Proc. 2004 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2004), Sendai, Japan, pp.1592-1597, 2004.

Human-Robot Collaboration Using Behavioral Primitives

Tetsuya OGATA[1)2)3)], Masaki MATSUNAGA[2)3)], Shigeki SUGANO[3)], *and* Jun TANI[2)]

1) Graduate School of Informatics, Kyoto University, Yoshida-honmachi, Sakyo, Kyoto, 606-8501, Japan, E-mail: ogata@i.kyoto-u.ac.jp

2) Brain Science Institute (BSI), RIKEN, Hirosawa 2-1, Wako, Saitama, 351-0198, Japan, E-mail: {ogata, tani}@brain.riken.jp

3) Humanoid Robotics Institute (HRI), Waseda University, Okubo 3-4-1, Shinjuku, Tokyo, 169-8555, Japan, E-mail: {ogata, sugano}@paradise.mech.waseda.ac.jp

Abstract

A novel approach to human-robot collaboration based on quasi-symbolic expressions is proposed. The target task is navigation in which a person with his or her covered and a humanoid robot collaborate in a context-dependent manner. The robot uses a recurrent neural net with parametric bias (RNNPB) model to acquire the behavioral primitives, which are sensory-motor units, composing the whole task. The robot expresses the PB dynamics as primitives using symbolic sounds, and the person influences these dynamics through tactile sensors attached to the robot. Experiments with six participants demonstrated that the level of influence the person has on the PB dynamics is strongly related to task performance, the person's subjective impressions, and the prediction error of the RNNPB model (task stability). Simulation experiments demonstrated that the subjective impressions of the correspondence between the utterance sounds (the PB values) and the motions were well reproduced by the rehearsal of the RNNPB model.

1. Introduction

Communication between people and robots requires an effective interface. The many kinds of interfaces that have been developed so far can be categorized into two types. The first type is those with "continuous interaction," and they include joysticks, master-slave interfaces, and other force-torque devices. Although users can control robot motions directly by using these devices, skill is needed to cooperate with the robot. The other type is those with "discrete interaction" based on language and/or symbolic expressions. While there have been many studies of human-robot speech communication, in which users share a task with robots explicitly, it is a major effort to implement the task model into the robot dialog efficiently due to the problem of "symbol grounding".

In this paper we discuss the possibility of a *quasi-symbolic* interface that uses the representations robots have acquired through experience. There have been many studies of machine learning, by which robots acquire representations of tasks, but they usually focused on only the analysis of the representations acquired and used to design and/or discuss robot intelligence. We focus on the "behavioral (motion) primitive" as an interface channel by which a robot can communicate and interact with a person. A behavioral primitive is a motion unit composes of various and complex motions inherent to biological systems [3][4].

We first describe the navigation task we used to investigate human-robot interaction. Then we describe our proposed approach, in which a person and a robot work together, using behavioral primitives self-organized in an artificial neural net. We then present some of the results of our trial experiments, and discuss the relationship between the results of recurrent neural net (RNN) learning and the person's subjective impressions.

2. Navigation Task

To investigate the essential mechanism of human-robot mutual interaction, we designed a navigation task [1] in which a humanoid robot called Robovie, developed at ATR [2], and a person work together to navigate a given workspace. Robovie has various features enabling it to interact with people: two arms with four degrees of freedom, a head with audiovisual sensors, and many tactile sensors attached to its body. Photographs of Robovie and of Robovie and a person performing the navigation task are shown in Figure 1. The experimental workspace was a 5x5-m course in which the outside walls were marked red and blue (Figure 2). Robovie and the person held their arms together and attempted to complete the course as quickly as possible without hitting the wall. Since the course had various branches, various kinds of experiments could be configured. The movement of the robot and the person was determined by a motor vector generated by the neural net in the robot. The person could affect the output of the neural net by using the tactile sensors on the arms of the robot, the detailed mechanism of which is described in Section 3. 3. The performance metric was by the time taken to complete the course.

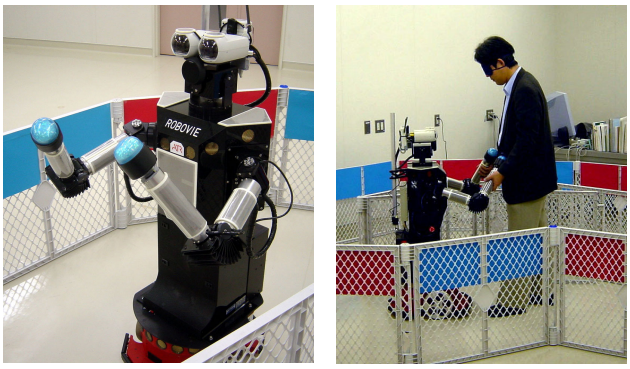

Figure 1 Robovie and Navigation Task Execution

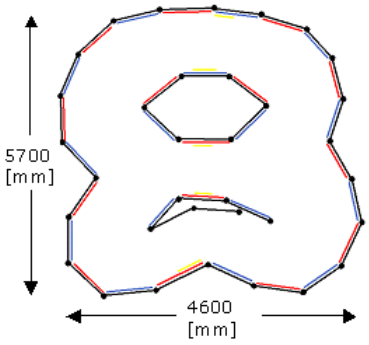

Figure 2 Experimental Course

It should be noted that the sensory information was quite limited for both the robot and the person in this collaboration task ("hidden state problem"). The robot could access only local sensory information from ultrasonic sensors and a poor vision system (which could only detect vague color information for its surroundings), not exact global position information. While the person was allowed to survey the course before the trial began, his eyes were covered during the entire trial. The person had to estimate his/her situation or position based on the image retained of the course geometry. During the trial the robot and the person had to help each other, utilizing the poor sensory information of different modalities and utilizing the history of the sensory-motor sequence (contextual information).

3. Proposed Approach

This section describes how the robot acquires motor primitives and uses them to interact with person. Some models which generate the motion primitives by articulating observed motions, have been proposed [3][4]. In our experimental tasks, because of the "hidden state problem" of the robot, we implemented the RNN, which can use and self-organize contextual information for the sensory-motor sequences, into the robot. We use the FF-model (forwarding forward model) proposed by Tani [3]. This model is also called the recurrent neural network with parametric bias (RNNPB) model. It articulates complex motion sequences into motion units, which are encoded as the limit cycling dynamics and/or the fixed-point dynamics of the RNN.

3. 1 RNNPB model

The RNNPB model has the same architecture as the conventional Jordan-type RNN model [5] except for the PB nodes in the input layer. Unlike the other input nodes, these PB nodes take a constant value throughout each time sequence and are used to implement a mapping between fixed length values and time sequences.

Like the Jordan-type RNN model, the RNNPB model learns data sequences in a supervised manner. The difference is that in the RNNPB model, the values that encode the sequences are self-organized in the PB nodes during the learning process. The common structural properties of the training data sequences are acquired as connection weights by using the back propagation through time (BPTT) algorithm [6], as used also in the conventional RNN. Meanwhile, the specific properties of each individual time sequence are simultaneously encoded as PB values. As a result, the RNNPB model self-organizes a mapping between the PB values and the time sequences.

The learning algorithm for the PB vectors is a variant of the BPTT algorithm. The step length of a sequence is denoted by l. For each of the sensory-motor outputs, the back-propagated errors with respect to the PB nodes are accumulated and used to update the PB values. The update equations for the ith unit of the parametric bias at the t in the sequence are as follows,

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

$$\Delta\rho_t = \varepsilon \cdot \delta\rho_t \quad (2)$$

$$p_t = sigmoid(\rho_t / \zeta) \quad (3)$$

In Eq. (1), the δ force for the update of the internal values of the PB p_t is obtained from the summation of two terms. The first term represents the delta error, $\delta_t^{bp,}$ back-propagated from the output nodes to the PB nodes: it is integrated over the period from the $t-l/2$ to the $t+l/2$ steps. Integrating the delta error prevents the local fluctuations in the output errors from significantly affecting the temporal PB values. The second term is a low-pass filter that inhibits frequent rapid changes of the PB values. Internal value ρ_t is updated using the delta force, as shown in Eq. (2). Then, the current PB values are obtained from the sigmoidal outputs of the internal values. After learning the sequences, the RNNPB model can generate a sequence from the corresponding PB values.

Furthermore, the RNNPB model can be used for recognition processes as well as for sequence generation processes. For a given sequence, the corresponding PB value can be obtained by using the update rules for the PB values (Eqs. (1) to (3)), without updating the connection weight values. This inverse operation for generation is regarded as recognition.

The other important characteristic of the RNNPB model is that relational structure among the training sequences can

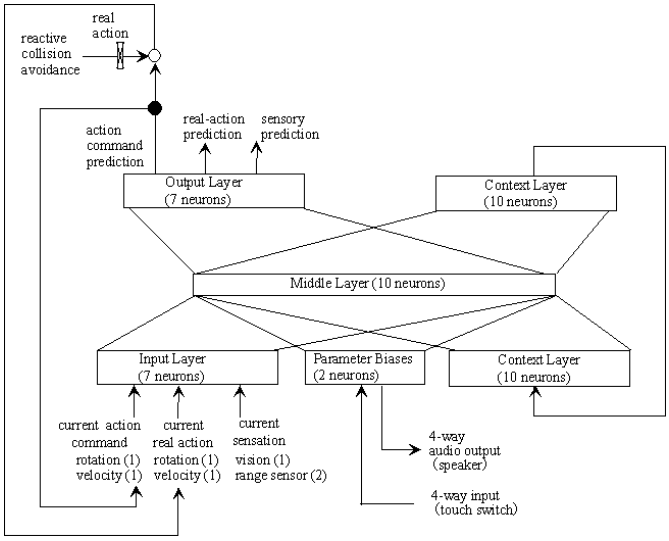

Figure 3 Architecture of RNNPB Model

Table 1 Correspondence between PB Conditions and Number

Parameter 1	Parameter 2	InPut & Output
Low	Low	"1"
Low	High	"2"
High	Low	"3"
High	High	"4"

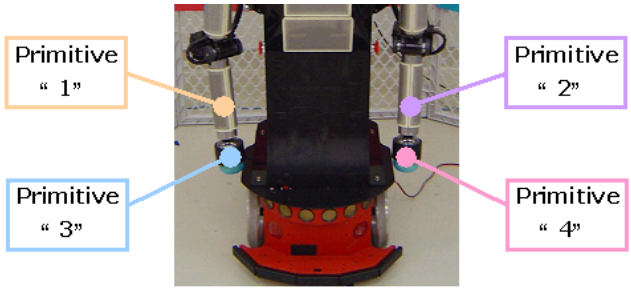

Figure 4 Correspondence between Primitive Numbers and Tactile Sensors

be acquired in the PB space through the learning process. This generation capability enables the RNNPB model to generate and recognize unseen sequences without any additional learning. For instance, by learning several cyclic time sequences of different frequencies, it can generate novel time sequences of intermediate frequencies.

3. 2 Implementation

Figure 3 shows the architecture of the RNNPB model used in the robot. The model has two parameter bias nodes, and operates in a discrete time manner by synchronizing each event.

The input layer of the RNNPB model consists of the current sensory inputs and the current motor values. The sensory inputs are comprised of the output of the ultrasonic range sensors and the color area acquired from an omni-direction camera mounted on the robot's back. The motor values are the forward velocity and the rotation velocity. The output layer is the prediction of the next sensory input and next action. The activations of the context outputs in the current time step are copied to those of the context inputs in the next time step. The context unit's activities are self-organized through learning processes. The robot obtains the color area, range sensor data, and vehicle conditions every 0.1 s. This data is compressed and filtered. The RNN receives this preprocessed data as input and generates the output with a time interval of 1 s.

3. 3 Interface

We designed the interface using the PB values, by which the person and robot interact.

Since the person's eyes were covered in the experiments, the robot had to inform the condition of the PB values using sounds. As described in the previous section, the RNNPB model has two parameter bias nodes. While it would be best if the person were informed of the analog values of these two nodes directly, it is quite difficult to express slight changes in the node values by sound. Therefore, we designed the robot to utter four different symbolic sounds (numbers) corresponding to the conditions of the PB nodes during navigation (Table 1). The activation of each parameter node was divided into two states (high and low) with the threshold set to 0.5. For example, if the perameter 1 is 0.7 and the parameter 2 is 0.3, then the output number becomes "3".

The person had to learn the relationship between these sounds (PB vectors) and the actual robot motions (RNNPB outputs). The person then adjusted the PB values by touching the appropriate tactile sensors attached to the robot so as to move around the course as quickly as possible. The four utterance numbers corresponded to the four tactile sensors on the forearms and the wrists of the robot, as shown in Figure 4. The PB value was switched to the value corresponding to the number of the tactile sensor touched by the person. This process was implemented by modifying of the Eq. (1) as follows.

$$\delta\rho_t^i = k_{bp} \cdot \sum_{t-l/2}^{t+l/2} \delta_t^{bp^i} + k_{nb}(\rho_{t+1}^i - 2\rho_t^i + \rho_{t-1}^i) + k_{input} \cdot \eta^i \quad (4)$$

Here, η^i is either +1 or −1 depending on the input from the person, and k_{input} is the influence level.

4 Experiments

4. 1 Pre-experiment on RNNPB model

A pre-experiment using only the robot was carried out to confirm the basic characteristics of the RNNPB model. In this experiment, the RNNPB model had only one PB node, facilitating observation of its change.

Figure 5 shows the sensory-motor data and the PB value

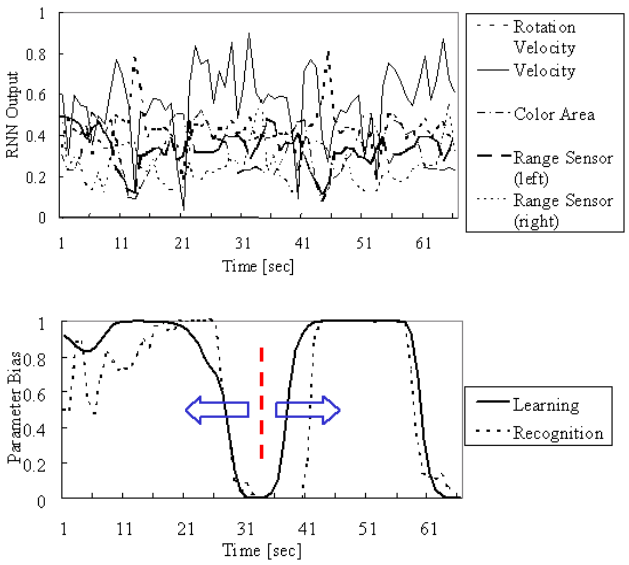

Figure 5 Sensory-Motor Data and Parameter Bias

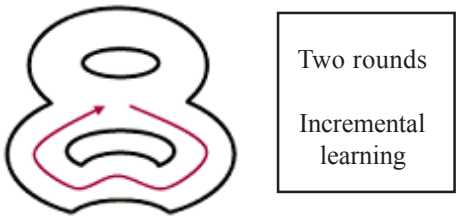

Figure 6 Course for Pre-Experment

output when the robot moved twice around the course shown in Figure 6. Here, "learning" curve means the result of PB learning, and "recognition" curve is the result of real-time PB identification. There was a delay between learning and recognition due to the time it took to calculate the updated PB value.

Although the dynamics of the sensory-motor data were quite complex, the curve of the PB vector showed that the actual motion could be clearly divided into two parts. The RNNPB model can thus convert complex dynamics into a combination of simple units.

4. 2 Experiments on Human-Robot Collaboration

We tested human-robot collaboration using six male participants to determine how well people can cooperate with the robot by using the PB interface and to determine how people might interpret the meanings of the four sounds (numbers) uttered by the robot. We also investigated the effect of the influence level (k_{input} in Eq. (4)).

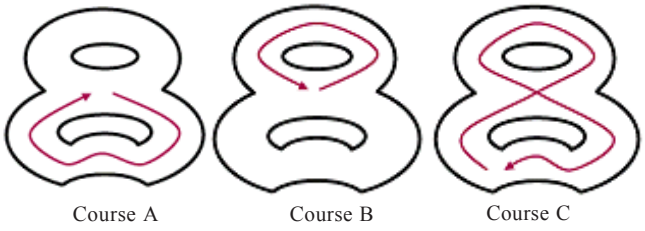

Figure 7 Three Courses for Pre-Learning and Experiment

Before the experiment, the robot's RNNPB model (Figure 3) was trained (acquired motion primitives) using courses A and B, which are shown schematically in Figure 7. The actual course used in the collaboration experiment was course C, which is also shown in Figure 7. Since the RNNPB model mainly uses the contextual information (sensory-motor sequence), the differences between the three courses are quite large from the robot's point of view. Furthermore, because the RNNPB model was not further trained during the collaboration experiment, the robot required the participant's support and had to "reuse" the acquired motion primitives to move around the unfamiliar course as quickly as possible.

The experiment had 14 trials and was divided to two parts 7 trials). In each part, k_{input} was set to either 0.05 or 0.01. After each trial, there was a break during which the participant completed a questionnaire based on NASA-TLX [7]. In total, we obtained 84 (14 trials x 6 subjects) sets of data (all-play-all). The parts were presented to the participants in random order to avoid the effects of a fixed-order presentation.

4. 3 Results

The average completion times are shown in Figure 8. The time was reduced when the robot received support from a person. However, the variance was greater with the higher influence level.

The average prediction errors of the RNNPB model are shown in Figure 9. While the errors were almost the same with only the robot and with human assistance (k_{input}=0.01), that with human assistance (k_{input}=0.05) was quite large. These

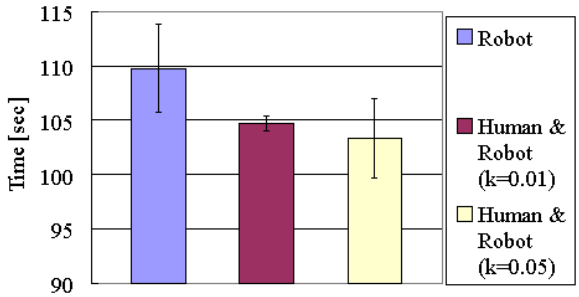

Figure 8 Transit Time

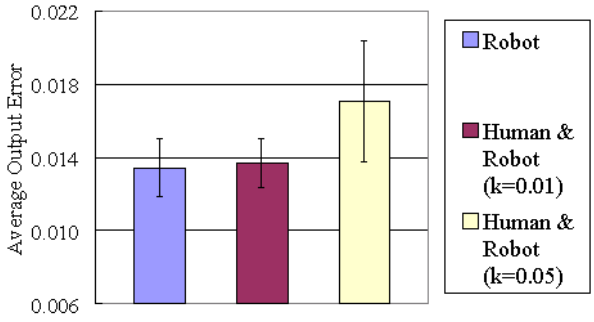

Figure 9 Comparison of the Prediction Errors of RNN

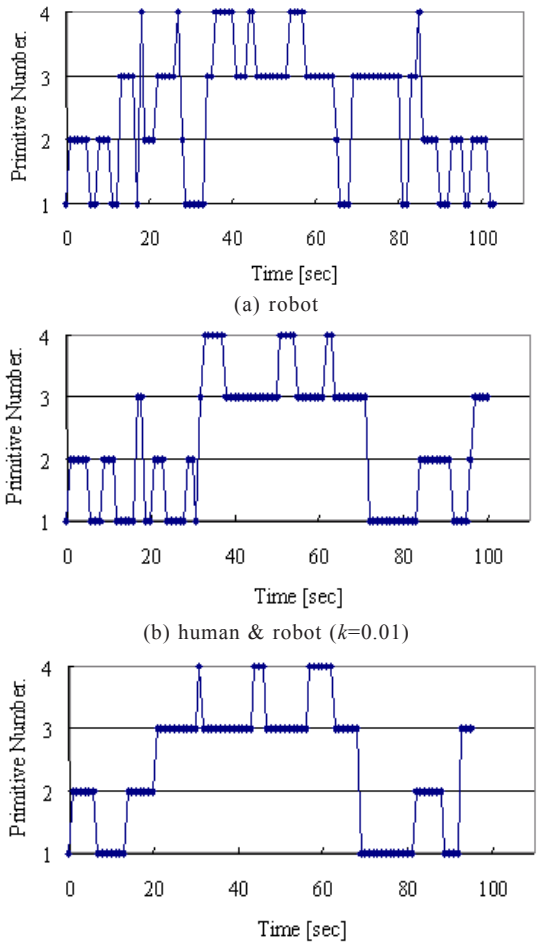

Figure 10 Comparison of Primitive Switching

results indicate that although a higher influence level can result in the better performance, the human-robot collaboration tends to be less stable.

Figure 10 shows representative examples of the transition in utterance number switching for the three cases. The switching frequency with only the robot" was much higher than with human assistance. In particular, the primitive switched quite frequently around the branch points in the course (at about 10-30, and 70-90 sec in Figure 10(a)).

4. 4 Subjective Impressions

The results of the NASA-TLX questionnaire are plotted in Figure 11. The significant values for the 1 and 5 % levels were calculated using a t-test. The higher influence level resulted in a higher evaluation for all items. Note, however, that only "mental load" showed a substantial difference ($p<0.01$). This could have been due to the instability in the performance when the influence level was high, as mentioned in Section 4.3.

Interviews with the participants about their impressions of the correspondence between the uttered numbers and actual motions of the robot revealed that all the participants had

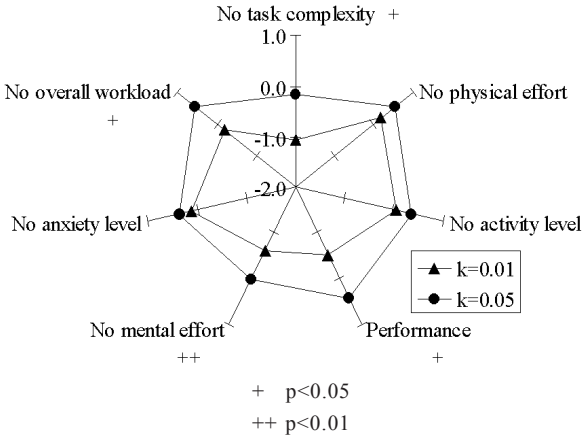

Figure 11 Results of NASA-TLX Questionnaire

Table 2 Correspondence between Number and Motion (Interview)

Number	Motion
"1"	Moving Straight
"2"	Turning Right
"3"	Moving Straight
"4"	Turning Left

almost the same impressions after the 14 trials. The impressions are shown in Table 2.

5. Discussion

5. 1 Fluctuation in PB Values at Branch Points

The uttered numbers tended to fluctuate at the branches in the course, as shown in Figure 10. Although the RNNPB model predicts the sensory-motor flow by using contextual information, it cannot predict the PB output by itself. Each branch in the course is thought to be a "saddle point" in the dynamical-system sense. Therefore, the robot requires higher level information concerning the PB dynamics to select the correct direction at a branch point. In our experiment, the person predicted the PB dynamics based on experience, and supported the RNNPB model's output.

5. 2 Influence Level

As described in Section 4.1, robot control was easier with the higher influence level k. However, the prediction error increased, and the performance became unstable, that is, the performance sometimes failed.

In our experiment, the person (with covered eyes) and the robot had to collaborate in a context dependent manner, because neither had enough sensory information to complete the task. A high influence level thus effectively improves task performance because a person can basically utilize contextual information better than a robot (RNN). However, once a prediction error occurred in the person's mental image, it

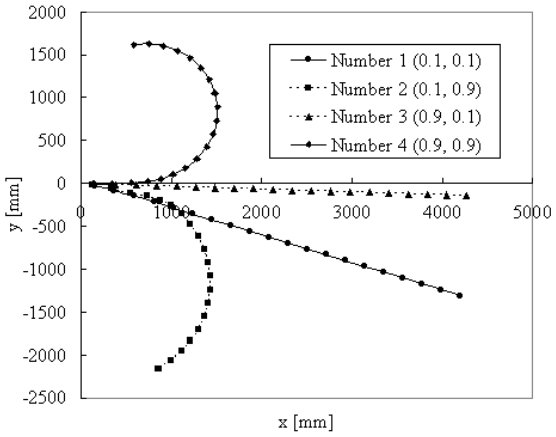

Figure 12 Simulated Trajectories of RNNPB model

was quite difficult to recover the contextual image without the support of the robot. This is a main reason why the variance in transit time was large. Efficient human-robot collaboration is thus achieved only when the influence level is set to an appropriate degree.

5. 3 Subjective Impression and RNNPB Learning

The dynamic properties of the RNNPB model we used were investigated in simulation experiments. Figure 12 shows the four trajectories generated when the RNNPB model rehearsed four times, each time with the PB values indicated in parentheses. For example, when the trajectory of number "2" was generated, the parameter 1 was set to "0.1" and the parameter 2 was set to "0.9" respectively. In rehearsal, copies of the current sensory-motor prediction outputs are fed back to the next inputs (closed loop). This enables RNN prediction for an arbitrary number of future steps.

The trajectories shown in Figure 12 correspond exactly to the subjective impressions listed in Table 2. Actually, it was not easy for the participants to establish the correspondence because the PB condition was not clearly categorized into one of the four states. For example, we observed that the robot sometimes changed the utterance number drastically, possibly due to the fluctuation in PB values around the threshold of 0.5. Nevertheless, the participants could still guess the meaning of the uttered numbers based on their experience. This shows the feasibility of human-robot collaboration based on quasi-symbolic expressions using behavioral primitives.

6. Conclusion

We have described a new approach to human-robot collaboration based on quasi-symbolic expressions. The target task is navigation in which a person (with his or her eyes covered) collaborates with a humanoid robot called Robovie in context dependent manner. The robot uses a recurrent neural net with parametric biases (RNNPB) to acquire the behavioral primitives, i.e., the sensory-motor units, composing the whole task. The robot expresses the PB dynamics as primitives using symbolic sounds, and the person influences the robot's dynamics by touching tactile sensors attached to the robot. Experiments carried out with six male participants demonstrated that the level of influence is strongly related to task performance, the subject's subjective impressions, and the prediction error of the RNNPB model (task stability). Simulation experiments demonstrated that the impression of the correspondence between the uttered sounds (the PB values) and the robot's motions were well reproduced by the rehearsal of the RNNPB model.

Our future work has two main objectives. One is to introduce a method for incremental learning. The RNNPB model we used was trained prior to the collaboration experiments, not during the experiment. When real-time incremental learning is introduced, we need to solve the problem of confliction between new memory and past memories [1]. The second is to apply human-robot verbal communication based on the proposed method. By preparing more expressions translated from the PB values, we should be able to use our method for more complex tasks. Examination of the binding between sentences and sensory-motor sequences (embodied language) [8] will thus be quite important.

References

[1] T. Ogata, N. Masago, S. Sugano, and J. Tani, "Interactive Learning in Human-Robot Collaboration", International Conference on Intelligent Robots and Systems (IROS2003), Vol.1, pp.162-167, 2003.

[2] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, and R. Nakatsu, "Robovie: an interactive humanoid robot." International Journal of Industrial Robotics, Vol. 28, No. 6, pp.498-503, 2001.

[3] J. Tani and M. Ito, "Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics: A Robot Experiment", IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, Vol.33, No.4, pp. 481-488, 2003.

[4] M. Haruno, D. Wolpert, and M. Kawato, "MOSAIC model for sensorimotor learning and control", Neural Computation 13, pp. 2201-2220, 2001.

[5] M. Jordan, "Attractor dynamics and parallelism in a connectionist sequential machine." in Proc. of the Eighth Annual Conference of the Cognitive Science Society (Erlbaum, Hillsdale, NJ), pp. 513-546, 1986.

[6] D. Rumelhart, G. Hinton, and R. Williams, "Learning internal representation by error propagation", in D.E. Rumelhart and J.L.McLelland, editors, Parallel Distributed Processing (Cambridge, MA: MIT Press), 1986.

[7] S.G. Hart et. al, "Development of NASA-TLX: Results of empirical and theoretical research." inP.A.Hancock and N.Meshkati(eds.), Human Mental Workload, North-Holland, pp. 139-183, 1988.

[8] Y. Sugita and J. Tani, "A Holistic Approach to Compositional Semantics: a connectionist model and robot experiments", Advances in Neural Information Processing Systems (NIPS), 2003.