Proceedings of the 2007 IEEE/RSJ International
Conference on Intelligent Robots and Systems
San Diego, CA, USA, Oct 29 - Nov 2, 2007

Discovery of Other Individuals by Projecting a Self-Model Through Imitation

Ryunosuke Yokoya, Tetsuya Ogata, Jun Tani, Kazunori Komatani, and Hiroshi G. Okuno

Abstract— **This paper proposes a novel model which enables a humanoid robot infant to discover other individual (e.g. human parent). In this work, the authors define "other individual" as an actor which can be predicted by a self-model. For modeling the developmental process of discovering ability, the following three approaches are employed. (i) Projection of a self-model for predicting other individual's actions. (ii) Mediation by a physical object between self and other individual. (iii) Introduction of infant imitation by parent. For creating the self-model of a robot, we apply Recurrent Neural Network with Parametric Bias (RNNPB) model which can learn the robot's body dynamics. For the other-model of a human, conventional hierarchical neural networks are attached to the RNNPB model as "conversion modules". Our target task is a moving an object. For evaluation of our model, human discovery experiments by the robot projecting its self-model were conducted. The results demonstrated that our method enabled the robot to predict the human's motions, and to estimate the human's position fairly accurately, which proved its adequacy.**

I. Introduction

There are many challenges in the development of human infants. One of these is how infants discover their parents in environments full of information. This paper proposes a constructive model that enables a robot to discover other individuals. Recently, cognitive developmental robotics [1] has gained attention as a novel approach for the design of humanoid robots and understanding infant development.

For modeling the development process, following three approaches are employed. (i) Projection of a self-model for predicting other individual's actions. (ii) Mediation by a physical object between self and other individual. (iii) Introduction of infant imitation by parent.

It is absolutely essential for creatures including humans to predict the environment around them so that they can adapt themselves to the real world. For adaptation, they need to construct models of environmental dynamics by learning, etc. However, it would not be effective for them to retain all the models discretely in their own mind because real environment has a great variety of dynamical properties.

Individual creatures fundamentally have their own model of dynamics, a "self-model", for generating behaviors in their own mind. As indicated by Makino et al. [2], if the external environment has certain properties similar to the creatures' internal models, they should be able to retain various environmental models, "other-models", by reusing their own self-models. When they discover another individual who has properties similar to their own in the external environment, they will recognize the individual as "another individual similar to me", and communication between them may eventually occur. In this work, the robot infant predicts human parent by reusing its self-model (approach (i)).

Humans acquire the ability to recognize themselves and others through interactions with their parents in early childhood. It has long been noted that "ternary relationship among self-object-other" [3] and "infant imitation by parent" [4] are important for infants to develop their cognitive faculty. Moreover, Jones empirically indicated the probability that infants notice being imitated by their parents [5]. We set a moving object on a table as a task (approach (ii)), and introduce "infant imitation by parent" to our model (approach (iii)) based on the knowledge about the development of human infants.

This paper proposes a novel model for discovery of other individuals by a humanoid robot based on predictions reusing the self-model. In this work, the authors define "other individual" as an actor which can be predicted by a self-model. The robot preliminarily has no correspondences between its own body and human body. It constructs the other-model and discovers the existence of other individual through some interactions focusing only on the object.

Section II describes our discovery method, and the details of the self-model and other-model. Section III describes the implementation of the robot hardware and the neural network model. Section IV describes the discovery experiments and the results obtained. Section V discusses the results and its relation with the brain science. Section VI concludes this paper.

II. Method for Discovering Other Individuals by Projecting a Self-Model

A. *Proposed Model*

1) Interaction between Parent and Infant: This section describes how the infant discovers his parent from the environment around him. We assume that a parent imitates an infant in the early interaction between them.

From before birth to after birth, an infant organizes his self-model through body babbling, which is an experiential process where the infant learns what muscle movements achieve a particular goal state [6]. Soon after birth, the infant begins to interact with the parent by focusing on specific information such as voice or sound resulting from his motions: the information is an object in this work. Through the interaction in which the parent imitates the infant, the

R. Yokoya, T. Ogata, K. Komatani and H. G. Okuno are with Graduate School of Informatics, Kyoto University, Kyoto, Japan {yokoya, ogata, komatani, okuno}@kuis.kyoto-u.ac.jp
J. Tani is with the Brain Science Institute, RIKEN, Saitama, Japan tani@brain.riken.jp

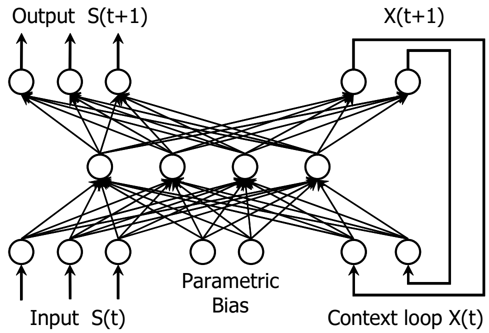


Fig. 1. Self-model: RNNPB.

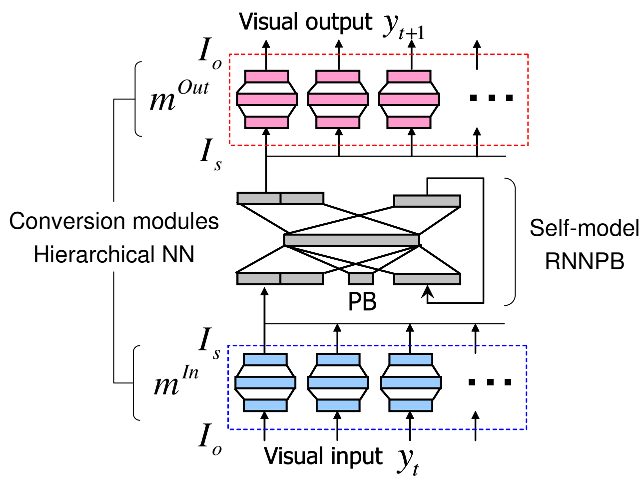


Fig. 2. Other-model: RNNPB with NN.

infant comes to be able to recognize the rough position of the parent who has a dynamical model similar to his own self-model.

2) Overview of Our Process: We present an overview of our method based on the situation described in section II-A.1.

The self-model for the robot infant and the other-model for the human parent consist of following two parts.

- Recurrent Neural Network with Parametric Bias
  RNNPB (See Fig. 1) is used for the self-model. It can learn the relations between the robot's motions and resulting object's motions.
- Hierarchical neural networks
  The other model is organized by attaching pairs of conventional hierarchical neural networks, m^{In} and m^{Out}, to input-output units of the RNNPB model (See Fig. 2). In this work, each pair of networks is called "conversion module". The conversion module plays the role of mutual conversion between the subjective and objective information; I_s and I_o. In this work, only object motion is a conversion target.

Our discovery process consists of three phases (See Fig. 3). Fundamentally, this process should be seamless; however, it is divided into these phases for convenience. Note that the robot focuses only on the object's motion all of the time. We overview it as follows.

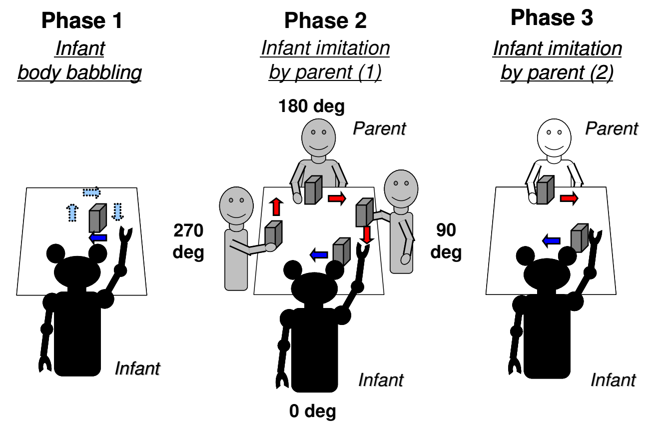


Fig. 3. Discovery process.

a) Phase 1: Learning of Self-Model (Infant Body Babbling): The robot collects the data of its arm motions and object motions while it manipulates the object. The self-model, RNNPB, is then trained with the data. The relations between the robot's motions and object's motions are acquired in this phase.

b) Phase 2: Learning of Conversion Modules (Infant Imitation by Parent (1)): Soon after the robot first manipulates the object, the human parent imitates it at multiple positions. While the human manipulates the object, the robot collects observed data. The conversion modules are trained with the data so that the other-model can predict the object's motion generated by the human. In this phase we assume that the robot knows being imitated by the human.

c) Phase 3: Estimation of Human Position (Infant Imitation by Parent (2)): In the same way as in the previous phase, the robot collects data while it observes the human's imitative motion at a certain position. The robot predicts the observed motions by using the other-model. The position of the parent is then estimated based on the prediction errors. We also assume that the robot knows being imitated by the human in this phase.

B. Self-Model and Other-Model

This section describes the learning models used in our method and their learning algorithm.

1) RNNPB and Its Learning Algorithm: This section describes the architecture and learning algorithm of the RNNPB model.

The RNNPB model is the FF-model (forwarding forward model) proposed by Tani and Ito. The RNNPB model works as a prediction system: its input data is the current sensory state S(t) and its output data is the predicted sensory state S(t+1) in the next step. This model has the same architecture as the conventional hierarchical neural network model except for the context layer and the PB nodes in the input layer. Unlike the other input nodes, these PB nodes take a constant value throughout each time sequence. The context layer has a loop that inputs the current output as input data in the next step. An advantage of this layer is that the RNNPB model can use it to learn the time sequences by leveraging past contexts.

After learning the time sequences, the RNNPB model self-organizes the PB values at which the specific properties of each time sequence are encoded.

An algorithm called the BPTT (back propagation through time) [8] is employed for RNNPB learning. Although the learning algorithm for the conventional hierarchical neural network is back propagation, the RNNPB model cannot learn with this algorithm because it does not have a teacher signal to the context layer.

The PB values are calculated during the learning process as follows.

$$\delta\rho_t = k_{bp} \cdot \sum_{0}^{T} \delta_t^{bp}, \quad (1)$$

$$p_t = sigmoid(\rho_t), \quad (2)$$

where k_{bp} is a constant; ρ_t is the internal value of the PB node at t; p_t is the PB value of the PB node at t; δ_t^{bp} is the delta error back-propagated from the output nodes to the PB nodes; and T is the sensory sequence length. In (1), the delta errors are integrated errors in all the steps. In (2), the current PB values are obtained from the sigmoidal outputs of the internal values. Based on these equations, a unique PB value is calculated for each time sequence.

In the first phase, the weights and the PB values of the RNNPB model are calculated simultaneously. The input data for the RNNPB model are motor information and visual information in this work.

2) Method for Organization of Conversion Modules: After the RNNPB model is organized using BPTT in the babbling phase, the RNNPB model is used in the second and third phases. This section describes how the conversion module is organized in the second phase.

The other-model is used for prediction of the object's motion generated by the parent. The weights of the conversion modules are updated using BPTT without updating the connection weights for the RNNPB model so that the other-model can predict the observed data. The PB values corresponding to the observed manipulation acquired in the first phase are input to the model. This is done based on the fact that the robot knows what kind of manipulation the parent did, due to the assumption that a parent imitates an infant.

The collected data include no arm motor data because the robot is just looking at the target and does not move, unlike in the first phase. The initial arm motor values are then input to the motion input layer in step 0, and the outputs are calculated forward in the closed looping mode from step 1; the outputs in the motion output layer in step t-1 are input to the motion input layer in step t (See Fig. 4). To put this simply, the motion input layer plays the same role as the context layer does.

3) Method for Discovery of Other Individuals: This section describes how the robot estimates the position of the parent by using the other-model in the third phase.

The robot predicts the observed data thorough the forward calculation of the other-model by using each conversion module and calculates output errors: the observed data are input to the other-model, and delta errors are calculated based on the output data and the real data (teacher data). The PB values corresponding to the observed manipulation are input to the model. The robot then estimates the position based on the errors calculated by each module. The estimation is described in detail in section IV-C.2.

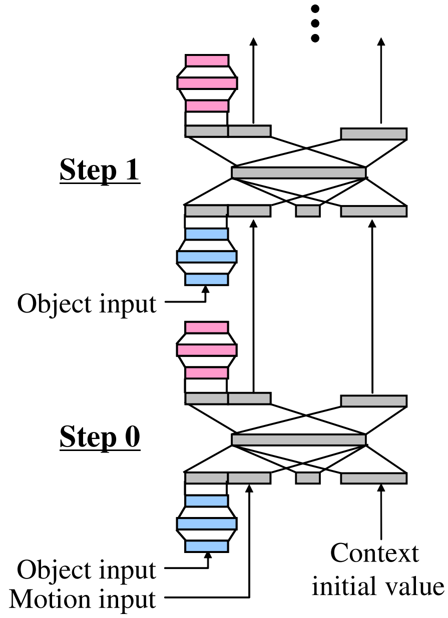


Fig. 4. Forward calculation of other-model.

III. Model and System

A. Humanoid Robot Robovie-IIs and Target Object

Our experimental platform was a humanoid robot, Robovie-IIs, a refined model of Robovie-II developed at ATR [9] (See Fig. 5). Robovie has three d.o.f. (degree of freedom) in its neck and four d.o.f. in each arm. Each motor angle value is measured with potentiometers. It also has stereoscopic CCD cameras on its head.

The manipulation target is a box-shaped object (See Fig. 5). The top of the object is separated into two colors; red and blue.

B. Experimental System

Figure 6 is the system diagram. After the camera and the motors collect data, the self-model and the conversion modules of the other-model are trained with the data off-line.

The self-model, the RNNPB model, consists of fifty-nine neurons: seven in the input layer, thirty-five in the middle layer, fifteen in the context layer, and two as parametric bias. The other-model has four conversion modules corresponding to four positions of the human: 0 deg., 90 deg., 180 deg. and 270 deg. Each conversion module consists of eight neurons: four in the input layer and four in the middle layer. The input for the modules is the visual data only.

The following sensory data were collected in the experiment for use in the model.

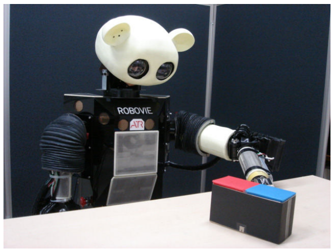


Fig. 5. Robovie-IIs and object.

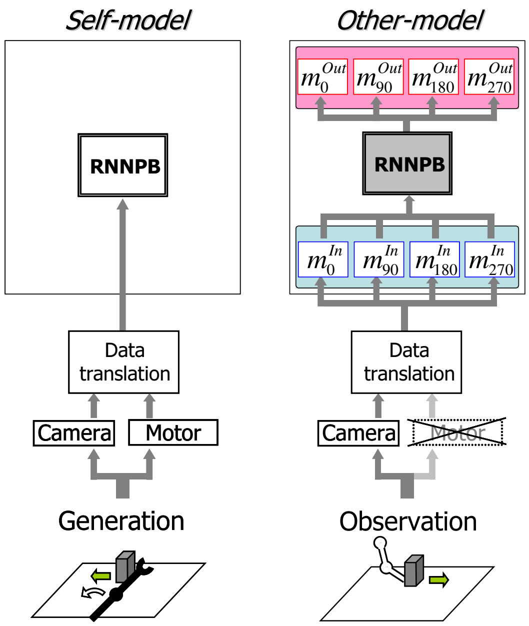


Fig. 6. System diagram.

a) Visual Information (four units): The trajectory of an object was collected from the image information by a CCD camera with a resolution of 500×400 pixels. The center position of each colored top face, the X-Y coordinates in the camera ([0-1]), was estimated by extracting the object from the color information: the visual information had four dimensions. The initial position of the object is the center of the camera view in all phases. Only the left eye camera was used.

b) Motor Information (three units): The left arm (3 d.o.f.: roll and yaw axes of the shoulder, and pitch axis of the elbow) was used. Note that the pitch axis of the shoulder and the neck motors were fixed in the experiment. In our experiments, the initial arm motor values were constant.

Those values were synchronized between different modalities, and were normalized in [0.1-0.9] based on the maximum and minimum values. The sensory data were stored every 800 msec for each manipulation, and their lengths were all ten steps: each trial duration was 8 sec.

IV. Experiments for Discovery of Others

A. Task

There were five kinds of manipulation: parallel translation from the left to the right ("L→R"), rotation to the right

TABLE I

Object manipulation.

	Moving direction
Motion 1	L→R
2	Rrot
3	Lrot1
4	Lrot2
5	F→B

("Rrot"), rotation to the left by pushing the right side of the object ("Lrot1"), rotation to the left by pulling the left side ("Lrot2") and moving from the front to the back ("F→B") (See Table I). Only motion 5 is not used for learning the conversion modules in the second phase.

B. Procedure

c) Phase 1: Learning of Self-Model: The robot first conducted the motions listed in Table I to manipulate the object. The robot began all motions from the four object postures (See Fig. 7), because this avoided the chance that the human's position was obvious only from the initial posture of the object. A total of twenty sets of data were collected in the first phase. The RNNPB model was then trained with the collected data simultaneously 200,000 times, which took approximately ten minutes.

In the second and third phases, both the robot and the human began the motions from the initial posture of the object 1 (See the upper left of Fig. 7).

d) Phase 2: Learning of Conversion Modules: The robot observed four types of manipulation, motions 1-4, presented by human parent. The parent did the motions at four positions, 0 deg., 90 deg., 180 deg. and 270 deg (See Fig. 3). The angles represent the parent's position in the counterclockwise direction relative to the initial position of the object. The robot collected sixteen sets of visual data. Four conversion modules were trained 100,000 times each with the collected sensory data, which took approximately ten minutes per module. Each pair of modules learned with four sets of data. Note that the PB values corresponding to the observed manipulations acquired in the first phase are input to the other-model.

e) Phase 3: Estimation of Human Position: Finally, the robot observed motions 1-5 presented by its parent. The parent did the motions at eight positions, 0 deg., 45 deg., 90 deg., 135 deg., 180 deg., 225 deg., 270 deg. and 315 deg. The robot collected visual data: there were five patterns for each parent position. The other-model with four pairs of conversion modules predicted the dynamic sequences by forward calculation with the collected sensory data. The robot then estimated the parent's position based on the calculated errors. The details of the estimation are described in section IV-C.2. As in the case of the previous phase, the PB values corresponding to the observed manipulations are input to the model.

Figure 8 shows an example of sequential photographs of a scene in the second or third phase: the manipulation is Rrot, and the human's position is 270 deg. The left photograph captures the robot manipulating the object, and the right one captures the human imitating the manipulation.

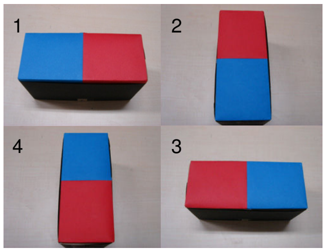


Fig. 7. Initial postures of the object.

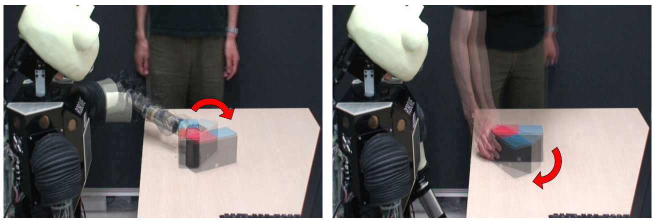


Fig. 8. Infant imitation by parent (Motion 2 at 270 deg).

C. Results

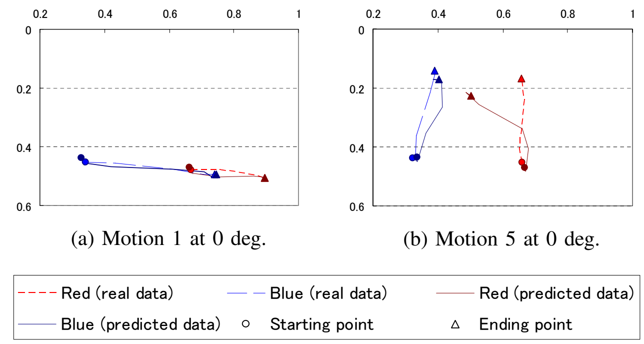
1) Prediction of Other's manipulations by Reusing Self-model: Figure 9 shows examples of the real trajectory of the object generated by the human, and object trajectory predicted by the conversion modules organized in the second phase. The solid lines describe the output (predicted data) of the other-model with the conversion modules corresponding to 0 deg, and the broken lines describe the input (real data). We confirmed that the other-model could predict the learned and un-learned sequences fairly accurately.

2) Estimation of Human Position: All the data observed in the third phase were input to the other-model with the conversion modules, and output errors were calculated for each module. A value L_i was then calculated as follows for all the data.

$$L_i = \frac{\frac{1}{e_i}}{\frac{1}{e_0} + \frac{1}{e_{90}} + \frac{1}{e_{180}} + \frac{1}{e_{270}}} \quad (i = 0,\ 90,\ 180,\ 270) \quad (3)$$

where $e_i(i = 0, 90, 180, 270)$ are errors calculated by each pair of modules. The errors were obtained through forward calculation of the other-model: the difference between the output of m^{Out} at the current step and the real data at the next step were accumulated for all the steps. Figure 10 shows X-Y coordinates in which each plotted point represents $(L_{90} - L_{270}, L_{180} - L_0)$. Thus, distance from the coordinate origin represents a value similar to the likelihood of the existence of the human at the corresponding position. The points resulting not only from the data used for the module learning but also from un-used data are plotted in fairly accurate positions.

3) Analysis of PB Space: Figure 11 shows the two-dimensional PB space acquired in the first phase, which consisted of pairs of PB values. The numbers in parentheses in the label box represent the initial postures of the object. The PB values corresponding to the initial posture 1, the rhomboid points, were used in the second and third phases.



(a) Motion 1 at 0 deg. (b) Motion 5 at 0 deg.

Fig. 9. Prediction of object trajectory.

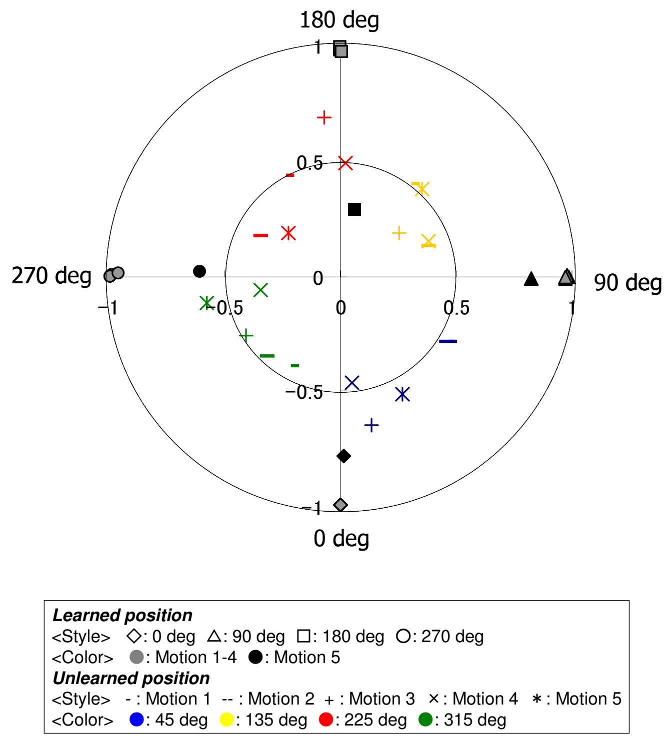


Fig. 10. Map of position likelihood.

The PB values were self-organized corresponding to the categories of object manipulations.

V. Discussion

A. Discovery of Other Individual

As can be seen from Fig. 9, the robot could predict the object motions that the parent generated by reusing its self-model. This confirms that the human was recognized as an individual who can be predicted by the robot's self-model.

Figure 10 indicates the motion used for the learning of the conversion modules was recognized quite accurately. Because even the unlearned motions (Motion 5) and other motions generated by the parent standing at unlearned positions (45 deg., 135 deg., 225 deg., and 315 deg.) were recognized as being at fairly accurate positions, it can be said that the conversion modules were accurately organized

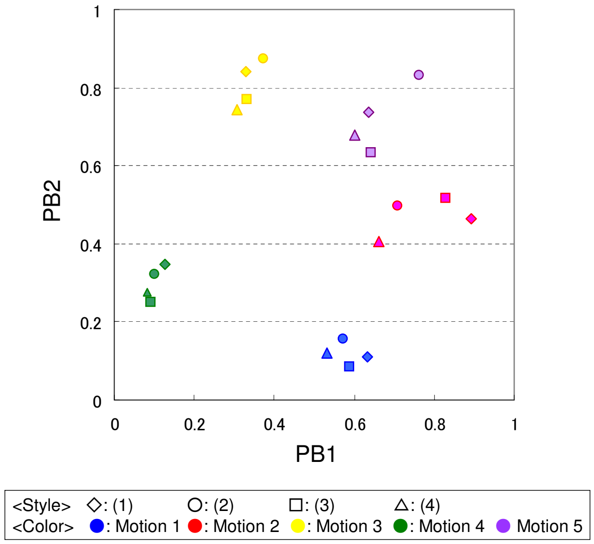


Fig. 11. PB space.

as viewing transformations. This proves the generalization capabilities of the proposed model.

B. Efficiency by Projecting Self-model

Our model enables the robot to conserve its resources. In our experiments, if the robot were to organize another RNNPB for the other-model from scratch, it would need to train 1,610 parameters of the weights. Instead, it could acquire the other-model by training only 64 parameters of the weights for each pair of conversion modules.

C. Relation with Mirror Neuron

The function of the PB values, which mediate the recognition of other's motions and the generation of self-motions, corresponds to that of mirror neurons [10]. The mirror neurons were originally discovered in area F5 of the monkey premotor cortex, which discharge both when the monkey makes a particular action and when it observes another making a similar action. This kind of system might be the basis of the mind-reading of other individuals.

VI. Conclusions

This paper proposed a constructive model which enabled a humanoid robot to discover other individuals by reusing its own self-model. For modeling the development process, following three approaches were employed. (i) Projection of a self-model for predicting other individual's actions. (ii) Mediation by a physical object between self and other individual. (iii) Introduction of infant imitation by parent. The task was moving objects on a table. The RNNPB model, which can learn temporal sequences, was used as the self-model of the robot, and the RNNPB model with hierarchical NN was used as the other-model for the human parent. The robot self-organized the relation between its own arm motions and the object motions in the babbling phase. The NN, conversion modules, for four human parent positions were then learned for predicting the object motions generated by the parent. Finally, the parent's position was estimated from the prediction errors by using the other-model. Both the learned and unlearned parent's positions were estimated fairly accurately both in the case of the prediction of the learned and unlearned motions.

This work is only the first step in modeling the development processes of infants. There still remain a lot of phenomena to be clarified: some of these are the discovery of the correspondence between self and other's body, and interactions by reading the other's mind after acquiring self-consciousness and other-consciousness. Our future work will validate the characteristics of our method in more detail, and evolve the method of interaction. Although the conversion target in this work was visual information only, there is other information to be converted such as contextual information. We believe that the process by which infants acquire imitative capabilities can be modeled by integrating the method proposed in this paper and the imitation method we have proposed [11].

VII. Acknowledgments

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (A) (No. 17680017, 2005-2007), Grant-in-Aid for Exploratory Research (No. 17650051, 2005-2006), and Kayamori Foundation of Informational Science Advancement.

References

[1] M. Asada, K. F. MacDorman, H. Ishiguro, and Y. Kuniyoshi, "Cognitive developmental robotics as a new paradigm for the design of humanoid robots," Robotics and Autonomous Systems, vol. 37, pp. 185-193, 2001.

[2] T. Makino and K. Aihara, "Self-observation Principle for Estimating the Other's Internal State," Mathematical Engineering Technical Reports METR 2003-36, the University of Tokyo, 2003.

[3] C. Trevarthen, P. Hubley, "Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year," Lock, A.(Ed.), Action, gesture, and symbol: The emergence of language, New York: Academic Press, pp.183-229, 1978.

[4] J. Piaget, "Play, Dreams, and Imitation in Childhood," New York: Norton, 1951.

[5] S. Jones, "Infants Learn to Imitate by Being Imitated," Proc. IEEE International Conference on Development and Learning, 2006.

[6] R. Rao, A. Shon, and A. Meltzoff, "A Bayesian Model of Imitation in Infants and Robots," Imitation and Social Learning in Robots, Humans, and Animals, Cambridge University Press, 2005.

[7] J. Tani and M. Ito, "Self-Organization of Behavioral Primitives as Multiple Attractor Motions: A Robot Experiment," IEEE Trans. on SMC Part A, Vol.33, No.4, pp.481-488, 2003.

[8] 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.

[9] H. Ishiguro et al., "Robovie: an Interactive humanoid robot," Int. Journal of Industrial Robotics, Vol.28, No.6, pp.498-503, 2001.

[10] G. Rizzolatti and M.A. Arbib, "Language within our grasp," Trends in Neuroscience, Vol.21, No.5, pp.188-194, 1998.

[11] R. Yokoya, T. Ogata, J. Tani, K. Komatani and H. G. Okuno, "Experience Based Imitation Using RNNPB," Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.3669-3674, 2006.