Proceedings of the 2011 IEEE
International Conference on Robotics and Biomimetics
December 7-11, 2011, Phuket, Thailand

Operational Feedback Considering Social Contingency for Robot Teleoperation

Toshimitsu Takahashi, Masahiko Morita, and Fumihide Tanaka

***Abstract*— In telerobotics systems for children, it is crucial that the teleoperation interface is intuitive and well-controllable in terms of communication delay so that the children do not become bored quickly. To this end, we consider an operational feedback design for the teleoperation interface that incorporates the idea of *social contingency detection*, which is borrowed from a developmental psychology literature. Because the idea itself is general and there has been no attempt to implement it in telerobotics systems to date, in this study, which aims at a better operational feedback design, we test some representative implementations in human-subject experiments and report the results.**

I. Introduction

One of the recent trends in telerobotics research is its application for social interaction purposes. A series of telexistence or telepresence robots [1], [2], [3], [4], [5], [6], [7] have been designed with the aim of offering distance communication among humans. Particularly we have strong interests in the application of telerobotics for early childhood education. Between 2004 and 2007, we conducted a field study about robotics supporting early childhood education (18-24 months old) by immersing ourselves into the daily classroom environment in Early Childhood Education Center at the University of California, San Diego [8], [9]. We were aware that remote communication applications such as Skype had already been used in this domain, and thus, it seems quite likely that people are waiting for the next step in development. However, we are of the opinion that there are a lot of practical issues that still remain unresolved, which might prevent this potentially popular technology from becoming a reality. To investigate, identify, and resolve these issues, we considered that conducting field trials here again would be the best approach.

In 2009, as part of the Japan Science and Technology Agency's (JST) PRESTO program [10], we started a research project with the ultimate goal of developing and operating a telerobotics infrastructure to link classrooms in Japan and other countries in real time [11]. We involved people working in and around classrooms in the project: nursery and elementary schools, companies that are in the education business, and of course, children and parents. This is important because we would not only like to demonstrate the results of our research but also develop a feasible model of the proposed system. The philosophy of designing robots by immersion (immersing ourselves into the target environment) is inspired by the RUBI project [12].

There are a lot of researches related to remote applications in the area of HCI [13], [14]. In a remote application, communication delay is one of the typical problems that inhibit an intuitive operation [15]. One significant issue we have already been facing is a delay in communications on the Internet. Preliminary investigation of packet transactions between Tsukuba (Japan) and San Diego (United States) showed that in addition to a regular delay in the packet transmission of more than 500 ms, there were issues such as the loss of packets and unpredictable fluctuations in the delay time. Furthermore, in our case, young children are the users of the system, and thus, it cannot be assumed that we would be able to give them detailed instruction on how to use the system. Children easily get bored with a system if they do not have a sense of mastering the system within their first few attempts to do so. In particular, we learned from our pilot trial [11] that the most critical interaction phase is the first 1-2 minutes in which a user starts using the system. If the user can have a sense of mastering the system (understanding the delay) within the initial phase, the user will continue using the system. Our tests showed that most children are able to persevere with the system even under conditions of a slight delay if the actuator being used is simple, such as a gripper.

In concrete terms, this means that it is crucial for a remote-controlling interface to be equipped with a function through which users can quickly recognize the delay property, and thus can have a sense of mastering how it works without being given verbal instruction. To this end, we investigate the social contingency detection of infants which is a well-studied human property in developmental psychology [16], [17], [18], [19], and apply the theory in the design for operational feedback of a remote-controlling device. In the next section, we explain the theory and how it is applied. Then, we report the results of experiments using human subjects, the goal of which is to gather knowledge on the impression the subjects had of different types of operational feedback incorporating social contingency. Because this was the first trial in which we test the theory behind our system, we decided to test it on adult subjects, who can later be surveyed on their impressions by means of a questionnaire. Based on the knowledge obtained from the study, we will go back into the field in the near future and conduct field trials using child subjects.

T. Takahashi is with the Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan taka@bcl.esys.tsukuba.ac.jp
M. Morita is with the Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan mor@bcl.esys.tsukuba.ac.jp
F. Tanaka is with the Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan fumihide@iit.tsukuba.ac.jp

978-1-4577-2137-3/11/$26.00

II. Social Contingency Detection and Operational Feedback

Watson [16] proposed that contingency detection plays a crucial role in the social and emotional development of infants, and it is also a fundamental source of information for the definition and recognition of caregivers. In fact, it has been known that even infants can detect contingency between their actions and an external event [16], [19]. For instance, if a sound occurs shortly after their actions, such as babbling, they instantly detect its contingency and pay attention to the sound source.

Operational feedback is commonly used to increase the usability/intuitiveness of a remote-controlling interface of a robotic system. For example, in the case of controlling a robot hand, haptic feedback is often given to the operator when the operator grasps an object using the robot hand [20]. However, with an unpredictable communication delay of more than 500 ms, an operator who had not been given instructions in advance as to how the robotic system works under conditions of delay, did not wait for the delayed robot movement, and therefore was easily confused. We observed this many times during our pilot trials [11].

To overcome this problem, here we consider ways to improve and enhance operational feedback based on the knowledge of social contingency detection. The basic idea is to introduce another feedback at a timing between the operator's action (e.g., clenching a hand) and the delayed robot movement (e.g., a remote-gripper closes) so that the operator can detect contingency between them (Fig. 1). Here there are some interesting properties we can exploit based on the knowledge of social contingency detection in developmental psychology. First, humans are known to be able to detect a contingency between different sensory modalities such as haptic feedback and audio feedback. Second, humans are known to be able to detect contingency across different spatial locations. For example, even if the source location of audio feedback is random and unknown to the operator (e.g., from behind), the operator will soon be able to detect its contingency.

Thanks to this generality in the theory of social contingency, there is a lot of flexibility in designing an operational feedback system that incorporates it. Therefore, in this paper, we implement several typical instances and conduct comparative experiments using human subjects, which are explained in the following sections.

III. Overview of our Teleoperational System

From our observations of classroom activities, we have found that most educational activities in children's classrooms include grasping behavior on the part of children. For example, flashcards are the most frequently used educational material in classrooms. Teachers use them in many ways in vocabulary learning lessons, but these lessons always contain a situation in which children grasp these flashcards (and read them, or hand them to others, etc.) so that the children do not get bored with lessons. In the case of lessons for children, it is very important to give them materials in a way so that they continue to concentrate and do not get bored too soon. Grasping is an important behavioral element for this purpose, and one which is commonly seen in classroom activities. Therefore, we developed a simple remote-hand controlling device with haptic and audio feedback (Fig. 2, Left), which can communicate with a PC through Bluetooth. It is equipped with a bend sensor (which slips over the finger), vibration motor (on a fingertip), electronic buzzer (worn on the wrist), and a thin rechargeable Li-polymer battery that lasts for more than 3 hours. We can replace the glove part to suit the operator's hand size. In this paper, we used one bend sensor and one vibration motor on the middle finger, but we can add sensors and motors to each finger for this device.

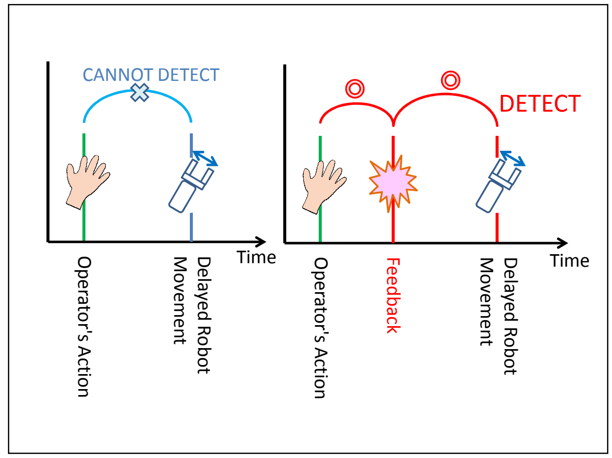

Fig. 1. In the left side of the diagram, the operator cannot detect the delayed robot movement. The right side shows a conceptual diagram of operational feedback incorporating social contingency. By means of hand movement, the operator obtains haptic—audio feedback prior to the robot starting to move.

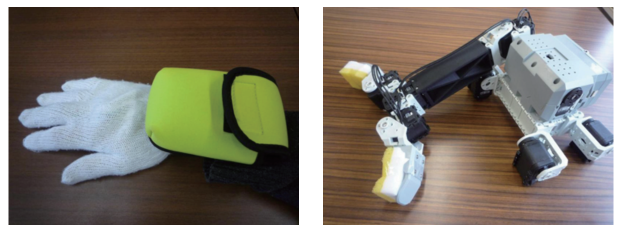

Fig. 2. (Left) Remote-hand controlling device equipped with a bend sensor, vibration motor, and buzzer circuit board. (Right) Robotic gripper consisting of Dynamixel servo motors.

In the experiments described in the next section, each subject is asked to remote control a robotic gripper (Fig. 2, Right) by using the controlling device explained in the previous paragraph and watching a TV monitor on which the subject can see the robotic gripper, which is sited in a remote classroom (Fig. 3). The video data stream is coded/decoded using a LifeSize HD video conference system.

IV. Experiments

The goal of the experiments described here is to test various representative implementations of operational feedback incorporating social contingency and collect knowledge

Fig. 3. Experimental environment: Subjects can only see the robotic gripper on the monitor. Later they answer questionnaires using a PC.

about their usability. Specifically, we conducted three experiments, each of which focuses on one particular aspect of the social contingency:

- Experiment-1: Timing of operational feedback
- Experiment-2: Source location of operational feedback
- Experiment-3: Combination of multiple feedbacks

Before the three experiments, we also conducted preliminary experiments to finalize their experimental protocols as well as the post-experiment questions. The general procedure for the experiments is as follows. A subject enters the test room, sits down in front of a TV monitor, and receives an overall description of the experiment. Then, the experiment starts and the subject is asked to grasp a bottle using a robot hand when an experimenter puts it in front of the robot hand. A delay of 2-3 seconds (uniformly random) is set between the subject's grasping motion and the actual movement of the robot hand. Each session lasts 30 seconds. As soon as each session finishes, the subject is asked to answer three questions (10-point scores each) as detailed below:

- Question-1: Could you control the robot hand as you wanted?
- Question-2: How much stress did you feel in controlling the robot hand?
- Question-3: How many unexpected robot behaviors were there during the trial?

Subjects chose a number and press an appropriate key on the PC.

In all three experiments, the subjects are faced with three different conditions in random order. Each condition is presented 10 times during each experiment, thus, each subject experiences a total of 30 sessions. Each subject can only participate in a single experimental session. When all the sessions have ended, the subject is asked to answer the questionnaire.

A. Experiment-1

The first experiment investigated the timing factor of social contingency during operational feedback. Here the following three conditions were tested:

- Condition-1a: No feedback signal (No FB)
- Condition-1b: Audio feedback when the subject clenches their hand (Same Time)
- Condition-1c: Audio feedback at 1.25 s after the subject clenches their hand (1.25 s)

In conditions-1b and 1c, audio feedback comes from the remote-hand controlling device.

The results for five subjects are shown in Fig. 4. One sample t-test showed that there was a statistically significant difference between Condition-1a (No FB) and Condition-1b (Same Time) in Question-1: $t(49) = -3.7$, $p < 0.01$. Also, post-test questionnaires were consistent with the results: all the subjects said Condition-1b (Same Time) was the best. However, there were no significant differences in any of the other questions.

One reason for this result is that the nature of the task (to grasp an object) made Condition-1b easy to understand how to control the system. Since the subjects only controlled the robotic gripper in this experiment, it was easy for them to understand the operational feedback when clenching a hand meant that the signal was transmitted to the robotic gripper.

B. Experiment-2

The second experiment investigated the spatial factor of social contingency during operational feedback. Specifically, we controlled the source location of an audio signal as follows:

- Condition-2a: No feedback signal (No FB)
- Condition-2b: Audio feedback comes from the remote-hand controlling device (Hand)
- Condition-2c: Audio feedback comes from the TV monitor (TV)

The results from another set of five subjects are shown in Fig. 5. There were statistically significant differences between Condition-2a (No FB) and Condition-2b (Hand) in Question-1 ($t(49) = -3.1$, $p < 0.01$) and Question-2 ($t(49) = 6.3$, $p < 0.01$). There were also statistically significant differences between Condition-2b (Hand) and Condition-2c (TV) in Question-1 ($t(49) = 2.4$, $p < 0.05$) and Question-2 ($t(49) = -2.3$, $p < 0.05$), and between Condition-2a (No FB) and Condition-2c (TV) in Question-2 ($t(49) = 3.9$, $p < 0.01$). Overall, the subjects seemed to feel that Condition-2b (Hand) was the most comfortable.

However, when filling the post-test questionnaires, many subjects could not distinguish where the audio feedback came from during Condition-2c (TV) because we positioned the buzzer (the audio feedback sound source) behind the TV monitor and the sound reflected off the walls, making it difficult to identify the source of the sound. Therefore, the subjects seemed to prefer Condition-2b (Hand).

C. Experiment-3

The results of Experiments-1 and 2 suggest that in case of solo operational feedback, the subjects prefer contingency with their actions both in timing and source location. Then, what if multiple operational feedbacks are introduced? We hypothesized that multiple operational feedbacks, each of

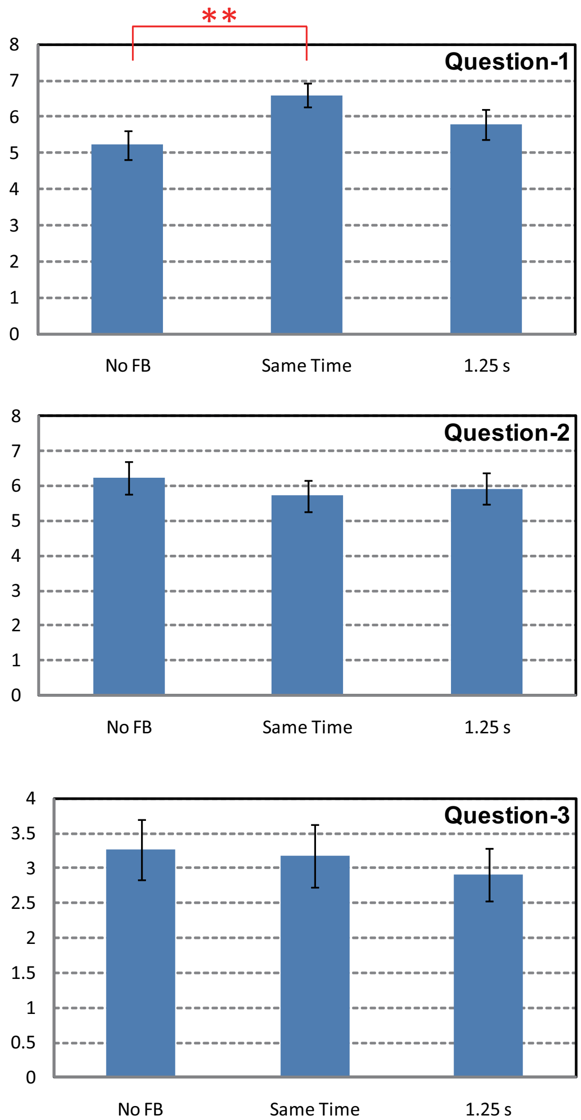

Fig. 4. Results of Experiment-1: Average scores from five subjects. In Question-1, higher scores mean better values. In Questions-2 and 3, lower scores mean better values. Error bars show standard errors. Double asterisks indicate a significance level: $p < 0.01$.

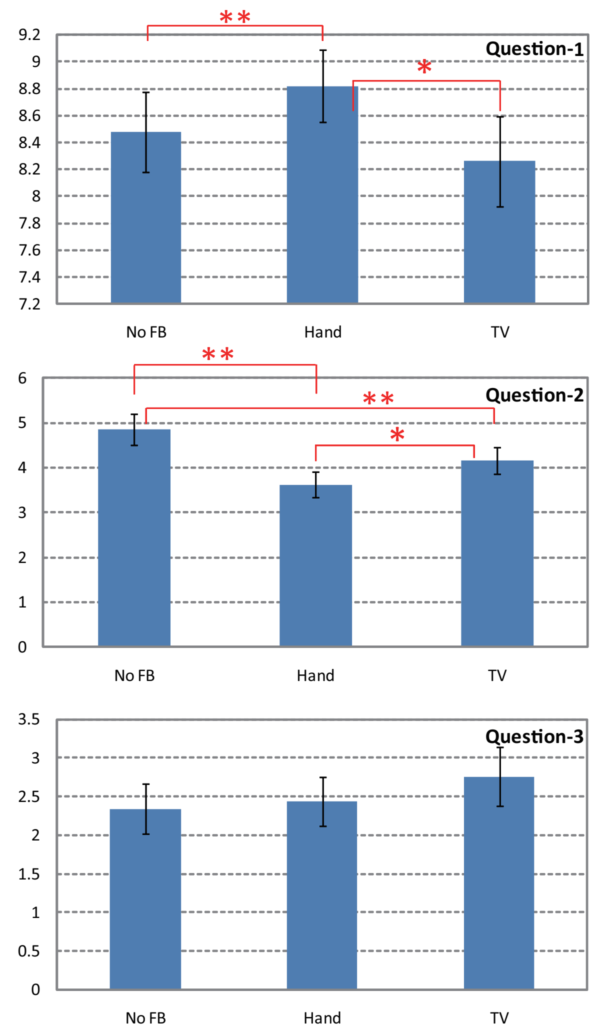

Fig. 5. Results of Experiment-2: Average scores from five subjects. In Question-1, higher scores mean better values. In Questions-2 and 3, lower scores mean better values. Error bars show standard errors. Double asterisks indicate a significance level: $p < 0.01$. Single asterisk denotes the significance level: $p < 0.05$.

which has a different contingency property, could be concatenated, which would bridge the delay period and thus offer more usability to the operator. The following three conditions were tested to investigate the hypothesis.

- Condition-3a: No feedback signal (No FB)
- Condition-3b: Haptic feedback when a subject clenches their hand (Haptic)
- Condition-3c: Condition-3b plus audio feedback 1.25 s after the subject clenches their hand (Haptic—Audio)

The results for five subjects are shown in Fig. 6. There were statistically significant differences between Condition-3a (No FB) and Condition-3c (Haptic-Audio) in Question-1 ($t(49) = -2.4$, $p < 0.01$) and Question-3 ($t(49) = 2.1$, $p < 0.05$). There was also a statistically significant difference between Condition-3b (Haptic) and Condition-3c (Haptic-Audio) in Question-1 ($t(49) = -1.9$, $p < 0.05$, one-side).

V. Discussion

Through Experiments-1 and 2, we investigated the spatiotemporal property of operational feedback incorporating social contingency. Overall, the results suggest that subjects prefer it at a near time and location from their own action. One reason for this could be due to the nature of the task, grasping an object. In the case of more complex operational tasks, subjects might need more information to self-understand how to control the system, and thus, the effect of introducing operational feedback incorporating social contingency could be made clearer. In fact, during our experiments, we observed that most subjects understood the delay effect by themselves by the middle of the session. Also, in terms of the speed of their understanding, it seems that

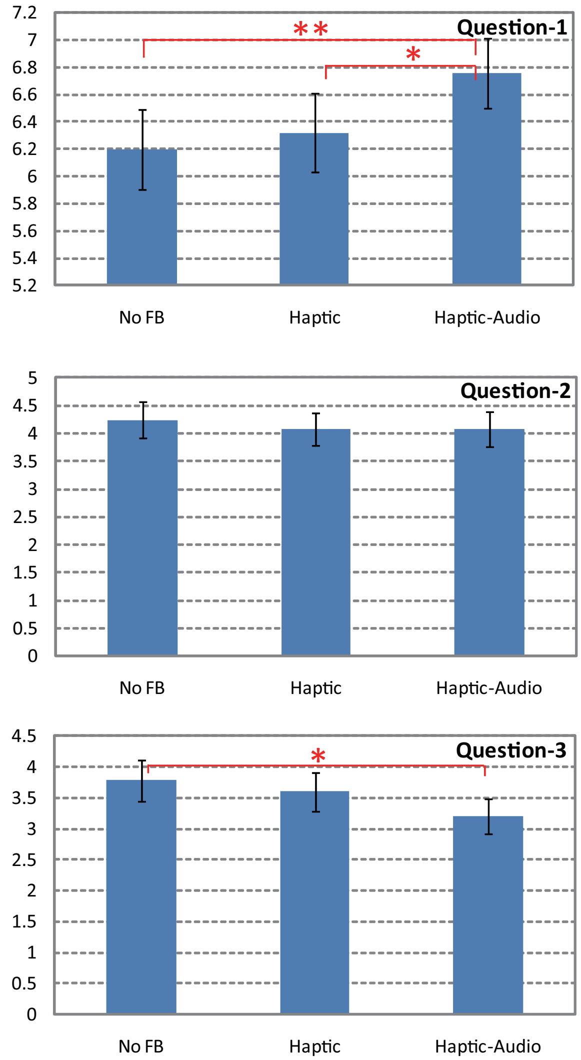

Fig. 6. Results of Experiment-3: Average scores from five subjects. In Question-1, higher scores mean better values. In Questions-2 and 3, lower scores mean better values. Error bars show standard errors. Double asterisks indicate a significance level: $p < 0.01$. Single asterisk denotes a significance level: $p < 0.05$.

the introduction of operational feedback incorporating social contingency clearly helped accelerate the speed.

The results of Experiment-3 are particularly encouraging because they showed that the operational feedback proposed here has the effect of enhancing conventional haptic feedback. Interestingly, four of the five post-test questionnaires cited the operators' preference for Condition-3b, which is inconsistent with the results obtained from questions asked during the experimental trial. This implies that haptic feedback is more vivid and easily memorized, although introducing another audio feedback actually enhances and improves usability even if it does not surface in the conscious mind of the subjects.

Haptic feedback was better than audio feedback because the nature of the task (to grasp an object) was advantageous to haptic feedback. People more easily understand that the glove vibrates when they want to grab something and clench their hand rather than listen to a sound. On the other hand, some of the subjects said that haptic feedback interfered with their movements because they felt that the feedback meant that some sort of error had occurred. Audio feedback has the same possibility for misunderstanding, but in this task, haptic feedback is more influential. Although we assemble a vibration motor on a fingertip in those experiments, some of the subjects said that it may be better that the vibration motor is placed on the palm of the hand. In the future, we try to make such devices and investigate its operability.

VI. Conclusion

The paper reported the results from human-subject experiments that investigated the effect of operational feedback incorporating social contingency. Three experiments were designed and conducted to reveal the spatiotemporal property of the operational feedback, as well as the effect of combining multiple operational feedbacks (haptic and audio). The knowledge obtained from the experiments is expected to be useful in designing a teleoperation interface subject to a communication delay.

VII. Acknowledgments

This work was supported by the JST PRESTO program. It was also supported in part by the Grant-in-Aid for the Global COE Program on "Cybernics: fusion of human, machine, and information systems" at the University of Tsukuba.

References

[1] S. Tachi, "Telecommunication, Teleimmersion and Telexistence," IOS Press, 2003.

[2] S. Tachi, "Telecommunication, Teleimmersion and Telexistence II," IOS Press, 2005.

[3] H. Ishiguro, "Android science: Conscious and subconscious recognition," *Connection Science*, Vol.18 No.4 p.319-332, 2006.

[4] S. Adalgeirsson and C. Breazeal, "Mebot: A robotic platform for socially embodied telepresence," *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*, p.15-22, 2010.

[5] Anybots, Inc.
http://www.anybots.com

[6] InTouch Health.
http://www.intouchhealth.com/index.html

[7] Willow Garage.
http://www.willowgarage.com/pages/texai/overview

[8] F. Tanaka, A. Cicourel, and J.R. Movellan, "Socialization between toddlers and robots at an early childhood education center," *Proceedings of the National Academy of Sciences of the U.S.A. (PNAS)*, Vol.104 No.46 p.17954-17958, 2007.

[9] F. Tanaka, J.R. Movellan, B. Fortenberry, and K. Aisaka, "Daily HRI evaluation at a classroom environment: reports from dance interaction experiments," *Proceedings of the 1st Annual Conference on Human-Robot Interaction*, p.3-9, Salt Lake City U.S.A., 2006.

[10] JST PRESTO.
http://www.human.jst.go.jp/en/index.html

[11] F. Tanaka and T. Takahashi, "Linking children by telerobotics: experimental field and the first target," *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*, p.267-268, Lausanne Switzerland, 2011.

[12] J.R. Movellan, F. Tanaka, B. Fortenberry, and K. Aisaka, "The RUBI/QRIO project: origins, principles, and first steps," *Proceedings of 4th IEEE International Conference on Development and Learning*, p.80-86, Osaka Japan, 2005.

[13] Y. Irie, S. Aoyagi, T. Takada, K. Hirata, K. Kaji, S. Katagiri, and M. Ohsaki, "Development of assistant system for ensemble in t-room," *IPSJ SIG Technical Report: Groupware and Network Service*, Vol.2009-GN-73 No.23 p.1-8, 2009.

[14] J. Fayolle, C. Gravier, and B. Jailly, ”Generic HCI and group awareness collaboration for remote laboratory in virtual world,” *International Journal of Education and Information Technologies*, Vol.5 No.1, 2011.
[15] H. Kuzuoka, G. Ishimoda, Y. Nishimura, K. Kondo, and R. Suzuki, ”GestureCam system: camera robot mediated remote education,” *Bulletin of the National Institute of Multimedia Education*, No.12, 1995.
[16] J. Watson, ”The perception of contingency as a determinant of social responsiveness,” *In E.B. Thoman (eds), Origins of the Infant’s Social Responsiveness*, p.33-64, 1979.
[17] A.E. Bigelow, ”Infants’ sensitivity to familiar imperfect contigencies in social interaction,” *Infant Behavior and Development*, Vol.21 Issue.1 p.149-162, 1998.
[18] M. Okanda, and S. Itakura, ”Young infants’ sensitivity to social contingency from mother and stranger: developmental changes,” *Proceedings of 2005 4th IEEE International Conference on Development and Learning*, p.165-165, Osaka Japan, 2005.
[19] M. Miyazaki and K. Hiraki, ”Delayed intermodal contingency affects young children’s recognition their current self,” *Child Development*, Vol.77 p.736-750, 2006.
[20] A.M. Okamura, ”Methods for haptic feedback in teleoperated robot-assisted surgery,” *Industrial Robot: An International Journal*, Vol.31, p.499-508, 2004.