"Lifelike" behavior of communication robots based on developmental psychology findings

Fumitaka Yamaoka[1,2], Takayuki Kanda[1], Hiroshi Ishiguro[1,2], and Norihiro Hagita[1]

[1]*ATR Intelligent Robotics and Communication Laboratories*
2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0228, Japan
[2]*Osaka University, Osaka, Japan*
2-1 Yamada-oka Suita Osaka 565-0871 Japan
E-mail:{yamaoka, kanda, ishiguro, hagita}@atr.jp

Abstract: **Researchers in developmental psychology have reported various findings about animate-inanimate distinctions, especially the decision criteria of human infants categorizing animate existence, such as distinguishing animals and humans from the inanimate. The purpose of our research is to develop animate or "lifelike" behavior for humanoid robots based on such categorizing, which would be expected to potentially make human-robot interaction more natural. Our approach is unique because we focus on the robot's mechanism and behavior design using environmental sensors. Consequently, we use a motion capturing system, which enables us to leapfrog technological developments for the robots' sensing abilities. The robot that we have developed reacts to human behavior, and it is implemented in order to satisfy the features of animate existence according to developmental psychology. Two experiments were conducted to verify its effects. The results of the first experiment verified the effectiveness of lifelike behavior and also showed that random behavior, which partly satisfies the features of animate existence, received a halfway evaluation for lifelikeness. The second experiment demonstrated the effectiveness of the developed robot for investigating the effect of a robot's cognitive ability for human-robot interaction.**

Index Terms: Human-robot interaction, communication robot, lifelike behavior, developmental psychology, motion capturing system

I. INTRODUCTION

Over the past several years, various interactive robots have been developed. We believe that in the near future robots will participate in our daily lives and interact with humans. However, before that, robots must acquire the ability to interact as naturally with humans as humans do.

In human-robot interaction, it is important to naturally approach humans as communication partners. For example, Arita et al. investigated whether infants identified a humanoid robot as a communicative agent. Their findings suggested that infants interpret a non-interactive robot as an object, but an interactive robot as a communicative agent [1]. Since our definition "interactive" object involves being "lifelike", such behavior in robots is necessary in interaction. In general, humans do not attempt to communicate with inanimate things because we unconsciously decide that an inanimate cannot be a communication partner. Moreover, the field of developmental psychology has shown that infants acquire the ability to distinguish animate from inanimate at an early stage [2]. Therefore, "lifelike" behavior is essential for robots that are to engage in natural communication.

Until now, several robots that try to imitate animals have been developed, for instance, AIBO [3], PARO [4], and Nakata's robot [5]. However, their "lifelikeness" has yet to be improved. According to Rakison and Poulin-Dubois, an infant distinguishes the animate from the inanimate based on seven characteristics [2], which we describe below. Based on this finding, we argue that the above robots lack several properties because of low cognitive abilities. While the same study states that it is important for the animate to react without making contact, those robots react to human behavior based mainly on touch. In addition, the study reflects that their range of reactions is too narrow (absence of contingent behavior) to compare them with living things. Consequently, those robots have not yet reached a "lifelike" state.

Our research approach is also unique in terms of methodology, which bypasses the difficulties of developing the cognitive ability of a robot. There are mainly two research directions for developing communication robots. One is to develop the cognitive abilities of the robot, with respect to visual, auditory, tactile, and any other sensor information. The other is to design robot behavior in order to make human-robot communication more natural. We believe that these two research directions are separable. However, until now the latter research has been waiting for the improvement of the former one. We take the latter approach by using a motion capturing system as a robot's visual sensor. We believe that this approach enables us to acquire important findings for the development of communication robots.

This paper reports the development of an interactive humanoid robot that autonomously performs "lifelike" behavior in human-robot interaction. As mentioned above, existing robots lack several fundamental properties because of low cognitive abilities. We solved this problem by utilizing a motion capturing system. Our robot can appropriately react to human behavior because it can accurately perceive human body movements. It selects an appropriate behavior from a set of several implemented ones that fit various situations. We conducted experiments with subjects to investigate the effect of the developed mechanism.

In this paper, although we limit our focus to "lifelike" behaviors and have not yet extended it to human-like communication, we believe that lifelike behavior with a human-like body is the basis of communication. This behavior includes facial
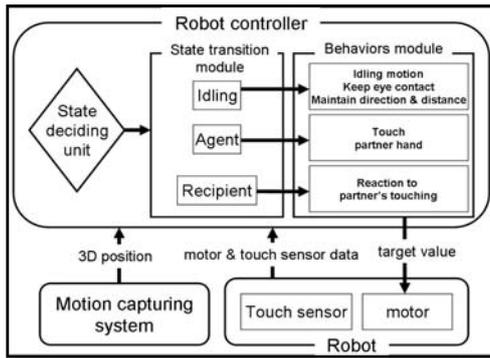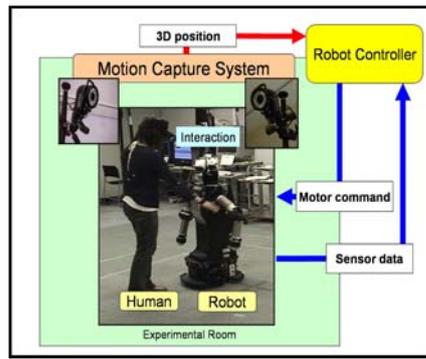
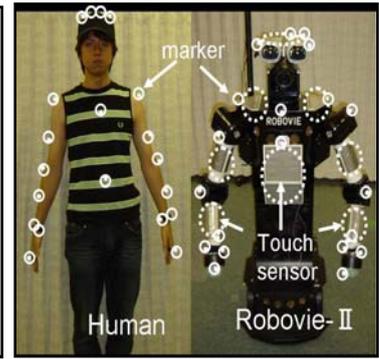Fig. 1 System configuration   Fig. 2 Overview of system   Fig. 3 Human and Robot

Table 1 Animate - Inanimate Distinction Characteristics

	Characteristic properties	Animate	Inanimate
a	Onset of motion	can move by itself	is moved by others
b	Line of trajectory	moves irregularly	moves smoothly
c	Form of causal action	can move at a distance	moves from contact
d	Pattern of Interaction	motion against other's approach	
		contingent	non-contingent
e	Type of causal role	role in interaction	
		agent or recipient	only recipient
f	Purpose of action	motion looks like	
		goal-directed	without goal
g	Influence of mental states	motion looks	
		intentional	accidental

expressions, gestures, eye gazing, and so on. Human-like behavior enables people to intuitively understand, causing them to unconsciously behave as if they were communicating. For example, Scassellati et al. developed a robot with a joint attention mechanism that follows the gaze of others to share attention [6]. Nakadai et al. developed a robot that tracks a speaking person [7]. Kanda et al. found that cooperative body movements, such as synchronized eye contact and body movements, cause entrainment in human-robot interaction [8][9]. We believe that the integration of such human-like communication with lifelike behaviors is an important area of future work.

II. "LIFELIKE" BASED ON DEVELOPMENTAL PSYCHOLOGICAL FINDINGS

In the field of developmental psychology, some researchers have been investigating animate-inanimate distinction in infants. From past findings, Rakison and Poulin-Dubois consider that infants distinguish the animate from the inanimate based on seven characteristics: five motion related (a to e) and two psychological (f, g) properties [2]. (Table 1)

We can easily distinguish the animate from the inanimate based on these properties. For example, consider the differences between a ball and a dog based on the findings. A ball cannot move by itself, while a dog moves autonomously without aid (a). A ball moves smoothly, for example, rolls on a line. On the other hand, a dog moves irregularly (b). A ball is moved after from contact with others, while a dog can move at a distance from others, for example, by running away before being touched (c). A ball never reacts to our approach, while a dog does, for example, by running away or approaching us (d). A ball always plays the recipient role in interaction with others, while a dog can play both recipient and agent roles. For instance, a dog runs away from our approach but sometimes bites us (e). Although psychological properties are subjective, we usually interpret a dog's purpose and mental condition more easily than a ball's. For example, "a dog moves toward food because it is hungry" (f) (g). Thus, according to these properties, balls are inanimate, and dogs are animate.

Existing robots lack several of these properties, in particular (c) and (d). Therefore we have attempted to develop a "life-like" robot by interpolating them.

III. SYSTEM CONFIGURATION

We developed a Robot-Controller for an autonomous humanoid robot that performs "lifelike" actions and autonomously selects appropriate behaviors according to human behavior (Fig. 1).

A. System Outline

The system consists of a humanoid robot, sensors, and the Robot-Controller. An overview of the system is shown in Fig. 2. A motion capturing system captures body motion data of humans and robots. Based on the motion data, the Robot-Controller recognizes human behavior and commands the robot to react to human's behavior.

B. Humanoid Robot "Robovie"

We used a humanoid robot named "Robovie," which is characterized by human-like body expressions (Fig. 3). Its human-like body consists of eyes, a head, and arms that generate complex body movements required for communication. Robovie has two 4-DOF arms, a 3-DOF head, and 2-DOF eyes. Thus, its body has sufficient expressive ability to make human-like gestures. In addition, it has two wheels to move (forward-reverse travel and rotation). Its height is 1.2 m, its radius is 0.5 m, and its weight is about 40 kg [8].

C. Sensors

We integrated touch sensors and a motion capturing system into the system. A total of eight touch sensors are attached to the robot's body (head, belly, right or left upper arm, lower
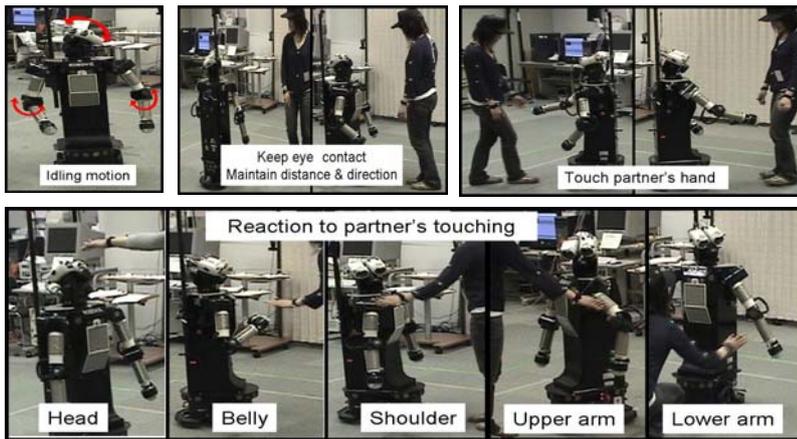
Fig. 4 Behaviors

Fig. 5 State transition

arm, shoulder), as shown by the broken line in Fig. 3, and used to determine whether a human touches the robot's body. The motion capturing system acquires 3-dimensional numerical data on human and robot body movements. It consists of 12 sets of infrared cameras with an infrared irradiation function and markers that reflect infrared rays. The motion capturing system calculates the 3-dimensional position of each marker based on the 2-dimensional positions on all of the cameras' images. In our experimental environment, the system's time resolution is 60 Hz and spatial resolution is about 1 mm. The position of each marker is shown by the solid lines in Fig. 3.

D. Robot-Controller

We used an external PC for the Robot-Controller. The connection from the Robot-Controller to the motion capturing system uses a wired LAN protocol and connection from the Robot-Controller to the robot uses a wireless LAN protocol. The Robot-Controller obtains human behavior from the sensors and make robot react to human behavior. The details of the Robot-Controller are described below:

[Cognitive abilities] The cognitive abilities for human behavior managed by the Robot-Controller are as follows:

➢ *Distance and direction recognition* - Based on the positions of markers attached to the right and left shoulders, the center positions of both the human and the robot body are calculated. The distance and direction between the human and robot are derived from the body's center position.

➢ *Face recognition* - Based on the positions of markers attached to the head, the position of the human face is estimated.

➢ *Hand recognition* - In the same way, based on the positions of markers attached to the hands, the positions of the human hands are estimated.

➢ *Touch recognition* - Based on data of touch sensors attached to the robot's body, the positions of human touch are recognized.

[Behaviors] The different behaviors managed by the Robot-Controller for our robot are implemented as follows:

➢ *Idling motion* - The robot performs an idling motion, moving its arm and neck at intervals (Fig. 4 upper left).

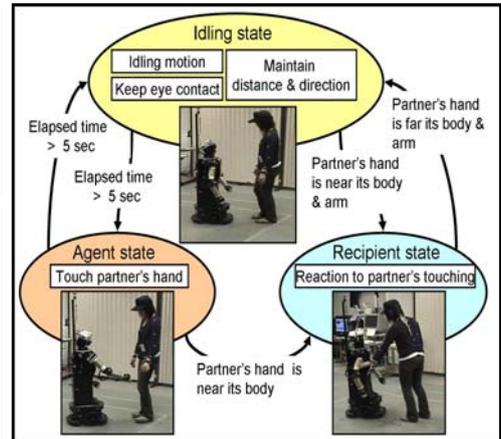➢ *Keep eye-contact* - The Robot-Controller forces the robot to track its partner's face by controlling its eye and neck motors (Fig. 4 upper middle).

➢ *Maintain distance and direction* - The Robot-Controller makes the robot maintain distance and direction with respect to a partner. The distance, about 60 cm, is selected so that the robot's arms cannot touch the partner's body. The direction is towards the face of the robot's partner (Fig. 4 upper middle).

➢ *Touching a partner's hand* - The Robot attempts to touch the partner's hand by making its hand approach the partner's hand by controlling its arms and wheels (Fig. 4 upper right).

➢ *Reaction to partner's touching* - When the distance between a partner's hand and each touch panel is less than 30 cm, the robot reacts to the nearer partner's hand and performs a behavior that represents its intention, which in our case, is that it dislikes being touched. Before the partner touches the robot's body, the robot tracks the partner's hand by controlling its eye and neck and dodges the partner's hand by controlling its arms and wheels. As shown in the lower part of Fig. 4, this dodging behavior is different for each touch part (eight patterns according to eight touch parts).

[State transition] The Robot-Controller selects from the following three states, according to the situation, and commands the robot to behave based on this state (Fig. 5).

➢ *Recipient State* - Robot performs "*Reaction to partner's touching.*"

➢ *Idling State* - The robot performs "*Idling motion*", "*Keep eye contact*", and "*Maintain distance and direction*" when the partner does not try to touch it.

➢ *Agent State* - If the Idling state continues more than five seconds, the robot performs "*Touching partner's hand*" for five seconds. But, if the partner tries to touch the robot's body or arm, except for the robot's hand that tries to touch the partner, the state changes from Agent to Recipient.

E. Considerations of Our Robot's Animate Properties

We developed a robot that can satisfy the seven properties of the findings, as described below.

◆ **Motion-related characteristics** - Concerning *a) onset of motion* and *b) line of trajectory*, these properties have already been satisfied by most existing robots. Our robot can move by
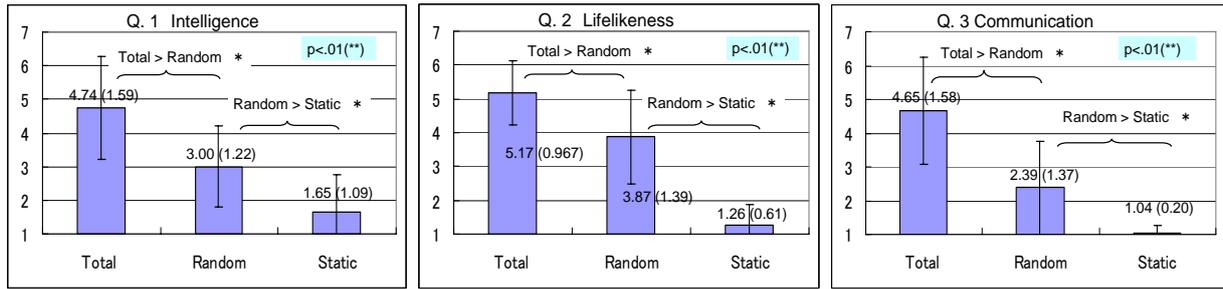
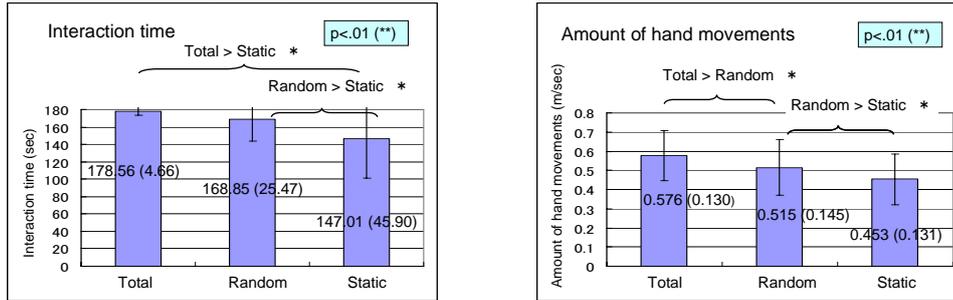Fig. 6 Experimental result 1 – Impression evaluation

Fig. 7 Experimental result 1 – Analysis of human body movements

Table 2 Questions in Experiment1

Q.1	Degree of intelligence
Q.2	Degree of lifelikeness
Q.3	Degree of communication

itself, with nonlinear motion. On the other hand, for stand-alone robots, it is difficult to satisfy *c) form of causal action*, *d) pattern of interaction*, and *e) type of causal role* properties because the cognitive abilities of the robots are too low. Our robot satisfies these properties by utilizing a motion capturing system. The robot can move without contact and react to its partner's motion. In addition, it can change its role from agent to recipient or from recipient to agent.

◆ **Psychological characteristics -** Although characteristics *f) purpose of action* and *g) influence of mental state* depend on individual impressions, we believe that our robot satisfies these properties because of the following two behaviors. First, it avoids human contact, which could be easily understood as the purpose of an action. Similarly, a robot trying to touch a human hand is easily understood as having a purpose. These behaviors seem to satisfy characteristic *f*. Second, the three internal states of the robot - idling, agent, and recipient - are easily recognized by humans and are assumed to reflect the mental state of the robot. Thus, we believe that this also satisfies characteristic *g*.

IV. EXPERIMENT1 - Effectiveness of our robot for "lifelikeness"

We believe that a robot can make humans think of it as "lifelike" if its behavior satisfies these seven animate features. Thus, we developed a robot to satisfy them as we explained in the third section. In this section, we explain an experiment that we conducted with subjects to verify the effect of our robot with regard to "lifelikeness". In our experiment, the robot didn't utter sounds because we believe that a robot without utterances can still be a "lifelike" object, like an animate being, other than a human.

A. Method

[Subjects] A total of 23 university students participated as subjects in the experiment (11 men, 12 women), whose average age was 19.2 years.

[Instructions to subjects] We instructed the subjects to interact with the robot. Experimental time per condition was a maximum of three minutes. Moreover, we instructed the subjects to stop interacting with the robot when they got bored.

[Experiment conditions] We set the following three conditions:

1. Total condition

The robot reacted to human behavior and moved by the commands of the Robot-Controller. During interaction, the interaction records, which are the input data of each motor of the robot, were recorded for the next subject's *Random* condition.

2. Random condition

The robot behaved randomly to human behavior. In other words, it didn't react to human behavior. The robot only replayed previous interaction records, recorded when it interacted with another subject in the Total condition. By utilizing this method, in the entire experiment, there is no difference between *Total* and *Random* conditions regarding the quality of the robot's movements (speed, travel distance, and so on).

3. Static condition

The robot stood still during this condition.

[Evaluation method] We administered a questionnaire to obtain subjective evaluations for every condition. The questions are shown in Table 2. They answered each question on a 1 to 7 scale, where 1 is the lowest evaluation and 7 is the highest.
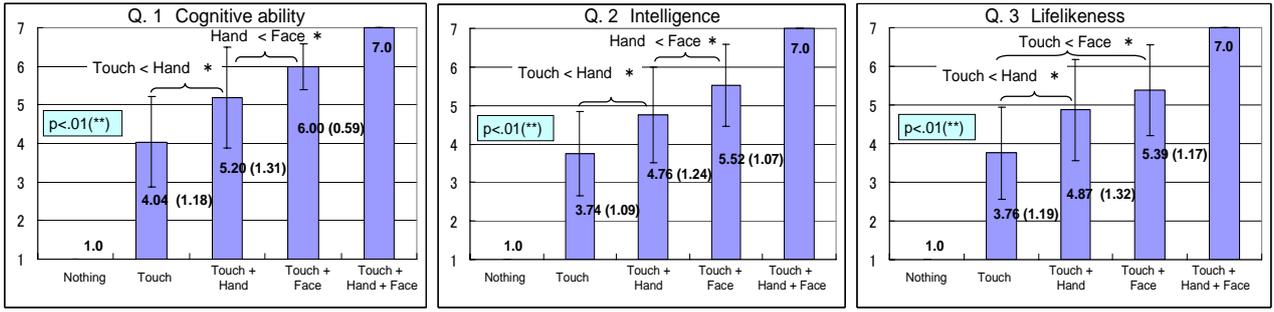
Fig. 8 Experimental result 2 – impression evaluation

Finally, we compared each subject's behavior when changing conditions. We analyzed the amount of hand movements, which were calculated based on the relative position of the hands from the center of the body.

B. Experimental results

[Subjects' impressions of the robot] Figure 6 shows the results. As a result of within-subject design analysis of variance (ANOVA), there was a significant difference in all questions. Moreover, for each of the significant items, a least significant difference (LSD) method provided a multiple comparison among all conditions as follows:

➤ **Q. 1 Intelligence** $(F(2,22)=34.38, p<.01)$
  Value of "*Total*" > Value of "*Random*" > Value of "*Static*" (MSe=1.602, p<.05)

➤ **Q. 2 Lifelikeness** $(F(2,22)=101.99, p<.01)$
  Value of "*Total*" > Value of "*Random*" > Value of "*Static*" (MSe=0.895, p<.05)

➤ **Q. 3 Communication** $(F(2,22)=58.99, p<.01)$
  Value of "*Total*" > Value of "*Random*" > Value of "*Static*" (MSe=1.296, p<.05)

These results proved that humans recognize the differences between each condition, and that the "*Total*" condition had the highest evaluation regarding "Intelligence," "Lifelikeness," and "Communication."

[The effect of the robot on subjects' behaviors] Figure 7 shows the results. As a result of within-subject design ANOVA, there is a significant difference in both "Interaction time" and "Hand movements." In addition, for each of the significant items, the LSD method provided a multiple comparison among all conditions as follows:

➤ **Interaction time** $(F(2,20)=7.96, p<.01)$
  Value of "*Total*" > Value of "*Static*", Value of "*Random*" > Value of "*Static*" (MSe = 688.2, p<.05)

➤ **Amount of hand movements** $(F(2,20)=9.12, p<.01)$
  Value of "*Total*" > Value of "*Random*" > Value of "*Static*" (MSe=0.0087, p<.05)

For interaction time, there was a significant difference between the *Static* and other two conditions, and there is no significant difference between the *Total* and *Random* conditions. For the amount of hand movements, *Total* got the largest amount in all three conditions, implying that the robot in the *Total* condition most encourages humans to try to interact.

As expected, for the robot's behavior, the amount of body movements did not show a significant difference between the

Table 3 Cognitive ability of the robot in Experiment 2

Condition	Face recognition	Hand recognition	Touch recognition
Nothing	×	×	×
Touch	×	×	O
Hand+Touch	×	O	O
Face+Touch	O	×	O
Face+Hand+Touch	O	O	O

Table 4 Questions in Experiment 2

Q. 1	Degree of cognitive ability
Q. 2	Degree of intelligence
Q. 3	Degree of lifelikeness

Total and *Random* conditions, which indicates that experimental control was correctly performed.

V. EXPERIMENT 2 - Effect of cognitive abilities of robots for human-robot interaction

The cognitive ability of our robot is superior to the other existing robots. By utilizing our robot, we can investigate the effect of the cognitive abilities of robots for human-robot interaction. In this section, we describe an experiment in which we investigated the relationship between subjective impressions and the robot's cognitive abilities by performing simple experiments.

A. Method

[Subjects] The subjects were the same as in Experiment 1.

[Instruction to subjects] The subjects interacted with the robot after Experiment 1. The interaction time was one minute per condition.

[Experiment conditions] The subjects interacted with the robot in five conditions (Table 3), where the cognitive abilities of the robot differed from one to the next. Then, the subjects answered questionnaires. First, they interacted with the robot in the "*Nothing*" and the "*Face+Hand+Touch*" conditions. Then, the subjects were requested to assume an evaluation value of the robot in the "*Nothing*" condition as one point and in the "*Face+Hand+Touch*" condition as seven points. Next, they interacted with the robot in another three conditions, which are better than "*Nothing*" but worse than "*Face+Hand+Touch*" in terms of the cognitive ability of the robot. After each session, the subjects evaluated the robot based on the evaluation value of the "*Nothing*" and "*Face+Hand+Touch*" conditions.

In this experiment, the robot was controlling its body and arms without moving its wheels.

[Evaluation method] We administered a questionnaire to obtain subjective evaluations. The subjects answered a questionnaire every time they interacted in one condition. We used the questions shown in Table 4. The subjects answered each question on a 1 to 7 scale, where 1 stands for the lowest evaluation and 7 for the highest.

B. Experimental results

Figure 8 shows the average and standard deviation values. It also describes the results of within-subject design ANOVA among the "*Face+Touch*", "*Hand+Touch*" and "*Touch*" conditions. As a result of within-subject design ANOVA, there was a significant difference shown by the results of all questions (Q. 1, Q. 2, and Q. 3). Moreover, for each of the significant items, the LSD method provided the following multiple comparison among the "*Face+Touch*," "*Hand+Touch*," and "*Touch*" conditions:

➢ **Q. 1 Cognitive ability** $(F(2,22)=22.32, p<.01)$
  Value of "*Face+Touch*" > Value of "*Hand+Touch*" >Value of "*Touch*" (MSe=0.9965, p<.05)

➢ **Q. 2 Intelligence** $(F(2,22)=17.72, p<.01)$
  Value of "*Face+Touch*" > Value of "*Hand+Touch*" > Value of "*Touch*" (MSe=1.0385, p<.05)

➢ **Q. 3 Lifelikeness** $(F(2,22)=15.91, p<.01)$
  Value of "*Face+Touch*" > Value of "*Touch*", Value of "*Hand+Touch*" > Value of "*Touch*" (MSe=1.0025, p<.05)

VI. DISCUSSION

A. Findings from Experiment 1

We have verified the effectiveness of a robot developed to satisfy animate features as described in the findings of developmental psychology. Even though the robot's motion in the "*Total* condition" is the same as that in the "*Random* condition" in terms of the quality of motion, experimental results showed a difference between the "*Total* condition" and the "*Random* condition" not only in terms of impression evaluations but also in human behavior. This result suggests that being contingent toward a partner's motion is an important property for making humans feel that a robot is "lifelike" and a "communication partner." It also revealed that the random behavior selection was responsible for a large difference in the subjects' evaluation of lifelikeness compared to the static behaviors in the *Static* condition. We believe that this behavior partly satisfies the animate features in terms of motion. To identify which properties are related to lifelikeness, we need to investigate each property one by one.

B. Findings from Experiment 2

These results will be one way to speculate on the effect of the cognitive ability of robots for human-robot interaction. By utilizing these findings, for example, we can estimate that robots with touch recognition and incomplete face recognition ability will be evaluated higher in the "*Touch*" condition but lower in the "*Face+Touch*" condition.

C. Perspective

In this research, the results imply that lifelike behavior entrains humans to interact with robots, suggesting that lifelikeness is important for communication. As mentioned in the introduction, we consider "lifelike" behavior to be essential for robots that will be engaged in natural communication with humans. While we anticipate that much research concerning communication robots will be conducted, researchers need to take this lifelikeness into account.

At the same time, we confirmed that our approach with a motion capturing system was useful for the design of the robot's behavior. First, we were able to verify the effectiveness of our lifelikeness mechanism inspired by developmental psychology. Secondly, we were able to investigate the effects of the robot's cognitive ability on the subjective impressions of humans, which will be developed in the future. We believe that these findings demonstrate the importance of our approach.

VII. CONCLUSION

This paper presented the development of "lifelike" behavior for a communication robot based on the findings of developmental psychology. We developed a robot that is constantly "aware" of human motion by utilizing a motion capturing system. Two experiments were conducted to investigate its effects. In the first experiment, we verified the effectiveness of lifelike behavior in interaction; in addition, this experiment revealed that random behavior selection was responsible for a large difference in the subjects' evaluations of lifelikeness compared to static behavior. The second experiment demonstrated the effectiveness of the robot for investigating the effect of adding cognitive abilities in human-robot interaction.

REFERENCES

[1] Arita, A., Hiraki, K., Kanda. T., and Ishiguro, H., Can we talk to robots? Ten-month-old infants expected interactive humanoid robots to be talked to by persons. *Cognition*, Vol. 95, pp. B49-B57, 2005

[2] David H. Rakison and Diane Poulin-Dubois, Developmental Origin of the Animate-Inanimate Distinction, *Psychological Bulletin*, Vol. 127,No. 2, pp. 209-228, 2001

[3] Fujita, M., and Kageyama, K., An Open Architecture for Robot Entertainment, *Proc. of Agent'97*, 1997.

[4] Shibata, T., Tashima, T., and Tanie, K., Emergence of Emotional Behavior through Physical Interaction between Human and Robot, *Proc. of the IEEE ICRA*, 1999.

[5] Nakata, T., Sato, T., and Mori, T., Producing Animal-likeness on Artifacts and Analyzing Its Effect on Human Behavioral Attitudes, *IROS'99*, pp. 549-554, 1999.

[6] Scassellati, B., Investigating Models of Social Development Using a Humanoid Robot, *Biorobotics*, MIT Press, 2000.

[7] Nakadai, K., Hidai, K., Mizoguchi, H., Okuno, H. G., and Kitano, H., Real-Time Auditory and Visual Multiple-Object Tracking for Robots, *International Joint Conference on Artificial Intelligence (IJCAI'01)*, pp.1425-1432, 2001.

[8] Kanda, T., Ishiguro, H., Imai, M., and Ono, T., Development and Evaluation of Interactive Humanoid Robots, *Proceedings of the IEEE*, Vol.92, No.11, pp. 1839-1850, 2004.

[9] Kanda, T., Ishiguro, H., Imai, M., and Ono, T., Body Movement Analysis of Human-Robot Interaction, *Proc. of International Joint Conference on Artificial Intelligence*, pp. 177-182, 2003.