

Research Article

How Can Brain Learn to Control a Nonholonomic System?

**Noriyasu Homma,¹ Shinpei Kato,² Takakuni Goto,³ Ivo Bukovsky,^{1,4}
Ryuta Kawashima,³ and Makoto Yoshizawa¹**

¹ *CyberScience Center, Tohoku University, 6-6-05 Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan*

² *Graduate School of Engineering, Tohoku University, 6-6-05 Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan*

³ *Institute of Development, Aging and Cancer, Tohoku University, Seiryō-machi 4-1, Aoba-ku, Sendai 980-8575, Japan*

⁴ *Faculty of Mechanical Engineering, Czech Technical University in Prague, Technická 4, 166 07 Prague 6, Czech Republic*

Correspondence should be addressed to Noriyasu Homma, homma@ieee.org

Received 14 December 2009; Accepted 5 March 2010

Academic Editor: Zeng-Guang Hou

Copyright © 2010 Noriyasu Homma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Humans can often conduct both linear and nonlinear control tasks after a sufficient number of trials, even if they initially do not have sufficient knowledge about the system's dynamics and the way to control it. Theoretically, it is well known that some nonlinear systems cannot be stabilized asymptotically by any linear controllers and we have reported by an f-MRI experiment that different types of information may be involved in linear and nonlinear control tasks, respectively, from a brain function mapping point of view. In this paper, from a controllability analysis, we still show a possibility that human may use a linear control scheme for such nonlinear control tasks by switching the linear controllers with a virtual constraint. It is suggested that the proposed virtual constraint can play an important role to overcome a limitation of the linear controllers and to mimic human control behavior.

1. Introduction

In the fields of control engineering, system engineering, and brain sciences, the excellent human's control abilities have been widely studied [1–6]. Among such studies, Wolpert and Kawato [7] have proposed a new model of human control mechanism called Modular Selection And Identification Control (MOSAIC) in which many elemental prediction and control system modules are combined in order to conduct the desired motion. Imamizu et al. [8] have evaluated the MOSAIC model using linear control tasks. They used a functional MRI (f-MRI) scanner to observe brain activities during the control tasks and reported plausibility of the model.

However, human ability is not limited to the linear control case. In fact, humans can often achieve difficult control tasks after a sufficient number of trials, even if they initially do not have sufficient knowledge about the system's dynamics and the way to control it. For example, Goto et al. [9] have reported manual control ability to operate a 2-link planar under actuated manipulator (2PUAM). The 2PUAM is a nonlinear system with a nonholonomic constraint and

cannot be stabilized asymptotically by any linear controller. In this case, to control such nonholonomic manipulator, the operators have to plan the object's trajectory first and then control the system to follow the trajectory. To do so, the operator must use shape information of the manipulator while the operator does not need such shape information to control linear manipulators.

To verify this difference between linear and nonlinear control tasks, we have conducted another f-MRI experiment using the 2PUAM [10]. The results suggest that the difference can be observed through the brain activities. In fact, a specific area involving the shape information processing was activated with a significant difference compared to the linear task only when subjects control the 2PUAM.

In this paper, to further clarify the experimental results, a new model-based analysis is conducted. The model is a multiple model-based reinforcement learning (MMRL) [11] for nonlinear controls but it is a linear controller. Using the MMRL, we attempt to explain the f-MRI results of the human information processing for the 2PUAM control task from a control theory point of view.

2. 2-Link Planer Under Actuated Manipulator (2PUAM)

In this study, we use the 2PUAM shown in Figure 1 as the control object. The 2PUAM moves in a horizontal plane. Thus, the 2PUAM is free from gravity, and arbitrary positions are equilibrium points. Motor is mounted on the first joint. Then the dynamic equations of the 2PUAM are

$$\begin{aligned} M_{11}\ddot{\theta}_1 + M_{12}\ddot{\theta}_2 + C_1 &= T_1 + T_{fr1}, \\ M_{21}\dot{\theta}_1 + M_{22}\ddot{\theta}_2 + C_2 &= T_2 + T_{fr2}, \end{aligned} \quad (1)$$

where

$$\begin{aligned} M_{11} &= (m_1 + m_2)l_1^2 + m_2l_2^2 + 2m_2l_1l_2 \cos \theta_2, \\ M_{12} &= M_{21} = m_2l_2^2 + m_2l_1l_2 \cos \theta_2, \\ M_{22} &= m_2l_2^2, \\ C_1 &= -m_2l_1l_2\dot{\theta}_1(2\dot{\theta}_1 + \dot{\theta}_2) \sin \theta_2 + c_1\dot{\theta}_1, \\ C_2 &= m_2l_1l_2\dot{\theta}_1^2 \sin \theta_2 + c_2\dot{\theta}_2. \end{aligned} \quad (2)$$

Here, as shown in Figure 1, m_1 and m_2 are the mass of the first and second links, l_1 and l_2 are the length, c_1 and c_2 are the resistance coefficient, respectively, θ_1 and θ_2 are angles, $\dot{\theta}_1$ and $\dot{\theta}_2$ are angular velocities, and $\ddot{\theta}_1$ and $\ddot{\theta}_2$ are angular accelerations of joints, respectively. The right-hand sides of (1) are the input and friction torques. In this study, we simplify the model as follows:

$$\begin{aligned} m_1 = m_2 = m, \quad l_1 = l_2 = l, \\ c_1 = c_2 = 0, \quad T_2 = T_{fr1} = T_{fr2} = 0. \end{aligned} \quad (3)$$

It is known that the 2PUAM is not controllable by the standard continuous feedback. This implies that reaching the 2PUAM and stopping it at some point is a difficult task. Although some indirect methods have been proposed to control the 2PUAM from the control engineering point of views [9], it is still unclear how human can control the nonlinear object such as the 2PUAM.

3. Summary of f-MRI Experiments

3.1. Control Tasks and Training. Here, human operator is required to feed an input torque $\tau \in [-1, 1]$ so that the end effector of the arm will be driven to the goal point and kept at the point. Since the position of the first joint is fixed, there are at least two objective positions of the second joint as shown in Figure 2: the upper position of the second joint p_1 and the lower position p_2 .

The manual control experiment was conducted by 6 neurologically normal subjects (19–24 years of age; six males) participated in the experiments. All subjects were right-handed. Informed written consent was obtained from each participant. They had no knowledge about the dynamic response of the system before the experiment. They observed the virtual manipulator's states (positions, angles, and velocities of the 2PUAM) through visual data displayed on an LCD

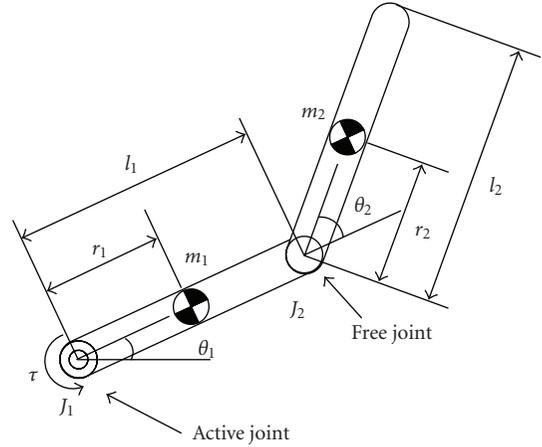


FIGURE 1: 2-link planner under actuated manipulator (2PUAM).

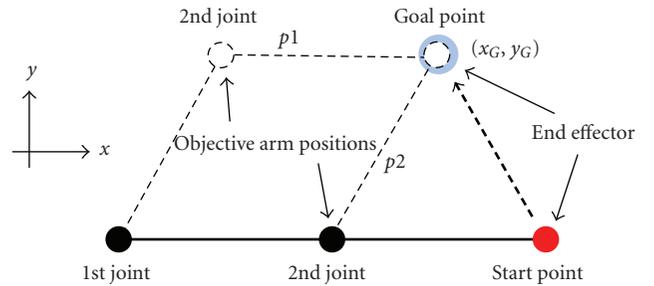


FIGURE 2: Positioning task.

monitor and fed the torque by using a joystick. The time limit was 30 seconds for each trial, and 300 training trials have been conducted outside of the MRI scanner before the following scanning sessions. The duration of 300 training trials depended on the subject's willingness and tiredness, but all the subjects completed them for 2-3 days.

For each trial, the performance index defined by the following equation was recorded:

$$J = \sum_{k=0}^T \sqrt{(x(k) - x_G)^2 + (y(k) - y_G)^2}. \quad (4)$$

Here $(x(k), y(k))$ denotes the position of the end effector at time $k\Delta t$ ($k = 0, 1, 2, \dots, T$), Δt is the interval of time step, (x_G, y_G) is the position of the goal point, and T is the maximum steps of each trial. The performance index J has been displayed on the monitor to guide subjects' criterion.

3.2. MRI Scanning Sessions. In scanning sessions, the trained subjects control two kinds of virtual manipulators whose shapes are projected on a screen in the MRI scanner as shown in Figure 3, by using an optical (magnet-free) joystick [10]. Both manipulators are the same in shape, but the first manipulator is the 2PUAM and the other is a manipulator which has a linear input and output relation (subjects can control it like a PC mouse). Controlling the 2PUAM is the main task while the linear control task is a comparative one,

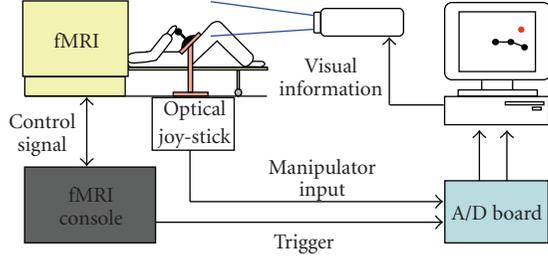


FIGURE 3: Experimental system.

and thus the former is called test trial and the latter is baseline trial, and their duration are called test and baseline periods, respectively. In other words, subjects can directly operate the coordinate of the end effector of the manipulator by the joystick in baseline trial. On the other hand, subjects operate only the torque of the first joint of the manipulator by the joystick and indirectly control the coordinate of the end effector in test trial. In each trial, subjects try to move the end effector of the manipulator to the goal point and keep it at the goal. The coordinates of the goal points are chosen randomly and displayed on the screen when trial starts.

MRI scanner (1.5T SIEMENS: Symphony 1.5T) was used to obtain blood oxygen level-dependant contrast functional images. Images weighted with the apparent transverse relaxation time were obtained with an echo-planner imaging sequence (repetition time: 3.0 s; echo time: 50 ms; flip angle: 90; field of view: 256×256 mm). High-resolution anatomical images of all subjects were also acquired with a T1-weighted sequence.

3.3. Results. Figure 4 indicates regions significantly more activated during test periods than baseline periods ($t > 5.89$, $P < .001$). Activities in the primary motor cortex, the somatosensory cortex, the somatosensory association cortex, the prefrontal cortex, the inferior temporal gyrus, and the fusiform gyrus were observed [10].

In the nonlinear control task, significant activities of the inferior temporal gyrus and the fusiform gyrus were observed. These areas are known to have an intimate involvement in recognizing the characteristic (color and shape) of the object. On the other hand, the prefrontal cortex is known as a region that receives the information that is necessary for action planning from both the temporal association area and the occipital association area and assembles complicated action plan. These suggest that in operating 2PUAM, subjects use the information about shape or position of 2PUAM and make trajectory planning of the positioning task based on that information. From a viewpoint of control theory, it may be worth to mention that such information is not needed for the linear control, but it is necessary for the nonholonomic systems control.

4. Controllability Analysis

The difference between linear and nonlinear control tasks observed through the significant brain activities can be a

reflection of the human control mechanism that can cognize the target nonlinear dynamics and use an appropriate piece of information. However, this is not sufficient to conclude that human does not use any linear control scheme. Instead, the hypothesis proposed in this paper is that human can use a linear control scheme by switching linear controllers responsible for specific regions where the linear approximation can work well for the target nonlinear task. The following linear model can then be employed to verify the hypothesis.

4.1. Multiple Model-Based Reinforcement Learning (MMRL). MMRL has multiple modules that are pairs of prediction model which predicts future state of the controlled objects and reinforcement learning controller which learns the control output. “Responsibility signals” are calculated from the softmax function of the prediction errors. The prediction model which outputs the more accurate prediction has the larger responsibility signal. By weighting control signal and learning of the each module with responsibility signals, these modules are adapted to the corresponding specific situations, respectively.

In this study, we use a multiple linear quadratic controller (MLQC) by using multiple linear prediction and quadratic reward models as an efficient implementation of the MMRL [11]. Change in the state vector of the target system, $x \in R^4$, is given by

$$\dot{x}(t) = f(x(t), u(t)), \quad (5)$$

where x is the state vector of the system, and u is the control output. Each variable of the vector for the 2PUAM can be defined as

$$x_1 = \theta_1, \quad x_2 = \dot{\theta}_1, \quad x_3 = \theta_2, \quad x_4 = \dot{\theta}_2. \quad (6)$$

Linear prediction models of the MMRL can be represented as follows:

$$\hat{\dot{x}}_i(t) = f_i(x(t), u(t)) = A_i x(t) + B_i u(t), \quad i = 1, 2, \dots, n. \quad (7)$$

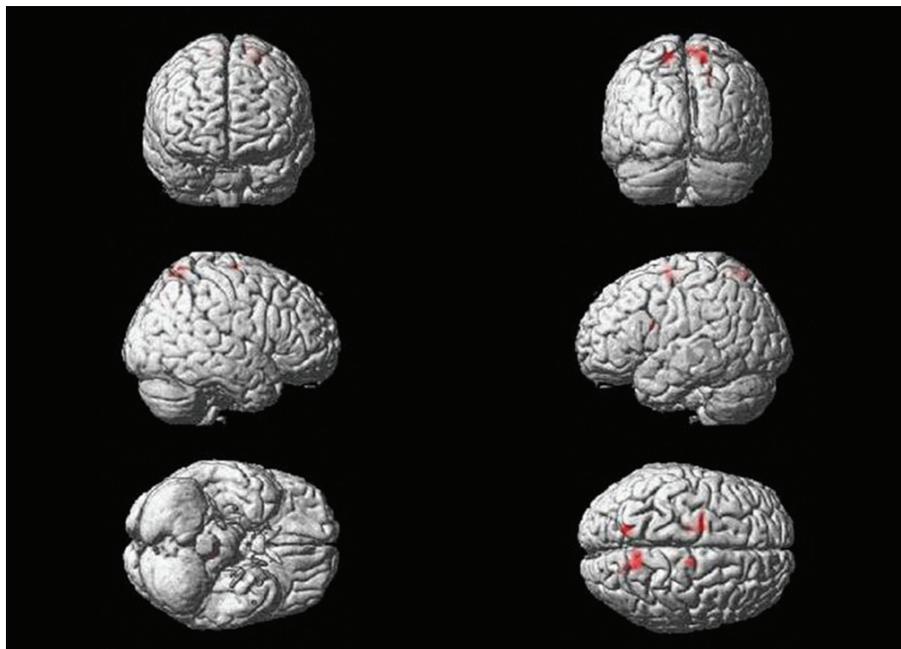
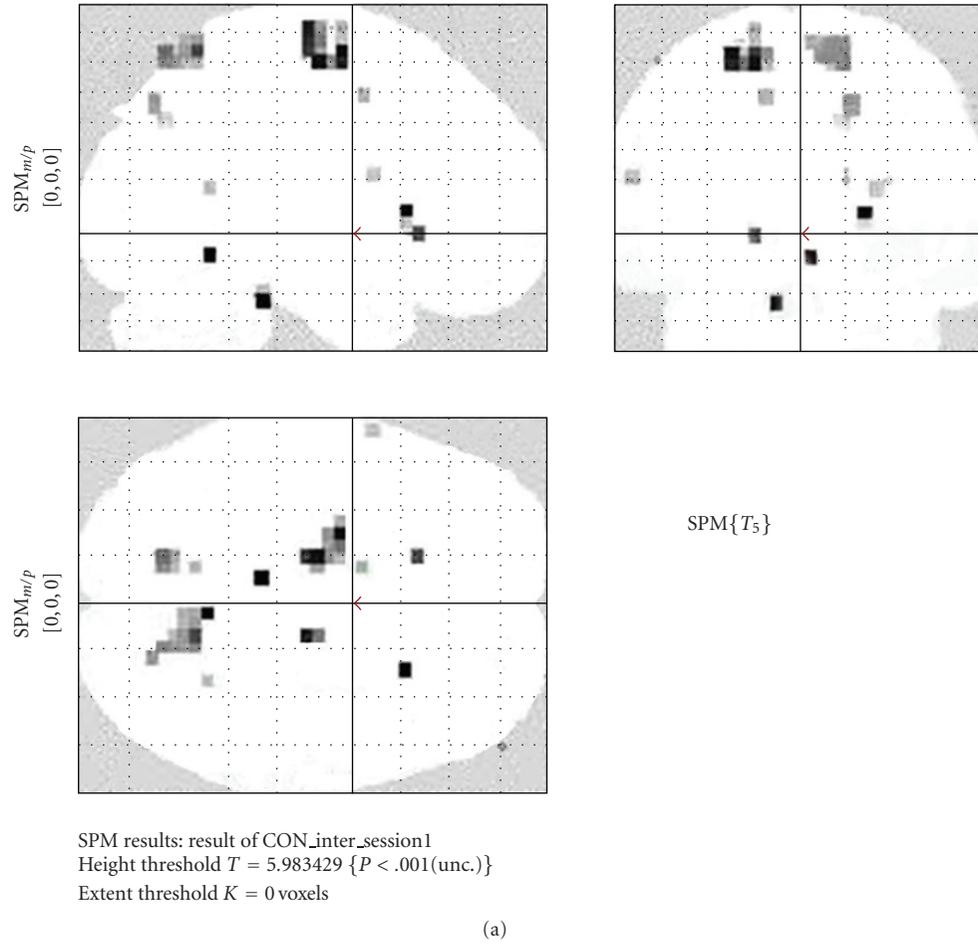
where n denotes the number of modules. State prediction \hat{x} is given by a weighting sum of prediction models with responsibility signals λ_i [11]:

$$\hat{\dot{x}}(t) = \sum_{i=1}^n \lambda_i(E_i(t)) \hat{\dot{x}}_i(t) = \sum_{i=1}^n \lambda_i(E_i(t)) (A_i x(t) + B_i u(t)), \quad (8)$$

where $E_i(t)$ is a short-time average of the prediction error $\dot{x}(t) - \hat{\dot{x}}(t)$. Learning of each prediction model in (7), $i \in \{1, 2, \dots, n\}$, is conducted by changing its parameter vector w_i consisting of all the elements of the matrices A_i and B_i as follows:

$$\dot{w}_i(t) = \eta_w \lambda_i(E_i(t)) \left(\frac{\partial f_i}{\partial w_i} \right)^T (\dot{x}(t) - \hat{\dot{x}}(t)), \quad (9)$$

where $\eta_w (> 0)$ is an update coefficient. Schematic diagram of the multiple predictor-controller pair architecture is shown in Figure 5.



(b)

FIGURE 4: Significant activated areas of test period compared to baseline period (for more details, see [10]).

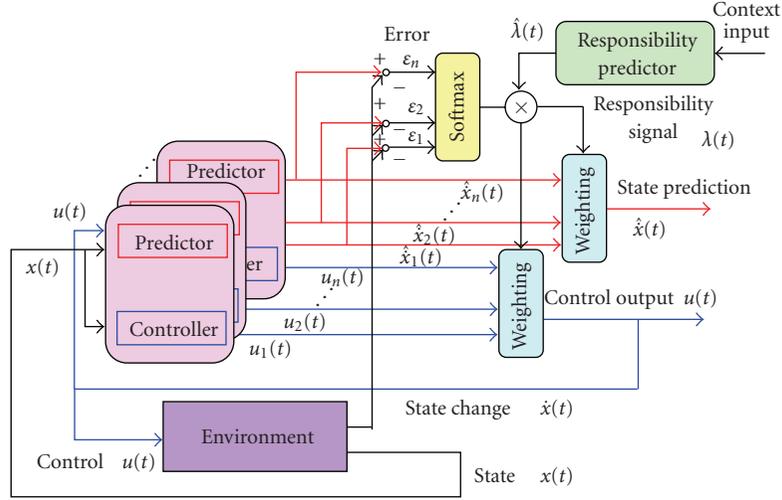


FIGURE 5: Schematic diagram of multiple predictor-controller pair architecture [11].

4.2. *Controllability of 2PUAM.* We obtain the following state equation of the 2PUAM from (1) and (6) under the condition in (3):

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \frac{1}{2 - \cos^2 x_3} \{ (3 + \cos x_3) x_2^2 \sin x_3 \\ &\quad + x_2 x_4 \sin x_3 + u/ml^2 \}, \end{aligned} \quad (10) \quad \text{where}$$

$$\begin{aligned} \dot{x}_3 &= x_4, \\ \dot{x}_4 &= -(1 + \cos x_3) \dot{x}_2 - x_2^2 \sin x_3. \end{aligned}$$

Equation (10) can be described using the vector form as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = f \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, u \right). \quad (11)$$

Here, we define the microscopic fluctuation at some positions x_a in phase space as $\delta x = x - x_a$, $\delta u = u - u_a$. Calculating the Taylor series of (11) and ignoring the higher-order terms from the second order, we get

$$\delta \dot{x} = A \delta x + B \delta u. \quad (12)$$

Here matrices A and B are

$$A = \frac{\partial f}{\partial x} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{\partial \dot{x}_2}{\partial x_2} & \frac{\partial \dot{x}_2}{\partial x_3} & \frac{\partial \dot{x}_2}{\partial x_4} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{\partial \dot{x}_4}{\partial x_2} & \frac{\partial \dot{x}_4}{\partial x_3} & \frac{\partial \dot{x}_4}{\partial x_4} \end{pmatrix},$$

$$B = \frac{\partial f}{\partial u} = \frac{1}{aml^2} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -(1 + \cos x_3) \end{pmatrix}, \quad (13)$$

$$\begin{aligned} \frac{\partial \dot{x}_2}{\partial x_2} &= \frac{(2x_2(3 + \cos x_3) + x_4) \sin x_3}{a}, \\ \frac{\partial \dot{x}_2}{\partial x_3} &= \frac{(6x_2 + x_4)x_2 \cos^2 x_3 - 2x_2^2}{a} + \frac{u \sin 2x_3}{ml^2 a}, \\ \frac{\partial \dot{x}_2}{\partial x_4} &= \frac{x_2 \sin x_3}{a}, \\ \frac{\partial \dot{x}_4}{\partial x_2} &= -(1 + \cos x_3) \frac{\partial \dot{x}_2}{\partial x_2} - 2x_2 \sin x_3, \\ \frac{\partial \dot{x}_4}{\partial x_3} &= \dot{x}_2 \sin x_3 - (1 + \cos x_3) \frac{\partial \dot{x}_2}{\partial x_3} - x_2^2 \cos x_3, \\ \frac{\partial \dot{x}_4}{\partial x_4} &= -(1 + \cos x_3) \frac{x_2 \sin x_3}{a}, \end{aligned} \quad (14)$$

$$a = 2 - \cos^2 x_3.$$

Since the MMRL is a linear controller, it cannot control the 2PUAM. Indeed, the rank of the controllability matrix $U_c = [B \ AB \ A^2B \ A^3B]$ at any equilibrium position x_0 becomes 2, which is not the full rank, 4. That is, let us denote an equilibrium point by using arbitrary values θ_{10} and θ_{20}

$$x_0 = (\theta_{10} \ 0 \ \theta_{20} \ 0)^T \quad (15)$$

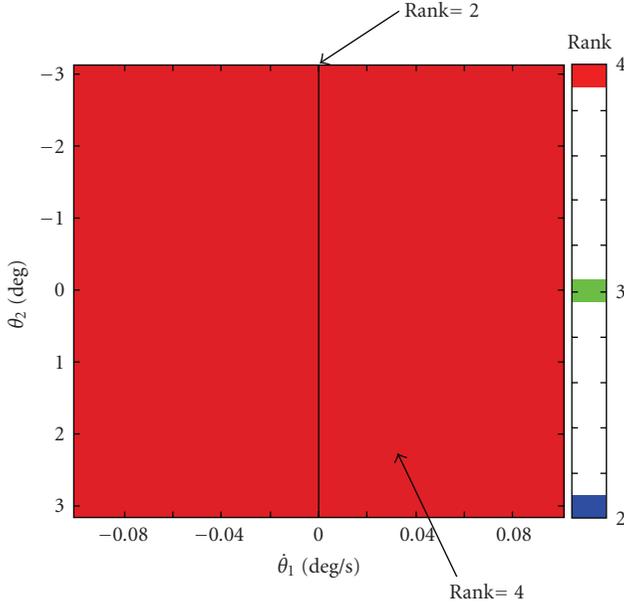


FIGURE 6: Rank of the controllability matrix U_c at some positions x_s ($\dot{\theta}_2 = 0, u = 0$).

with no input $u = 0$. Then, from (14),

$$A = \left. \frac{\partial f}{\partial x} \right|_{u=0}^{x=x_0} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (16)$$

$$B = \left. \frac{\partial f}{\partial x} \right|_{u=0}^{x=x_0} = \frac{1}{a_0 ml^2} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -(1 + \cos \theta_{20}) \end{pmatrix},$$

where $a_0 = 2 - \cos^2 \theta_{20}$. In this case, the rank of the U_c is not equal to the full rank, 4, but is 2 as follows:

$$\text{rank} [U_c] = \text{rank} [B \ AB \ A^2 B \ A^3 B] = 2. \quad (17)$$

4.3. Possible Control Strategy. The rank of the controllability matrix can, however, be the full rank, 4, at some positions x_s in the phase space where angular velocity of the first joint is not exactly zero, $\dot{\theta}_1 \neq 0$, even if the $\dot{\theta}_1$ is very close to zero, as shown in Figure 6.

To verify the controllability, let us denote positions x_s in the phase space by using arbitrary values θ_{1s} , θ_{2s} , and $\dot{\theta}_{2s}$:

$$x_s = (\theta_{1s} \ \dot{\theta}_{1s} \ \theta_{2s} \ \dot{\theta}_{2s})^T, \quad \dot{\theta}_{1s} \neq 0. \quad (18)$$

Then, from (14),

$$A_s = \left. \frac{\partial f}{\partial x} \right|_{u=u_s}^{x=x_s} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{\partial \dot{x}_2}{\partial x_2} & \frac{\partial \dot{x}_2}{\partial x_3} & \frac{\partial \dot{x}_2}{\partial x_4} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{\partial \dot{x}_4}{\partial x_2} & \frac{\partial \dot{x}_4}{\partial x_3} & \frac{\partial \dot{x}_4}{\partial x_4} \end{pmatrix}, \quad (19)$$

$$B_s = \left. \frac{\partial f}{\partial x} \right|_{u=u_s}^{x=x_s} = \frac{1}{a_s ml^2} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -(1 + \cos \theta_{2s}) \end{pmatrix},$$

where $a_s = 2 - \cos^2 \theta_{2s}$. In this case, the rank of the controllability matrix is 4, the full rank:

$$\text{rank} [U_{c_s}] = \text{rank} [B_s \ A_s B_s \ A_s^2 B_s \ A_s^3 B_s] = 4. \quad (20)$$

Thus, it is confirmed that 2PUAM cannot be stabilized at any equilibrium point, but if $\dot{\theta}_{1s} \neq 0$, the end effector of the 2PUAM can approach any geometrical point with any slow speed.

4.4. Discussions. The slow speed approach verified above might be an interesting result because, nevertheless human subjects can control the 2PUAM very well, it is often very hard even for human subjects to stop the 2PUAM completely in the manual control tasks [9, 10]. The linear controller can be responsible only for a small region in which the linear approximation can work well. Thus, by switching multiple linear models, there is a possibility to control the 2PUAM with the MMRL.

Different from (16), we assumed the condition in (19) where $\partial \dot{x}_i / \partial x_j \neq 0$, $i = 2, 4$, and $j = 2, 3, 4$ to make the rank of controllability matrix be the full rank. The condition implies a virtual constraint of the manipulator's shape (joint movement) because, for example, $\partial \dot{x}_2 / \partial x_3 \neq 0$ implies the virtual existence of relation between $\dot{\theta}_1$ and θ_2 that makes a constraint on the angular acceleration depended on the angle of the manipulator shape. Interestingly, if human subjects could feel and realize such virtual constraint on the 2PUAM control (joint movement), the control task can sometimes be achieved relatively easier [9].

According to the f-MRI results, subjects may use the shape and position information of the 2PUAM. The virtual constraint discussed above can further be created in order to make the control easier. Unfortunately, the controllability analysis does not prove this hypothesis, but there is a possibility that in operating the 2PUAM, subjects use the shape information to switch the controllers [10] and create the virtual constraint to make the control easier [9]. In this sense, MMRL can be regarded as a linear model for controlling the 2PUAM. Then, if human's superior learning ability to control the complex nonlinear system could be based on such linear control schemes, its implementation on a robot system might be easier than we expected.

5. Conclusions

In this paper, by using a controllability analysis, we have revised the previous f-MRI experimental results that reveal significant activation areas for the nonlinear control task compared to the linear one. Even the useful pieces of information for the linear task may be different from nonlinear ones, the analysis suggests some possibility to control the 2PUAM with a set of linear control models in a similar way by which human subjects can control it. In fact, to stop the 2PUAM at an equilibrium point completely seems very hard or almost impossible, but to approach there with any arbitrary slow speed seems an easier task. The additional information of shape and position of the 2PUAM can then be used for switching the linear controllers. Although the internal relation between the virtual constraint and the controllability of the nonlinear task is still unclear and should be clarified further from both computational and brain sciences point of views, the hypothesis proposed in this paper implies that it could be possible to design or realize robots with the human-level learning ability in an easier way.

References

- [1] R. Coulom, "High-accuracy value-function approximation with neural networks applied to the acrobat," in *Proceedings of the 12th European Symposium on Artificial Neural Networks (ESANN '04)*, pp. 7–12, Bruges, Belgium, April 2004.
- [2] K. Doya, "Reinforcement learning in continuous time and space," *Neural Computation*, vol. 12, no. 1, pp. 219–245, 2000.
- [3] D. P. Bertsekas and J. N. Tsitsiklis, *Nuero-Dynamic Programming*, Athena Scientific, Belmont, Calif, USA, 1996.
- [4] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, UK, 1998.
- [5] E. Wiewiora, "Potential-based shaping and Q-value initialization are equivalent," *Journal of Artificial Intelligence Research*, vol. 19, pp. 205–208, 2003.
- [6] A. Karniel and F. A. Mussa-Ivaldi, "Sequence, time, or state representation: how does the motor control system adapt to variable environments?" *Biological Cybernetics*, vol. 89, no. 1, pp. 10–21, 2003.
- [7] D. M. Wolpert and M. Kawato, "Multiple paired forward and inverse models for motor control," *Neural Networks*, vol. 11, no. 7-8, pp. 1317–1329, 1998.
- [8] H. Imamizu, T. Kuroda, T. Yoshioka, and M. Kawato, "Functional magnetic resonance imaging examination of two modular architectures for switching multiple internal models," *Journal of Neuroscience*, vol. 24, no. 5, pp. 1173–1181, 2004.
- [9] T. Goto, N. Homma, M. Yoshizawa, and K. Abe, "An analysis of human learning process on manual control of complex systems," *Ergonomics*, vol. 42, no. 5, pp. 287–294, 2006 (Japanese).
- [10] N. Homma, S. Kato, T. Goto, et al., "Human brain activities related to manual control of a nonholonomic system: an f-MRI study," *International Journal of Advanced Computer Engineering*, vol. 2, no. 2, pp. 129–133, 2009.
- [11] K. Samejima, K. Katagiri, K. Doya, and M. Kawato, "Multiple model-based reinforcement learning of nonlinear control," *The Journal of The Institute of Electronics, Information and Communication Engineers*, vol. 84, no. 9, pp. 2092–2106, 2001 (Japanese).



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

