\end{aligned}\quad (\text{A.1})$$

where $\kappa \in R$ and $(\sqrt{(n + \kappa)\mathbf{P}_{xx}})_i$ represent the i -th row or column of the matrix square root of $(n + \kappa)\mathbf{P}_{xx}$.

The weights are given by

$$\begin{aligned}W_0 &= \kappa / (n + \kappa) \\ W_i &= 1 / (n + \kappa)\end{aligned}$$

$$W_{i+n} = 1/(n + \kappa) \quad (\text{A.2})$$

We set $n + \kappa = 3$ in this paper.

2.3. Algorithm of unscented filter

The following procedure is repeated for each time step.

(1) The state vector is rewritten as an $n^a = n + q$ -dimensional vector, including the process noise term:

$$\mathbf{x}^a(k) = \begin{pmatrix} \mathbf{x}(k) \\ \mathbf{v}(k) \end{pmatrix}$$

The process model is rewritten as follows, as a function of \mathbf{x}_i^a :

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}_i^a, \mathbf{u}(k), k)$$

The mean and variance of $\mathbf{x}^a(k)$ are

$$\hat{\mathbf{x}}^a(k|k) = \begin{pmatrix} \hat{\mathbf{x}}(k|k) \\ 0_{q \times 1} \end{pmatrix} \quad (\text{A.3})$$

$$\mathbf{P}^a(k|k) = \begin{pmatrix} \mathbf{P}(k|k) & 0 \\ 0 & \mathbf{Q}(k) \end{pmatrix} \quad (\text{A.4})$$

(2) $n^a + 1$ sigma points are generated from Eqs. (A.1) to (A.4) and are transformed by the process model:

$$\mathcal{X}_i(k+1|k) = \mathbf{f}(\mathcal{X}_i^a(k|k), \mathbf{u}(k), k)$$

(3) The mean and variance of the state vector are estimated:

$$\hat{\mathbf{x}}(k+1|k) = \sum_{i=0}^{2n^a} W_i \mathcal{X}_i^a(k+1|k),$$

$$\begin{aligned} \mathbf{P}(k+1|k) &= \sum_{i=0}^{2n^a} W_i \{ \mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k) \} \\ &\quad \cdot \{ \mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k) \}^T. \end{aligned}$$

(4) All sigma points are transformed by the observation model.

$$\mathcal{Z}_i(k+1|k) = \mathbf{h}(\mathcal{X}_i(k+1|k), \mathbf{u}(k), k)$$

(5) The mean and variance of the observation vector are estimated.

$$\hat{\mathbf{z}}(k+1|k) = \sum_{i=1}^{2n^a} W_i \mathcal{Z}_i(k+1|k)$$

$$\begin{aligned} \mathbf{P}_{\nu\nu}(k+1|k) &= \sum_{i=0}^{2n^a} W_i \{ \mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k) \} \\ &\quad \cdot \{ \mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k) \}^T \end{aligned}$$

AUTHORS



Hiroyuki Miyamoto (member) received his B.S. degree from the Department of Biological Engineering, Osaka University, in 1985. Before completing the doctoral program, he served as a visiting researcher at ATR Human Information Communications Research Institute in 1994. Since 1996, he has been affiliated with the Japan Science and Technology Corporation, Kawato Dynamic Brain Project.

AUTHORS (continued) (from left to right)



Eri Nakano (member) completed the second half of her doctoral program at Kobe University in 1999. She was assigned as a research student to ATR Human Information Communications Research Institute in 1996. Since 1999, she has been a visiting researcher at ATR International Electronic Communications Basic Technology Research Institute. She is engaged in psychological and computational research on human motion control and learning mechanisms. She holds a Ph.D. degree, and is a member of the Japan Psychological Society, Japan Neural Science Society, and North American Neural Science Society.

Daniel M. Wolpert received his Ph.D. degree (physiology) from Oxford University in 1992 and served as a postdoctoral fellow at Massachusetts Institute of Technology from 1992 to 1995. Since then, he has been a lecturer at the University of London. He is a member of the North American Neural Science Society and Physiological Society.

Mitsu Kawato (member) received his B.S. degree from the Department of Physics, University of Tokyo, in 1976 and completed the doctoral program at Osaka University in 1981. After serving as a research associate and lecturer there, he joined the ATR Audio-Visual Mechanics Research Laboratory in 1988. He became head of the 3rd Laboratory, ATR Human Information Communications Research Institute in 1992. He served as a visiting professor at the Research Institute of Electronic Science, Hokkaido University, from 1992 to 1995; at the University of Genoa, Italy, in 1993; at Kanazawa Institute of Technology in 1994; and with the Collaborative Project, Nara Institute of Advanced Science and Technology, in 2000. He held an adjunct appointment as manager, Kawato Dynamic Brain Project, Japan Science and Technology Corporation, in 1996. He has been engaged in research on computation-theoretic neural science. He has received a Yonezawa Award, Osaka Science Award, Science and Technology Agency Minister's Award, and Tsukahara Award. He is the author of *Mechanisms of the Brain*, *Computational Theory of the Brain*, and other books. Editor, *J. Neural Networks*. Board member, Japan Neural Science Society. Board Member and Member of Award Committee, Japan Neural Network Society. Executive Committee Member, International Association for the Study of Attention and Performance.