Proc. 3rd Int. Conf. on Development and Learning (ICDL'04), La Jolla, USA, 2004.

Joint attention between a humanoid robot and users in imitation game

Masato Ito
Sony Corporation
6-7-35 Kitashinagawa, Shinagawa-ku
Tokyo, 141-0001, Japan
masato@pdp.crl.sony.co.jp

Jun Tani
Brain Science Institute, RIKEN
2-1 Hirosawa, Wako-shi
Saitama, 351-0198, Japan
E-mail tani@brain.riken.go.jp

Abstract

This paper studies a synthetic model for joint attentions and turn taking by conducting experiments of robot-user imitative interactions. A Sony humanoid robot learns multiple cyclic movement patterns as embedded in a neural network model proposed by us and each of memorized movement patterns can be retrieved and generated by means of entrainment by external inputs in terms of users' movement patterns perceived. Our experiments on a simple imitation game showed that multiple movement patterns are generated as synchronized between the robot and users while the shared patterns shift spontaneously from one to another. Based on the experimental results, we show possible dynamical systems accounts for the underlying mechanisms for joint attentions and turn taking.

1 Introduction

In entertainment robotics, achieving natural interactions between robots and their users is one of the most essential issues to be solved. The ultimate goal for this is to develop communicative abilities on robots like humans. Human communications involve dynamic processes such as joint attention and turn taking with others. Joint attention is to share behaviors, events, interests and contexts in the world among agents from time to time. It requires mutual awareness of companion's attentions. On the other hand, turn taking is to switch the initiatives in interactions among agents spontaneously. Turn taking is considered to be prerequisite for joint attention. Speculating that such dynamic processes might appear as a consequence of mutual adaptation among agents, we develop a synthetic model for experiments of interactions between robots and their users.

Recent research on robotics have implemented a model of joint visual attention [3] between robots and humans [9, 11]. In such models, the robot guess the human's attentional target by detecting their gazing and pointing, and also pays attention to it. And then joint attention can be archived by the recognition of the robot's attention by human. However, in human communications, it seems that there are more complex situations of joint attention that can never be achieved by simply using such static and explicit cues [8]. For example, to share topics in dialogues and to share goals in collaborative works. It seems that the targets of such joint attention are determined in the flow of ongoing interactions in contextual ways. We speculate that such context dependent communicative interactions could emerge in terms of a class of dynamical structures appeared in the mutual adaptation processes between robots and humans.

In this study, we assume simple movement imitation game between a robot and human subjects where the problems of imitations are simplified from the reality. The imitation in our robot platform is not yet goal-oriented ones as have been discussed by [17]. Also the correspondence problems [4] between the perceptual space for others and motor space of own in learning are simplified. Rather, our focus is to observe dynamic interaction processes which take place in the coupling between robots and human in the simplified imitation game.

Firstly in this paper, we will introduce our neural network model: recurrent neural network with parametric biases (RNNPB) [15, 16]. The robot learns multiple cyclic movement patterns as embedded distributedly in self-organized dynamic structures of the RNNPB. Then, each of memorized movement patterns can be regenerated by means of entrainment by a corresponding users' movement pattern perceived. This is done by the on-line adaptation scheme of the parametric biases (PB). Then, two types of imitative interactions will be demonstrated using this model. In the first experiment, the on-line adaptation only in the robot side is considered in the imitation game. In the second experiment, the on-line adaptation in both of the robot and the users is performed. By going through the close examinations of different aspects in these two experiments, a novel theory for joint attentions and turn taking will be elucidated.

2 Task setting and neural network modeling

2.1 Task setting

In the current study, the Sony humanoid robot QRIO (SDR-4X II) [5] was used as the experimental platform (see Figure 1). In this experiment, only movement patterns of both arms were considered. Other movements were frozen.

Figure 1. A user is interacting with the Sony humanoid robot QRIO SDR-4XII.

Before interaction experiments with users, the robot learns a set of robot movement patterns with different profile with associated with the corresponding user's visually perceived hand movement patterns as off-line in the learning phase. 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 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

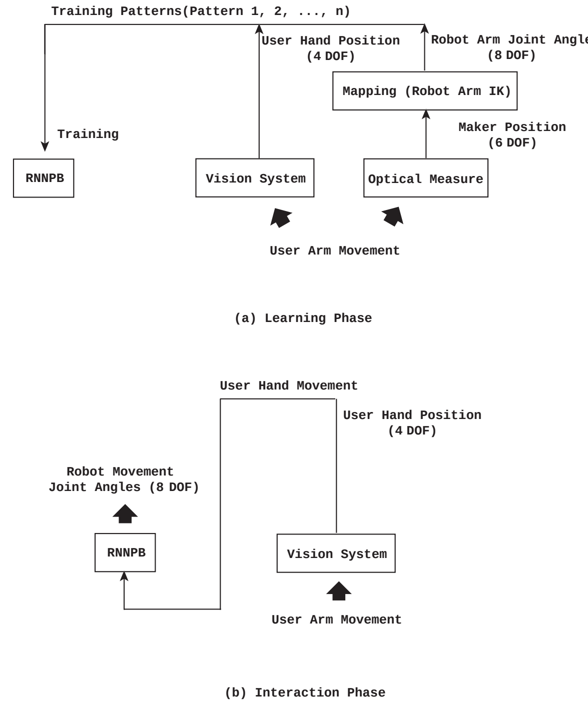

Figure 2. System configurations in learning phase (a) and interaction phase (b).

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.

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

2.2 RNNPB modeling

RNNPB is a version of the Jordan-type RNN [7] 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, 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 [12] 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

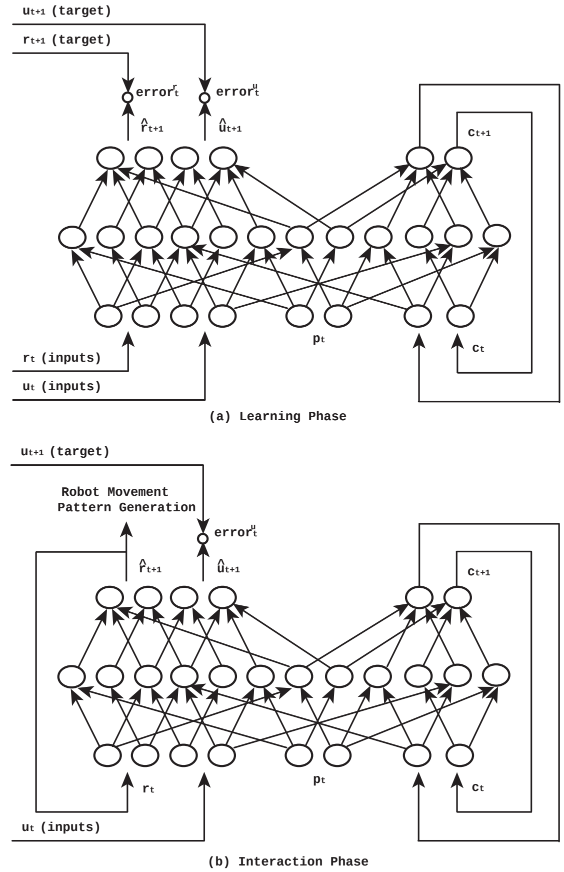

Figure 3. The system flow of RNNPB in learning phase (a) and interaction phase (b).

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 system.

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 [12]. 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:

$$\delta\rho_t{}^i = k_{bp} \cdot \sum_{step=t-l/2}^{t+l/2} \delta_{step}^{bp}{}^i + k_{nb}(\rho_{t+1}^i - 2\rho_t^i + \rho_{t-1}^i) \quad (1)$$

$$\triangle\rho_t^i = \epsilon \cdot \delta\rho_t{}^i + \eta \cdot \triangle\rho_{t-1} \quad (2)$$

$$p_t^i = sigmoid(\rho_t) \quad (3)$$

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 Experiments on imitation game

Two types of imitation game experiments are conducted using the proposed model. In the Experiment-1, the imitation game is set such that the on-line adaptation is conducted only in the robot side. In the Experiment-2, mutual adaptations between the robot and human subjects are conducted in the imitation game.

3.1 Experiment-1: robot adaptation only

In the Experiment-1, the robot learns three movement patterns shown by user's hand movements in the learning phase. In the interaction phase, we examined how the robot could follow target patterns while the user switched to demonstrate among various learned patterns.

The results of the experiment are plotted in Figure 4. It is observed that when the user hand movement pattern is switched from one pattern to another, the patterns in the sensory prediction and the motor outputs are also switched correspondingly by accompanying substantial shifts in the PB vector. Although the synchronization between the user hand movement pattern and the robot movement pattern is lost once during the transitions, the robot movement pattern is re-synchronized to the user hand movement pattern within several steps.

3.2 Experiment-2: mutual imitation game

The previous experiments focused mainly on the adaptation in the robot side. We conducted another experiment which focus on bi-directional adaptation in mutual interaction between the robot and users. In this new experimental set-up, after the robot learns 4 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.

One example of the interaction in imitation game is plotted in Figure 5.

It is observed that joint attention to a certain movement pattern between the robot and the subject as synchronization is achieved after some exploratory phase. It is also observed that this joint attentional state is break down once but joint attention to another pattern is achieved again.

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.

4 Discussion

The authors speculate that appropriate analysis of these observed phenomena might shed a ray of light on the mechanism of joint attention [2, 10] as well as turn taking behaviors [18]. 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 [1], 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 [14]. 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 [14] 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 [19]. Furthermore, Ikegami and Iizuka [6] 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 [13] 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.

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. It is speculated that the spontaneity originated from the total system dynamics including the users in the loop might play an important role in attracting people to play with entertainment robots.

5 Summary

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. Our experiments of the mutual adaptation suggest that the essential mechanism for autonomous shifts in joint attention and turn taking behavior could be explained by the open dynamic structures in which stability, in terms of rational goal-directedness, and instability, caused by unpredictability of others coexist.

References

[1] P. Andry, P. Gaussier, S. Moga, J. Banquet, and J. Nadel. 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, 2001.

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

[3] G. E. Butterworth and N. L. M. Jarrett. What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9:55–72, 1991.

[4] K. Dautenhahn and C. L. Nehaniv, editors. *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, 2002.

[5] M. Fujita, Y. Kuroki, T. Ishida, and T. Doi. A small humonoid robot sdr-4x for entertainment applications. In *Proceedings of International Conference on Advanced Intelligent Mechatronics*, 2003.

[6] T. Ikegami and H. Iizuka. Joint attention and dynamics repertoire in coupled dynamical recognizers. In *the AISB 03: the Second International Symposium on Imitation in Animals and Artifacts*, pages 125–130, 2003.

[7] M.I. Jordan and D.E. Rumelhart. Forward models: supervised learning with a distal teacher. *Cognitive Science*, 16:307–354, 1992.

[8] F. Kaplan and V.V. Hafner. The challenges of joint attention. In *Proceedings of 4th International Workshop on Epigenetic Robotics*, 2004.

[9] H. Kozima and H. Yano. A robot that learns to communicate with human caregivers. In *First International Workshop on Epigenetic Robotics*, 2001.

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

[11] Y. Nagai. *Understanding the Development of Joint Attention from a Viewpoint of Cognitive Developmental Robotics*. PhD thesis, Osaka University, Japan, 2004.

[12] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning internal representations by error propagation. In *Parallel Distributed Processing*. MIT Press, 1986.

[13] Y. Sato and T. Ikegami. Undecidability in the imitation game. *Minds and machines*, 14:133–143, 2004.

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

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

[16] J. Tani and M. Ito. 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, 2003.

[17] M. Tomasello. *The caltural origins of human cognition*. Harvard University Press., 1999.

[18] C. Trevarthen. Descriptive analyses of infant communicative behaviour. In H.R. Schaffer, editor, *Studies in Mother-Infant Interaction*. Academic Press, London, 1977.

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

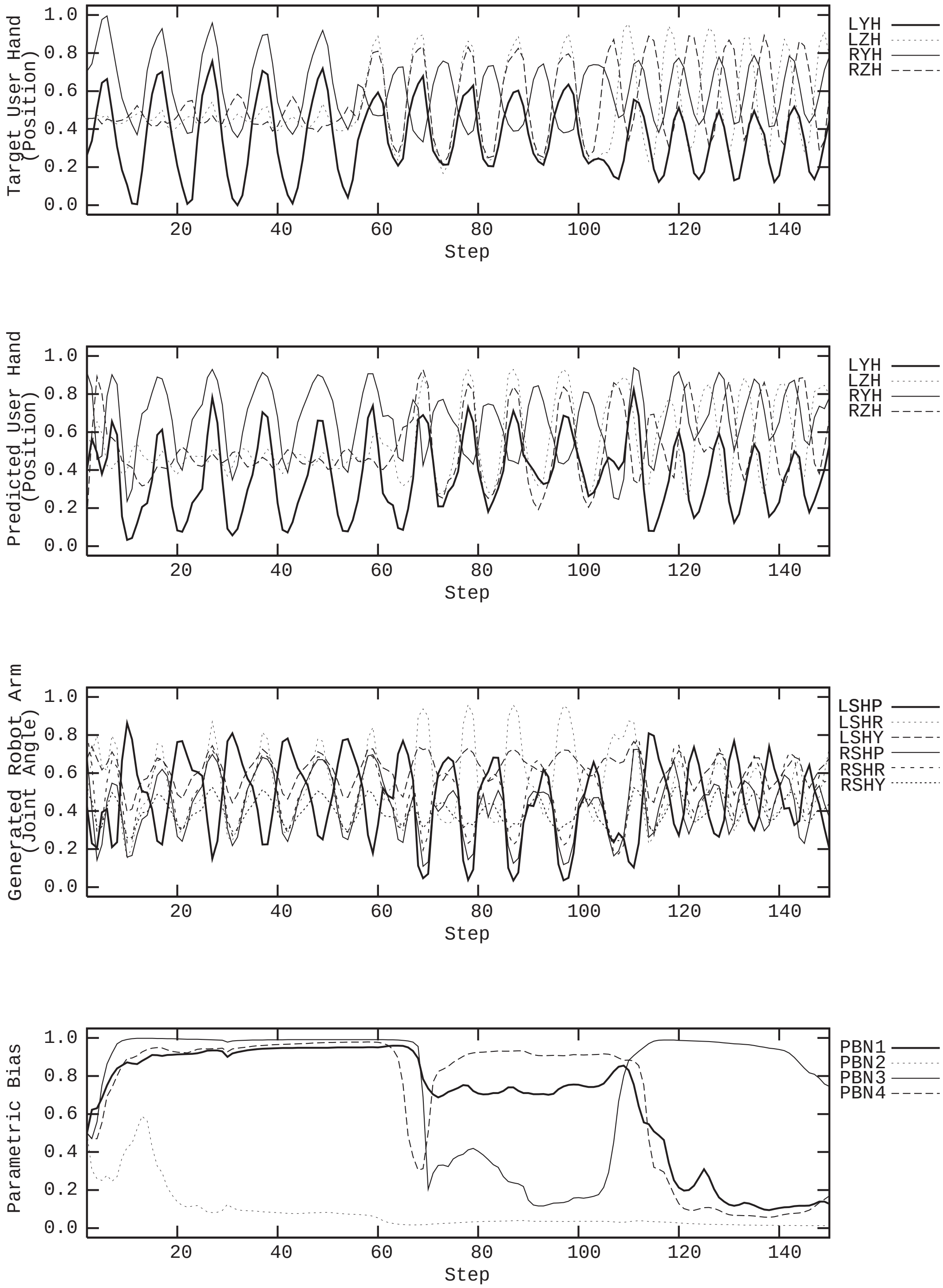

Figure 4. Switching of the robot movement pattern among three learned patterns as initiated by switching of user hand movement. User hand position and its prediction by the robot are shown in the first and the second row, respectively. The third row an the fourth row show motor outputs and PB vector, respectively.

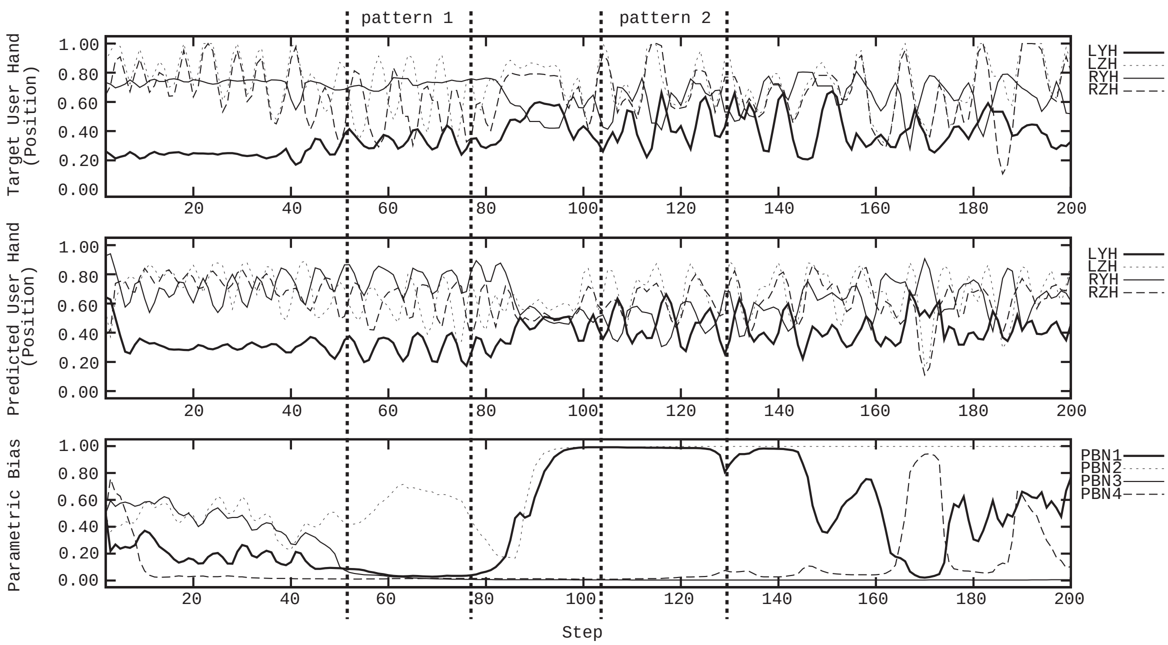

Figure 5. Joint attention as synchronization between the robot and the subject in imitation game. User hand position and its prediction by the robot are shown in the first and the second row, respectively. The third row shows PB vectors of the RNNPB.