Proceedings of the 2003 IEEE/RSJ
Intl. Conference on Intelligent Robots and Systems
Las Vegas, Nevada · October 2003

Interactive Learning in Human-Robot Collaboration

Tetsuya OGATA[1)2)], Noritaka MASAGO[2)], Shigeki SUGANO[2)] *and* Jun TANI[1)]

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

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

Abstract

In this paper, we investigated interactive learning between human subjects and robot experimentally, and its essential characteristics are examined using the dynamical systems approach. Our research concentrated on the navigation system of a specially developed humanoid robot called Robovie and seven human subjects whose eyes were covered, making them dependent on the robot for directions. We compared the usual feed-forward neural network (FFNN) without recursive connections and the recurrent neural network (RNN). Although the performances obtained with both the RNN and the FFNN improved in the early stages of learning, as the subject changed the operation by learning on its own, all performances gradually became unstable and failed. Results of a questionnaire given to the subjects confirmed that the FFNN gives better mental impressions, especially from the aspect of operability. When the robot used a consolidation-learning algorithm using the rehearsal outputs of the RNN, the performance improved even when interactive learning continued for a long time. The questionnaire results then also confirmed that the subject's mental impressions of the RNN improved significantly. The dynamical systems analysis of RNNs support these differences.

1. Introduction

Many kinds of the mechanical systems that cooperate with human beings have recently been studied in efforts to improve task performance and the mental impression of the persons using the system. A humanoid robot, for example, will not only have to help people work but also have to establish a new relationship with people in daily life.

We focused on interactive learning between a human operator and a robot system, in a fundamental form to design a natural human-robot collaboration. It consists of the robot system, which learns the task including a human operator, and the human, who learns the task including the robot system. However, it is usually difficult to stabilize the system for a long period of time of operation because the incremental learning of such coupled and nested systems between humans and robots tends to generate quite complex dynamics.

Although there have already been some studies of learning systems in man-machine cooperation [1][2], most of them only focused on short period operations in which the cooperation relation between the person and the machine is organized. Therefore, they did not discuss important aspects such as the mutual interaction after the relation organization, the collapse and modification of the relation, and the long process of development from a beginner to an expert.

Miwa et al. performed an experimental study [3] exploring the collapse and modification of relationships between people, but such phenomena are hard to analyze because human learning and cognitive processes cannot be measured directly. Miyake et al. studied the walking cooperation between a person and a robot model [4], but because of their simple modeling using a nonlinear oscillator, their analysis was limited to some simple phenomena such as the synchronization and revision of the walking rhythm.

To investigate interactive learning for a long time, we developed a navigation task performed by a humanoid robot. This paper describes the results of our experiments and the validity of the "consolidation learning" method implemented to ensure the robustness of neural network output.

2. Navigation Task

A navigation task is employed in which a humanoid robot, Robovie, developed in ATR [5], and a human subject navigate together in a given workspace. The Robovie is a small robot, 1200 mm in height and weighing about 60 kg. It has various features enabling it to interact with human beings: two arms with four degrees of freedom, a humanlike head with audiovisual sensors, and many tactile sensors attached to its body. Photographs of Robovie and the navigation task are shown in Figure 1. The experimental environment used was a 4x4-m L shaped course whose outside walls were marked red and blue for every block (Figure 2). Robovie and the human subject held their arms together and attempted to travel clockwise in the workspace as quickly as possible without hitting obstacles. The actual movement of the robot and the subject is determined by adding two motor forces; one is the motor vector determined from a neural network in the robot and the other is the subject's directional control force exerted to the robot's arms. The neural network in the robot is adapted incrementally after each trial of travel based on the travel performance. The performance is measured by the travel time period at each trial.

An interesting point of this collaboration task is that the sensory information is quite limited for both the robot and

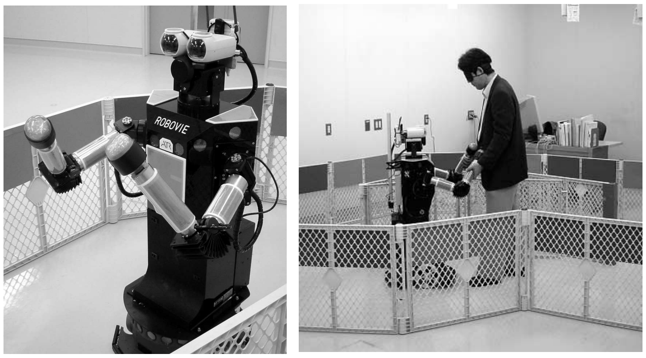

Figure 1 Robovie and Navigation Task

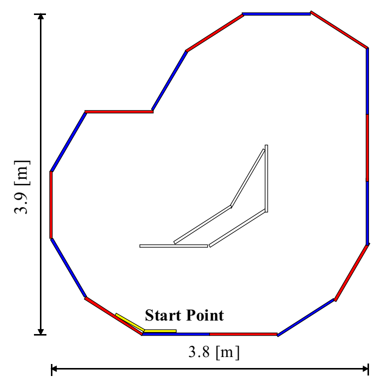

Figure 2 Experiment and Workspace

the subject. The robot can access only local sensory information such as ultrasonic sensors and a poor vision system (it only detects vague color information of its surroundings), but not for exact global position information. The subject's eyes are covered during the navigation task, however the subject is allowed to look around the workspace before the experiments begin. The subject has to guess his/her situation or position by means of the interactive force felt between the robot and his/her arms utilizing his/her memorized image of the workspace geometry. Both sides attempt to acquire the forward model of anticipating the future sensory image as well as the inverse model of generating the next motor commands utilizing the poor sensory information of different modalities from past experiences.

3 The Model and System

3.1 Neural Network Architecture

In many cases the actual states of the systems cannot be identified explicitly just by looking at the current sensory cues, but they do through more context-dependent manners by utilizing the history of the sensory-motor experiences. In our experiment case, the current sensory inputs may not tell the exact position of the robot due to the sensory aliasing problems. This is called the hidden state problem. Long-Ji Lin [6] as well as Tani [7] have shown that the recurrent neural network (RNN) can solve this problem where the so-called context units are self-organized to store the contextual information. We applied the Jordan type recurrent neural network (RNN) [8] in which context units are added to the usual-feed forward neural network (FFNN) .

Figure 3 shows the RNN architecture design of the robot. The RNN operated in a discrete time manner with the synchronizing of each event, and the input layer of the RNN consisted of the current sensory input and the current motor values. The sensory inputs are comprised of the output of the ultrasonic range-sensors and the color area acquired from the omni-direction camera mounted on the robot's back. The motors consist of the current forward velocity and rotation velocity. The input layer has only seven units. The output layer also has seven units, and its outputs are the prediction of the next sensory input and the next action. This is the implementation of the paired forward and inverse model proposed by Kawato [9]. There are forty context units in the input and output layers. The activations of the context outputs in the current time step is copied to those of the context inputs in the next time step. It is noted that the context units activities are self-organized through learning processes such that they can represent the current state of the system corresponding to the past input sequences.

In our application, the RNN is utilized not only as a mapping function from inputs to outputs but also as an autonomous dynamical system. Concretely, the RNN can have two modes of operations as shown in Figure 4. The first mode is the open-loop mode where one-step prediction of the sensory-motor prediction is made using the inputs of the current sensory-motor values. The second mode is the close-loop mode in which the output re-entered the input layer through the feed back connection. By iterating this with the closed loop, the RNN can generate an arbitrary length of the look-

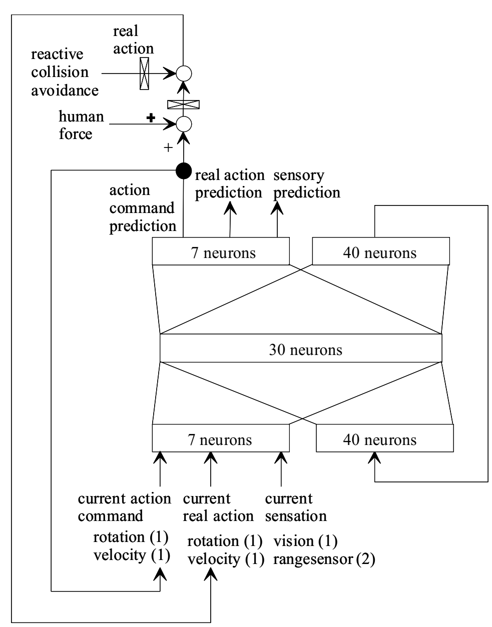

Figure 3 Neural Network Architecture

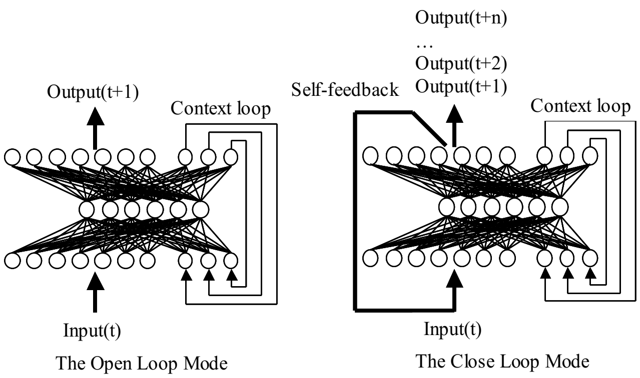

Figure 4 Open loop mode and Close loop mode

ahead prediction for future sequences with given initial states in the input layer. This function for the look-ahead prediction of the sensory-motor sequences can achieve the mental rehearsal [10] that will be described later in the explanations of the consolidation learning. The middle layer had thirty neurons. The RNN was trained by using the back propagation through the time (BPTT) learning method [11].

Here, a pre-experimental result demonstrates how these operations of the open-loop and the close-loop modes work. Figure 5 shows the comparisons between the actual rotation velocity and its prediction in the open-loop and the close-loop modes while the robot travels in the workspace three times after the off-line training of the RNN is completed with 30,000 learning steps.

3.2 Consolidation Learning

It is generally observed that if the RNN attempts to learn a new sequence, the contents of the current memory are severely damaged. One way to avoid this problem is to save all the past teaching data in a database, add new data, and use all the data to retrain the network. The problem with this method, however, is that the learning time of the RNN is increased by increasing the amount of stored data.

Therefore, we used the consolidation-learning algorithm proposed by Tani [10]. Observations in biology show that some animals use the hippocampus for temporary storage of episodic memory and consolidate them into neocortical systems as long-term memory during sleep. Tani modeled this process by using an RNN and a database. In this method the newly obtained sequence pattern is stored in the "hippocampal" database. The RNN, which corresponds to the neocortex rehearses the memory patterns, and these patterns are also saved in the database. The rehearsal can be performed in the close-loop mode described in the previous section. Various sequence patterns can be generated by setting the initial state of the RNN differently. The ensembles of such various rehearsed sequences actually represent the structure of the past memory in the dynamical systems sense. The RNN is trained using both the rehearsed sequential patterns which correspond to the former memory and the current sequence of the new experience.

It is expected that this method enables the RNN to carry out incremental learning while maintaining the structure as much as possible. Although some robot studies using this algorithm have been performed, its detailed characteristics have not yet been clarified.

3.3 Navigation system

Figure 6 shows the navigation system developed in this study. Since the robot moves using the RNN, its performance is inadequate in the initial stage of learning. We therefore implemented a collision avoidance system which overrides the RNN commands when the minimum output of a range sensor falls below the threshold value. This system is just the reflection system tuned up by the designer. In our experiments, the more overrides made by this man-made collision avoidance system means the less performance of the RNN. The robot obtained the color area, range sensor data, and vehicle conditions every 0.1 s. This data was compressed and filtered. The RNN receives this preprocessed data as input and generates the output with a time interval of 2 seconds.

A simplified reinforcement-learning method was employed for the RNN learning as follows. At each trial, the robot and the subject go around the workspace together for a fixed number of times. Then the time period taken for this travel is measured. If the performance in terms of the time period is

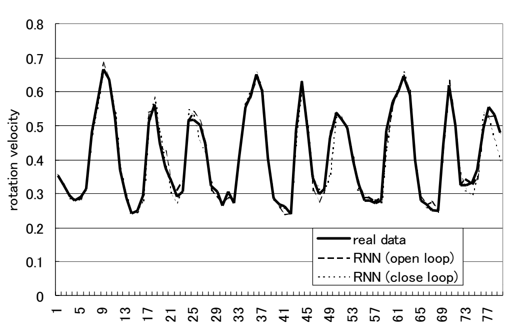

Figure 5 An Example of the result of RNN learning

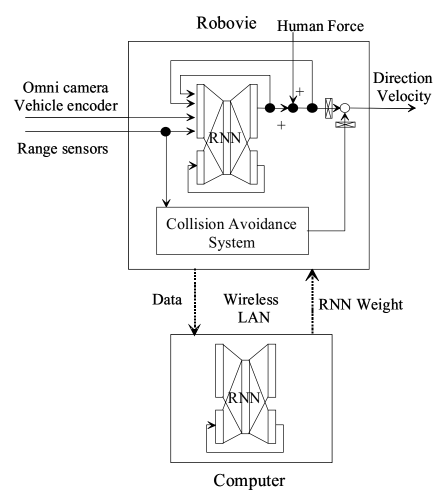

Figure 6 Navigation system diagram

better (less period) than the previous trial, the RNN is trained with the sensory-motor sequence experienced with this trial (with rehearsed ones in the consolidation learning). Otherwise, no training is conducted on the RNN. At each trial, 3,000 steps of iterative learning is conducted off-line using the external computer. In this way, the learning in the robot side is conducted incrementally depending on the performance achieved at each trial.

3.4 Pre-experiments for Comparing RNN and FFNN

To compare the adaptability of the RNN and the FFNN, we carried out experiments using only the robot. The RNN used in this experiment was the same as that shown in Figure 3. In this experiment we trained the RNN using the usual method rather than the consolidation method described in the previous section. The FFNN had no context layer and had a middle layer of 110 neurons. The total number of the neurons of bothneural networks was the same (124 neurons).

Figure 7 shows the results of the 15 trials, in each of which the robot went around the workspace three times. The vertical axis shows the travel time the robot needed for one trial. It was confirmed that the RNN performed significantly better than the FFNN. This result shows that in this task the acquisition of the context information of the environment was effective. It also shows that as long as only the robot learns the environment, instability of the learning process does not cause the learning method to become a problem.

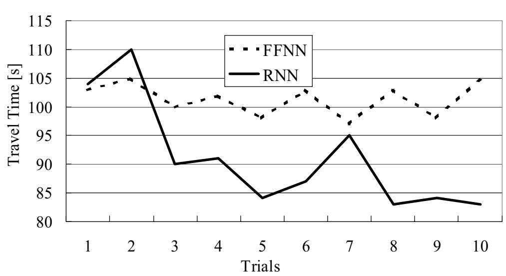

Figure 7 Performance comparison between FFNN and RNN

4 Experiments

The learning algorithms were evaluated and compared in 15-trial navigation experiments with seven male subjects. In each trial, the subject and the robot went around the workspace two times. After each trial there was a one-minute break for the questionnaire, which consisted of 11 items [1]. Additionally, at the end of each experiment, the subjects filled out the questionnaire based on NASA-TLX [12].

Three neural networks, the FFNN, the RNN with a usual learning method, and the RNN with consolidation learning explained in section 3.2 were compared in the experiments involving the seven subjects. In consolidation learning, the teaching data consisted of the current sequence pattern and the three rehearsal patterns. These three rehearsal patterns were generated in the close-loop mode with changes in the initial value of the context units randomly. The order of the experiments was changed with the subjects to avoid presenting subjects with a fixed order that might influence the results.

4.1 Comparison of the Performances

Figure 8 shows the transitions of the travel time of three neural networks. Each travel time is the average of the seven subjects. The goal of this task was to decrease the travel time. Although all performances improved in the first half of the learning, differences appeared in the second half. The performance of the FFNN gradually deteriorated and that of the RNN with the usual learning method stagnated. Only the performance of the RNN with the consolidation-learning algorithm continued to improve.

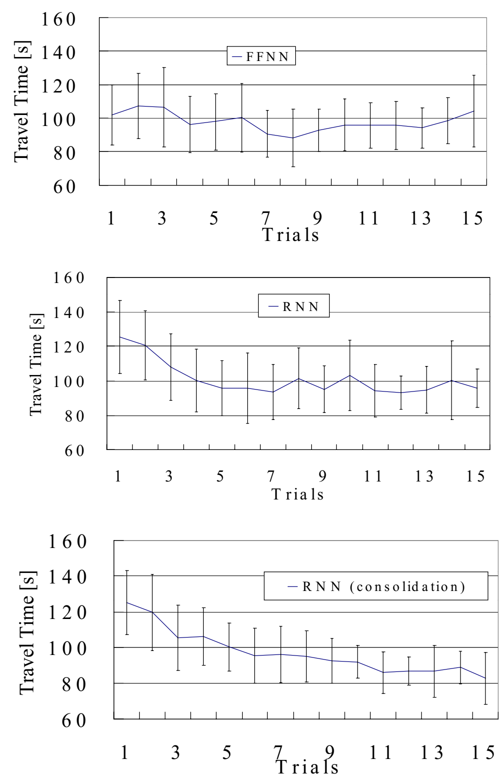

Figure 8 Performance comparison between neural networks

4.2 Mental Impression

The results of the NASA-TLX questionnaire and the 11-item questionnaire are shown in Figures 9 and 10. In each questionnaire, the significant values of the levels of 1 % and 5 % were calculated by a Scheffe test. It is easy to see that in both questionnaires the RNN with the consolidation-learning algorithm gave the best mental impressions.

It is interesting here to compare the results between the FFNN and the RNN with the usual learning method. In the robot experiments described in Section 3, the performance of the RNN was better than that of the FFNN. In the navigation experiments, however, the FFNN tended to give the subjects a better impression than the RNN especially in “result of work”, “fatigue free”, and “operability” in the 11-item ques-

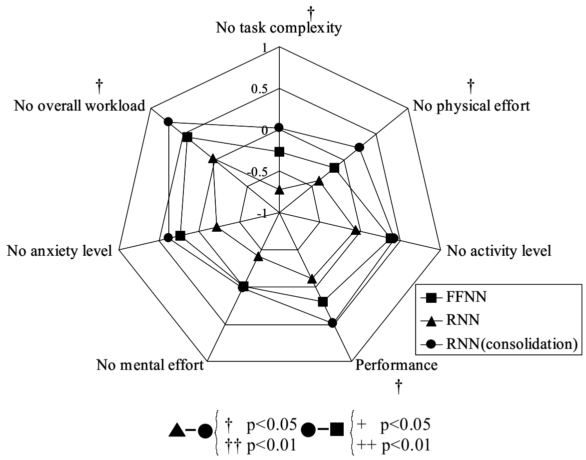

Figure 9 Results of Quesionnaire of NASA-TLX

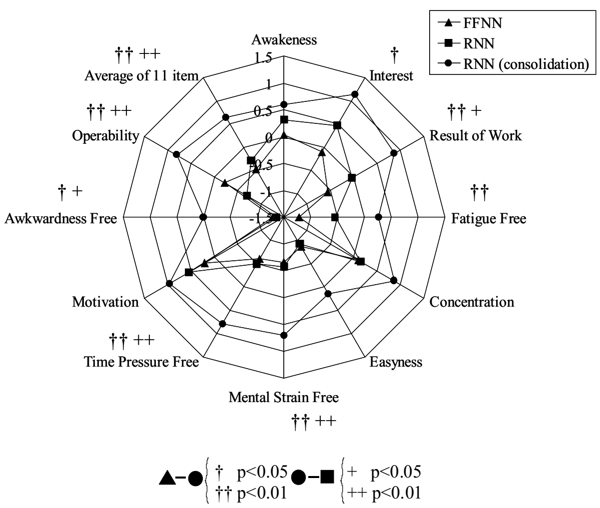

Figure 10 Results of Quesionnaire of 11-item Evaluation

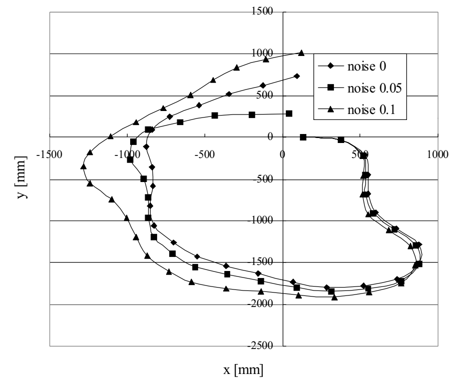

Figure 11 Examples of Rehearsed Trajectory with Input Noises
(The RNN with usual-learning method)

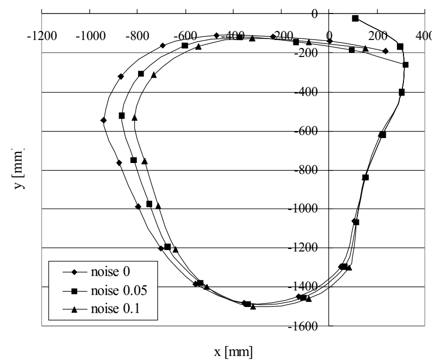

Figure 12 Examples of Rehearsed Trajectory with Input Noises
(The RNN with consolidation-learning method)

tionnaire.

5. Discussion

In the experiment with only the robot, the RNN performed effectively, because the robot could decide the action using not only the sensor input including the noises but also the information of the context layer. In the human-robot cooperation, however, the performance of the RNN with the usual learning was worse than that of the FFNN. It is thought that this is due to interactive learning including "incremental learning" which damaged the memory of the RNN. As the result, the "operability" became worse because the robustness of the RNN to the input noise decreased.

To analyze the effect of consolidation learning, we examined the robustness of the RNN dynamics by looking at its initial sensitivity characteristics. Both RNNs obtained after the usual learning and consolidation learning in our experiments were tested to generate the output sequences in the close-loop mode with the addition of three different sizes of noise in the initial input values. Figures 11 and 12 shows the motion trajectories of the robot re-constructed from the rehearsal motor output of the RNN. These represent how the output trajectories developed with small differences in the initial input conditions for the RNNs with usual learning and consolidation learning. It is observed that output trajectories of the RNN by the usual learning tends to diverge more than those by consolidation learning. This implies that the usual learning scheme tends to generate more unstable dynamic structures in the trained RNN.

The following analysis was carried out to investigate this property in detail. A sequence corresponding to one trial (about 100 steps) was rehearsed in the close-loop mode by the RNN. In this process, the random noises were added to seven units in the input layer for all steps. Ten sequences were rehearsed by this method. Figure 13 shows the step error obtained from the difference between the average of these ten sequences and the other sequence which was also rehearsed by the same RNN without noise. The horizontal axis is the maximum width of the noise at each step, and the vertical axis is the average error of seven RNNs corresponding to the subjects. This result also shows that the RNN with a consolidation-learning algorithm is more robust than that

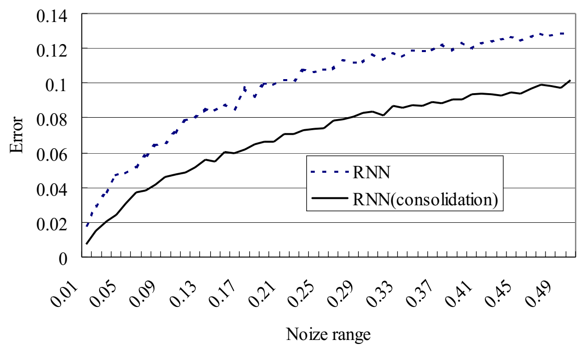

Figure 13 Robustness Comparison of RNNs

with the usual learning algorithm.

This robustness characteristic of the RNN seems to be directly related to the "operability" in the mental impressions. If the RNN tends to diverge largely even with small deviations in the input sequences from the learned ones in the past, it would be difficult for the subjects to harness the robot to keep it in the right directions. In the usual RNN learning, if the contents of the current memory conflict with the one to be newly learned, the internal structure of the RNN could be deflected severely and undesired pseudo memories could be generated. Consolidation learning is beneficial in this aspect since this scheme allows iterative rehearsing of past experiences. Also, training with enough number of rehearsed sequences could achieve a sort of generalization while attaining the global structures in the internal representation.

6. Conclusion

In this paper, we showed that interactive learning is essential and important for man-machine cooperation. We also pointed out that it is difficult to actualize context dependence learning. The RNN was introduced as a learning algorithm system which can treat the hidden state problem. The target task was the human navigation by a humanoid robot called Robovie. The FFNN and the RNN were compared as the learning algorithm of Robovie. Although the performances of the RNN and the FFNN both showed improvements in the early stages of learning, they gradually became unstable as the subject changed the operation by learning on its own. Finally, all performances failed. The results obtained when the consolidation-learning algorithm, which uses the rehearsal outputs of RNN, was applied confirmed that this algorithm's performance was improved even when interactive learning continues for a long time and that the subject's mental impressions were better. The analyzing and comparing the characteristics of RNNs produced results which support these differences.

Two further studies should be carried out. One should be a more detailed analysis of the characteristics of the consolidation-learning algorithm. Although the robustness of the RNN have been compared in this paper, the relation between the dynamic structure of the RNN and the learning algorithm have not be analyzed mathematically yet. The other study that should be carried out is a transition structure analysis of the cooperation form between a human and a machine in the interactive learning process. Although we showed that the RNN with the consolidation-learning algorithm could continue to improve the performance, the adaptation phenomenon that might be the changing process of the cooperation form has not been analyzed in detail yet. The correspondence between the transition of the RNN structure and the development of human operation should be investigated.

References

[1] Y. Hayakawa, I. Kitagishi, Y. Kira, K. Satake, T. Ogata, and S. Sugano, "An Assembling Support System based on a Human Model -Provision of Physical Support According to Implicit Desire for Support," Journal of Robotics and Mechatronics, Vol.12, No.2, pp.118-125, 2000.

[2] T. Sawaragi, T. Kudoh, and S. Ozawa, "Extracting Motion Skills from Expert's Proficient Operation Records Using Recurrent Neural Network," Reprints of 14th World Congress of IFAC, Beijing, Vol.M, pp.359-364, 1999.

[3] Y. Miwa, S. Wesugi, C. Ishibiki, and S. Itai, "Embodied interface for emergence and co-share of 'Ba', Usability Evaluation and Interface Design." in Proc. of HCI International 2001, pp. 248-252, 2001.

[4] Y. Miyake, and T.Minagawa, "Internal observation and co-generative interface," in Proc. of IEEE International Conference on Systems, Man, and Cybernetics (SMC'99), I229/I-237, 1999.

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

[6] L. Lin, and T. Mitchell, "Efficient Learning and Planning within the Dyna Framework", Proc. of the Second International Conference on Simulation of Adaptive Behavior (SAB '92), pp.281-290, 1992.

[7] J. Tani, "Model-based Learning for Mobile Robot Navigation from the Dynamical Systems Perspective", IEEE Trans. on System, Man and Cybernetics Part B (Special Issue on Robot Learning), Vol.26, No.3, pp.421-436, 1996.

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

[9] D. Wolpert, and M. Kawato, "Multiple paired forward and inverse models for motor control". Neural Networks Vol. 11, pp.1317-1329.

[10] J. Tani, "An Interpretation of the `Self` from the Dynamical Systems Perspective: A Constructivist Approach." Journal of Consciousness Studies, Vol.5 No.5-6, 1998.

[11] 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.

[12] 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.