Adaptive Behavior, Vol.12, No.2, pp.93-115, 2004

On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system.

Masato Ito
Sony Corporation
6-7-35 Kitashinagawa, Shinagawa-ku, Tokyo, 141-0001, Japan
Tel +81-3-5448-5901, Fax +81-3-5448-6833
masato@pdp.crl.sony.co.jp

Jun Tani
Brain Science Institute, RIKEN
2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan
Tel +81-48-467-6467, FAX +81-48-467-7248
E-mail tani@brain.riken.go.jp

(In press. Adaptive Behavior)

Abstract

This study presents experiments on the imitative interactions between a small humanoid robot and a user. A dynamic neural network model of a mirror system was implemented in a humanoid robot, based on the recurrent neural network model with parametric bias (RNNPB). The experiments showed that after the robot learns multiple cyclic movement patterns as embedded in the RNNPB, it can regenerate each pattern synchronously with the movements of a human who is demonstrating the corresponding movement pattern in front of the robot. Further, the robot exhibits diverse interactive responses when the user demonstrates novel cyclic movement patterns. Those responses were analyzed and categorized. We propose that the dynamics of coherence and incoherence between the robot's and the user's movements could enhance close interactions between them, and also that they could explain the essential psychological mechanism of joint attention.

keywords

entertainment robot, imitation learning, mirror system, recurrent neural network, dynamical systems approach

1 Introduction

Recently, entertainment robotics has attracted much attention not only in the research community, but also in commercial markets. In entertainment robotics, user interactions are one of the most essential issues, and industry has made substantial efforts to implement various interaction schemes in designing their robots. Such examples may be seen in applications of Sony's "AIBO" (Sony, 1999), NEC's "PaPeRo" (NEC, 1999), ATR's "Robovie" (Ishiguro et al., 2001) and so on. Those robots can identify their owners' faces and generate friendly gestures, or understand simple words spoken by the owners and respond to them. However, such interactions might seem to be a bit artificial in the sense that their behaviors are pre-described by the designers. The behavior patterns are mostly pre-programmed, resulting in a lack of behavioral adaptability and flexibility. The authors considered that some abilities for behavior learning might enhance the entertainment features of those robots. It would be fun if behavior or movement patterns could be acquired adaptively through user interactions, rather than merely generating predefined movements. Further, it would be attractive if the robots could generate learned behaviors synchronously with users. The interaction would be akin to a pair of salsa dancers who improvise and switch from one dancing pattern to another while maintaining their joint attention to the dance patterns. How can this sort of dynamic adaptation and coherence in generating movement patterns between robots and their human users be achieved? The authors consider this problem in the context of imitation learning, focusing on its background in developmental psychology as well as neuroscience.

Studies of imitation learning in robotics have been conducted actively for several years, as was recently summarized in (Schaal, 1999; Breazeal & Scassellati, 2002; Dautenhahn & Nehaniv, 2002). In developmental psychological studies, various imitation behaviors of infants have been observed and they are considered to have several developmental stages (Piaget, 1962). Furthermore, it seems that imitation is roughly divided into two stages. Immediate imitation appears early, and is then followed by deferred imitation. Most robot imitation studies can be categorized by these two types.

Many studies have observed that young infants produce imitative behaviors in an immediate manner. Meltzoff and Moore (1977) first discovered that young infants imitate facial and manual gestures of their mothers synchronously. He also proposed the active inter modal (AIM) mapping model to explain the underlying mechanisms of such behaviors (Meltzoff & Moore, 1989). This model assumes a direct mapping between perception and generation modalities. It enables infants to immediately follow arbitrary patterns without memory of each pattern. In some robotic imitation studies, this kind of mechanism has been utilized for producing proto-imitation behaviors. Some studies (Hayes & Demiris, 1994; Dautenhahn, 1995; Gaussier, Moga, Banquet, & Quoy, 1998) showed that a mobile robot could produce following behavior by keeping the distance from a teacher robot constant. Some other studies (Andry, Gaussier, Moga, Banquet, & Nadel, 2001; Kuniyoshi, Yorozu, Inaba, & Inoue, 2003)showed that an arm robot can follow a user's hand movements by learning a mapping between movement perception and generation. In this way, these mechanisms of immediate imitation enable robots to produce intuitive interaction behavior. Further, it enables them to share a context with teacher robots or users for learning their behavior patterns (Hayes & Demiris, 1994; Dautenhahn, 1995; Gaussier et al., 1998).

In later developmental stages, remarkable deferred imitation behaviors are observed. The definition of deferred imitation is that previously perceived behaviors are reproduced after long time intervals. This process involves certain memories for reproducing the observed behavior patterns. If the situation requires reproduction of complex behavior patterns based on learning, the memories should be organized with certain compositional structures. Arbib (1991) proposed in his motor schemata theory that biological agents might have a set of behavior primitives, combinations of which can generate a variety of behavior patterns. Regarding the representation of motor or behavior primitives, several models have been proposed. Amit and Mataric (2002) as well as Inamura, Nakamura, Ezaki, and Toshima (2001) proposed a model of deferred imitation in which behavior primitives are represented through sequential combinations of cluster points (Amit & Mataric, 2002) or a mixture of Gaussian models (Inamura et al., 2001) in the vector space of movement. Ijspeert, Nakanishi, and Schaal (2003) envision motor primitives that consist of oscillatory movements and discrete movements. Each of these two types of primitives can be represented by a specific differential equation, where movement profiles can be modified by changing the equation parameters. Tani and Nolfi (1998); Wolpert and Kawato (1998) proposed a general model of sensory-motor learning in which multiple pairs of forward and inverse models are separately stored in local module networks by utilizing winner-take-all dynamics among the networks. Demiris and Hayes (2002) utilized a similar network architecture for a movement imitation task.

All of these models are characterized by a local representation scheme. Tani and Ito (2003) proposed a distributed representation scheme in a sensory-motor supervised learning context by introducing a recurrent neural network (RNN) model which is characterized by its additional control neural units, called the parametric biases (PB). It was shown that a set of movement patterns can be learned distributedly in a network in which those patterns are recalled by modulating the values of the PB. Furthermore, Tani (2003) extended this model with level structures where it was shown that the lower level performs autonomous segmentation of the continuous sensory-motor flow into movement primitives and the higher level learns to recombine those primitives into sequences.

A considerable problem in reproducing deferred imitation in artificial systems is how to determine the timing of mode switching between observing others' behaviors and reproducing one's own as suggested by Andry et al. (2001); Dautenhahn and Nehaniv (2002). We often observe with surprise that young children spontaneously perform turn taking behaviors, switching between being imitated by and imitating others without explicit end signals to segment the on-going flow of behaviors. Also, the infants often show spontaneous shifts of their joint attention (Trevarthen, 1977). They share behaviors, events and interests in the world with others occasionally. However, the contents of joint attention flips from moment to moment (Baron-Cohen, 1996; Moore & Corkum, 1994). Andry et al. (2001) studied the synchronization in the interaction between two agents that have the ability to modulate the timing of motor generation based on the prediction of a companion's response. It was shown that the dynamic mechanism of synchronization can be utilized for reproducing memorized behavior patterns with appropriate timing. More interestingly, Ikegami and Iizuka (2003) recently explained the underlying mechanisms of spontaneous shifts in turn taking behaviors by the pseudo-stability of the coupled dynamics in paired agents. Dautenhahn and Nehaniv (2002) discussed in their ideas of imitation-in-context that deciding when to imitate or when to learn should be determined in relation to the environment, including other agents. Such aspects were actually tested by Breazeal and Scassellati (1999) using an upper-torso humanoid robot. Breazeal and Scassellati (1999) demonstrated that robot interactions with humans were dynamically changed as the attention shifted with emotional states, which were accumulated in the past history of the interactions.

The idea of mirror neurons by Rizzolatti, Fadiga, Galles, and Fogassi (1996) is indispensable when the mechanisms of deferred imitation are contemplated from the neuroscience perspective. Rizzolatti et al. (1996) discovered that there is a "mirror system" in which those neurons active when the monkey executes a specific object handling behavior are also active when the monkey observes other monkeys or humans carrying out the same behavior. Arbib (2002) speculated in the "Mirror System Hypothesis" that the crucial mechanisms for imitation and language might be routed back to the mirror systems of apes in the evolutional pathway. Further studies on human subjects showed that different regions in the inferior parietal cortex are activated depending on when human subjects imitate others' actions or when they have their actions imitated by others (Georgieff & Jeannerod, 1998; Decety & Grezes, 1999; Decety, Chaminade, Grezes, & Meltzoff, 2002). This suggests that, although the mirror neuron responds to both perception and generation of identical behavior patterns, there is a distinction between actions produced by oneself and those generated by others. The idea of mirror neurons has been utilized in robot imitation learning using hidden markov models (Inamura et al., 2001; Amit & Mataric, 2002) or using neural networks (Billard, 2002).

The current paper presents a novel modeling of imitation using our previously proposed model of the RNN with PB (RNNPB) (Tani, 2003; Tani & Ito, 2003) as a mirror system. Note, that in our work mirroring refers to movements, or patterns of movements, not actions. The model is implemented on a small Sony humanoid robot QRIO (SDR-4X II) through which experiments of imitative interactions with human users are conducted. The imitation shown here is categorized as deferred imitation for given movement patterns. It is not goal-directed or goal-understanding imitation, which is suggested to be unique to humans by Tomasello

(1999).

The robot learns multiple movement patterns, demonstrated by the human users, encoded in different PB vector values in terms of sensory-motor prediction learning (Tani, 1996; Demiris & Hayes, 2002). The learning is conducted in a supervised manner, avoiding correspondence problems (Nehaniv & Dautenhahn, 2001) between the joint angle coordinate system and the view-centered coordinate system of the robot. After learning, we examine how the learned patterns can be re-activated, one by one, by means of entrainment by visually received inputs of the users movements. Although this idea of reproduction of movements by entrainment is related to the studies by Andry et al. (2001); Miyake (2002), our studies investigate different task situations. The difference is the fact that multiple movement patterns are embedded in the RNNPB distributedly (Tani & Ito, 2003) in the current experiments. This feature of the RNNPB would provide highly nonlinear dynamics characteristics to the interactions between the human subjects and the robot. Both coherent and incoherent dynamics would be generated in the interaction between the robot and the user's movements. The current paper attempts to describe such diverse nature of imitative interactions through dynamical systems language (Schoner & Kelso, 1988; Beer, 1995; Gelder, 1998; Tani, 1996; Berthouze, Shigematsu, & Kuniyoshi, 1998; Dauce, Quoy, & Doyon, 2002; Miyake, 2002) and will discuss its psychological meaning as well as its impact on entertainment robotics.

The next section will introduce the imitation model and review the employed neural network architecture of the RNNPB. Then, the experiments using the Sony small humanoid robot QRIO (SDR-4X II) are described. Finally, the current experimental results, their psychological correspondences, and future research directions will be discussed.

2 Humanoid Robot Imitation Learning

2.1 Overview

In the current study, the Sony humanoid robot QRIO (SDR-4X II) (Fujita, Kuroki, Ishida, & Doi, 2003) was used as the experimental platform (see Figure 1). In this experiment, only movement patterns of both arms were considered. Other movements were frozen.

The robot task consists of the learning and interaction phases. In the learning phase, a set of robot movement patterns with different profile is learned and associated with the corresponding user's visually perceived hand movement patterns as off-line. It is actually done in the form of sequence prediction learning for sensory-motor flow as will be detailed in later.

In the learning phase, the target trajectories of the robot are obtained by mapping the user's arm position to the robot joint angles. This mapping was conducted using the following engineering scheme. First, the user's hands' spatial coordinates were optically measured by tracking colored balls on the user's hands. (The depth information was obtained by using the

Figure 1: A user is interacting with Sony humanoid robot QRIO (SDR-4X II).

measured size of the ball.) The obtained spatial coordinates of the user's hands are simply mapped to the robot's hand's 3-D positions in robot centered cartesian coordinates. Next, they are mapped to the robot joint angles (shoulder roll, pitch, yaw, and elbow pitch for each arm) by solving the inverse kinematics of the robot arm, assuming the constraint that elbow pitch is dependent on shoulder pitch. Note that this 3-D measuring is utilized only for generating the motor trajectories for the training data, and not used in the interaction phase.

As summarized in Figure 2(a), the learning process utilizes the paired trajectories of the robot joint angles, obtained by the mapping, and the user's hand positions, as visually perceived by the robot. The training of the employed neural network model (RNNPB) is conducted by using a set of training patterns, corresponding to multiple robot and user movement patterns.

In the interaction phase, the robot attempts to follow synchronously the user's hand movement patterns with predicting their sequences. As shown in Figure2(b), the robot perceives the user's hand movement patterns visually and generates its corresponding movement patterns in robot joint angles. The robot's ability to follow the user depend on the degree to which the user patterns are familiar to the robot, based on prior learning.

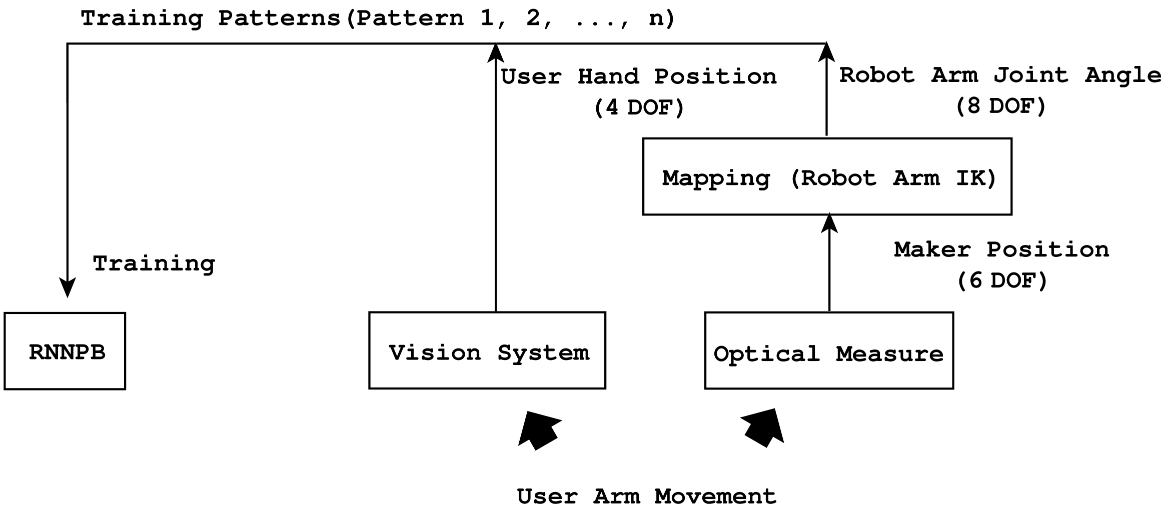
It is important to note that the robot learns not just a static mapping from user hand positions to robot joint angles. Instead, the robot learns the intrinsic dynamics hidden in the target movement patterns. Actually, the robot can generate its movement patterns autonomously without perceiving the user's hand movements, but by imagining it by means of a prediction mechanism, as will be described later. The perception of the hand movement patterns just triggers regeneration of the corresponding dynamic patterns of the robot movement. The underlying mechanism of how the perceptual patterns can trigger the generation of the motor patterns will be the focus of the current study.

It is also noted that the current implementation employs the position feedback control in the joint angles space. Although it might be better to employ the joint torque control scheme in the case of performing more precise control tasks including object handling or quick movements, such implementations might require more computational power in the on-line interaction process of the current implementation which requires the iterative computation of the back-propagation through time (BPTT) algorithm (Rumelhart, Hinton, & Williams, 1986), as will be described later.

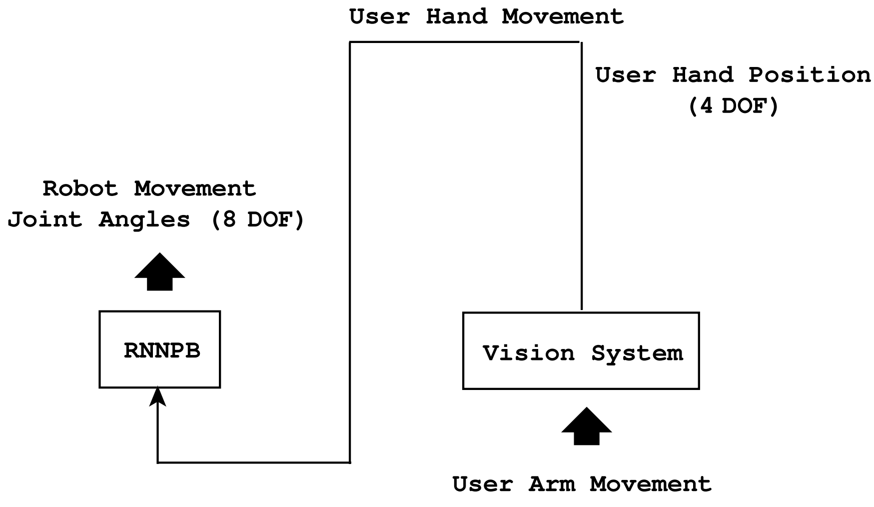
Next, the employed neural network model (RNNPB) is explained.

2.2 Model description

RNNPB is a version of the Jordan-type RNN (Jordan & Rumelhart, 1992) where the PB units allocated in the input layer play the roles of mirror neurons since their values encode both of generating and recognizing the same movement patterns. In generating patterns,



(a) Learning Phase



(b) Interaction Phase

Figure 2: System configurations in learning phase (a) and interaction phase (b).

the PB values function as control parameters for modulating the forward dynamics of the RNN. On the other hand in recognizing patterns, the corresponding PB values for currently perceiving patterns can be dynamically obtained by using the inverse dynamics of the RNN. It is, however, important to note that these recognition and generation processes are conducted simultaneously in the interaction phase i.e.– the robot generates corresponding patterns while recognizing the user's movement patterns. These ideas are detailed in the following associated with descriptions of the learning scheme.

A set of movement patterns is learned, in terms of the forward dynamics of the RNNPB, by self-determining both the PB values, that are differently assigned for each movement pattern, and a synaptic weight matrix that is common for all patterns. The information flow of the RNNPB in the learning phase is shown in Figure 3(a). This learning is conducted using both target sequences of the robot joint angles r_t and the user's hand positions u_t. With given r_t and u_t in the input layer, the network predicts their values at the next time step in the output layer as $\hat{r_{t+1}}$ and $\hat{u_{t+1}}$. The outputs are compared with their target values r_{t+1} and u_{t+1} and the error generated is back-propagated (Rumelhart et al., 1986) for the purpose of updating both the synaptic weights and PB values. Note that the determined synaptic weights are common to all learning patterns, but the PB values are differently determined for each pattern. This scheme will be described in more detail later. c_t represents the context units where the self-feedback loop is established from c_{t+1} in the output layer to c_t in the input layer. The context unit activations represent the internal state of the network.

In the interaction phase, the pre-learned network is utilized without updating the synaptic weights. While the forward dynamics of the RNNPB generates the prediction of the sensory-motor sequences, the PB values are inversely computed by utilizing the error information obtained between the sensory prediction and the outcome. See Figure 3(b) for the information flow of the network in the interaction phase. The visually perceived hand positions are fed into the RNNPB as the target sequences. The RNNPB, when receiving u_t, attempts to predict its next value $\hat{u_{t+1}}$ in the outputs. The generated prediction error from the target value u_{t+1} in the outputs is back-propagated to the PB units and the PB values are updated in the direction of minimizing the error. Note that although the PB plays the role of the inputs for the forward computation, its values are slowly modulated in order to adapt to the current target sequence patterns. If pre-learned hand movement patterns are perceived, the PB values tend to converge to the values that have been determined in the learning phase while minimizing the prediction error. It is guaranteed that by minimizing the prediction error to zero the forward dynamics does not modulate anymore since the PB values converge. Then, the network becomes able to generate the associated motor patterns $\hat{r_{t+1}}$ as previously learned. The robot movement patterns are generated based on the PB values while these values are adapted by perceiving the hand movement patterns. An interesting feature of this model is that generation and perception are performed simultaneously in one neural dynamic

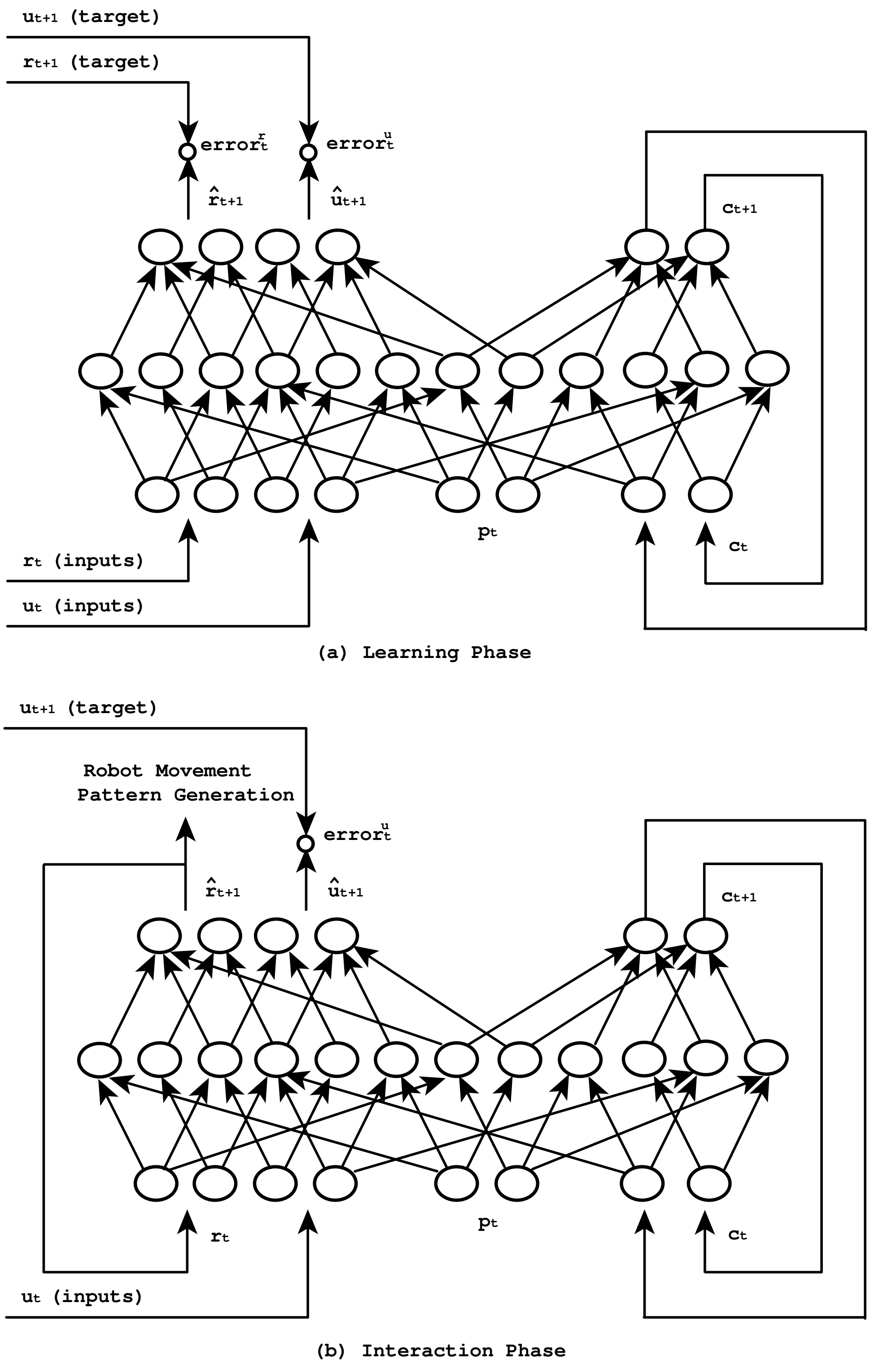


Figure 3: The system flow of RNNPB in learning phase (a) and interaction phase (b).

system.

Note that there could be a local minimum problem in iteratively computing the optimal PB values in the learning phase and in the interaction phase. Our experience is that, although the learning processes can be trapped by a local minimum occasionally as with standard RNNs, it has not happened in the interaction phase so far. The local minimum problem should be related to the complexity and number of embedded patterns. In the next section, the computational algorithm for modifying the PB values is reviewed.

2.3 Computational algorithm

The PB values are determined through regression of the past sequence pattern. In the interaction phase, the regression is applied for the immediate past window steps L and the temporal profile of p_t from L steps before to the current step ct is updated. Then, the current time motor outputs r_{ct} are generated by using the p_{ct-1} determined by this regression process. The window for the regression shifts as time goes by while p_t is updated through the iterations. In the learning phase the regression is conducted for all steps of the training sequence patterns. (This means that the window contains the whole sequence and it does not shift.)

The temporal profile of p_t in the sequence is computed via the back-propagation through time (BPTT) algorithm (Rumelhart et al., 1986). In this computation ρ_t, the internal value of the parametric bias, is obtained first. The internal value ρ_t changes due to the update computed by means of the error back-propagated to this parametric bias unit, which is integrated for a specific step length in the sequence. Then the parametric bias, p_t, is obtained by a sigmoid function of the output of the internal value. The utilization of the sigmoid function is just a computational device to bound the value of the parametric bias to a range of 0.0 to 1.0. In this way, the parametric bias is updated to minimize the error between the target and the output sequence.

For each iteration in the regression of the window, L steps of look-ahead prediction, starting from the onset step of the window, are computed by the forward dynamics of the RNN. Once the L steps of the prediction sequence are generated, the errors between the targets and the prediction outputs are computed and then back-propagated through time. The error back-propagation updates both the values of the parametric bias at each step and the synaptic weights. The update equations for the ith unit of the parametric bias at time t in the sequence are:

$$\begin{aligned} \delta{\rho_t}^i &= k_{bp} \cdot \sum_{step=t-l/2}^{t+l/2} {\delta_{step}^{bp}}^i + k_{nb}(\rho_{t+1}^i - 2\rho_t^i + \rho_{t-1}^i) && (1) \\ \triangle\rho_t^i &= \epsilon \cdot \delta{\rho_t}^i + \eta \cdot \triangle\rho_{t-1} && (2) \\ p_t^i &= sigmoid(\rho_t) && (3) \end{aligned}$$

In Eq. (1), $\delta\rho_t$, the delta component of the internal value of the parametric bias unit, is obtained from the summation of two terms. The first term represents the summation of the delta error, $\delta_t^{bp^i}$, in the parametric bias units for a fixed time duration l. $\delta_t^{bp^i}$, which is the error back-propagated from the output units to the ith parametric bias unit, is summed over the period from $t - l/2$ to $t + l/2$ time step. By summing the delta error, the local fluctuations of the output errors will not affect the temporal profile of the parametric bias significantly. The parametric bias should vary only with structural changes in the target sequence. Otherwise it should become flat, or constant, over time. The integration period, l, is taken as 20 steps in the experiment which is close to the time constant of the movement patterns in the training set.

The second term plays the role of a low pass filter through which frequent rapid changes of the parametric bias are inhibited. k_{nb} is the coefficient for this filtering effect. ρ_t is updated based on $\delta\rho_t$ obtained in Eq. (1). The actual update $\triangle\rho_t$ is computed by utilizing a momentum term to accelerate convergence as shown in Eq. (2). Then, the current parametric bias p_t is obtained by means of the sigmoidal outputs of the internal values ρ_t in Eq. (3).

In the interaction phase, the window step length for the regression L is taken as 30 steps. The regression, by means of the forward computation, and the error back-propagation iterates about 100 times in the real-time computation while the window shifts one step ahead. In the learning phase, the regression is iterated 50000 times for the fixed window containing the whole training sequence.

3 User-Robot Interaction Experiments

In this experiment, 4 rhythmic movement patterns were created for the learning sample, which differ greatly in their profiles. Figure 4 shows the learning sample, which is composed of the target trajectories of user hands (upper plot) and the target trajectories of robot arms (lower plot). The 4 different movement patterns were pattern 1: horizontally swinging both arms in opposite phase; pattern 2: vertically swinging both arms in phase; pattern 3: rotating both arms in opposite phase; and pattern 4: rotating both arms in phase. Moreover, two sets of sample data were prepared for each pattern, for robust learning. Each set of sample data was obtained by sampling the user's movement each 0.4 seconds for 50 steps. Note that the rhythmic pattern in each learning sample was not always completely cyclic, although the users tried to generate precise cyclic patterns. The actual learning was conducted using such noisy data.

The RNNPB used in these experiments has 12 input nodes and 12 prediction output nodes for learning the forward dynamics of movement patterns. The input as well as the output vectors are composed of 4 sensory values and 8 motor values. The sensory informations are just the center positions of balls in the user's both hands perceived by the robot camera. The

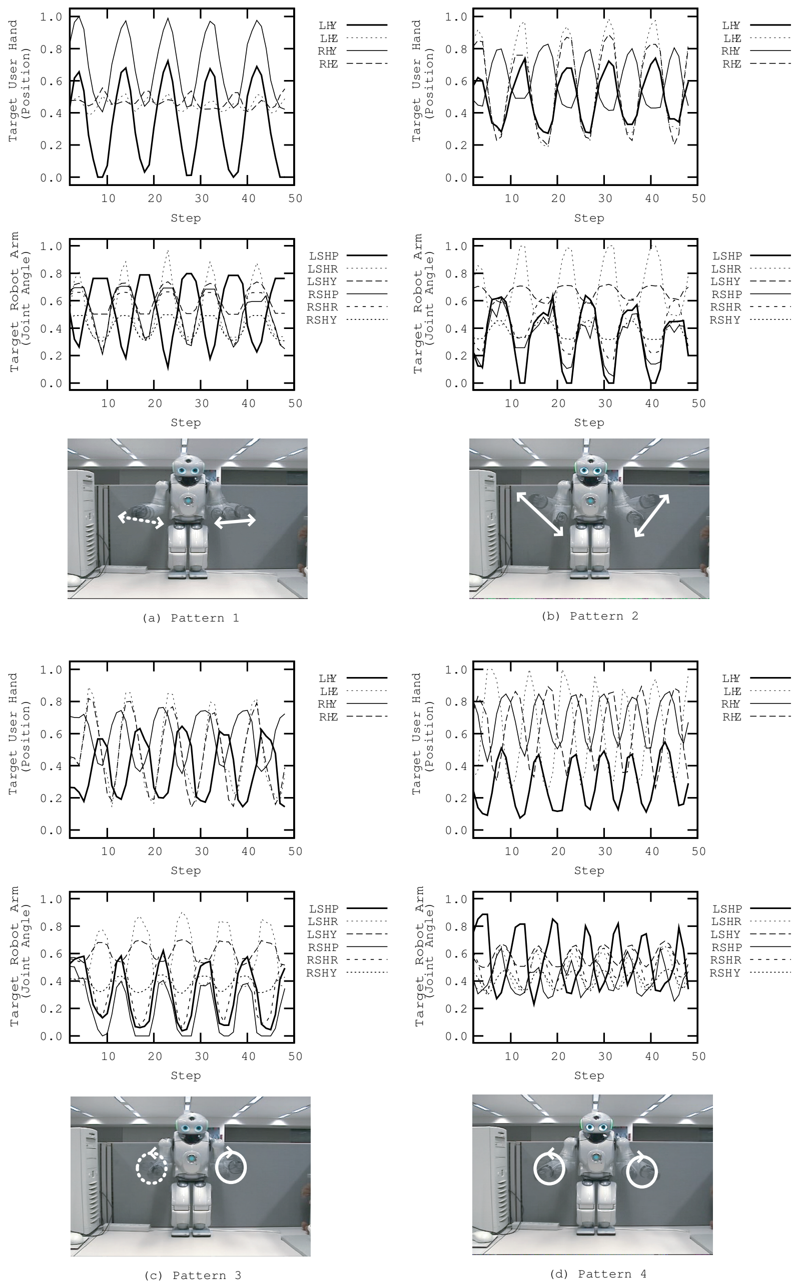


Figure 4: 4 different movement patterns employed for learning. Plots in the top row and the middle row represent the target user hands movement patterns in their visually perceived positions, and the target robot movement pattern in robot joint angles, respectively. The photo in the bottom row represents the corresponding movement pattern of the QRIO.

sensory vector consists of the positions of user hands (LHY,Z: left hand position in Y,Z axis and RHY,Z: right hand position in Y,Z) where YZ is two axis of the robot camera view plane facing the user. (It is noted that the depth information of X is not obtained.) The motor state represents angles of the robot arms (L,RSHP: left,right shoulder pitch joint, L,RSHR: left,right shoulder roll joint, L,RSHY: left,right shoulder yow joint and L,RELP: left,right elbow pitch joint). It also has 4 parametric nodes, 40 hidden nodes, and 30 context nodes.

In the learning phase, 4 learning sets were prepared. Permutations of three out of the 4 movement patterns were formed in order to repeat the same experiments while using different combinations of patterns. Learning set 1: pattern 1, 2, 3; Learning set 2: pattern 2, 3, 4; Learning set 3: pattern 1, 3, 4; Learning set 4: pattern 1, 2, 4. The RNNPB learns 6 sample data points of each learning set in a parallel manner. The remaining movement pattern is used as a novel pattern in the interaction phase. For 4 learning sets, the learning is iterated for 50000 steps, starting from randomly set initial synaptic weights. The final root-mean-square error of the output nodes was less than 0.0003 over all learning results.

Next, we conducted the user-robot interaction experiments. We examine the dynamic structure of interactions between the robot and the user, based on the memory structure self-organized through the learning phase. In the following experiments, we examine (1) how the robot can synchronously follow the learned movement pattern performed by the user; (2) how the robot can follow switching of the user's pattern among three learned patterns; and (3) how the robot responds to the novel movement pattern.

3.1 Synchronization with learned patterns

At the interaction phase of the first experiment, we examined how the robot can follow one of the target patterns which were learned by the RNNPB in the training phase. In this phase, the robot has to modulate the PB to regenerate its corresponding 8 degree-of-freedom (DOF) movements by perceiving the user's 4 DOF hand movements. It is important to observe not only the profiles of the regenerated patterns but also their synchrony to the user's movements.

In these interaction experiments, first the user continued to demonstrate pattern 1 as a target movement pattern in front of the robot. The robot had learned this pattern already in learning set 1. Subsequently, the user stopped moving for about 2 seconds and then restarted the same movement pattern again. The responses of the robot were observed.

The course of the interaction and the parametric bias of the RNNPB are plotted in Figure 5. In Figure 5, the plot at the top shows the target position of user hands. The second row plot shows the user hands positions predicted by the RNNPB. The third row plot shows the robot joint angle generated by the RNNPB (Only 6 DOF are plotted among a total of 8 DOF), the plot at the bottom shows the parametric bias of the RNNPB.

First, it was observed, up to step 75 of the user-repeated movement pattern, that the RNNPB-predicted pattern became synchronized with it in the initial 10 steps and then re-

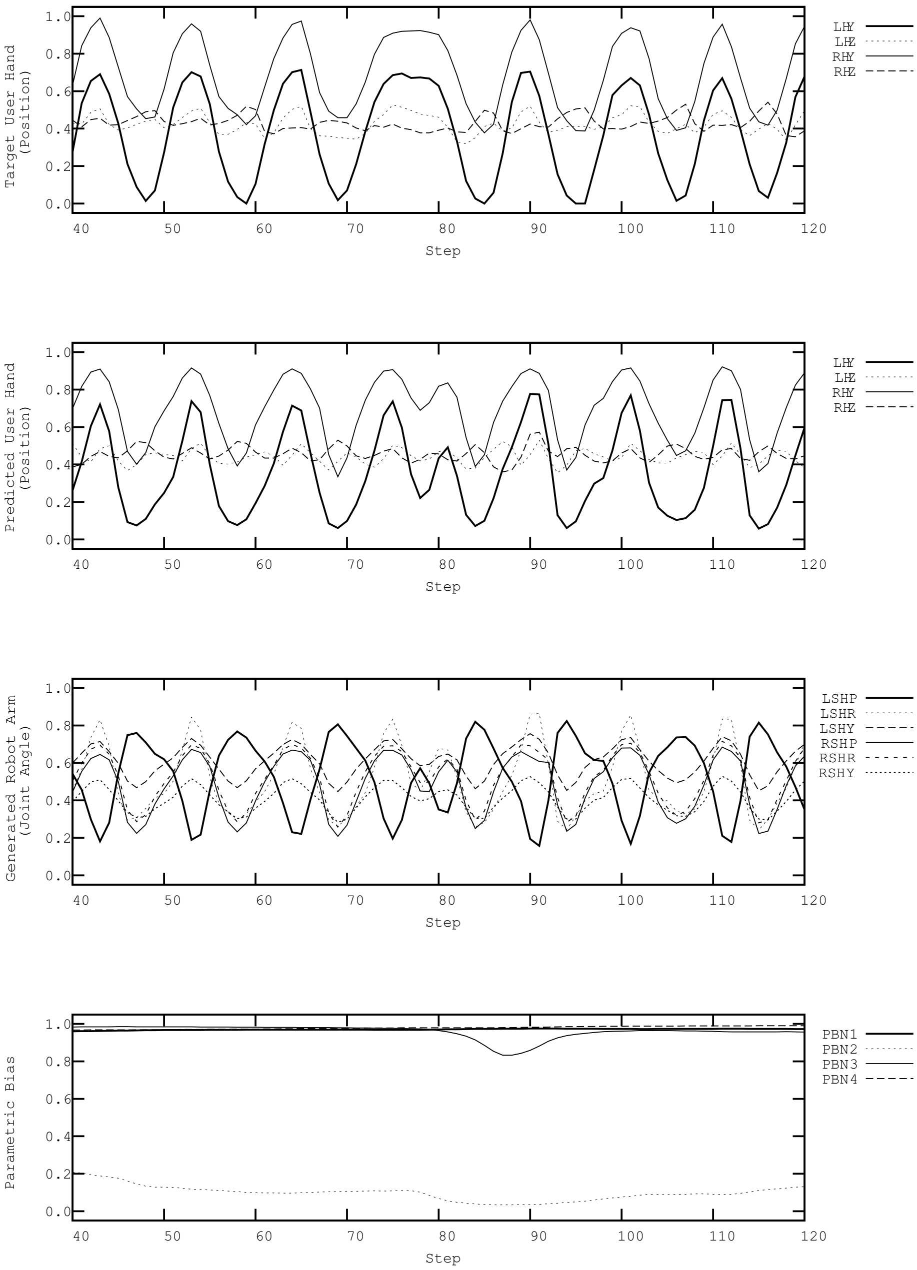


Figure 5: Synchronization of robot movement to a learned user hand movement pattern. Plot in the top row represents the target position of user hands visually perceived. Plots in the second and the third row represent the user hands position predicted and the robot joint angles generated by the RNNPB. (The values of position and angle are normalized for inputs of the RNNPB.) Plot in the bottom row represents the parametric bias of the RNNPB.

mained synchronized. Further, the RNNPB also generated a mostly synchronous corresponding robot movement pattern. In other words, a coherent state was achieved between the user and the robot. During this period, the parametric bias values changed rapidly in the initial 20 steps, because of the large prediction error, and then became almost constant. Next, we observed that, while the user stopped his hand movements from steps 75 to 80, the RNNPB continued to generate the same movement pattern. When the user restarted the same hand movement pattern, synchronization was achieved again in a few steps by recovering from the phase lag. During this perturbation period, one of the parametric biases changed slightly due to the prediction error, and then returned to the previous value. Similar results were obtained for all patterns over all other cases. As will be shown later in the analysis, all patterns are embedded in the RNNPB through a limit cycle attractor. The observed coherent state between the robot and the user can be explained by the entrainment (Beer, 1995; Tani, 1996; Andry et al., 2001) of the limit cycle dynamics of the RNNPB by the corresponding user's hand movement pattern. Note also that the synchronization was not always complete, since the user's hand movement was not precisely periodic. However, the important point is that a certain range of coherence could be maintained even when there were slight perturbations in the inputs to the system.

3.2 Switching among multiple learned patterns

In the second interaction experiment, we examined how the robot could follow target patterns while the user switched to demonstrate among various learned patterns. More specifically, the user demonstrated patterns 1, 2, and 3 sequentially for about 20 seconds each. These patterns had been learned by the robot in learning set 1.

The time course of the interaction in terms of the user hand movements, its prediction, the motor movements in the joint angles, and the PB of the RNNPB are plotted in Figure 6. Figure 7 shows the enlargement of Figure 6 with focusing on the transition periods.

We observed that the robot could follow the target movement patterns well while the user performed switching among multiple patterns. First, we observed that, while the user repeated pattern 1 during the period from steps 0 to 55, the robot generated its corresponding pattern synchronously. Subsequently, when the user switched to pattern 2 at step 55, the RNNPB continued to predict pattern 1 for several steps. It then switched to predict pattern 2 and generated the corresponding robot movement pattern synchronously. It is also observed that the PB values adapt stepwisely at each transition. The same stable transitions were observed for all the switching cases in other learning sets. In summary, multiple coherent states were achieved between the robot and the user, based on the learned patterns. Further, the state could be switched easily from one coherent state to another by the user's actions.

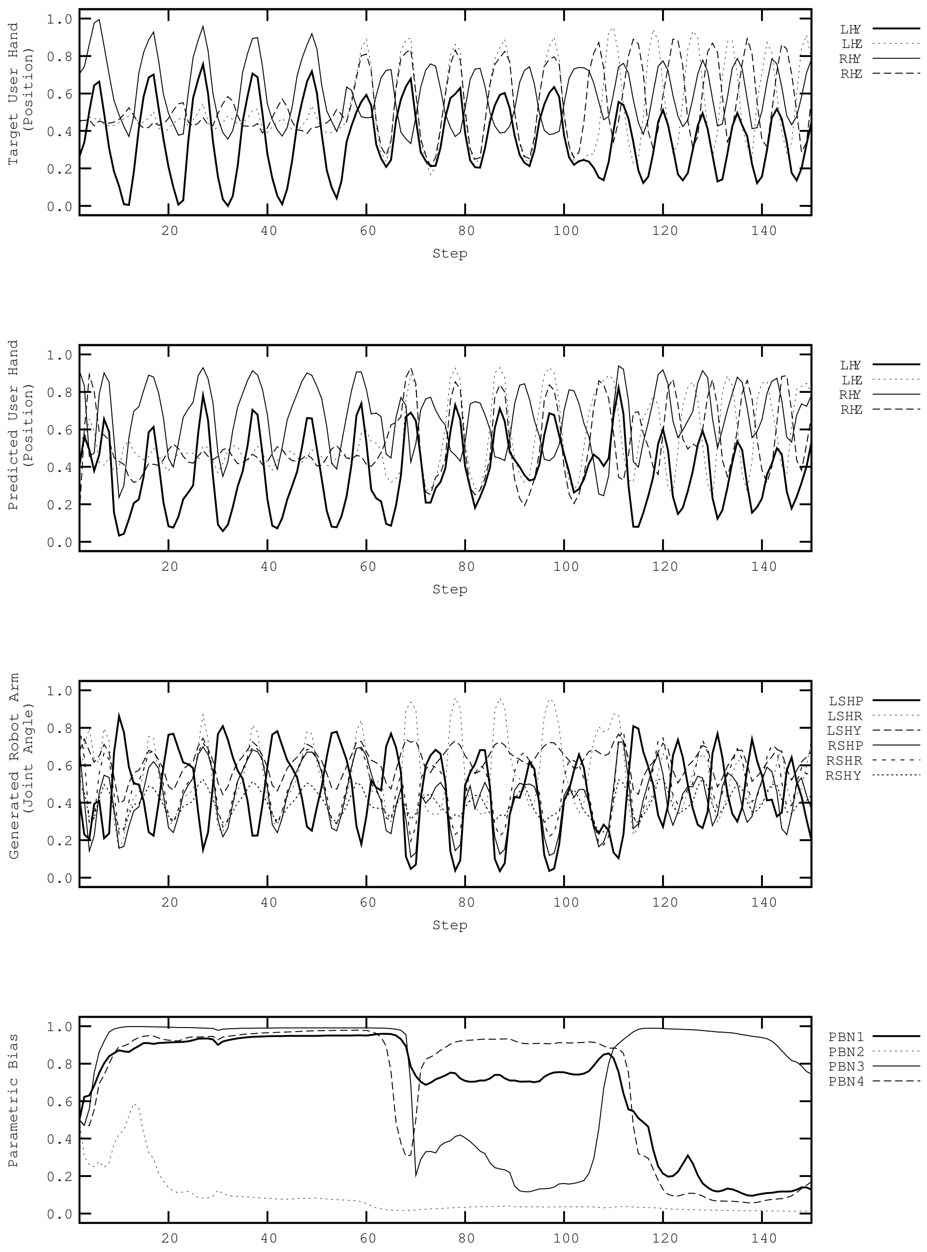
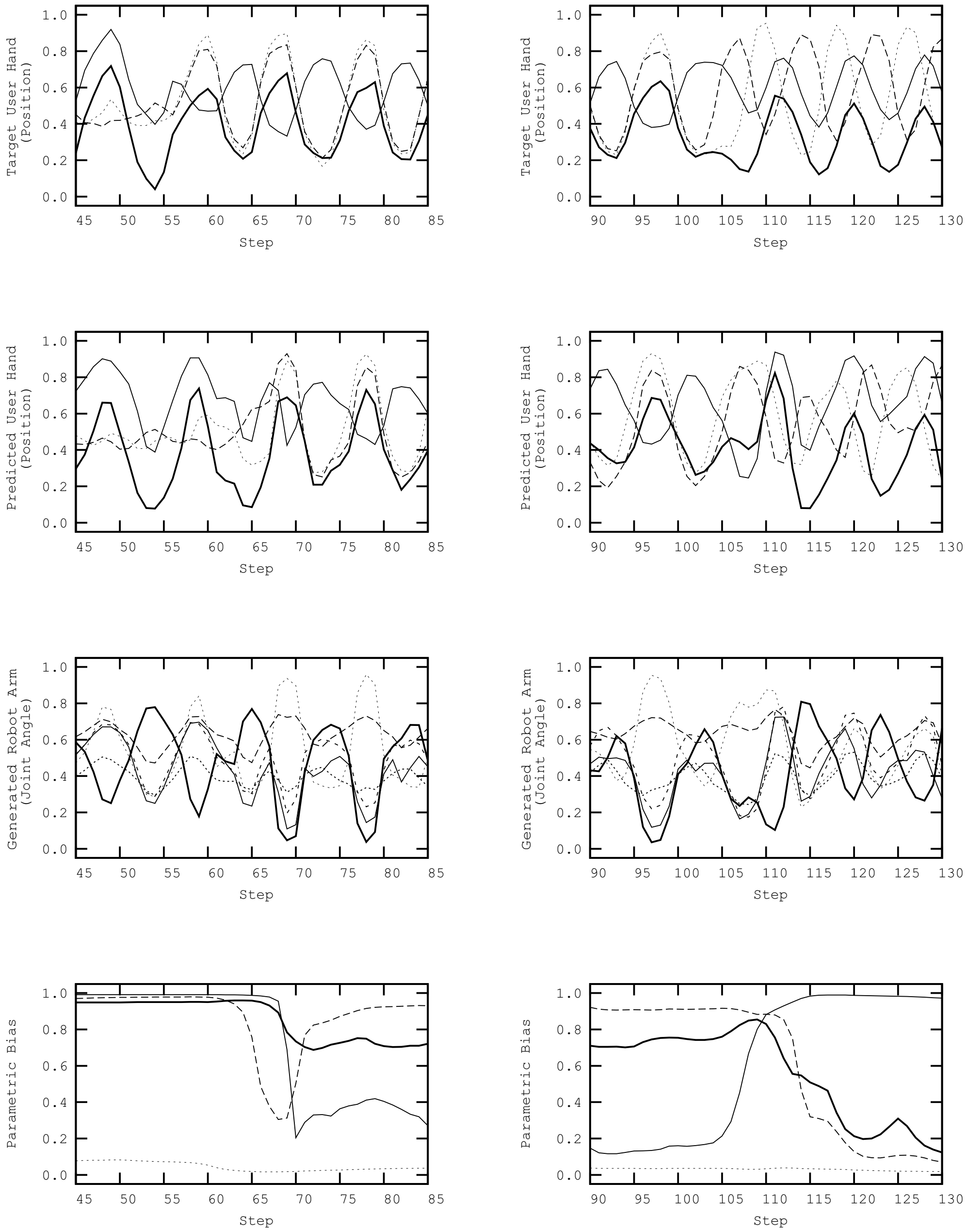


Figure 6: Switching of the robot movement pattern among three learned patterns as initiated by switching of user hand movement.



(a) steps 45 - 85

(b) steps 90 - 120

Figure 7: Enlargements of the plots in Figure 6 with focusing on the switching periods, (a)steps 45 to 85 (b) steps 90 to 130.

3.3 Response to novel patterns

In the third interaction experiments, we examined how the robot responded to novel movement patterns which it had never learned in the learning phase. More specifically, the user demonstrated a pattern that was not included in the learning set. For instance, if the robot had learned patterns 1, 2 and 3 in learning set 1, then pattern 4 was demonstrated by the user. We conducted several interaction experiments for each learning set, where the robot responses were classified into 3 typical cases.

Examples of those 3 typical cases are shown in Figure 8. In this figure, the plots in the

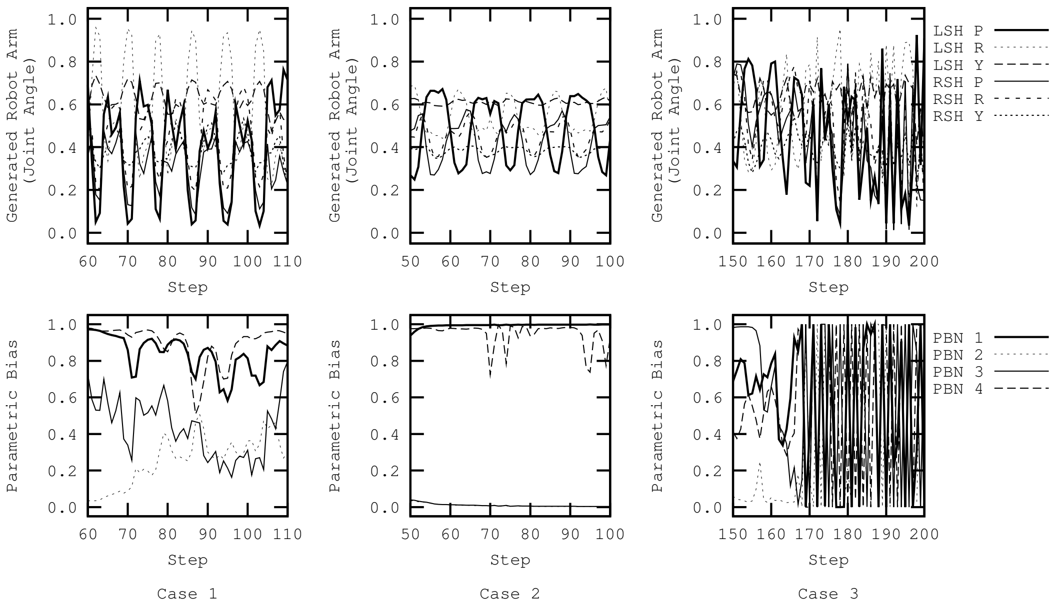


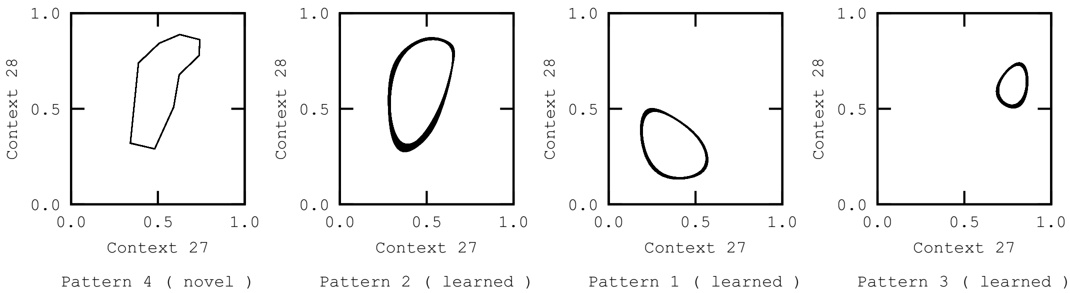
Figure 8: Three representative cases of the robot responses to novel hand movement patterns. Plots in the top and bottom rows represent robot movement patterns and corresponding PB time profiles, respectively.

upper row show the time courses of the robot joint angles generated by the RNNPB for the 3 cases(Only 6 DOF are plotted of a total of 8 DOF). The plots in the lower row show the corresponding PB values over time. In case 1 (the user demonstrated pattern 4 as a novel movement pattern, while the robot had learned set 1), the robot generated a pattern similar to pattern 2, which had been learned in learning set 1. The phase of the robot movement was stable and synchronized with the user, although the PB values seem to be modulated slightly. In the case 2, (the user demonstrated pattern 1 while the robot had learned set 2), the robot responded by generating a stably cycling pattern that was different from the learned patterns. This pattern, however, seemed to be synchronized and in phase with the user's movements. The PB values became almost constant in this case. In case 3 (the user demonstrated pattern 3 while the robot had learned set 3), the robot generated an unsteady movement pattern while the PB values fluctuated greatly.

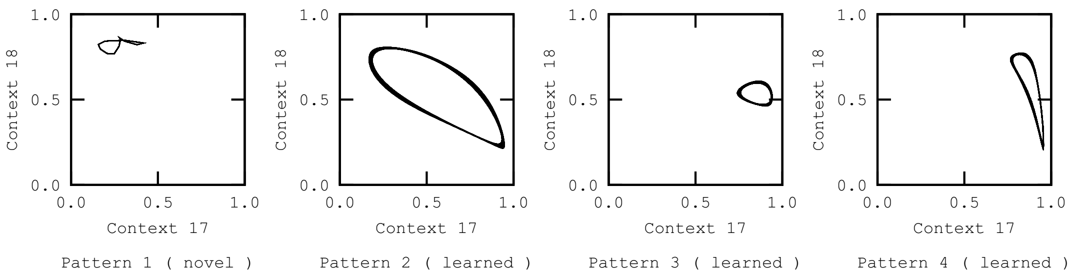
In order to examine the relationship between the memory dynamics organized in the RNNPB and the patterns that appeared in the interactions with the user, phase-space analyses were conducted, both for the autonomous RNNPB dynamics without inputs and the RNNPB dynamics coupled with the user's movements. The RNNPB can generate each of the learned patterns autonomously if the prediction outputs of the sensory-motor values fed back to the inputs. This mode is called as the closed-loop mode. In order to examine the attractor corresponding to each memory pattern in the RNNPB, the RNNPB was set to the closed-loop mode and its forward dynamics was computed for 1000 steps for the PB value determined for each learned pattern. Its trajectory, in terms of the values of two arbitrarily selected context units, was then plotted in the two dimensional state space. The initial transient trajectory was excluded. For conducting the phase-space analysis of the interaction dynamics, the same two context unit values were plotted with the RNNPB in the open-loop mode. One cycle of the novel hand movement pattern was sampled and repeatedly fed into the inputs of the RNNPB for 1000 steps. The forward computation and regression for updating the PB were simultaneously conducted. By this means, the invariant set (attractor) of the RNNPB with completely cyclic input patterns could be plotted.

Figure 9 shows the comparisons among the three cases. For each case shown in Figure 9 (a), (b) and (c), the left plot shows an attractor obtained from interaction with the novel pattern in the open-loop mode. The other three plots to the right show attractors corresponding to three memorized patterns plotted in the closed-loop mode. First, it is observed that the closed-loop memory dynamics exhibits three differently shaped one dimensional closed trajectories. This denotes that a differently shaped limit cycling attractor is generated for each learning pattern in the closed memory dynamics. It is observed that the open-loop dynamics converges to simple limit cycling attractors in cases 1 and 2, but not in case 3. The attractor shape obtained in the interaction is similar to that for a memorized pattern in case 1. Both have similarly shaped limit cycling dynamics. In case 2 on the other hand, although the interaction generates limit cycling dynamics, its shape is different from those of the memorized patterns. Finally in case 3, the plots that appeared in the interaction were quite complex and not from a simple limit cycling attractor. The shape is quite different from those of the memorized patterns.

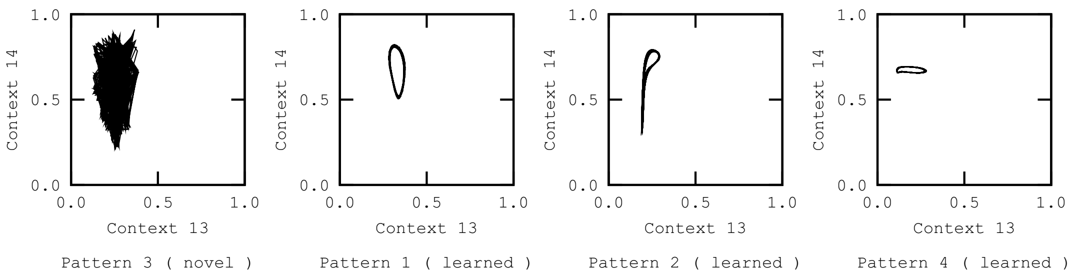
These analyses suggest that the interactions for novel patterns generate various consequences. They could generate stable limit cycling dynamics, which could be similar to the learned patterns, or entirely different dynamics. In the former case, the learned attractor dynamics is assumed to be resistant to perturbations by the interactions with the novel patterns. On the other hand, in the latter case a stable novel attractor is generated through collaboration between the user's movements and the RNNPB memory dynamics. It could



(a) Case 1



(b) Case 2



(c) Case 3

Figure 9: The attractor analysis for three representative cases of interacting with novel patterns as shown in (a), (b) and (c). The left plot represents a phase plot in the interaction mode, and the remaining three plots to the right represent phase plots for the closed RNNPB dynamics, corresponding to three learned patterns.

also generate unstable dynamic patterns that are quite complex and totally different from those previously learned.

3.4 Larger learning set

We conducted further learning and interaction experiments for larger learning sets: patterns 1, 2, 3 and 4. In these experiments, the learning is iterated for 100000 steps 10 times, starting from randomly set initial synaptic weights. The final root-mean-square error of the output nodes was less than 0.0003 over all the learning results. In the interaction phase, while similar results were obtained in 9 cases out of 10, the robot could not follow one of the target patterns switched by the user. In this case the robot could not generate the movement pattern autonomously. It is also interesting to note that the responses to novel patterns become more diverse in this case. These preliminary results might come from the fact that the nonlinearity of the RNNPB dynamic structure increases as the number of embedded patterns is increased.

4 Discussion & summary

4.1 The RNNPB and mirror systems

The current paper has shown that the mirror system, by means of the RNNPB, can both recognize and generate multiple learned movement patterns by using its dynamic characteristics. The PB activations can correspond to the mirror neuron firing. Especially our results that recognition of a movement pattern of the user can trigger the corresponding movement pattern of the robot reminds us the human experiments by Fadiga, Fogassi, Pavesi, and Rizzolatti (1995) in which observation of an experimenter grasping objects enhance motor evoked potentials of the subject's hands under conditions of Transcranic Magnetic Stimulation of the subject's premotor cortex.

Some may argue that an equivalent mechanism of the mirror system could be achieved by instead using discrete representations such as graphical modeling, including hidden markov models (HMMs) (Amit & Mataric, 2002; Inamura et al., 2001). In such formulations, a set of hidden markov models is prepared and each is tuned to represent a specific movement pattern, both in generation and recognition. When a user's movement pattern is perceived, the corresponding hidden markov model is selected by computing its likelihood. The selected model then generates the corresponding movement pattern. Although this scheme appears similar to ours at first glance, there are essential differences between them.

First, it might be difficult for the HMMs to achieve the spontaneous synchronization we have described in our scheme because the perception and generation processes must be conducted sequentially in the HMM computations. During the process of selecting the greatest likelihood model of perception, the generation process must be frozen. In contrast, in our proposed dynamic systems scheme, perception and generation can be performed simultaneously by means of the entrainment of the RNNPB dynamics by the user movement patterns. By this entrainment mechanism, even when the synchronization is lost accidentally, it can be autonomously recovered while the interactions continue.

Second, what would happen if the user demonstrated certain distorted patterns relative to the learned ones? There could be two cases for the HMM. If the distortion were less than a certain threshold, the model of greatest likelihood would be selected and patterns would be generated exactly following the model. There would be no effect of the distortion in the user movement patterns on the robot movement pattern generations. On the other hand, if the distortion were larger than the threshold, a relevant model that matched the prepared pattern set would not be found and the generation process would freeze. The situation is significantly different in our proposed scheme. If the user movement patterns are distorted slightly, it will slightly modulate the generation of robot movement patterns with accompanying slight modulations of the PB vector. If the distortion is much larger, it simply modulates the pattern generation more, with larger modulations of the PB vector. The response of the robot in our scheme is not a binary choice of "accept" or "do not accept". Instead, the response is graded between the coherent and incoherent states.

There is one concern that the incoherence might generate undesired or harmful behaviors. There has always been a conflict between autonomy and safety in dealing with autonomous robots. It is suggested that this problem could be solved by making the robot learn about its own behavior characteristics. More specifically, it learns from past experience in which situations safe or unsafe behavior patterns could be generated. This can be done by the "meta-learning" of the PB mapping. Another network is prepared which learns safe or unsafe regions in the PB space from experience. The network can be used to prevent the PB values from coming too close to the learned undesired regions. This sort of the meta-learning could be utilized by the robot for assessing its own possible behavior patterns in the PB space in various aspects. It would tell whether the behavior patterns generated by a specific PB vector had actually been experienced or not, whether they were good or bad based on some motivating goals, and so on. It would be interesting to know if such meta-levels of assessing mirror neurons' activities actually exist in animals or humans. For example, how do actual mirror neurons respond when unsafe behavior patterns are perceived? It might be that corresponding mirror neuron populations are about to fire, but they are somehow suppressed from complete firing. If this happens, it could be explained that the mirror neurons recognize the perception of an unsafe pattern, and the meta-level inhibits the animal from actually regenerating the pattern. This scheme of meta-level learning of the PB is open for future research.

An interesting characteristic of the RNNPB is that it has two levels of adaptation. Off-line learning generates the global PB mapping structure from the PB vector to movement patterns. Then, adaptation of the PB vector in real time can generate not only learned patterns, but also various novel patterns (Tani & Ito, 2003). Recall, however, that the various novel patterns do not necessarily include particular newly desired patterns. In such situations, the desired patterns should be learned incrementally. However, there is a problem in the current scheme. If the RNNPB attempted to learn novel patterns, it would damage the current memory structure severely since each memory is distributedly represented in the network, sharing the same synaptic weights with other memories. This is called the memory interference problem (McCloskey & Cohen, 1989) and is one disadvantage of employing a distributed representation. One possible solution to the problem is rehearsing and consolidation, proposed by Tani (1998) for incremental learning of RNNs. The idea comes from the biological hypothesis (Squire, Cohen, & Nadel, 1984) that novel episodic events are first stored in the hippocampus as short-term memory. They are then consolidated during sleep in the neocortex as long-term memory, and are rehearsed along with prior memories. A possible application of these ideas to the RNNPB might be as follows. When novel movement patterns are acquired, they are first stored in a certain database as the short-term memory. Then the RNNPB, functioning as the long-term memory, regenerates the learned patterns in a closed-loop mode, without external inputs, by setting the PB with appropriate values. These "rehearsed patterns" are also stored in the short-term memory database. Then, the RNNPB is trained with both the novel and rehearsed patterns stored in the short-term memory. This learning corresponds to the consolidation process in neocortex. The memory interference problems should be reduced significantly since the network's incremental training utilizes both the novel and previously learned patterns. Future research should examine how well this scheme works in the RNNPB as compared to other localist schemes, especially those using modular network architectures, shown in (Tani & Nolfi, 1998; Wolpert & Kawato, 1998; Demiris & Hayes, 2002), in the incremental learning. Also, future studies of incremental learning should be conducted as related to its associated changes in the nonlinearity of the RNNPB dynamics. As speculated in the preliminary experiments, the diversity of incoherent patterns increases with the number of learned patterns. Detailed examinations of these characteristics are expected in future research.

One hard problem related to incremental learning is how to determine autonomously when and what to learn in the current context of the interactions (Dautenhahn & Nehaniv, 2002; Andry et al., 2001). The problem is how the agents can attend to particular segments of interest in the on-going flow of their experiences. One possible future direction to tackle this problem might be extensions of the novelty detectors shown in (Andry et al., 2001; Demiris & Hayes, 2002). The novelty detector detects novel sequence patterns by monitoring their prediction errors. Then, the learner learns those detected sequences. The drawback of this idea, however, is that it also detects random patterns without meaning and structure. Future study should involve how to detect not just novel but also meaningful patterns in the on-going experiences, possibly by employing the ideas of meta-learning of the PB space as described previously.

Although the current implementation has achieved only trajectory level repetitions of given movement patterns, its extensions to imitation through understanding others' goals as well as one's own (Tomasello, 1999) are important future research topics. It is also true that many mirror neurons are found in rather goal-directed task settings, where they seem to encode not exact movement patterns, but their abstraction or goals (Rizzolatti et al., 1996). In order to achieve goal-directed imitation, certain hierarchical architectures which are capable of manipulating movement patterns in an abstract manner might be required. The abstraction of the sensory-motor flow can be done by employing graph representation using HMM in the higher level for sequencing the lower level movement primitives as shown in (Amit & Mataric, 2002; Inamura et al., 2001), or by using the multiple levels of the RNNPB. The higher level RNNPB manipulates movement pattern generation in the lower RNNPB by sending the PB vector sequences to the lower level in a top-down manner as shown in (Tani, 2003). Dautenhahn and Nehaniv (2002); Nehaniv and Dautenhahn (2001) argued that the correspondence problems have to be addressed in order to achieve goal-directed imitation. When imitating someone holding a cup in his right hand, one might use one's left hand to hold the cup instead if one's right hand is unavailable. However, one would not use one's right foot. It is not yet fully understood whether the abstraction capability in hierarchical architectures could make this sort of flexible and generalized correspondence between imitatees and imitators. It is also interesting to extend the current study to include object manipulation tasks since most of the mirror neurons were found in such animal task settings (Rizzolatti et al., 1996). Future research should address such higher order cognition involving the tasks of goal-oriented or manipulating objects including tools since the mirror neurons are considered to play important roles in acquiring the higher order concepts based on actions.

4.2 Mutual interactions and psychological implications

The experiments introduced in the current paper have focused mainly on the robot's dynamic adaptation. We are currently conducting new experiments which focus on bi-directional adaptation in mutual interaction between the robot and users. Although complete results have not yet been obtained, the experiments are briefly introduced here for the purpose of enhancing the discussion of future research directions. In this new experimental set-up, after the robot learns multiple movement patterns in the same way as described previously, subjects who are ignorant of what the robot learned are faced with the robot. The subjects are then asked to find as many movement patterns as possible for which they and the robot can synchronize together by going through exploratory interactions. Five subjects participated in the experiments. The settings of the network and the robot were exactly the same as those in the previous interaction experiments. Each subject was allowed to explore the interactions with the robot for one hour, including four 5 minute breaks. Although most of the subjects could find all movement patterns by the end, the exploration processes were not trivial for the subjects. If the subjects merely attempted to follow the robot movement patterns, they could not converge in most situations since the PB values fluctuated when receiving unpredictable subject hand movement patterns as the inputs. If the subjects attempted to execute their desired movement patterns regardless of the robot movements, the robot could not follow them unless the movement patterns of the subjects corresponded with the ones learned by the robot.

There are interesting points in this new experiment as compared to the previous one. First, the master-slave relation, which was fixed between the subjects and the robot in the previous experiments, is no longer fixed but is instead spontaneously switched between the two sides. (Recall that the subjects initiated new movement patterns while also switching among multiple learned patterns in the previous experiments.) When the subjects feel that the robot movement patterns become close to theirs, they just keep following the robot movement patterns passively in order to stabilize the patterns. However, when the subjects feel that they and the robot cannot match each other's movements, they often initiate new patterns, hoping that the robot will start to follow them and become synchronized. Second, there are autonomous shifts between the coherent phase and the incoherent phase after the subjects become familiar with the robot responses to some extent. When the subjects happen to find synchronized movement patterns, they tend to keep the achieved synchronization for a moment in order to memorize the patterns. However, this coherence can break down after a while through various uncertainties in the mutual interactions. Even small perturbations in the synchronization could confuse the subjects if they are not yet fully confident of the repertoire of the robot's movement patterns. Also, the subjects' explorations of new movement patterns makes it difficult for the robot to predict and follow their movements.

The authors speculate that appropriate analysis of these observed phenomena might shed a ray of light on the mechanism of joint attention (Baron-Cohen, 1996; Moore & Corkum, 1994) as well as turn taking behaviors (Trevarthen, 1977). In our new experiment, when movement patterns of the robot and human are synchronized, joint attention is assumed to have been achieved for the pattern. However, the current joint attention can break down and another joint attention (attending to another movement pattern) can emerge after a while. Although joint attention itself might be explained simply by synchronization (Andry et al., 2001; Ijspeert et al., 2003), a more interesting question is how a joint attention can break down and flip to another one spontaneously. This sort of spontaneity is also essential in turn taking behaviors. It was observed that the initiatives leading to synchronization switch spontaneously between the robot and the subjects. The essential question here is how the spontaneous shifts in turn taking behaviors can emerge.

Although extensive analysis of the observed data is required for further reasoning of the underlying mechanisms, the authors speculate that they might be closely related to the so-called open dynamic structure (Tani, 1998). It was argued that the system state tends to flip between the coherent and the incoherent phases if stability, in terms of rational goal-directedness, and instability, caused by unpredictability of the open environment, coexist in cognitive systems. Tani (1998) proposed one possible explanation of the spontaneous breakdown of self-consciousness through dynamic system characteristics. A more theoretical framework of this idea has been explained by the chaotic itinerary (Tsuda & Umemura, 2003). Furthermore, Ikegami and Iizuka (2003) recently showed that spontaneous turn taking behaviors can emerge by evolving the coupled-dynamics for a simulated pair of agents. Their analysis indicated that both stable and unstable manifolds are generated in the evolved coupled dynamics. In our experiments of mutual interactions, the stability originated from the synchronization mechanisms for shared memories of movement patterns between the robot and the subjects. The instability arose from the potential uncertainty in predicting each other's movements. It is likely that the coexistence of stable and unstable characteristics in the system dynamics might be the main cause for the spontaneous shifts. Recently, Sato and Ikegami (2004) related this characteristics to the undecidability of the turing test in the theoretical analysis of imitation game, although further examination is required in this part of the analysis. Future collaborative research among developmental psychology, synthetic modeling studies, and theoretical nonlinear dynamics studies would gain further understanding of the essential mechanisms in joint attention and turn taking behaviors.

4.3 Entertainment robotics

Before summarizing the paper, some comments are made on the implications of the current research to entertainment robotics. One of the most important goals in entertainment robotics is to keep users interacting with robots without getting bored. Spontaneity is a crucial element in this aspect. In the mutual interaction experiments, most of the subjects reported that they occasionally felt as if the robot had its own "will" because of the spontaneity in the generated interactions.

Some might consider that this sort of spontaneity could be mimicked by adequately designing fluctuations in the system by adding external stochasticity in a rather artificial way. The authors, however, argue that the spontaneous fluctuations are not programmable, but are self-generated depending on the contextual flow of the on-going interactions. Different fluctuations should occur in interactions with different individuals, or even with the same individual in different situations. When the fluctuations are generated diversely, as is inherent to the system dynamics, the resultant interactions should be felt to be spontaneous. If the fluctuation processes are pre-defined, interactions with the robot may not be felt to be unique, and the users would likely get bored with the robot quickly.

We can learn a lot about how such spontaneity in cognitive behaviors emerges from psychological observations as well as from complex adaptive systems studies. This approach would enhance our understanding of how close communicative relationships can be maintained between robots and their users, and it would thus improve the entertainment value of the robots.

4.4 Summary

We have formulated deferred imitation by using the RNNPB scheme, in which memories of multiple movement patterns are distributedly represented in an RNN. The PB vector encodes movement patterns both for their generation and recognition analogously to a mirror neuron system. Our experiments with a humanoid robot have shown that diverse dynamic interactions can emerge in the form of either coherence or incoherence between the robot and the user. The robot can follow the learned user movement patterns synchronously by generating coherent dynamic states. It can be said that joint attention is accomplished for the current movement pattern shared in both the memories of the robot and the user. The robot cannot follow novel patterns, but responds to them in various ways with incoherent dynamic states. Joint attention is apparently lost in this situation. Our preliminary experiments of mutual interactions suggest that the essential mechanism for autonomous shifts in joint attention and turn taking behavior would be better understood in the future through collaborations between developmental psychology, synthetic modeling, and theoretical nonlinear dynamics studies.

References

Amit, R., & Mataric, M. (2002). Learning movement sequences from demonstration. In *Proceedings of the ieee international conference on development and learning (icdl-2002)* (pp. 203–208).

Andry, P., Gaussier, P., Moga, S., Banquet, J., & Nadel, J. (2001). Learning and communication in imitation: An autnomous robot perspective. *IEEE Transaction on Systems, Man and Cybernetics. Part A : Systems and Humans*, *31*(5), 431–444.

Arbib, M. (1991). Perceptual structures and distributed motor control. In *Handbook of physiology: The nervous system, ii. motor control* (pp. 1448–1480).

Arbib, M. (2002). The mirror system, imitation, and the evolution of language. In *Imitation in animals and artefacts* (pp. 229–280). Cambridge: MIT Press.

Baron-Cohen, S. (1996). *Mindblindness: An essay on autism and theory of mind.* MIT Press.

Beer, R. (1995). A dynamic systems perspective on agent-environment interaction. *Artificial Intelligence*, *72*(1), 173–215.

Berthouze, L., Shigematsu, Y., & Kuniyoshi, Y. (1998). Dynamic categorization of explorative behaviors for emergence of stable sensorimotor configuration. In *Proceeding of the fifth international conference on simulation of adaptive behaviour* (pp. 67–72). Hitsuji Press.

Billard, A. (2002). Imitation: a means to enhance learning of a synthetic proto-language in an autonomous robot. In *Imitation in animals and artifacts* (pp. 281–311).

Breazeal, C., & Scassellati, B. (1999). A context-dependent attention system for a social robot. In *Proceedings of sixteenth international joint conference on artifical intelligence (ijcai1999).*

Breazeal, C., & Scassellati, B. (2002). Robots that imitate humans. *Trends in Cognitive Science*, *6*(11), 481–487.

Dauce, E., Quoy, M., & Doyon, B. (2002). Resonant spatio-temporal learning in large random neural networks. *Biological Cybernetics*, *87*, 185–198.

Dautenhahn, K. (1995). Getting to know each other - artificial social intelligence for autonomous robots. *Robotics and Autonomous System*, *16*(2-4), 333–356.

Dautenhahn, K., & Nehaniv, C. L. (Eds.). (2002). *Imitation in animals and artifacts.* Cambridge, MA: MIT Press.

Decety, J., Chaminade, T., Grezes, J., & Meltzoff, A. (2002). A pet exploration of the neural mechanisms involved in reciprocal imitation. *Neuroimage*, *15*(1), 265–272.

Decety, J., & Grezes, J. (1999). Neural mechanisms sub serving the perception of human actions. *Trends in Cognitive Sciences*, *3*, 172–178.

Demiris, Y., & Hayes, G. (2002). Imitation as a dual-route process featuring predictive and learning components: a biologically-plausible computational model. In *Imitation in animals and artefacts* (pp. 1448–1480). Cambridge: MIT Press.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.

Fujita, M., Kuroki, Y., Ishida, T., & Doi, T. (2003). A small humonoid robot sdr-4x for entertainment applications. In *Proceedings of international conference on advanced intelligent mechatronics.*

Gaussier, P., Moga, S., Banquet, J., & Quoy, M. (1998). From perception-action loops to imitation processes : A bottom-up approach of learning by imitation. *Applied Artificial Intelligence*, *12*(7-8), 701–727.

Gelder, T. van. (1998). The dynamical hypothesis in cognitive science. *Behavior and Brain Sciences*, *21*, 1–14.

Georgieff, N., & Jeannerod, M. (1998). Beyond consciousness of external reality: A who system for consciousness of action and self-consciousness. *Consciousness and Cognition*, *7*, 465–477.

Hayes, G., & Demiris, J. (1994). A robot controller using learning by imitation. In *Proceedings of the 2nd international symposium on intelligent robotic systems.*

Ijspeert, A., Nakanishi, J., & Schaal, S. (2003). Learning attractor landscapes for learning motor primitives. In *Advances in neural information processing systems 16.*

Ikegami, T., & Iizuka, H. (2003). Joint attention and dynamics repertoire in coupled dynamical recognizers. In *the aisb 03: the second international symposium on imitation in animals and artifacts* (pp. 125–130).

Inamura, T., Nakamura, N., Ezaki, H., & Toshima, I. (2001). Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop. In *Proceedings of international conference on robotics and automation* (pp. 4208–4213).

Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., & Nakatsu, R. (2001). Robovie: an interactive humanoid robot. *International Journal of Industrial Robot*, *28*(6), 498–503.

Jordan, M., & Rumelhart, D. (1992). Forward models: supervised learning with a distal teacher. *Cognitive Science*, *16*, 307–354.

Kuniyoshi, Y., Yorozu, Y., Inaba, M., & Inoue, H. (2003). From visuo-motor self learning to early imitation – a neural architecture for humanoid learning. In *Proceedings of ieee international conference on robotics and automation* (pp. 3132–3139).

McCloskey, M., & Cohen, N. (1989). Catastrophic interference in connectionist network. *Psych. Learning Motivat.*, *24*, 109–165.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78.

Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: exploring the range of gestures imitated and the underlying mechanisms. *Developmental psychology*, *25*(6), 954–962.

Miyake, Y. (2002). Co-creation systems. *Cognitive Processing*, *3*, 131–136.

Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, *14*(4), 349–450.

NEC. (1999). (http://www.incx.nec.co.jp/robot/)

Nehaniv, C., & Dautenhahn, K. (2001). Like me? - measures of correspondence and imitation. *Cybernetics and Systems*, *32*(1-2), 11–51.

Piaget, J. (1962). *Play, dreams and imitation in childhood.* Routledge.

Rizzolatti, G., Fadiga, L., Galles, B., & Fogassi, L. (1996). Promotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In *Parallel distributed processing.* MIT Press.

Sato, Y., & Ikegami, T. (2004). Undecidability in the imitation game. *Minds and machines*, *14*, 133–143.

Schaal, S. (1999). Is imitation learning the route to humanoid robots ? *Trends in Cognitive Science*, *3*(6), 233–342.

Schoner, S., & Kelso, S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science*, *239*, 1513-1519.

Sony. (1999). (http://www.aibo.com/)

Squire, L., Cohen, N., & Nadel, L. (1984). The medial temporal region and memory consolidation: A new hypothesis. In H. Weingartner & E. Parker (Eds.), *Memory consolidation* (pp. 185–210). Hillsdale, N.J.: Erlbaum.

Tani, J. (1996). Model based learning for mobile robot navigation from the dynamic systems perspective. *IEEE Transactions on System, Man and Cybernetics Part B, Special issue on learning autonomous robots*, *26*(3), 421–436.

Tani, J. (1998). An interpretation of the "self" from the dynamical systems perspective: a constructivist approach. *Journal of Consciousness Studies*, *5*(5-6), 516–42.

Tani, J. (2003). Learning to generate articulated behavior through the bottom-up and the top-down interaction process. *Neural Networks*, *16*, 11–23.

Tani, J., & Ito, M. (2003). Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment. *IEEE Transactions on System, Man and Cybernetics Part A*, *33*(4), 481–488.

Tani, J., & Nolfi, S. (1998). Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. In R. Pfeifer, B. Blumberg, J. Meyer, & S. Wilson (Eds.), *From animals to animats 5.* Cambridge, MA: MIT Press. (later published in Neural Networks, vol12, pp1131–1141, 1999)

Tomasello, M. (1999). *The caltural origins of human cognition.* Harvard University Press.

Trevarthen, C. (1977). Descriptive analyses of infant communicative behaviour. In H. Schaffer (Ed.), *Studies in mother-infant interaction.* London: Academic Press.

Tsuda, I., & Umemura, T. (2003). Chaotic itinerancy generated by coupling of milnor attractors. *Chaos*, *13*(3), 937–946.

Wolpert, D., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317–1329.