

Development of Body Mapping from Human Demonstrator to Inverted-Pendulum Mobile Robot for Imitation

Sataya TAKAHASHI*, Yasutake TAKAHASHI,
and Yoichiro MAEDA

*Faculty of Engineering, University of Fukui
Graduate School of Engineering, University of Fukui
3-9-1, Bunkyo, Fukui, Fukui, 910-8507, Japan
Email: {stakahasi,yasutake,maeda}@ir.his.u-fukui.ac.jp

Takayuki NAKAMURA

Faculty of Systems Engineering, Wakayama University
Sakaetani 930, Wakayama 640-8510, Japan
Email: ntakayuk@sys.wakayama-u.ac.jp

Abstract—This paper addresses development of body mapping from a human demonstrator to an inverted-pendulum mobile robot for imitation. An inverted-pendulum mobile robot with torso and body links learns a dynamic kicking motion shown by a human. The robot observes the human demonstration with a camera, extracts the human region in each of images, maps the region to its own two links, estimates the link posture trajectories, and starts kicking motion learning based on the trajectory parameters for imitation. The mapping parameter gives an important role for successive imitation. A reasonable and feasible procedure of learning from observation for an inverted-pendulum robot and development of the body mapping is proposed and investigated in this paper.

Keywords—imitation; corresponding problem; reinforcement learning; policy gradient method; inverted pendulum wheeled robot

I. INTRODUCTION

In order to develop skills, actions, behavior in a human symbiotic environment, a robot is going to learn something from behavior observation of predecessors or humans. Observation of others makes the behavior learning rapid and therefore much more efficient [1], [2], [3]. Recently, many approaches for robotics imitation have been proposed (for example, [4], [5]). It is desirable to acquire various unfamiliar behavior with some instructions from others, for example, surrounding robots and/or humans in real environment because of huge exploration space and enormous learning time to learn without any a priori knowledge. Therefore, behavior learning through observation has been more important.

However, a robot imitator does not usually share the same embodiment, for example, link configurations, constraints, dynamics, and so on, with a demonstrator. The imitator has to solve a problem of body mapping between the demonstrator and itself, so called, the correspondence problem[6], in an appropriate way. In order to avoid the problem, some conventional approaches focus not on body movement imitation but on target control imitation. For example, Bentivegna et al. [3] shows imitation of playing air hockey and marble maze. The robot does not imitate the motion of the body but the motion of the target, a paddle for the air hockey and knobs on the board for the marble maze. Another approach for the

body mapping problem is to use a robot that has a similar embodiment with a demonstrator. Humanoid robots are popular for imitation of human demonstration[4], [5]. They use motion capture systems to capture human body postures and map them to the body postures of a humanoid robot. They have few body mapping problems and actually it is easy to map the human body links to humanoid robot ones because the demonstrator and the imitator share similar embodiment.

For example, a dolphin imitates human motion as discussed in [6] and a child imitates a long trunk of an elephant with his/her arm. They have different embodiments with the demonstrators and show successful imitation based on mapping the body of a demonstrator to the one of the imitator. The body mapping problem becomes more difficult and important if the robot imitator does not share the embodiment with a demonstrator. Therefore, an appropriate approach is necessary for the problem to realize body movement imitation learning in a human symbiotic environment.

On the other hand, two-wheeled inverted pendulum mobile robots have been developed as a realistic solution for the human symbiotic environment so far[7][8]. A inverted pendulum mobile robot has many advantages over the statically stable wheeled robots and any other biped robots. It requires a smaller amount of space to stand and stay upright than the other wheeled robots and smaller number of actuators than the conventional biped robots. Furthermore, it can easily generate more dynamic motions only by its body balancing control. The two-wheeled inverted pendulum robot has a possibility to be standard platform as a human partner robot. Most of researches regarding to the two-wheeled inverted pendulum robot focuses on the control theory for achieving stable navigation. There are few researches to realize more dynamic motions such as throwing, jumping, pushing, kicking, and so on. Takahashi et al.[9] introduced a reinforcement learning method for generating a dynamic motion such that a two-wheeled inverted pendulum robot kicks a ball far away utilizing its own body dynamics while it keeps standing. However, the learning time is too long to apply the learning method for a real mobile inverted-

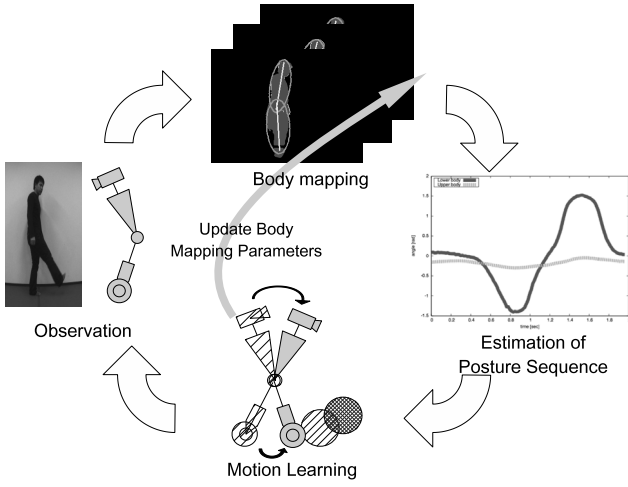


Fig. 1. Imitation Learning

pendulum wheeled robot. Furthermore, they designed initial motions by hand before learning in order to reduce unnecessary useless exploration as much as possible. In order to produce feasible initial motion for efficient learning, imitation is one of the most convincing approaches.

Tamura et al.[10] have proposed a method for recognition of demonstrator's posture for imitation based on reinforcement learning. They focus on how a two-link inverted-pendulum wheeled robot updates body mapping parameters while the demonstrator shows throwing motion after the robot learns the throwing motion. Takahashi et al.[11] investigated imitation learning performance for the two-link inverted-pendulum wheeled robot. A human demonstrator shows kicking motion and a robot learns the motion from observed data. They proposed a reasonable and feasible procedure of learning from observation for the inverted-pendulum wheeled robot. However, the body mapping parameters are tuned by hand and fixed during the imitation.

In this paper, we propose a novel learning method of body mapping parameters for imitation based on previous work[11]. A human demonstrates a ball kicking motion and a two-link inverted-pendulum mobile robot imitates the motion. Updating body mapping parameters based on gradient of evaluation of imitated motion enables the imitator to reasonably recognize demonstrator's body links for imitation and shows better performance of imitated motion than hand-tuned ones.

II. IMITATION LEARNING

Figure 1 shows a flow of imitation learning. First, a two-link inverted-pendulum mobile robot as an imitator observes human demonstrations with a camera. A human region in the captured image is extracted and mapped to two links of the robot with a certain body mapping parameter. A sequence of link postures during the motion is estimated for imitation learning. The robot starts learning the motion from the estimated link posture sequence as initial motion parameters. After the learning, the

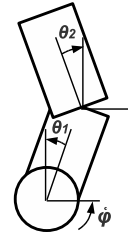


Fig. 2. Model of Two-Link Inverted-Pendulum Mobile Robot

body mapping parameter is updated based on the performance of the learned motion. This procedure repeats until the robot finds a local optimum body mapping parameter. Details of the each step are described in following sections.

III. INVERTED PENDULUM WHEELED ROBOT

Figure 2 shows a model of our inverted-pendulum wheeled robot. The robot body size is 20cm length, 30.5cm width, and 70cm height. The weight is 8kg including the torso, body, wheels, batteries, motors, gears, encoders, an accelerogymometer unit, and control unit. The wheel radius is 8.5 cm and its weight is 500g each. The ball kicked by the robot is a soccer ball, roughly 22cm in diameter, and about 450g in weight. In order to perform comprehensive experiments for learning a kicking motion with the robot, we develop a computer simulation of the robot and the ball using the ODE(open dynamics engine) library.

IV. DESIGN FOR LEARNING KICKING MOTION

The kicking motion is designed with two control layers, a low-level posture controller and a posture generator. The posture controller follows a conventional torque control theory. Torque for the wheels u_1 is calculated as follows:

$$u_1 = -k_1^l(\theta_1 - \theta_d^l) - k_2^l\dot{\theta}_2 - k_3^l\dot{\varphi} - k_4^l\sum(\dot{\varphi} - \dot{\varphi}_d), \quad (1)$$

where θ_1 , θ_d^l , $\dot{\varphi}$, $\dot{\varphi}_d$ are error of body angle, desired body angle, wheel angular velocity, and desired wheel angular velocity, respectively. k_1^l , k_2^l , k_3^l , k_4^l are gains for the body angle, body angular velocity, wheel angular velocity, and accumulated error of wheel angular velocity, respectively. For simplicity, torque for a joint between the body and the torso u_2 is designed as follows:

$$u_2 = -k_1^u(\theta_2 - \theta_d^u) - k_2^u\dot{\theta}_2 - k_3^u\sum(\theta_2 - \theta_d^u), \quad (2)$$

where k_1^u , k_2^u , and k_3^u are control gains.

In this work, we design the kicking motion in combination with two primitive motions (Fig.3). Each primitive motion defines the desired body and torso angles, desired wheel angular velocity, gain parameters, and a period of time. The posture generator sends the control parameters to the posture controller within the period defined by the primitive motion, one by one. The first primitive motion tries to lean forward to kick the ball and the second tries to make the robot kick the ball actually. Before and after the kicking motion, the robot controls itself to stay upright avoiding falling-down, autonomously.

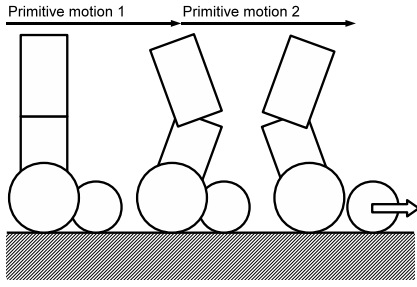
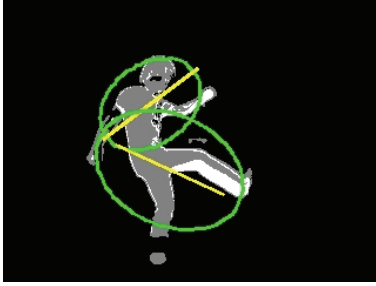


Fig. 3. Two Primitive Motions to Kick a Ball



(a)



(b)

Fig. 4. View of Human Kicking Motion (a) and Upper and Lower Body Recognition (b)

V. OBSERVATION AND INITIAL LEARNING PARAMETERS

Using a camera, a robot observes a human demonstration of kicking a ball. Figure 4(a) shows a picture of observed motion of human kicking behavior. It captures pictures with 30fps and extracts regions of the human in the images with simple background subtraction and frame difference methods. Then, it estimates postures of upper and lower bodies by calculating covariances of the regions in the image. First, it divides the whole human region into two regions, upper body and lower body, with a body mapping parameter r_{bm} . The body mapping parameter r_{bm} specifies a ratio of upper and lower bodies. For example, a whole human region is divided at the center of the region when $r_{bm} = 0.5$.

Then, a region moment M_{pq} is calculated for each region:

$$M_{pq} = \sum_i \sum_j (i - i_G)^p (j - j_G)^q f(i, j) \quad (3)$$

(i_G, j_G) is the center of the region and $f(i, j)$ is pixel value on the point (i, j) which is 1 if the point in the region, otherwise

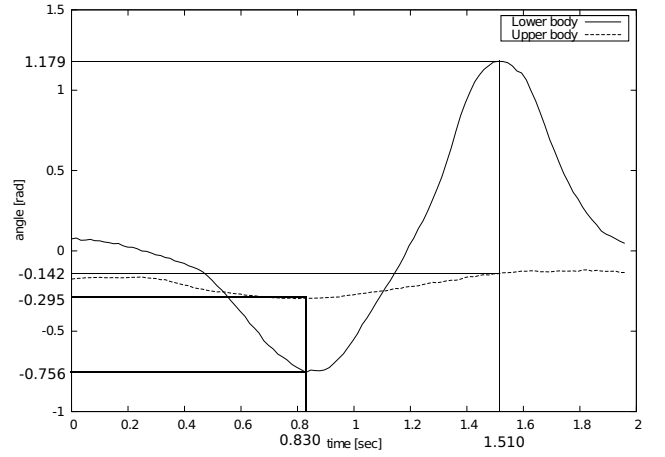


Fig. 5. Posture Estimation during Human Kicking Motion

zero. The posture θ of the region (as the upper or lower body) can be calculated as:

$$\theta = \frac{1}{2} \arctan\left(\frac{2M_{11}}{M_{20} - M_{02}}\right), \quad (4)$$

where p and q are indexes of order of moment for x and y coordinates, respectively. Figure 4(b) shows a extracted human body region and estimated posture of upper and lower bodies.¹ Figure 5 shows a sequence of estimated postures during the observation of human kicking motion.

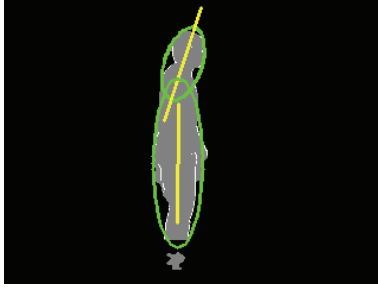
Primitive motion parameters are initially set from the observed postures from Fig.5. In general, a human kicking motion does not consist of two primitive motions as described in Section IV. A learning robot has to extract useful parameter values from observation of human kicking motion. To be concrete, the observed human motion divided into two primitive motions for generating learning initial parameter values. A sequence of an observed trajectory is divided at points each of which lower body posture acceleration is the highest. 0.830 [sec] and 1.510 [sec] are the points to be divided from Fig.5. 0.830 [sec] and 0.680 [sec] are set as periods of time for primitive motion 1 and 2. Posture angles of lower and upper body at 0.830[sec] are -0.126 [rad] and -0.0492 [rad], respectively. Then, desired posture angles θ_d^l , θ_d^u for primitive motion 1 are set -0.126 [rad] and -0.0492 [rad]. In turn, posture angles of lower and upper body at 1.510 [sec] are -0.808 [rad] and 1.30 [rad], respectively. Then, desired posture angles θ_d^l , θ_d^u for primitive motion 2 are set 0.197 [rad] and 0.135 [rad]. Eventually, the initial parameter values with body mapping parameter $r_{bm} = 0.5$ are shown in TABLE I. The table also shows wheel angular velocity ϕ_d and initial control gains for experiments later.

The body mapping parameter r_{bm} affects not only the division of the human region into upper and lower bodies but also the estimation of body postures of the observed motion. The parameter should be tuned appropriately in order to acquire

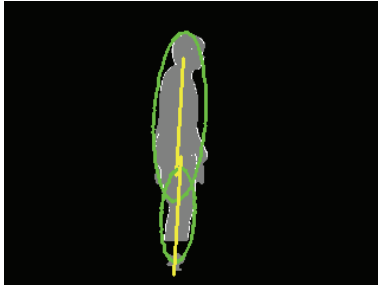
¹The part of ankle is failed to be extracted because the part has similar color with background. However, the effect of the extraction failure is small enough to estimate the posture of the lower body.

TABLE I
INITIAL PARAMETERS OF PRIMITIVE MOTION 1 AND 2 FROM
OBSERVATION WITH BODY MAPPING PARAMETER $r_{bm} = 0.5$ AND
INITIAL CONTROL GAINS

learning parameter	primitive motion 1	primitive motion 2
T [sec]	0.830	0.680
θ_d^l [rad]	-0.126	0.197
θ_d^u [rad]	-0.0492	0.0236
ϕ_d [rad/sec]	0	0
k_1^l	15.0	15.0
k_2^l	-0.3	-0.3
k_3^w	0.01	0.01
k_4^w	0.0	0.0
k_1^u	25.0	25.0
k_2^u	-0.1	-0.1
k_3^u	0.01	0.01



(a) $r_{bm} = 0.3$



(b) $r_{bm} = 0.7$

Fig. 6. Effect of Body Mapping Parameter r_{bm}

proper initial motion parameters for imitation learning. Figure 6 shows an example. Two pictures in Fig.6 are exactly same with different body mapping parameters $r_{bm} = 0.3$ and 0.7. TABLES II and III show extracted primitive motion parameters with body mapping parameter $r_{bm} = 0.3$ and 0.7, respectively. They indicate that different body mapping parameters on a same observed image result in different body postures of the observed motion.

VI. MOTION LEARNING

A simple policy gradient method introduced by Kohl and Stone[12] enables the posture generator to learn the kicking motion, to say more precisely, improve the control parameters of each primitive motion by trail and error. First, it has a initial control parameter and prepares a set of similar parameters by adding small disturbances. It evaluates the control parameters one by one and estimates gradient of the evaluation of the control parameters. Finally, it updates the control parameters

TABLE II
EXTRACTED PRIMITIVE MOTION PARAMETERS WITH BODY MAPPING
PARAMETER $r_{bm} = 0.3$

learning parameter	primitive motion 1	primitive motion 2
T [sec]	0.896	0.674
θ_d^l [rad]	-0.0784	0.0915
θ_d^u [rad]	-0.0770	-0.0638

TABLE III
EXTRACTED PRIMITIVE MOTION PARAMETERS WITH BODY MAPPING
PARAMETER $r_{bm} = 0.7$

learning parameter	primitive motion 1	primitive motion 2
T [sec]	0.830	0.714
θ_d^l [rad]	-0.285	0.288
θ_d^u [rad]	-0.0533	0.0172

according to the estimated gradient of the evaluation. The learning system repeats this procedure and updates the motion parameters to reach the local maximum of evaluation. For further details, please refer to the original paper by Kohl and Stone[12] or our previous work[11]. Figure 7 shows the

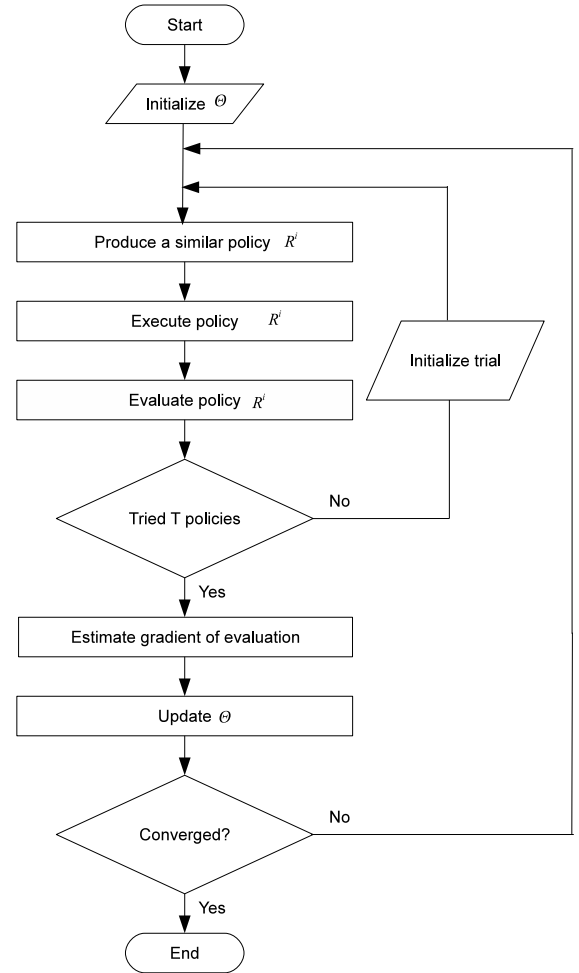


Fig. 7. Learning Procedure

learning procedure.

TABLE IV
STEP SIZE OF LEARNING PARAMETERS

learning parameter	step size
T [sec]	0.001
θ_d^l [rad]	0.01
θ_d^u [rad]	0.01
ϕ_d [rad/sec]	0
k_1	0.01
k_2	0.001
k_3	0.0001
k_4	0.00001

Each primitive motion is defined by the desired body and torso angles θ_d^l , θ_d^u , the desired wheel angular velocity ϕ_d , the gain parameters k_1 , k_2 , k_3 , and k_4 , and a period of time t . The learning parameters are those primitive motion parameters for each, therefore, 16 parameters in all. We determine the evaluation criterion of the kicking motion while learning process in consideration of the following points:

- Traveling distance of the robot body should be small.
- Velocity of the kicked ball should be large.
- Robot should avoid falling down by kicking the ball.

The actual evaluation is defined as follows:

$$E = \begin{cases} w_1 v_b - w_2 l_b^2 - w_f & \text{if it falls down} \\ w_1 v_b - w_2 l_b^2 & \text{else} \end{cases} \quad (5)$$

where l_b , v_b are traveling distance of the robot body during kicking motion and velocity of the kicked ball, respectively. w_1 , w_2 , w_f are penalty for falling down and weights and set as 2.0, 1.0, and 10.0, respectively, in the following experiments.

VII. UPDATE OF BODY MAPPING PARAMETER

Evaluation of the acquired motion $E^f(r_{bm})$ is checked after learning with body mapping parameter r_{bm} . Then, a random disturbance Δr_{bm} is added to the current body mapping parameter, estimation of sequence of upper and lower body postures and recalculation of initial motion parameters are conducted again, and motion learning is restarted. After the learning, the robot evaluates the acquired motions as $E^f(r_{bm} + \Delta r_{bm})$ and $E^f(r_{bm} - \Delta r_{bm})$. Finally, it updates the body mapping parameter to imitate the motion with higher evaluation as follows:

$$r_{bm} \leftarrow \begin{cases} r_{bm} + \Delta r_{bm} & \text{if } E^f(r_{bm} + \Delta r_{bm}) > E^f(r_{bm}) \text{ and} \\ & E^f(r_{bm} + \Delta r_{bm}) > E^f(r_{bm} - \Delta r_{bm}) \\ r_{bm} - \Delta r_{bm} & \text{if } E^f(r_{bm} - \Delta r_{bm}) > E^f(r_{bm}) \text{ and} \\ & E^f(r_{bm} - \Delta r_{bm}) > E^f(r_{bm} + \Delta r_{bm}) \\ r_{bm} & \text{else} \end{cases} \quad (6)$$

$|\Delta r_{bm}|$ is set to 0.1 in this paper.

VIII. EXPERIMENTS

A. Experiment Setup

The learning step size parameters are shown in TABLE IV. The number of deviated policies for estimating gradient of the evaluation T is set to 50 in the experiments.

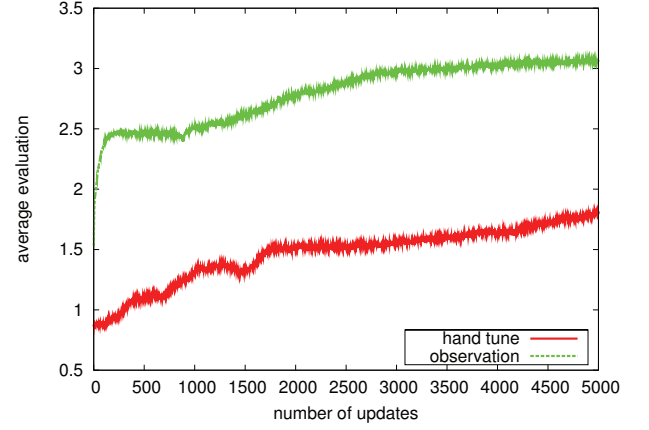


Fig. 8. Motion Evaluation during Learning with Initial Motion Parameters from Observation and Tuned by Hand

TABLE V
HAND TUNED INITIAL PARAMETERS OF PRIMITIVE MOTION 1 AND 2

learning parameter	primitive motion 1	primitive motion 2
period of time [sec]	0.500	0.500
θ_d^l [rad]	-0.052	0.052
θ_d^u [rad]	0.052	-0.052
ϕ_d [rad/sec]	0	0
k_1	15.0	15.0
k_2	-0.3	-0.3
k_3	0.01	0.01
k_4	0.0	0.0

B. usefulness of Imitation learning

Figure 8 shows an example sequence of motion evaluation during learning with initial motion parameters from observation and tuned by hand. Acquired kicking motion with learning from observation is twice as good as the one with hand-tuned initial motion parameters. The hand-tuned parameters are shown in Table V. This result shows usefulness of the imitation learning. Figure 9 shows an example sequence of the learned kicking motion.

C. Effect of Body Mapping Parameter

We investigated motion evaluation after learning while the body mapping parameter r_{bm} is changed gradually from 0.1 to 0.9. Figure 10 shows the result. It shows the acquired motion evaluation at $r_{bm} = 0.6$ is the best and distributed like a mountain. This means that a simple gradient method enable the imitator to find one of the best body mapping parameters by updating the parameters gradually. Figure 10 suggests that the body mapping parameter converges to the best with a simple gradient method if it starts from relatively large range (from 0.4 to 0.9) although the evaluation at $r_{bm} = 0.2$ is local maxima and it might prevent the imitator to find the best parameter value. However, a proper parameter exploration method such as simulated annealing easily avoids the local maxima.

IX. CONCLUSION

This paper addressed development of body mapping from a human demonstrator to an inverted-pendulum mobile robot

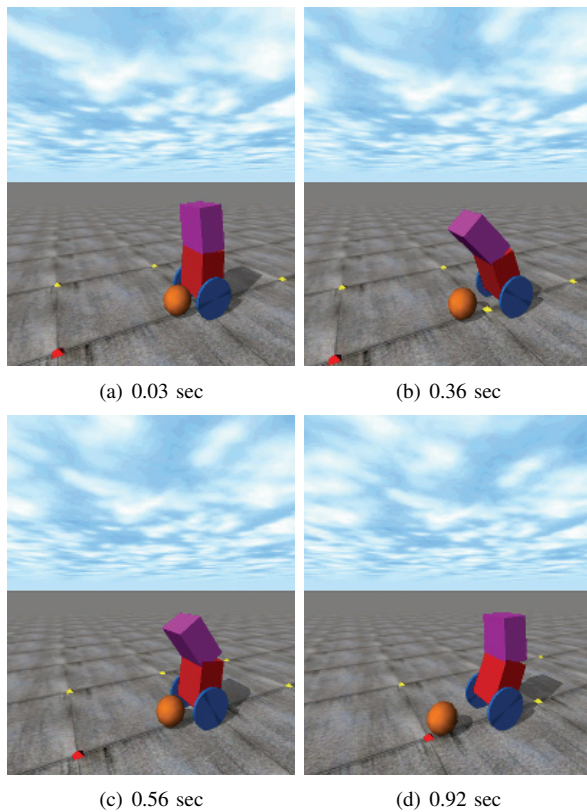


Fig. 9. A Sequence of Ball Kicking Motion after Learning

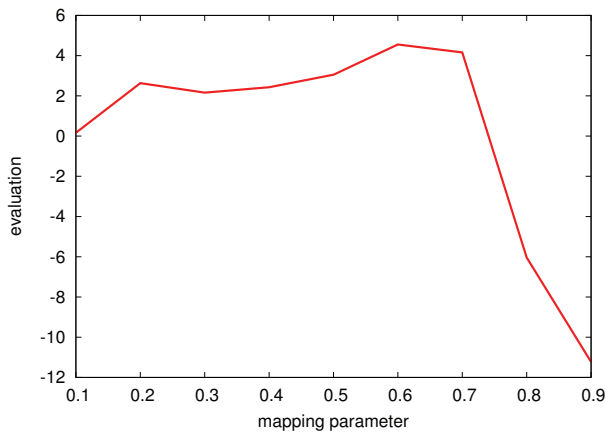


Fig. 10. Acquired Motion Evaluation with Variety of Body Mapping Parameter Values

for imitation. A human demonstrator shows kicking motion and a robot learns the motion from the observed data. In order to imitate an observed motion demonstrated by a human, the robot has to map the body of the demonstrator to its own body links. We proposed a reasonable and feasible method that enables a two-link inverted-pendulum mobile robot to learn a human motion from observation with a simple body mapping technique. Experimental results indicate that imitation gives better initial motion parameters for learning and body mapping parameter for the imitation can be updated based on the evaluation of the acquired motion after the learning. In order to show the proposed method has some generality, we

are planning to apply to not only a kicking motion but also, for example, ball throwing, trapping, and so on.

ACKNOWLEDGMENT

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B), 22700201, 2010.

REFERENCES

- [1] S. D. Whitehead, "Complexity and cooperation in q-learning," in *Proceedings Eighth International Workshop on Machine Learning (M 91)*, 1991, pp. 363–367.
- [2] B. Price and C. Boutilier, "Accelerating reinforcement learning through implicit imitation," *Journal of Artificial Intelligence Research*, vol. 19, pp. 569–629, Dec 2003.
- [3] D. C. Bentivegna, C. G. Atkeson, and G. Chenga, "Learning tasks from observation and practice," *Robotics and Autonomous Systems*, vol. 47, pp. 163–169, 2004.
- [4] S. Schaal, A. Ijspeert, and A. Billard, *Computational approaches to motor learning by imitation*. Oxford University Press, 2004, no. 1431, pp. 199–218.
- [5] T. Inamura, Y. Nakamura, and I. Toshima, "Embodied symbol emergence based on mimesis theory," *International Journal of Robotics Research*, vol. 23, no. 4, pp. 363–377, 2004.
- [6] C. L. Nehaniv and K. Dautenhahn, *The correspondence problem*. Cambridge, MA, USA: MIT Press, 2002, pp. 41–61. [Online]. Available: <http://portal.acm.org/citation.cfm?id=762896.762899>
- [7] S. Ha, Y. Yuta, "Trajectory tracking control for navigation of self-contained mobile inverse pendulum," in *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems '94*, vol. 3, 1994, pp. 1875–1882.
- [8] B. Browning, J. Searock, P. E. Rybski, and M. Veloso, "Turning segways into soccer robots," *Industrial Robot: An International Journal*, vol. 32, no. 2, pp. 149 – 156, 2005.
- [9] Y. Takahashi, H. Nonoshita, T. Nakamura, and Y. Maeda, "Behavioral development of ball kicking motion of a two-wheeled inverted pendulum mobile robot," in *Proceedings of 2010 IEEE World Congress on Computational Intelligence*, vol. DVD-ROM, Jul 2010, pp. 830–835.
- [10] Y. Tamura, Y. Takahashi, and M. Asada, "Clustering observed body image for imitation based on value system," in *Proceedings of 2010 IEEE World Congress on Computational Intelligence*, vol. DVD-ROM, Jul 2010, pp. 917–923.
- [11] S. Takahashi, H. Nonoshita, Y. Takahashi, Y. Maeda, and T. Nakamura, "Inverted-pendulum mobile robot motion learning from human player observation," in *Proceedings of Joint 5th International Conference on Soft Computing and Intelligent Systems and 11th International Symposium on Advanced Intelligent Systems*, vol. CD-ROM, December 2010, pp. 211–216.
- [12] N. Kohl and P. Stone, "Policy gradient reinforcement learning for fast quadrupedal locomotion," in *Proceedings of the IEEE International Conference on Robotics and Automation*, May 2004, pp. 2619–2624.