Dance Step Estimation Method Based on HMM for Dance Partner Robot

Takahiro Takeda, *Student Member, IEEE*, Yasuhisa Hirata, *Member, IEEE*, and Kazuhiro Kosuge, *Fellow, IEEE*

***Abstract*—The main purpose of this paper is to realize an effective human–robot coordination with physical interaction. A dance partner robot has been proposed as a platform for it. To realize the effective human–robot coordination, recognizing human intention would be one of the key issues. This paper focuses on an estimation method for dance steps, which estimates a next dance step intended by a human. In estimating the dance step, time series data of force/moment applied by the human to the robot are used. The time series data of force/moment measured in dancing include uncertainty such as time lag and variations for repeated trials because the human could not always exactly apply the same force/moment to the robot. In order to treat the time series data including such uncertainty, hidden Markov models are utilized for designing the dance step estimation method. With the proposed method, the robot successfully estimates a next dance step based on human intention.**

***Index Terms*—Ballroom dances, dance step estimation, hidden Markov models (HMMs), human intention, human–robot cooperation, mobile robot.**

I. Introduction

In most human–robot coordination systems that have been developed by several researchers, the control architecture is designed so that the robots move passively against force/moment applied by a human and execute tasks in cooperation with a human [1]–[3]. These systems are effective in executing simple tasks such as handling an object. On the other hand, some researchers have proposed pet robots [4], [5], which move actively against interactions among humans and themselves. With information such as sound, light, and a simple interaction by using the touch sensors, etc., these robots could move actively for entertainment or human mental healing. However, human–robot coordination for realizing tasks is not considered in their systems.

If robots could move not only passively but also actively based on human intentions, environments, knowledge of tasks, etc., we could realize more effective human–robot coordination system than the conventional one. Considering the case of coordination among humans, each human would move not only passively but also actively based on such information. In this paper, human–robot coordination with physical interaction between a human and a robot is discussed to execute tasks more effectively, in which the robot moves not only passively but also actively based on such information.

In this paper, a dance partner robot executing ballroom dances with a human is focused as an example of human–robot coordination with physical interaction. In the previous research [6], the concept of the dance partner robot was proposed, and the robot, which was referred to as the Mobile Smart Dance Robot (MS DanceR), and its control architecture, which was referred to as Control Architecture based on Step Transition (CAST), were developed. CAST is composed of three modules, namely: 1) "Knowledge;" 2) "Step Estimator;" and 3) "Motion Generator." Knowledge stores the information on dancing such as basic step trajectories and transition rules for dance steps. Step Estimator estimates a next step based on the rules and human's intention, which is mainly communicated to the robot by interactive force/moment applied between the human and the robot. Motion Generator generates actual motions of the robot based on the trajectories and the physical interactions with the human.

The human–robot coordination would be more successful and effective if the robot could estimate human's intention and behave actively so that the robot helps the human to execute tasks according to the intention. Therefore, recognition of human's intention would be one of the essential robot's functions for realizing the coordination. This paper focuses on the step estimation problem in CAST, i.e., an estimation method for a dance step intended by the human. Two step estimation methods have been developed in the previous research, and the robot could realize dancing with a human successfully [6], [7]. In these step estimation methods, only the instantaneous force/moment information at the step transitions are utilized. Considering the case of dancing among humans, however, it might be difficult for humans to estimate his/her partner's intention using such instantaneous information because a male dancer could not always apply the same lead to his partner, and the lead would include uncertainty such as a time lag and variation for repeated trials. Therefore, it would be more successful to model tendencies of dancer's leads with such uncertainty by using time series of force/moment information. Hidden Markov models (HMMs) [8] are utilized to model the time series data with human's uncertainty because HMMs can stochastically model tendencies of time series data with uncertainty.

HMMs are used successfully in the fields of speech recognition, bioinformatics, etc. In these days, the applications on gesture recognition, control of robots, etc., using HMMs are

Manuscript received December 1, 2005; revised August 10, 2006. Abstract published on the Internet January 14, 2007. This work was supported in part by the Japan Society for the Promotion of Science under Grant-in-Aid for Scientific Research A-2-15206049.

The authors are with the Kosuge and Hirata Laboratory, Department of Bioengineering and Robotics, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan (e-mail: taketaka@irs.mech.tohoku.ac.jp; hirata@irs.mech.tohoku.ac.jp; kosuge@irs.mech.tohoku.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2007.891642

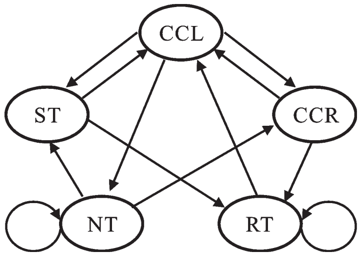

Fig. 1. Step transition in waltz.

studied by some researchers. Lee and Xu have developed a gesture recognition system, which is interfaced to a Cyberglove for use in recognition of gestures from the sign language alphabet [9]. Inamura *et al.* have studied motion generations for robots imitating human motions [10]. Yamada *et al.* have proposed a method for preventing hazardous accidents due to operators' action slip in their use of a Skill-Assist [11].

Communication among partners using physical interactions, which is focused in this paper, would be one of the key factors enhancing efficiency of tasks not only in coordination among humans but also in human–robot coordination. In this paper, HMMs are designed so as to model the physical interactions changing with time, and an estimation method for a dance step intended by a human is proposed using the models. With the proposed method, the communication between a human and the dance partner robot could be realized for more effective coordination. The product of our study could be available for estimating human's intention in human–robot cooperating systems and human assist systems for welfare fields, in which machines would be required to behave not only passively but also actively in coordination with a human based on physical interactions.

In the following part of this paper, first, the brief of step estimation is described. Next, the new step estimation system is designed. In addition, its main module "Calculator" is modeled using HMM. Finally, the estimation system is applied to MS DanceR, and experiments are performed in order to illustrate the validity of the proposed system.

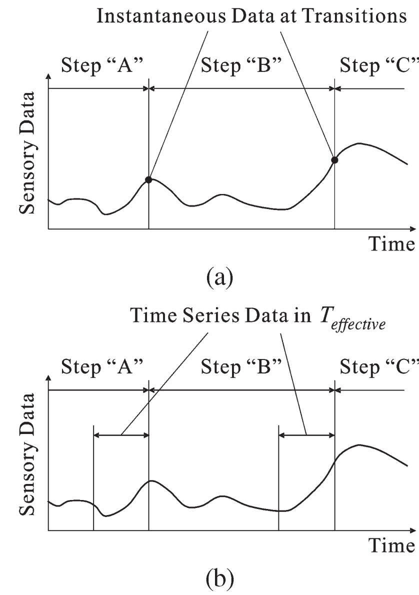

Fig. 2. Sensory data used in Step Estimator. (a) Instantaneous data for step estimations. (b) Time-series data for step estimations.

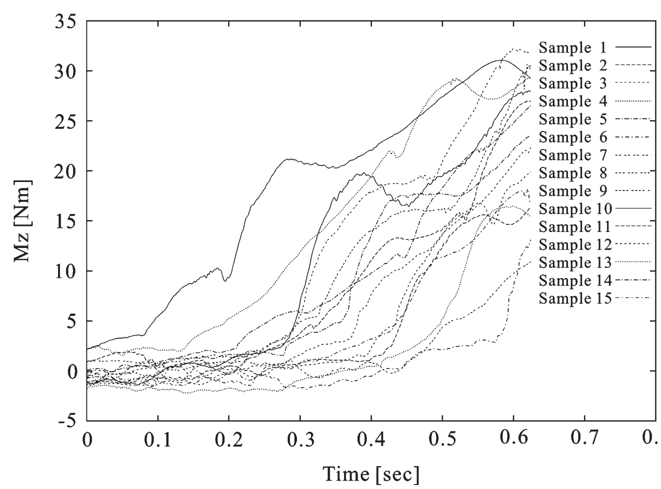

Fig. 3. Sensory data with time lag and variation for repeated trials.

II. Dance Step Estimation

In this paper, a waltz is selected as an example of ballroom dances. For the simplicity of modeling the waltz, five basic steps in the waltz are used, namely: 1) closed change left (CCL); 2) closed change right (CCR); 3) natural turn (NT); 4) reverse turn (RT); and 5) square turn (ST). Transition rules for these steps, which are referred to as "Step Transition," are shown in Fig. 1. A human selects a step according to Step Transition, and the robot stochastically estimates the step. Step Estimator estimates a next step based on Step Transition and human's intention. In this paper, it is assumed that human's intention is mainly communicated to the robot by force/moment applied between the human and the robot.

In the previous research, two methods for step estimations have been proposed, i.e., 1) a method based on production rules using force/moment thresholds [6] and 2) a method based on neural networks (NNs) using force/moment patterns [7]. MS DanceR could dance together with a human using these methods. In these methods, the instantaneous force/moment data at transitions of steps illustrated in Fig. 2(a) have been utilized.

Considering the case of dancing among humans, however, it would be difficult for humans to estimate his/her partner's intention using such instantaneous information. It would be more successful to use the time series of force/moment information illustrated in Fig. 2(b). In this paper, a step estimation system is designed, in which the time series of force/moment information are utilized. In treating the time series data for estimations, the uncertainty of data such as time lag and variation for repeated trials (Fig. 3) has to be considered because the human could not always exactly apply the same force/moment to the robot in dancing. In order to estimate the next step more successfully, the estimation models have to be designed so that they allow such human's uncertainty.

For designing the new Step Estimator, it is assumed that the features of male dancer's leads in dancing are observed effectively during $T_{\mathrm{effective}}$, which is a short time in the

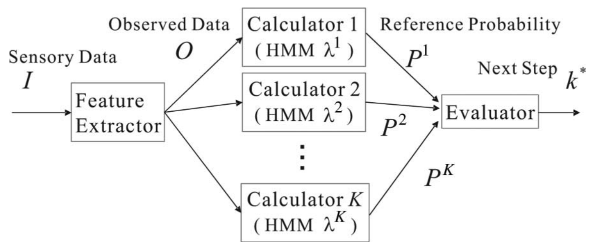

Fig. 4. Step Estimator.

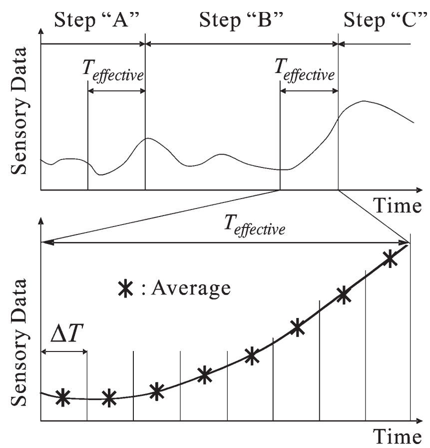

Fig. 5. Feature Extractor.

latter part of a step shown in Fig. 2(b). In the new step estimations, not only the instantaneous sensory data but also the sensory data in the short time $T_{\text{effective}}$ are utilized. HMM is used to design the new Step Estimator. In the new Step Estimator, HMM models tendencies of male dancer's leads with uncertainty, i.e., force/moment applied by a male dancer to the robot, by using time series of force/moment information.

III. Designing Step Estimation System

The Step Estimator shown in Fig. 4 consists of three modules, namely: 1) "Feature Extractor;" 2) "Calculator;" and 3) "Evaluator."

A. Feature Extractor

Feature Extractor outputs features of the time series data $I \in \mathbf{R}^D$ in $T_{\text{effective}}$, where D is the dimension of sensory data. For feature extraction of the data, the averages of the data in each time segment ΔT shown in Fig. 5 are utilized. Feature Extractor outputs the features of the data as the observation sequences $O = \{o(t) \in \mathbf{R}^D | t = 1, 2, \ldots, T\}$, where T is the number of the time segments.

B. Calculator

Calculator outputs the reference probability P^k, which is the kth indicator corresponding to the kth step of all possible K steps limited by Step Transition. This value is used for a selection of the most valid step expected to be intended by a human. In order to treat the time series data including the human's uncertainty, Calculator is designed using HMM. The detail of modeling and an expression of the reference probability P^k are described in Section IV.

C. Evaluator

Evaluator outputs a next step. Two processes are executed in Evaluator. The first process searches the largest reference probability $P^{k^{\max}}$ and the second largest reference probability $P^{k^{\text{2nd}}}$ from $P^k (k = 1, 2, \ldots, K)$, i.e.,

$$P^{k^{\max}} = \max_{1 \le k \le K} P^k \tag{1}$$

$$P^{k^{\text{2nd}}} = \max_{1 \le k \le K, k \ne k^{\max}} P^k. \tag{2}$$

In the second process (3), the fraction $P^{k^{\max}} / P^{k^{\text{2nd}}}$ is compared with a constant κ. Step $k^{\max}$ is outputted as the next step if $P^{k^{\max}} / P^{k^{\text{2nd}}} > \kappa$. If this condition is not true, STOP is outputted, which means that Step Estimator cannot estimate a next step and stops dancing, i.e.,

$$k^* = \begin{cases} k^{\max}, & \text{if } P^{k^{\max}} / P^{k^{\text{2nd}}} > \kappa \\ \text{STOP}, & \text{else} \end{cases}. \tag{3}$$

IV. Designing Estimation Module Using HMM

The sensory data measured in dancing include the uncertainty such as time lag and variation for repeated trials, which arise from the fact that a human cannot always apply the same force/moment to the robot. In order to execute step estimations more successfully, it is needed to model the estimation system that considers the influence of these errors. In this section, Calculator is designed using HMM, which is a main module of Step Estimator. Next, the reference probability P^k outputted by Calculator is expressed.

A. Expression of HMM

HMM is a stochastic method for modeling observed sequences including uncertainty. HMM has three sets of probabilistic parameters, i.e., 1) the probability distribution for the initial state Π, 2) the probability distribution for state transitions A, and 3) the probability distribution for observed sequences B. For convenience, a compact notation is used to indicate these sets, i.e.,

$$\lambda = (\Pi, A, B). \tag{4}$$

These sets are expressed as follows:

$$\Pi = \{\pi_i | i = 1, 2, \ldots, N\}, \qquad \pi_i = P(q(1) = s_i) \tag{5}$$

$$A = \{a_{ij} | i, j = 1, 2, \ldots, N\}, \quad a_{ij} = P(q(t+1) = s_j | q(t) = s_i) \tag{6}$$

$$B = \{b_i(t) | i = 1, 2, \ldots, N, \ t = 1, 2, \ldots, T\}, \quad b_i(t) = P(o(t) | q(t) = s_i) \tag{7}$$

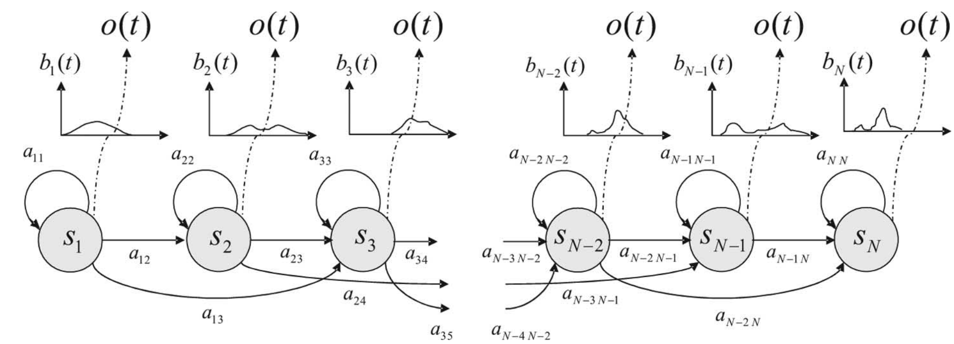

Fig. 6. Continuous left-to-right HMM.

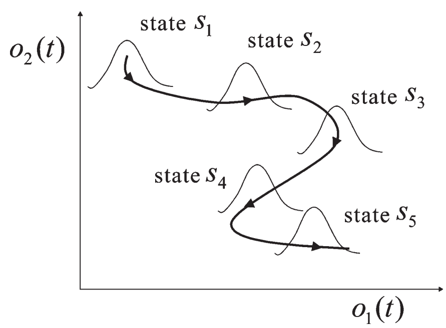

Fig. 7. State space and observation sequence.

where

- N number of states;
- T number of time segments;
- S $= \{s_i |\ i = 1, 2, \ldots, N\}$, set of states;
- Q $= \{q(t)|t = 1, 2, \ldots, T\}$, time series of states;
- O $= \{o(t)|t = 1, 2, \ldots, T\}$, time series of observed data.

In this paper, the continuous left-to-right HMM shown in Fig. 6 is used for modeling Calculator. The HMM models behaviors of training samples, i.e., force/moment measurements in dancing, obtained from repeated trials. In this model, each state shows the pattern of observed data, which are features of sensory data. In other words, the probability that the current state is state s_i is large if observed data are similar to the specific values with respect to the state s_i. The relationship between states and observed data in the case of dimension $D = 2$ is illustrated in Fig. 7, where $o_1(t)$ and $o_2(t)$ denote the first and second components of $o(t) \in \mathbf{R}^2$, respectively. According to a specific continuous probability density included in each state, i.e., $b_i(t)$, as introduced in Section IV-B, observation sequences are outputted from each state.

B. Initial Setting for HMM Parameters

HMM is a stochastic method, and its probabilistic parameters are computed by Baum–Welch algorithm. Initial parameter settings are very important and difficult issues because Baum–Welch algorithm, which is one of the expectation–maximization algorithms, increases the objective function $P(O|\lambda)$ to local maximum.

Considering the left-to-right HMM, at first, the initial values of parameters (Π, A) are set by

$$\pi_i = \begin{cases} 1, & \text{if } i = 1 \\ 0, & \text{else} \end{cases} \tag{8}$$

$$a_{ij} \begin{cases} \neq 0, & \text{if } j \leq i + 2 \\ = 0, & \text{else} \end{cases}. \tag{9}$$

Equations (8) and (9) show that the initial state is always state s_1 and that state transition is limited to three, i.e., $s_i \Rightarrow s_i$, $s_i \Rightarrow s_{i+1}$, and $s_i \Rightarrow s_{i+2}$.

The continuous observed sequence probability $b_i(t)$ is expressed by mixed Gaussian distributions, i.e.,

$$b_i(t) = \sum_{m=1}^{M} c_{im} b_{im}(t) \tag{10}$$

$$b_{im}(t) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_{im}|^{\frac{1}{2}}} \times \exp\left(-\frac{1}{2}\left(o(t) - \mu_{im}\right)^{\mathbf{T}} \Sigma_{im}^{-1} \left(o(t) - \mu_{im}\right)\right) \tag{11}$$

where M is the number of mixed Gaussian distributions; c_{im} is the mixture coefficient for the mth mixture in state s_i; $b_{im}(t)$ is the mth mixture component of mixed Gaussian distributions in state s_i; D is the dimension of sensory data; $o(t) \in \mathbf{R}^D$ are observed data outputted by Feature Extractor; and $\mu_{im} \in \mathbf{R}^D$ and $\Sigma_{im} \in \mathbf{R}^{D \times D}$ are Gaussian mean vector and variance matrix for the mth mixture component in state s_i, respectively. With respect to time t and state s_i, $b_i(t)$ expresses the degree of similarity between sensory data and training samples.

The initial values of c_{im} are set by

$$c_{im} = \frac{1}{M}, \qquad \text{for all } i \text{ and } m. \tag{12}$$

In order to simplify the initial parameter settings for μ_{im} and Σ_{im}, the number of HMM states N is set to be equal to the number of time segments T. In addition, a time index t corresponds to a state index i in the initial parameter settings. The

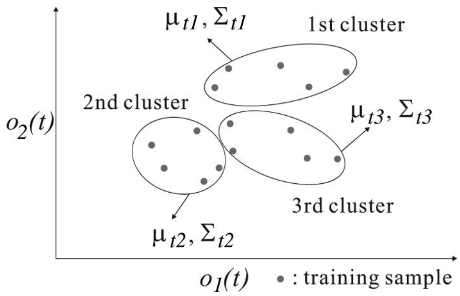

Fig. 8. Initial value settings for parameters μ_{im} and Σ_{im}.

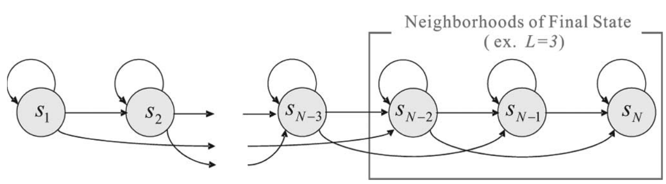

Fig. 9. Neighborhoods of final state.

initial values of μ_{im} and Σ_{im} are calculated based on sets of training samples, i.e., $O^1, O^2, \ldots, O^V$, where V is the number of sets of training samples. With respect to time t, training samples $o^1(t), o^2(t), \ldots, o^V(t)$ are divided into M clusters. In the clustering process, samples are divided so that each cluster has the same number of samples and that the sum of Euclidean distances between samples $o^{v1}(t)$ and $o^{v2}(t)(1 \leq v1, v2 \leq V/M)$ included in each cluster is minimized. Then, $\mu_{im}|_{i=t}$ and $\Sigma_{im}|_{i=t}$ are calculated as averages and variances of samples included in each cluster, respectively. Fig. 8 illustrates the idea in the case of $D = 2$, $M = 3$, and $V = 15$. After initial value settings, the parameters of HMM are learned by the Baum–Welch algorithm.

C. Expression of Reference Probability

The reference probability P^k outputted by Calculator is computed by forward algorithm [8] and expressed as follows:

$$P^k = P\left(O|\lambda^k\right) \sum_{l=1}^{L} P\left(q(T) = s_{N-(L-l)}|O, \lambda^k\right). \qquad (13)$$

The reference probability P^k is a product of two probabilities. One is the probability that HMM λ^k outputs the observed sequences O, which confirms the validity of data. The other is the probability that a state at the last time $q(T)$ exists at any one of the neighborhoods of the final state s_N, i.e., states $s_{N-(L-1)}, s_{N-(L-2)}, \ldots, s_N$, where L is the number of neighborhoods. The idea of neighborhoods is illustrated in Fig. 9. This probability evaluates the approach of the state $q(T)$ to the neighborhoods of the final state s_N. Considering the second term, the observed data are evaluated more strictly because Calculator is designed so that a close relationship between a time t and a state s_i is generated in initial parameter settings.

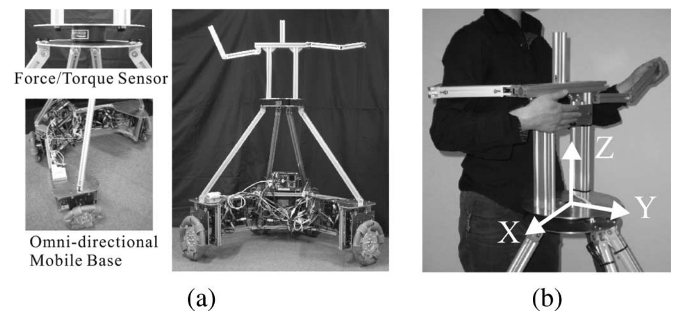

Fig. 10. Dance partner robot "MS DanceR." (a) Robot structure. (b) Interaction with human.

V. Experiments

In this section, experiments on step estimations are performed in order to illustrate the validity of the proposed system.

A. Condition

The robot used in the experiments is shown in Fig. 10(a). An omnidirectional mobile base is used for executing various motions of dance steps. In addition, a force/torque sensor is installed between the upper body and lower body of the robot. A human affixes his/her own body to the robot's upper body, as shown in Fig. 10(b), and applies force/moment to the robot through the interaction. The force/moment applied by the human to the upper body of the robot is aggregated into the sensor, and the force/moment measurements are used for step estimations (the dimension of sensory data $D = 6$). The real-time operating system QNX is used to control the robot, whose control frequency is 1 kHz.

In order to evaluate HMM-based estimations, experiments on step estimations with the previous method based on NNs [7] are performed. Three subjects perform these experiments. Together with the robot, each subject intends to dance the following two step sequences. Needless to say, the robot does not know the step sequences.

Step sequence 1:

$$\text{CCL} \Rightarrow \text{NT} \Rightarrow \text{CCR} \Rightarrow \text{RT} \Rightarrow \text{CCL} \Rightarrow \text{ST} \Rightarrow \text{CCL} \Rightarrow \text{CCR} \Rightarrow \text{CCL}.$$

Step sequence 2:

$$\text{NT} \Rightarrow \text{NT} \Rightarrow \text{ST} \Rightarrow \text{RT} \Rightarrow \text{RT}.$$

All step transitions shown in Fig. 1 are executed by dancing these step sequences. Fifteen trials are repeated in each experiment.

B. Result and Evaluation

Experimental results of subject A are shown in Fig. 11, where each sign of circle (◯), triangle (△), and cross (×) means success, STOP, and a mistake, respectively. Experimental results of the other subjects are obtained in the same manner.

For experimental results of all subjects, the number of successes (◯), the number of STOPs (△), and the number of

Trial	1 CCL	2 NT	3 CCR	4 RT	5 CCL	6 ST	7 CCL	8 CCR	9 CCL
1	=>	○	○	○	○	○	○	○	○
2	=>	○	○	○	○	○	○	○	○
3	=>	○	○	○	○	○	○	○	○
4	=>	○	○	○	○	○	○	○	○
5	=>	○	○	○	○	○	○	○	○
6	=>	○	○	○	○	○	○	○	○
7	=>	○	○	○	○	○	○	○	△
8	=>	○	○	○	○	○	○	○	○
9	=>	○	○	○	○	○	○	○	○
10	=>	○	○	○	○	○	○	○	○
11	=>	○	○	○	○	○	○	○	○
12	=>	○	○	○	○	○	○	○	○
13	=>	○	○	○	○	○	○	○	○
14	=>	○	○	○	○	○	○	○	○
15	=>	○	○	○	○	○	○	○	○

Trial	1 NT	2 NT	3 ST	4 RT	5 RT
1	=>	○	○	○	○
2	=>	○	○	○	○
3	=>	○	○	○	○
4	=>	○	○	○	○
5	=>	○	○	○	○
6	=>	○	○	○	○
7	=>	○	○	○	○
8	=>	○	○	○	○
9	=>	○	○	○	○
10	=>	○	○	○	○
11	=>	○	○	○	○
12	=>	○	○	○	○
13	=>	○	○	○	○
14	=>	○	×	—	—
15	=>	○	○	○	○

(a)

Trial	1 CCL	2 NT	3 CCR	4 RT	5 CCL	6 ST	7 CCL	8 CCR	9 CCL
1	=>	○	○	○	○	○	○	○	○
2	=>	○	○	○	○	○	○	○	○
3	=>	○	×	—	—	—	—	—	—
4	=>	○	○	○	○	○	○	○	○
5	=>	○	○	○	○	○	○	○	○
6	=>	○	○	○	○	○	○	○	○
7	=>	×	—	—	—	—	—	—	—
8	=>	○	○	○	○	○	○	○	○
9	=>	○	○	○	○	○	○	○	△
10	=>	○	○	○	○	○	○	○	○
11	=>	○	○	○	○	○	○	○	○
12	=>	○	○	○	○	○	○	○	○
13	=>	○	○	○	○	○	×	—	—
14	=>	○	○	○	○	○	×	—	—
15	=>	○	○	○	○	○	○	○	○

Trial	1 NT	2 NT	3 ST	4 RT	5 RT
1	=>	○	○	○	○
2	=>	○	○	×	—
3	=>	○	○	○	○
4	=>	○	○	○	○
5	=>	○	○	○	○
6	=>	○	○	○	○
7	=>	○	○	○	○
8	=>	×	—	—	—
9	=>	○	○	○	○
10	=>	○	○	○	○
11	=>	○	○	○	△
12	=>	○	○	○	○
13	=>	○	○	○	○
14	=>	○	○	○	○
15	=>	×	—	—	—

(b)

Fig. 11. Experimental results of subject A. (a) With HMMs. (b) With NNs.

mistakes $(\times)$ are counted and expressed in Table I. In order to evaluate experimental results, the following two methods for calculating success rates are used.

Evaluation method 1:

$$\text{Success Rate} = (\text{Num. of Successful Step Transitions}/\text{Num. of Trials of Step Transitions}) \times 100[\%].$$

Evaluation method 2:

$$\text{Success Rate} = (\text{Num. of Successful Step Sequences}/\text{Num. of Trials of Step Sequences}) \times 100[\%].$$

Success rates evaluated by these methods are expressed in Table II.

TABLE I
EXPERIMENTAL RESULTS OF THE THREE SUBJECTS. (a) WITH HMMS. (b) WITH NNS

(a)

Subject	Trials	Successes(○)	STOPs(△)	Mistakes(×)
A	178	176	1	1
B	177	172	1	4
C	140	125	10	5

(b)

Subject	Trials	Successes(○)	STOPs(△)	Mistakes(×)
A	156	147	2	7
B	142	134	0	9
C	94	67	0	27

C. Consideration

According to Table II, the success rates with HMM-based step estimation method are higher than those with the previous estimation method for each subject. These successes are achieved by using the time series data for step estimations and treating the data with human's uncertainty as stochastic models.

TABLE II
SUCCESS RATE

Subject	Evaluation Method 1 HMM : NN	Evaluation Method 2 HMM : NN
A	98.88 [%] : 94.23 [%]	93.33 [%] : 70.00 [%]
B	97.18 [%] : 94.36 [%]	83.33 [%] : 70.00 [%]
C	89.29 [%] : 71.28 [%]	50.00 [%] : 10.00 [%]

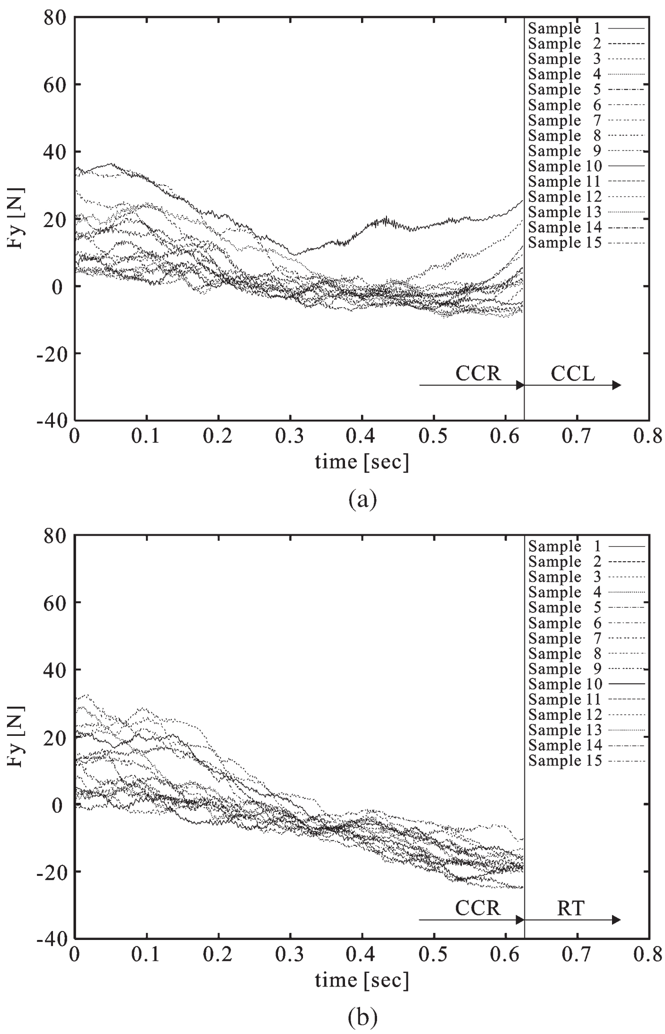

Fig. 12. Force data (F_y: force along the y-axis) applied by subject A. (a) Transition from step CCR to step CCL. (b) Transition from step CCR to step RT.

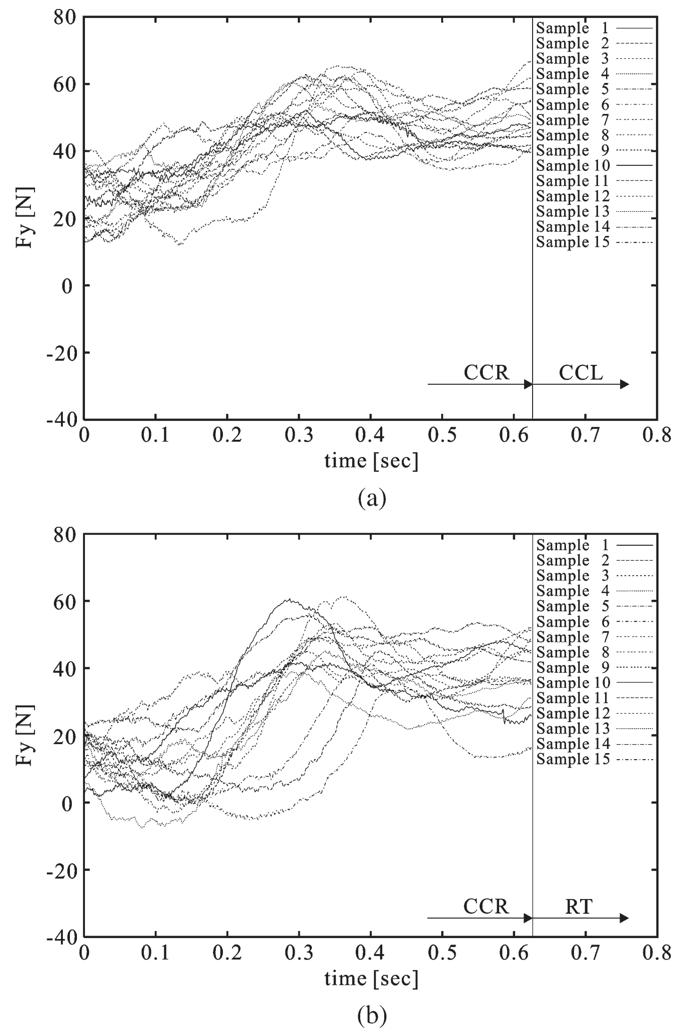

Fig. 13. Force data (F_y: force along the y-axis) applied by subject C. (a) Transition from step CCR to step CCL. (b) Transition from step CCR to step RT.

With respect to subject C, success rates are less than those of the other subjects because subject C is not an experienced dancer. Dance instructors sometimes say that beginners are liable to apply vague leads to their partner because of lack of experience. As an example of leads applied by subjects, force data along the y-axis, i.e., F_y, at the transition to step CCL and step RT are focused. Compared with the leads applied by subject A, which are shown in Fig. 12(a) and (b), variation of the leads applied by subject C is large according to Fig. 13(a) and (b). In addition, a range of instantaneous data at the transition to CCL, i.e., from 40 to 60 N, overlaps with the one at the transition to RT, i.e., from 30 to 50 N. The leads applied by subject C would not be recognizable due to the overlap. This also appears in other components of force/moment. Such vague leads applied by subject C could make step estimations difficult.

According to Table II, however, differences of success rates between HMM-based method and NN-based method about subject C, i.e., $89.29\% - 71.28\% = 18.01\%$, $50.00\% - 10.00\% = 40.00\%$, are much larger than those about the other subjects. Then, we consider the following. When instantaneous data at the transition are focused, with respect to subject C, the overlap mentioned in the previous paragraph could not give the NN-based method a high performance. When time series data are focused, however, overlap between time series data for the transition to step CCL and those for the transition to step RT is not too large to be recognized. From 0.1 to 0.2 s in Fig. 13(a) and (b), for example, differences between the data for the transition to step CCL and those for the transition to step RT are large. For the aforementioned reason, HMM-based method could estimate the next step more successfully than the previous method for all subjects, especially subject C.

These facts illustrate the validity of the HMM-based step estimation method proposed in this paper.

VI. CONCLUSION

In this paper, human–robot coordination with physical interaction was discussed. As an example of the effective human–robot coordination, ballroom dances were taken up. A dance partner robot, which was referred to as MS DanceR, and its control architecture, which was referred to as CAST, were introduced. In CAST, the step estimation system had to behave successfully because the robot's motions were mainly decided by the estimations. In this paper, the step estimation system was improved, and its main modules were designed using HMMs. Experiments on the step estimations were performed using the HMM-based method and the previous method. The validity of the estimation method proposed in this paper was confirmed by experimental results.

Although the proposed estimation method works successfully, the experimental results also describe that completely estimating the behavior intended by a human is difficult. Considering the case of coordination among humans, however, they could not always correctly estimate his/her partner's intention. The more important issues for continuing the coordination would be detecting mistakes as soon as possible and changing and adapting his/her behavior to the correct behavior by perceiving his/her partner's behavior. The future works will address the error recovery problems in order to realize the more effective human–robot coordination.

REFERENCES

[1] H. Kazerooni, "Human robot interaction via the transfer of power and information signals. I. Dynamics and control analysis," in *Proc. IEEE Int. Conf. Robot. and Autom.*, 1989, pp. 1632–1640.

[2] O. M. Al-Jarrah and Y. F. Zheng, "Arm manipulator coordination for load sharing using variable compliance control," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, 1993, pp. 49–54.

[3] Y. Hirata and K. Kosuge, "Distributed robot helpers handling a single object in cooperation with a human," in *Proc. IEEE Int. Conf. Robot. and Autom.*, 2000, pp. 458–463.

[4] T. Shibata, "Emergence of emotional behavior through physical interaction between human and robot," in *Proc. IEEE Int. Conf. Robot. and Autom.*, 1999, pp. 2868–2873.

[5] M. Fujita and K. Kageyama, "An open architecture for robot entertainment," in *Proc. 1st Int. Conf. Auton. Agents*, 1997, pp. 234–242.

[6] K. Kosuge, T. Hayashi, Y. Hirata, and R. Tobiyama, "Dance partner robot - MS DanceR-," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, 2003, pp. 3459–3464.

[7] Y. Hirata, T. Hayashi, T. Takeda, K. Kosuge, and Z. Wang, "Step estimation method for dance partner robot "MS DanceR" using neural network," in *Proc. IEEE Int. Conf. Robot. and Biomimetics*, 2005, pp. 523–528.

[8] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–285, Feb. 1989.

[9] C. Lee and Y. Xu, "Online, interactive learning of gestures for human/robot interfaces," in *Proc. IEEE Int. Conf. Robot. and Autom.*, 1996, pp. 2982–2987.

[10] T. Inamura, I. Toshima, H. Tanie, and Y. Nakamura, "Embodied symbol emergence based on mimesis theory," *Int. J. Robot. Res.*, vol. 23, no. 4/5, pp. 363–377, 2004.

[11] Y. Yamada, T. Morizono, and Y. Umetani, "A method for preventing accidents due to human action slip utilizing HMM-based Dempster–Shafer theory," in *Proc. IEEE Int. Conf. Robot. and Autom.*, 2003, pp. 1490–1496.

Takahiro Takeda (S'05) was born in 1981. He received the M.E. degree from Tohoku University, Sendai, Japan, in 2005. He is currently working toward the Ph.D. degree in the Kosuge and Hirata Laboratory, Department of Bioengineering and Robotics, Graduate School of Engineering, Tohoku University.

His research interests include human–robot cooperation system.

Mr. Takeda is a Student Member of the Robotics Society of Japan. He received the Miura Award from the Japan Society of Mechanical Engineers in 2005 and the IEEE Robotics and Automation Society Japan Chapter Young Award (IROS) from the IEEE Robotics and Automation Society Japan Chapter in 2005.

Yasuhisa Hirata (M'04) was born in 1975. He received the B.E., M.E., and Ph.D. degrees in mechanical engineering from Tohoku University, Sendai, Japan, in 1998, 2000, and 2004, respectively.

From 2000 to 2006, he was a Research Associate in the Department of Bioengineering and Robotics, Tohoku University. From 2002 to 2004, he was a Researcher with the Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency. Since 2006, he has been an Associate Professor in the Department of Bioengineering and Robotics, Tohoku University. His research interests include intelligent control of multiple mobile robots in coordination, human–robot cooperation systems, power assist systems, and walking support systems.

Prof. Hirata is a member of the Robotics Society of Japan and the Japan Society of Mechanical Engineers (JSME). He received the Young Investigator Excellence Award from the Robotics Society of Japan in 2001, the Best Paper in Robotics Award from ROBIO in 2004, the Research Promotion Award from the Aoba Foundation for the Promotion of Engineering in 2004, the JSME Best Paper Award in 2005, the Best Paper Award from the Robotics Society of Japan in 2005, and the Original Paper Award from the FANUC FA and Robot Foundation in 2006.

Kazuhiro Kosuge (SM'00–F'06) received the B.S., M.S., and Ph.D. degrees in control engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 1978, 1980, and 1988, respectively.

He is a Professor in the Department of Bioengineering and Robotics, Tohoku University, Sendai, Japan. From 1980 to 1982, he was a Research Staff Member in the Production Engineering Department, Nippon Denso Company, Ltd. (currently DENSO Company, Ltd.). From 1982 to 1990, he was a Research Associate in the Department of Control Engineering, Tokyo Institute of Technology. From 1990 to 1995, he was an Associate Professor at Nagoya University, Nagoya, Japan. Since 1995, he has been with Tohoku University.

Dr. Kosuge is a Fellow of the Japan Society of Mechanical Engineers (JSME). He is currently an AdCom Member of the IEEE Robotics and Automation Society and the Editor-in-Chief of *Advanced Robotics*. He was Vice President of the IEEE Robotics and Automation Society (1998–2001) and a member of the Board of Trustees of the Robotics Society of Japan (1993–1994 and 2001–2002). He also served on committees for several academic meetings, which include ICRA'95 in Nagoya as a Steering Committee Vice-Co-Chair and IROS2004 in Sendai as the General Chair. He received JSME Best Paper Awards in 2002 and 2005 and the Robotics Society of Japan Best Paper Award in 2005.