Association of Whole Body Motion from Tool Knowledge for Humanoid Robots

Dongheui Lee, Hirotoshi Kunori and Yoshihiko Nakamura
Department of Mechano-Informatics
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN
{dhlee, kunori, nakamura}@ynl.t.u-tokyo.ac.jp

Abstract—Since humanoid robots have similar body structures to humans, they are expected to perform various tasks including tool-use manipulation tasks instead of humans. This research studies on learning and performing tool-use manipulation tasks. For tool-use manipulations, understanding the relation between tool motion and whole body motion is crucial. In this paper, a tool-use motion model is designed with tool knowledge and body motion knowledge. The authors propose a method which enables a humanoid robot to associate whole body motion from tool knowledge by adopting the mimesis method from partial observations [1]. When a specific tool trajectory of a tool-use motion is given, appropriate hand motion is associated from the calculated hand motion, appropriate whole body motion is associated successively. The proposed algorithm is implemented on a humanoid robot.

Index Terms—Tools, Body schema, Mimesis Model

I. INTRODUCTION

In order to coexist with humans in human society, robots need to understand human motions. In particular, for humanoid robots which have similar body structures to humans, the ability to learn motion patterns by imitating humans and to perform tasks instead of humans is highly desirable. Because tasks in daily life often involve tool-usage, this paper studies on humanoid robot’s intelligence for tool-use manipulation tasks.

The imitation learning mechanism provides a means of automatic programming of complex systems without extensive trials or complex programming. Since the neuroscience evidence of motor primitives and mirror neurons in humans and other primates have been discovered [2], a number of researchers have developed models for robot imitation learning, inspired by the neuroscience evidences. Bentivegna and Atkeson [3] used the idea of primitives for motor learning to play marble maze. Billard and Mataric [4] used connectionist-based approaches to represent movements. Fod et al. [5] automatically derived primitives through an off-line process of segmentation and application of the principal component analysis. Inamura et al. [6] proposed the mimesis model which is inspired by the bidirectional structure of mirror neurons for motion recognition and generation.

Most previous works on imitation learning and dynamic motion modeling focus on humanoid body motion only [6] [7] or they are highly task-specific [3] or tool-specific. Relatively small amount of research has been carried out for humanoid tool-use. Ogura et al. [8] presented geometric models of tools for humanoid motion generation. The information for grasping and for attention are embedded in the tool models. Nabeshima and Kuniyoshi [9] constructed a tool-use model for a blind retrieving manipulation task.

Recent neuroscience research [10] revealed that, when humans use a tool to reach for a distant object, the extended motor capability is followed by changes in specific neural networks that hold an updated map of body shape and posture. This changes are compatible with the notion of the inclusion of tools in the ‘body schema’, as if our own effector (e.g. the hand) were elongated to the tip of the tool.

In this paper, tool-use motion knowledge consists of two models; tool manipulation knowledge and body motion knowledge. Tool manipulation knowledge contains information of the tool and grasping hand. Body motion knowledge contains motion information of full body including grasping hand without tool information. These models are learned from real human data of tool-use motions. By separating tool-use motion into two models, changes of the body schema with and without a tool can be implemented easily by including and excluding the tool manipulation model.

The main contribution of this research is development of a method to associate whole body motion from tool knowledge. Based on the neuroscience evidence of tool-to-body assimilation [10], a tool can be considered as an effector of body and trigger the appropriate whole body motion for a specific tool-usage. When playing tennis (tool-use motion), in order to hit a ball (task in environment), a trajectory of a tennis racket (tool) is decided. When the tool trajectory is known, the tool trajectory becomes a trigger to generate appropriate hand motion. The hand motion becomes the other trigger to associate whole-body motion. By adopting the mimesis method from partial observations [1], consistent whole body motion is generated from partial information, such as the tool trajectory. Compared to the inverse kinematics method, the proposed method provides natural motion because tool-use motions are learned from real human motion patterns via the imitation mechanism.
II. MIMESIS FROM PARTIAL OBSERVATION

A. Motion Representation by Hidden Markov Models

The mimesis model as shown in Fig. 1 was developed as a mathematical model, inspired by the mirror neuron system [6]. It is a bidirectional computational model which performs three functions: motion learning, motion recognition and motion generation with the concept of proto-symbols.

![Fig. 1. The conceptual diagram of the mimesis model.](image)

In the motion learning procedure, time series of an observed human motion are abstracted as HMM parameters $\lambda = (A, B, \pi)$ via the Expectation Maximization (EM) technique. The trained HMM $\lambda = (A, B, \pi)$ is called a proto-symbol. $A$ is the state transition probability matrix. $\pi$ is the initial state probability vector. $B$ is the observation emission probability distribution and is represented with a Gaussian distribution. $B = \{\mu, \Sigma\}$. $\mu$ and $\Sigma$ are the mean vector and covariance matrix. Once HMMs are trained, the trained HMMs are used both for motion recognition and motion generation.

B. Whole Body Motion Imitation from Partial Data

The hidden observation is an innate and essential problem in imitation. In the extended mimesis model [1] shown in Fig. 3, algorithms of motion recognition from a partial observation and proto-symbol based duplication of an observed motion are developed. By combining the two algorithms, imitation of whole-body motion from occluded observations is achieved.

![Fig. 2. Periodic continuous HMM](image)

When the human motion is partially invisible, only visible parts are compared to the corresponding parts of HMMs in the motion database. The human motion is recognized as the HMM that has the highest likelihood. Then, the humanoid generates its whole-body motion from the HMM by taking account of the observed motion in computing state transition [1].

The decision of which part is visible or invisible is not easy problem to solve in the real world. Our previous work [11] showed a 3D motion recovery from monocular image sequences without prior knowledge of visible parts. In this paper, since we study on association of whole body motion from tool information, it is clear which part is visible: only tool information is given.

C. Recognition of motion patterns with missing data

Motion recognition is a problem to find the most probable proto-symbol for the input observation sequences $x$.

$$\lambda^* = \arg \max_{\lambda} P(x|\lambda)$$  \hspace{1cm} (1)

Log-likelihood, $\log P(x|\lambda)$, for HMMs is calculated by the forward-backward algorithm. When there are missing motion elements in the input observation sequence $x_t$, the emission probability distribution $b_i(x_t)$, which is represented with a Gaussian distribution, is modified.

$$b_i(x_t) = \frac{\exp\{-\frac{1}{2}(x_t - \mu_i)^T\Sigma_i^{-1}(x_t - \mu_i)\}}{\sqrt{(2\pi)^M \det \Sigma_i}}$$  \hspace{1cm} (2)

If $k$-th element $\{x_k\}_t$ of observation $x_t$ is missing, a corresponding element $\{\mu_k\}_i$ of a mean vector $\mu_i$ or that $\{\Sigma_{kk}\}_i$ of a covariance matrix $\Sigma_i$ is set as following equations. Either eq. (3) or eq. (4) is substituted into eq. (2),

$$\{x_k\}_t = \{\mu_k\}_i = *$$  \hspace{1cm} (3)

$$\{\Sigma_{kk}\}_i = \infty$$  \hspace{1cm} (4)

so that the invisible motion elements do not affect the output probability density function. In eq. (3), $*$ indicates a constant
value. In our experiments ∗ is set to zero, and any other constant values are also fine.

D. Proto-symbol based Motion Duplication

Motion patterns are decoded using the expectation operator in the stochastic model. The general motion generation is a two-stage stochastic process; state transition generation and motion output generation from the state sequence. Invisible data in the observed motion sequences can be estimated by applying previous knowledge and current observations. By estimating the optimal state sequence using the Viterbi algorithm [12], the imitated motion pattern is temporally synchronized with the observation. Different motion patterns (e.g., fast and normal-speed walk motions) corresponding to the same proto-symbol (e.g., walk proto-symbol) can be imitated with different temporal sequences. This allows situated motion generation by temporal synchronization.

When calculating the optimal state-sequence, the output probability density function is required. In order for the invisible motion elements not to affect the output probability density function, for the invisible motion elements, either \( x - \mu = * \) or \( \Sigma = \infty \) is applied, where ∗ indicates a constant value. After the optimal state sequence for each chain is obtained, the output observation sequence is calculated according to its output emission probability distribution in state \( i \), i.e., \( b_i(x) \). The simplest way to generate the output is by taking the mean vectors of each gaussian for the state-sequence. After the motion trajectory is generated, post processing like low-pass filtering or averaging over repetition is adopted to eliminate the artifacts caused by discrete state switching and generate a smooth trajectory for use as a motor command input.

III. TOOL-TO-BODY ASSIMILATION

A. Tool Information

When using a tool, humans feel as if our own effector (e.g., the hand) were elongated to the tip of the tool [10]. From experiences of tool-use motions, humans learn how to grasp the tool and how to move whole body for a specific tool usage. In order to satisfy the desired tool information for a specific tool-use motion, the motion of our own effector and whole-body is associated naturally. For instance, in order to hit a ball with a tennis racket, the trajectory of position and orientation of the tennis racket is estimated. In order to follow the tennis racket trajectory, appropriate hands’ motions and even whole body motion are estimated.

Tool information is presented with following nine dimensional vectors.

- \( p_{rep} \): Tool position (3D) in human’s basebody coordinates.
- \( R_{tool} \): Tool orientation (3D) in global coordinates.
- \( p_{grasp} \): Tool grasping position (3D) in tool coordinates.

The reasons to choose the above variables are as follows:

- Tool position is described in human basebody coordinates in order to address the inclusion of a tool in the ‘body schema’.
- Tool should also be addressed in global coordinates because it is a mediator to interact with environment.
- Tool grasping is a mediator between a tool and a human body. How and where to hold the tool is necessary information for tool-use.

The human basebody coordinates and global coordinates are displayed in Fig. 6. The tool coordinates are displayed in Fig. 5.

In order to present the tool information, three artificial markers are attached to each tool and they are not on a same line, as shown in Fig. 4. Detail definition of tool information is explained with an example of a tennis racket shown in Fig. 5. Tool information of the other tools is described in the same manner.

Tool position \( p_{rep} \) is the mean position of three markers.
attached to the tool in demonstrator’s basebody coordinates.

\[ p_m = \frac{p_1 + p_2 + p_3}{3} \]
\[ p_{rep} = R_{\text{hip}}^T(p_m - p_{\text{hip}}) \]

(5)

where \( p_1, p_2 \) and \( p_3 \) are absolute positions of three markers attached to the tool. \( R_{\text{hip}} \) and \( p_{\text{hip}} \) are the human basebody orientation matrix and position vector.

Tool orientation \( R_{\text{tool}} \) is described by eq. 6. Let \( p_1 \) be the origin of the tool coordinates.

\[ e_x = \frac{\begin{pmatrix} p_1 - p_1 \\ p_2 - p_1 \\ p_3 - p_1 \end{pmatrix}}{\left\| \begin{pmatrix} p_1 - p_1 \\ p_2 - p_1 \\ p_3 - p_1 \end{pmatrix} \right\|} \]
\[ e_z = \frac{\begin{pmatrix} p_2 - p_1 \\ p_3 - p_1 \end{pmatrix}}{\left\| \begin{pmatrix} p_2 - p_1 \\ p_3 - p_1 \end{pmatrix} \right\|} \]
\[ e_y = e_x \times e_z \]
\[ R_{\text{tool}} = \begin{bmatrix} e_x & e_y & e_z \end{bmatrix} \]

(6)

Tool grasping position \( p_{\text{grasp}} \) is the position of the right hand in the tool coordinates.

\[ p_{\text{grasp}} = R_{\text{tool}}^T(p_{\text{right hand}} - p_1) \]

(7)

where \( p_{\text{right hand}} \) is the absolute position of the right hand. For simplicity, grasping only by the right hand is considered.

B. Knowledge of Tool Manipulation and Body Motion

Tool Manipulation Knowledge: Tool manipulation knowledge contains information of a tool and a hand during tool use. Time sequence of the tool information and the grasping hand for each motion pattern is abstracted through a hidden Markov model (HMM). The acquired HMM is called a Tool-Hand HMM (TH-HMM). The hand information consists of right hand position in human basebody coordinates and orientation in global coordinates. The body information consists of basebody’s height in global coordinates (1D), basebody orientation in global coordinates (3D), basebody translational velocity (3D), and the other joint angles. The basebody translational velocity is presented as translation vector with respect to the one-step prior basebody coordinates.

C. Whole Body Motion Estimation From Tool Motion

The main concept of the proposed method is illustrated in Fig. 7. To achieve a task in an environment during a tool-use motion, the trajectory of a tool is estimated. Let say this tool trajectory is known. The tool trajectory becomes a trigger in associating whole-body motion. The best matching TH-HMM for the desired tool trajectory is found among dataset of TH-HMMs by the algorithm of motion recognition from partial observation, whose detail is addressed in section II-C. Appropriate effector (a grasping hand) motion for the desired tool trajectory is estimated by the algorithm of proto-symbol based duplication of an observed motion, whose detail is explained in section II-D.

The best matching BH-HMM for the estimated hand motion is found among dataset of BH-HMMs by the algorithm of motion recognition from partial observation. Appropriate full body motion is generated from the best BH-HMM by the algorithm of proto-symbol based duplication of an observed motion. Shortly saying, the associated whole body motion with the hand trajectory is generated.

If estimating a whole body motion which satisfy the constrained trajectory only by inverse kinematics, there might be many possible solutions because human/humanoid model is highly redundant. Some solutions might not be natural-looking. Yamane et al. [13] addressed a similar problem in animation of manipulation tasks of human figures. In order to satisfy geometric constraints during manipulation tasks and to generate natural-looking motions, they proposed to combine an iterative inverse kinematics method and a data-driven method. In a similar sense, our proposed algorithm provides a natural whole body motion, because TH-HMMs and BH-HMMs are learned via both a model-driven and data-driven way. Motion patterns are captured dataset of real human motion patterns (data-driven). The motion patterns are converted into joint angles and basebody position/orientation by inverse kinematics (model-driven). The dataset of inverse kinematics results are abstracted into stochastic models.
With the proposed method, after building database of motions, there will be no need to attach markers to a human subject. Markers attached to a tool are only required in order a humanoid to recognize a motion and to imitate the subject. This relieves some inconvenience of experimenting setting.

![Diagram of motion estimation from tool motion]

**IV. EXPERIMENTS**

The proposed method is tested on a data of 12 different human tool-use motion types, obtained through a motion capture system: WoodenSword_Dou, WoodenSword_Kesagiri, WoodenSword_Men, WoodenSword_Tsuki, GolfClub_FullSwing, GolfClub_HalfSwing, CoffeeCup_Drink, CoffeeCup_Move, CoffeeCup_Reach, TennisRacket_BackhandStroke, TennisRacket_ForehandStroke, and TennisRacket_Serve. Each motion type contains multiple observations, as shown in table I. The sampling time of motion data is 30 msec and the average length of each motion pattern is 2.61 second. Figure 8 shows selected frames of captured motions from the dataset. From arbitrarily selected observations among the dataset of 12 motion patterns, 12 TH-HMMs (tool manipulation knowledge) and 12 BH-HMMs (body motion knowledge) are learned via the EM algorithm. The number of observations used for learning is shown in table I.

### A. Implementation to a Skeleton Model

The proposed algorithm is implemented on a skeleton model shown in Fig. 6. The skeleton model has 39 joint angles, consisting of 3 joints actuating the head, 9 joints in each of the arms, 6 joints in each of the legs, 3 joints in the lumbar vertebra and 3 joints in the rib vertebra. When using the skeleton model, the output vector of BH-HMMs is composed of 52 dimensional elements. The output vector descriptions of TH-HMMs is independent to a model.

The first experiment validates estimation of hand motion from a given tool trajectory. The desired tool trajectory is a WoodenSword_Kesagiri motion pattern. The given tool trajectory is selected from dataset which are not used for TH-HMM learning and BH-HMM learning. The log-likelihood to generate the given tool trajectory from each TH-HMM is calculated and shown in Table II. The table shows that WoodenSword_Kesagiri TH-HMM has the highest likelihood. It is shown that the correct tool manipulation symbol is found from the partial observation.

From the best TH-HMM, associated right hand motion with the given tool trajectory is generated by the algorithm of *proto-symbol based duplication of observed motion*. In Fig. 9, the generated hand motion is drawn with dotted lines and compared to the true hand motion which is drawn with solid lines. Note that sequences of true whole body motion are shown.

### TABLE I

<table>
<thead>
<tr>
<th>Tool</th>
<th>Motion</th>
<th>No. of observations</th>
<th>No. of Data for Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>WoodenSword_Dou</td>
<td>Kesagiri</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>WoodenSword_Men</td>
<td>Tsuki</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>GolfClub_FullSwing</td>
<td></td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>GolfClub_HalfSwing</td>
<td></td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>CoffeeCup_Drink</td>
<td></td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>CoffeeCup_Move</td>
<td></td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>CoffeeCup_Reach</td>
<td></td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>TennisRacket</td>
<td>BackhandStroke</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>TennisRacket</td>
<td>ForehandStroke</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>TennisRacket</td>
<td>Serve</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
TABLE II
TH-HMM RECOGNITION RESULTS (Kesagiri)

<table>
<thead>
<tr>
<th>TH-HMM name</th>
<th>log-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>WoodenSword_Dou</td>
<td>-∞</td>
</tr>
<tr>
<td>WoodenSword_Kesagiri</td>
<td>-∞</td>
</tr>
<tr>
<td>WoodenSword_Men</td>
<td>-∞</td>
</tr>
<tr>
<td>GolfClub_HalfSwing</td>
<td>-1.04 × 10^2</td>
</tr>
<tr>
<td>GolfClub_FullSwing</td>
<td>-1.48 × 10^2</td>
</tr>
<tr>
<td>CoffeeCup_FullSwing</td>
<td>-∞</td>
</tr>
<tr>
<td>CoffeeCup_HalfSwing</td>
<td>-∞</td>
</tr>
<tr>
<td>TennisRacket_Move</td>
<td>-8.12 × 10^3</td>
</tr>
<tr>
<td>TennisRacket_Reach</td>
<td>-9.44 × 10^3</td>
</tr>
<tr>
<td>TennisRacket_ForehandStroke</td>
<td>-9.86 × 10^4</td>
</tr>
</tbody>
</table>

TABLE III
GENERATION ERRORS, AFTER ESTIMATING HAND MOTION (Kesagiri)

<table>
<thead>
<tr>
<th>Variable [unit]</th>
<th>Average Error</th>
<th>Max Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right Hand x pos [m]</td>
<td>0.022</td>
<td>0.073</td>
</tr>
<tr>
<td>Right Hand y pos [m]</td>
<td>0.060</td>
<td>0.145</td>
</tr>
<tr>
<td>Right Hand z pos [m]</td>
<td>0.056</td>
<td>0.149</td>
</tr>
<tr>
<td>Right Hand x euler angle [rad]</td>
<td>0.144</td>
<td>0.474</td>
</tr>
<tr>
<td>Right Hand y euler angle [rad]</td>
<td>0.052</td>
<td>0.136</td>
</tr>
<tr>
<td>Right Hand z euler angle [rad]</td>
<td>0.062</td>
<td>0.180</td>
</tr>
</tbody>
</table>

Fig. 9. Generated right hand motion from wooden sword motion (Kesagiri). The generated trajectories (dotted) is compared with the true trajectories (solid).

Fig. 10. Association of whole body motion from wooden sword motion (Kesagiri). The generated trajectories (dotted) are compared with the true (solid): joint angle of the left (left figure) and right (right figure) shoulder.

motions are captured by a motion capture system but not used for the experiments. From the fact that the generated position/orientation are well matched to the true values, we can say that the trajectory of the right hand is correctly associated from the given wooden sword motion. Table III shows the quantitative generation results. In the table, errors denote the differences between the true and generated right hand trajectory. Positional errors in basebody coordinates and euler angular errors in global coordinates are calculated. Mean errors of position and orientation are around 5 [cm] and 5 [deg] and they can be considered as acceptable errors for kesagiri motion.

Then, whole body motion association with the generated right hand motion is investigated. Probability to generate the hand motion from each BH-HMM is calculated by the algorithm of motion recognition from partial observation. WoodenSword_Kesagiri BH-HMM has the highest likelihood. From the best BH-HMM, whole body motion is generated by the algorithm of proto-symbol based duplication of observed motion. The generated motions (dotted lines) of selected body parts are compared to their true motions (solid lines) in Fig. 10. It is shown that the generated trajectories are well matched to the true trajectories. Table IV shows the quantitative generation results. Errors denote the differences between the true and generated trajectory. Hip pitch angular error and right hand positional errors are shown. Figure 11 shows snapshots from the animation of the generated whole body kesagiri motion. The maximum error of hip pitch angle is around 3.55 [deg] and is considered as an acceptable error for kesagiri motion.

In the next experiment, association of whole body motion from a tool trajectory of an untrained tool-use motion is surveyed. For this survey, intentionally the TennisRacket_ForehandStroke BH-HMM is deleted. In this experiment, a tennis racket trajectory of forehand stroke is given as a target tool trajectory. The given tennis racket trajectory is selected from dataset which are not used for TH-HMM learning. Probability to generate the given tool trajectory from each TH-HMM is calculated by the algorithm of motion recognition from partial observation. TennisRacket_ForehandStroke TH-HMM has the highest likelihood. From the best TH-
TABLE IV
GENERATION ERRORS, AFTER ESTIMATING BODY MOTION (Kesagiri)

<table>
<thead>
<tr>
<th>Variable [unit]</th>
<th>Average Error</th>
<th>Max Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip pitch angle [rad]</td>
<td>0.022</td>
<td>0.062</td>
</tr>
<tr>
<td>Right Hand x pos [m]</td>
<td>0.024</td>
<td>0.113</td>
</tr>
<tr>
<td>Right Hand y pos [m]</td>
<td>0.074</td>
<td>0.191</td>
</tr>
<tr>
<td>Right Hand z pos [m]</td>
<td>0.059</td>
<td>0.172</td>
</tr>
</tbody>
</table>

Fig. 11. Generated whole body motion from a sword trajectory (Kesagiri)

HMM, appropriate right hand motion for the given tool trajectory is generated by the algorithm of *proto-symbol based duplication of observed motion*.

With the generated hand motion, log-likelihood against each BH-HMM is calculated and shown in table V. GolfClub_FullSwing is found as the best motion symbol. This recognition result coincides with our expectation that golf full swing motion is the closest to tennis forehand stroke motion among the eleven BH-HMMs. By using the GolfClub_FullSwing BH-HMM, whole body motion is associated and shown in Fig. 13. The associated trajectories of selected body parts are displayed in Fig. 12. Solid and dotted curves are true and associated trajectories respectively. The associated motion is not exactly matched to the true human motion. However, it shows that reasonable whole body motion can be associated with the target tool trajectory, even for an unlearned tool-use motion, by applying an alternative symbol, without learning from scratch.

TABLE V
BH-HMM RECOGNITION RESULTS (FOREHAND STROKE)

<table>
<thead>
<tr>
<th>BH-HMM name</th>
<th>log-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>WoodenSword_Dou</td>
<td>∞</td>
</tr>
<tr>
<td>WoodenSword_Kesagiri</td>
<td>-1.14 × 10⁴</td>
</tr>
<tr>
<td>WoodenSword_Tsuki</td>
<td>-∞</td>
</tr>
<tr>
<td>GolfClub_FullSwing</td>
<td>-3.59 × 10⁴</td>
</tr>
<tr>
<td>GolfClub_HalfSwing</td>
<td>-6.26 × 10⁴</td>
</tr>
<tr>
<td>CoffeeCup_Drink</td>
<td>-∞</td>
</tr>
<tr>
<td>CoffeeCup_Move</td>
<td>-2.91 × 10⁴</td>
</tr>
<tr>
<td>CoffeeCup_Reach</td>
<td>-∞</td>
</tr>
<tr>
<td>TennisRacket_BackHand</td>
<td>-4.73 × 10⁴</td>
</tr>
<tr>
<td>TennisRacket_Serve</td>
<td>-1.23 × 10⁴</td>
</tr>
</tbody>
</table>

Fig. 12. Generation results, after estimating whole body motion (Forehand Stroke). The generated trajectories (dotted) are compared with true trajectories (solid). Right hand position and orientation are shown: x (top-left), y (top-right), z (bottom-left), and roll (bottom-right).

Fig. 13. Generated whole body motion from a tennis racket trajectory of an unknown motion (Forehand Stroke). The motion is generated by using TennisRacket_ForehandStroke TH-HMM and GolfClub_FullSwing BH-HMM, because TennisRacket_ForehandStroke BH-HMM has not been trained.

B. **Implementation to a Humanoid Robot**

The proposed algorithm is implemented on a humanoid robot. The robot has 38 degrees of freedom, consisting of 3 joints actuating the head, 7 joints in each of the arms, 6 joints in each of the legs, 1 joint at the waist, 3 joints in each of the hands actuating the fingers, and 1 joint for each toe. During this experiment, degrees of freedom in the fingers were not used after grasping tools. Namely, 32 degrees of freedom are used. When using the humanoid robot, the output vector of BH-HMMs is composed of 45 dimensional elements.

From tool trajectories, associated whole body motions of the humanoid robot are generated. Implementation of three tool-use motions are shown: wooden sword *Dou*, tennis forehand stroke and tennis backhand stroke. The whole body motions are associated from each tool motion. For the implementation, only upper body motions including hip height and orientation are commanded to the robot. Constraints of
Associated whole body motion with a specific tool knowledge is generated in a stochastic way. When a tool trajectory is given, an associated effector trajectory with the given tool trajectory is generated from TH-HMMs. Again, by taking the effector trajectory, a whole body motion is associated. In order to associate a tool motion with a consistent body motion, the algorithms of motion recognition from partial observation and proto-symbol based duplication of observed motion are adopted. By applying the same strategy, whole body motion can be estimated even for an untrained tool-use motion, without learning it from scratch. The proposed method is validated on the human motion data set and implemented to a humanoid robot.

ACKNOWLEDGMENT

This research is supported by Japan Society for the Promotion of Science, “Category S of Grant-in-Aid for Scientist Research” and by Special Coordination Funds for Promoting Science and Technology, “IRI Foundation to Support Man and Aging Society”.

REFERENCES


V. CONCLUSIONS

This paper proposes a method to associate tool knowledge with appropriate whole body motion of the tool-usage. A tool-usage motion is designed with two separate models; tool manipulation knowledge (TH-HMM) and body motion knowledge (BH-HMM). The TH-HMM is learned from time sequence data of tool and effector trajectory. The BH-HMM is learned from time series of whole body motion patterns including an effector trajectory. The changes of the body-schema with and without a tool can be implemented by including and excluding the BH-HMM. Combination of different tools (TH-HMMs) and whole-body motions (BH-HMMs) is possible for a variety of tool-use motion representation.