PALADYN Journal of Behavioral Robotics

VERSITA

Research Article · DOI: 10.2478/pjbr-2013-0008 JBR · 4(2) · 2013 · 113-122

Attractive, Informative, and Communicative Robot System on Guide Plate as an Attendant with Awareness of User's Gaze

Tomoko Yonezawa[1*], Hirotake Yamazoe[2†], Akira Utsumi[3‡], Shinji Abe[4§]

1 Kansai University, 2-1-1 Ryozenji, Takatsuki, Osaka 569-1095, Japan

2 Osaka University, 1-31 Machikaneyama, Toyonaka, Osaka 560-0043, Japan

3 ATR IRC, 2-2-2 Hikaridai, Seika, Soraku, Kyoto 619-0288, Japan

4 Hiroshima Institute of Technology, 2-1-1 Miyake, Saeki, Hiroshima 731-5193, Japan

Received 01-05-2013

Accepted 08-10-2013

Abstract

In this paper, we introduce an interactive guide plate system by adopting a gaze-communicative stuffed-toy robot and a gaze-interactive display board. An attached stuffed-toy robot on the system naturally show anthropomorphic guidance corresponding to the user's gaze orientation. The guidance is presented through gaze-communicative behaviors of the stuffed-toy robot using joint attention and eye-contact reactions to virtually express its own mind in conjunction with b) vocal guidance and c) projection on the guide plate. We adopted our image-based remote gaze-tracking method to detect the user's gazing orientation. The results from both empirical studies by subjective / objective evaluations and observations of our demonstration experiments in a semipublic space show i) the total operation of the system, ii) the elicitation of user's interest by gaze behaviors of the robot, and iii) the effectiveness of the gaze-communicative guide adopting the anthropomorphic robot.

Keywords

gaze-correspondence · anthropomorphic media · stuffed-toy robot · gazing behaviors · gaze-tracking

1. Introduction

While communicative robots and artificial agents show anthropomorphic presences and multimodal representations, they have been proposed and developed over the past decade. Their anthropomorphism is expected to provide familiar and natural expressions evoking emotional communication. Anthropomorphism is an effective method in representing not only life-like appearances but also parallel and nonverbal expressions [4, 24, etc.]. Among various channels of multimodal expressions, the gazing behaviors of these anthropomorphic media (e.g. virtual agents [5] and humanoid robots [6, 26]) provide strong cues to help people guess the media's internal states, just as in human-human communication [9]. For instance, Gaze behaviors have been tentatively applied to the artificial behaviors of robots [10] to express their internal intentions or emotions. Joint attention is one of the gaze behaviors to look at a same object with the gaze of the other presence. The joint attention has been adopted to various anthropomorphic media as a fundamental expression of internal state of life-like mind as discussed for communicative learning between mothers and toddlers [12].

We have proposed a gaze-communicative guide system using anthropomorphic guidance by a stuffed-toy robot and an interactive display board. Our proposed system provides the user with sensitively attractive information as though there were a human attendant in front of the guide plate. The system is expected to be applied not only for intuitive guide system for healthy adult individuals but also for dementia, vocally-challenged, and children in kindergartens. Especially, the system can become one solution to dementia prowls. Gaze is one of the few communicative channels available for dementia sufferers who have difficulties in verbal communication. The gaze-communicative system is expected i) to refrain from stimulating the patients through efficient style and to provide them relaxed and relieved feeling by attracting them with natural gaze-communication from a stuffed-toy, ii) to remind them of, or give them, the purpose of the stroll, and iii) to naturally lead them toward the destination of their purpose. Children and elderly people wandering in shopping plazas are also expected to be helped by the proposed system. When the stuffed toy reacts to the conscious or underconscious gaze of the user, she/he can intuitively communicate with the system and find their destinations.

Stuffed toys have been applied to many uses in the care of dementia patients [13] and traumatized children [16] since they offer familiar and unforced communication for people in various situations. This adaptability is caused by their flexible anthropomorphic characteristics in both *avatars* and *partners* as seen in children's playing house. Es-

*E-mail: yone@kansai-u.ac.jp
†E-mail: yamazoe@osipp.osaka-u.ac.jp
‡E-mail: utsumi@atr.jp
§E-mail: s.abe.pk@it-hiroshima.ac.jp

pecially, the communicative gazing behavior of stuffed toys should be noteworthy as an indirect controlling method to get and lead user's subconscious attention with the familiar anthropomorphism. For instance, there are many information guide plates with voice guidance, however, they are not so attractive because of the characteristic impression of the automatic and artificial guidance regardless of the user's behavior.

The eyes of animal-like stuffed toys are composed almost completely of irises without whites. This characteristic causes a distinctive ambiguity of the gaze direction. Therefore, gaze direction should be expressed by the direction of their heads. For maximizing effectiveness of comprehensive gaze-communication, stuffed toys must behave in a simple and appropriate manner.

In this research, we focused on how the gaze-interactive guide might serve triggers for i) finding of a new destination with a clear purpose, or ii) recalling of the original purpose of the user's stroll. Photographs or graphic drawings that convey concrete and visual expressions become important clue-givers especially for people who have information shortfalls such as dementia patients suffering from memory disorder. However, such patients have difficulty in understanding a map or figures with their positional relationships oriented from the present location to the destination. Accordingly, we propose the "GazeRoboard," an information guide plate system with a guide plate and an attached stuffed-toy in a natural communication flow corresponding to the user's gaze, which may indicate her/his potential purpose. We adopted our remote gaze-tracking method [23], which requires no fixing or attachment worn by the user, for intuitive guide interaction.

In Section 2, we introduce some of the related works, especially the works on guide systems not mentioned in this section. In Section 3, the system structure and its components are described to explain how the stuffed-toy robot guides the user corresponding to her/his gaze. In Section 4, we describe experiments on the system's operations, including its gaze-tracking accuracy, and we report a demonstration experiment conducted in a semi-public space. In Section 5, we show the results of our evaluations for a) basic effectiveness of gaze behaviors in eliciting user's interests and b) gaze-communicative guide in association with the robot's behaviors. Based on the results in Section 4 and 5, we discuss the effectiveness of the system with this configuration in Section 6. Finally, we summarize this paper in Section 7.

2. Related Research

2.1. Gaze-interactive Systems

There have been many researches on communication robots utilizing social gaze. Several studies and research efforts have examined social interactions or recognitions between humans and robots. Breazeal et al. [2] suggested that imitative behaviors of a robot are effective for social communication. Sidner et al. [18] focused on the concept of engagement gestures to discuss gaze and various interaction between humans and robots. Thomaz et al. [21] focused on a social referencing model when they examined the affective and cognitive aspects of communication with using an embodied robot. Although these works mainly discussed the imitative interactions of the gazing or facial behaviors based on each particular model, joint attention has not been sufficiently focused from the viewpoint of combined use with the other gazing behaviors. Accordingly, this research is motivated to clarify not only the independent effectiveness of eye contact and joint attention between human and robot but also the interaction of these gaze behaviors for exploitation of them in the guide system.

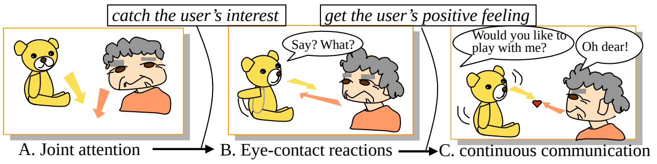


Figure 1. Stepwise Gaze-communication Model

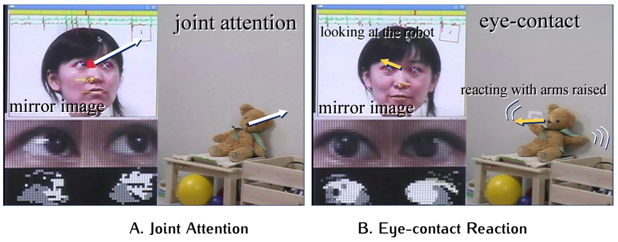


Figure 2. Examples of Gazing Behaviors of Stuffed-toy Robot

2.2. Guide-agent Systems for Guide or Assistance

The importance of anthropomorphic agents, in natural/intuitive guidance and assistance, have been shown by many research works for various purposes such as a personal exhibition-guidance system [19] and a virtual training guide [7]. Katagiri et al. verified that the different behaviors of an agent differently affect the user's knowledge and performance [8]. We regard their results as the robot's gazing behavior as a kind of persuasive power used in the same way as the behavior of a virtual agent.

Embodied guide agents and partner systems have also been proposed. An interactive musium tour-guidance [3], a nursing and daily-life guide agent [20], and an animation guide puppet [22] showed the effectiveness of the anthropomorphic guidance. Some of their physical interactions are appropriate for multiple persons, allowing the system to roughly react to the people; however, our approach is quite different from them by adopting the subtle interactions of gaze. Pineau et al. and Pollack et al. have proposed a guide system with a mobile robotic assistant using map differencing to detects users to provide various guidance / helps [14, 15]. The concept of their systems are very practical and similar to our purpose, but their robots does not offer any intuitive, casually familiar, or subtle interaction as seen in gaze communication. On the other hand, Taggart et al. [20] showed that part of the elderly subjects engaged with the interactive robot Paro [17], while others avoided Paro when the experimenter turned it on. It is presumed that an anthropomorphic interface should include a simple escapable mechanism. When the elderly people subconsciously feel unwilling, they can easily refrain from interaction without straining to express their wills. Our proposed gaze-communicative interaction would be useful for capturing the user's gaze even from their defective interest.

3. System Design of Gaze-communicative Guide

We hypothesized a stepwise model of gaze-communicative level and constructed a gaze-communication system [25] with a stuffed-toy robot

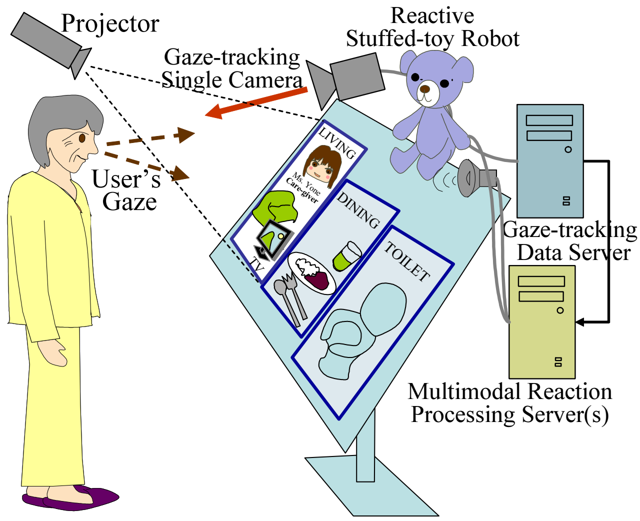


Figure 3. The *GazeRoboard* System Structure

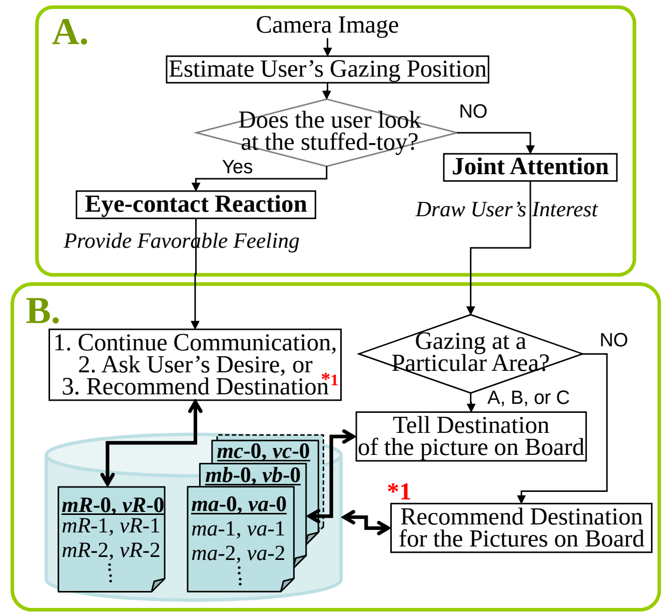


Figure 4. System Flow's Implementation

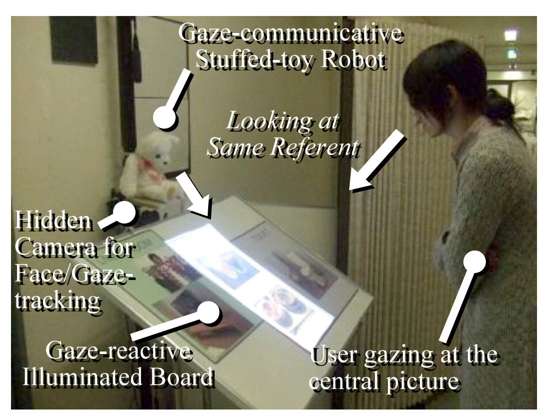


Figure 5. System View of Gaze-communicative Guide

that offers communicative reaction, especially focusing on the facing behavior of the stuffed-toy, corresponding to the user's gaze. To evoke the user's desire and possibility to communicate, the guide system should be aware of the user's conscious / subconscious gaze. We developed GazeRoboard, a gaze-communicative guide plate system based on a stepwise model of gaze communication using the stuffed-toy robot (Figure 1). The joint attention behavior is expected to indirectly evoke the user's interest as a communicative trigger. We adopted the behavior based on the direction of the robot's head as shown in Figure 2-A. The eye-contact reactions of the robot (Figure 2-B) were also adopted to give the user a favorable feeling for sustainable communication. Thus the reactive stuffed-toy robot was designed to be capable of making expressions in voice and gestures with its internal speaker and two-axis degrees of freedom in each joint, i.e. the head and arms.

3.1. Gaze-communicative Guide System

Figure 3 shows the configuration of our system. The attached stuffed-toy robot is controlled by a PC. In order to make reaction corresponding to the user's gazing position, a camera is installed to estimate the user's gaze. A gaze-tracking server calculates the directions of the user's gaze and sends the angles to another PC which processes multimodal reactions for the stuffed-toy robot.
The multimodal reaction-processing server determines and controls 1) the robot's gazing behaviors, 2) the robot's utterances, and 3) the illumination in a particular region on guide board according to the relative orientation of the user's gaze, which is calculated from the 3D positions of the stuffed-toy robot and the user. Accordingly, the multimodal processing server consists of i) the gaze-communicative reaction part (Figure 4-A) as the basis of the anthropomorphic attitude for communicative guide and ii) the detailed guide reaction part (Figure 4-B) with approximating the user's subconscious request. At first, the robot reacts to the user's gaze behaviors, and next the robot begins the guidance of the daily-life event shown by the picture at which she/he gazes at least for one second. When the user's gaze is unstable or staying on for more than three seconds to a vacant space, the robot recommends information of somewhere to go or a place of interest. These tentative values of the durations are defined assuming a proactive guide for wandering people.
Multiple data sets for the robot's motion and voice are allocated to each contents-area on the interactive guide plate in relation to the contents of the picture. Picture *A* has been assigned the corresponding motion data from ma-0 to ma-4 and voice data from va-0 to va-4. When the user looks at the picture for the first time, the robot behaves ma-0 and utter va-0 after joint attention. While the user is continuing to gaze at the same picture or the same region in the guide plate, the robot behaves from ma-1 to ma-4 and utters from va-1 to $vXa-4$ stepwise. The structured data sets are used to provide the user with explanations and information on the place (e.g. how to go there) in stepwise depths of detail corresponding to the time length of the user's gaze at the same

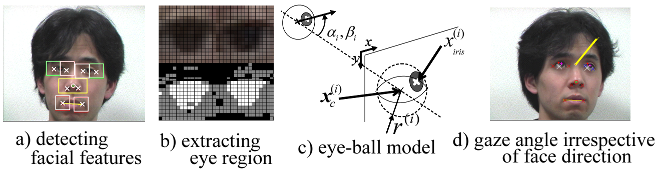


Figure 6. Results of Gaze-tracking

region. The robot itself has also the assigned multimodal data sets for eye-contact reactions.

Here, we show a sample use. When the user is wandering by losing her/his original purpose or way around the system, the system detects her/his presence and start estimating her/his gazing direction. In order to attract the user, the robot starts talking with reactions of eye-contact or joint attention according to the user's gazing object. The gaze-communication using both of joint attention and eye contact is expected to provide a relieved feeling through natural interactions. Next when the user still seems to be wandering, the gaze of the robot at the pictures draws the user's interest toward a particular region on the board as shown in Figure 4-B. Finally when the user looks at the region of a particular picture on the board, the robot 1) behaves joint attention at first and 2) explains the information about the picture while looking at both the user and the picture. At the same time, a projector illuminates the region of the picture. Thus the gaze-communicative guide is associated with the pictures at which the user gazes. From this configuration, it is expected to i) make her/him remember the original purpose of her/his stroll or to ii) recommend her/him a destination when she/he feels insecure in aimlessly wandering. A view of our guide plate system is shown in Figure 5.

3.2. Remote Gaze-tracking Method

For realizing natural and familiar interaction, users should be unaware of gaze-tracking. However, most of conventional gaze-tracking systems need fixtures to mount [1, etc.], or need special processes to perform calibration [11, etc.]. To reduce these burdens, we have proposed a gaze-tracking method using a remote camera [23]. This ambient gaze-tracking is expected to realize subconscious gaze-communication between a user and the stuffed-toy robot.

In the gaze-tracking system [23], gaze directions are estimated as the 3-D vectors that connect the centers of the eye ball and iris (Figure 6-b). Since we cannot directly observe the eye-ball center from images, these positions are calculated by using observed positions of facial features (Figure 6-a). The relationships between the eye-ball centers and facial features are automatically calibrated from the initial images. Differently from conventional gaze-tracking methods, our method does not need special calibration process such as instruction to users to look at several specific points for her/his own model of the eye and head. This can be realized by combination of a 3-D face-model reconstruction based on factorization and the head/eye model estimation by nonlinear optimization. Iris centers are calculated by fitting an ellipse into the observed image (Figure 6-c), and finally gaze direction can be determined (Figure 6-d). The estimation accuracy of the gaze orientation is about 5° in the horizontal angle and 7° in the vertical angle by our proposed method.

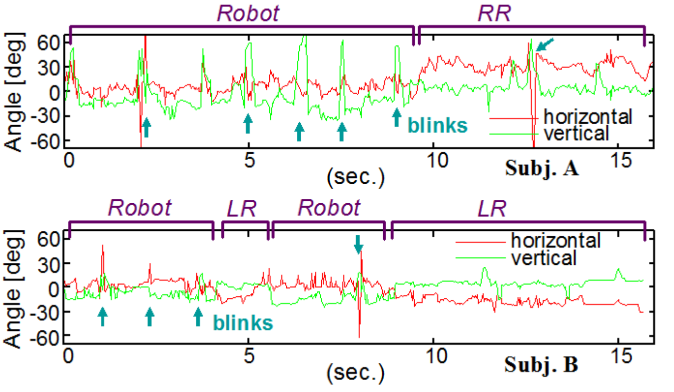


Figure 7. Examples of Estimated Gaze Angles

Table 1. Accuracies of Gaze-tracking for Each Region on the Board

(average)	*LR*	*CR*	*RR*	robot
horizontal	3.81°	3.51°	2.43°	1.40°
vertical	6.89°	5.44°	5.79°	9.11°

4. Verification of System Operation

Based on the evaluation of gaze-communicative effectiveness, we tested the system operation for gaze-tracking accuracies on the display board to consider whether the system settings are appropriate for our target interactive architecture.

4.1. Region-based Gaze-tracking Evaluation

First, we evaluated the accuracy of the system operation by the three regions composing the constructed system's display board.

Subjects: Five people aged from 21 to 30 years (2 females and 3 males).

Procedures: The board size is B1 (1030 mm × 728 mm), and the width of each region is 300 mm. The height of both the left region (*LR*) and the right region (*RR*) is 650 mm, and the center region's (*CR*) height is 600 mm, as in Figure 3. The gaze-tracking camera is placed between the robot and the board. We prepared various markers on the board to detect the accuracy of the gaze-tracking. The markers were set at the center and corners of the region on the board and around the stuffed-toy robot. The experimenter instructed the subjects to gaze at the directed marker. The subjects stood about 110 cm from the camera.

Results of Accuracy Tests in Gaze-tracking: Table 1 shows the gaze-tracking accuracy. The results show characteristics of our gaze-tracking method: it is more accurate at horizontal angles than at vertical angles. From the results, the pictures should be desirable to be placed near the center of the each region. On the other hand, the horizontal accuracies are high enough to detect the user's gazed region. The detailed results shown in Table 2 indicate the possibility of wrong judgment of the gazing region. Judgment errors are assumed especially between the lower part in the robot's region and the upper part in the *CR* region (underlined in Table 2) due to the low vertical accuracy. In

Table 2. Detailed Accuracies of Gaze-tracking for Each Region

region	center	corner(*lt*)	corner(*rt*)	corner(*lb*)	corner(*rb*)
LR	(4.45, 7.48)	(1.22, 3.22)	(2.67, 2.93)	(5.05, 11.3)	(5.67, 9.51)
CR	(4.01, 5.65)	(3.41, 2.2)	(2.18, 1.98)	(4.5, 8.41)	(3.45, 8.95)
RR	(2.63, 5.73)	(1.52,1.82)	(1.14, 1.73)	(2.73, 9.17)	(4.12, 10.5)
robot	(1.65, 9.00)	(1.14, 12.9)	(1.73, 12.5)	(0.96, 5.8)	(1.49, 5.38)

lt: left top, *rt*: right top, *lb*: left bottom, *rb*: right bottom
Described with (horizontal°, vertical°) in degrees of the angles

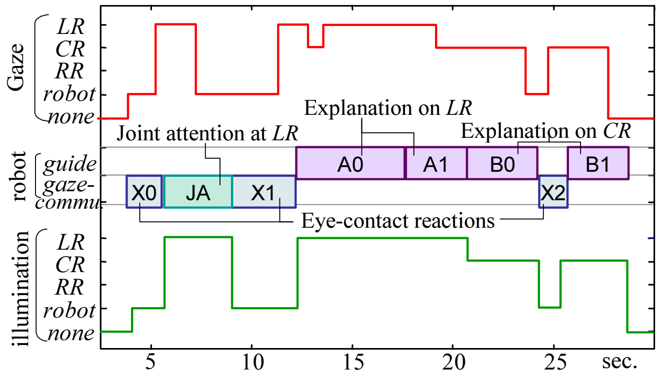
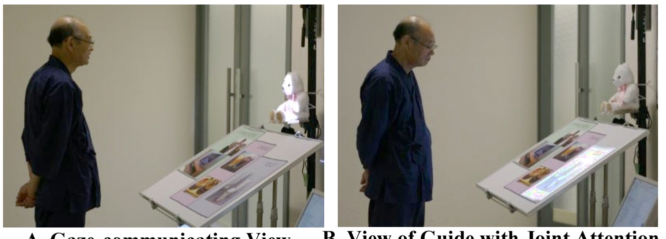


Figure 8. Example of System Operation Flow



A. Gaze-communicating View
Communicating with the illuminated robot

B. View of Guide with Joint Attention
Joint attention at illuminated *RR*

Figure 9. System Operation View

this configuration, the robot was placed at the position sufficiently distant from *CR*.

4.2. Operation of Guide System

We tested the operation of the three-region guide system of daily-life situations while tracking the user's gaze at:
a) region with pictures of TV and sofa (*LR*),
b) region with pictures of meals and drinks (*CR*),
c) region with pictures related to toiletry (*RR*),
and the region around the stuffed-toy robot.
One of the results is shown in Figure 8. The system generated both gaze-communicative reactions and guidance on each daily-life object. The robot and the illuminative board could react to the user's gaze at sufficient response speed for natural interaction of gaze. It was also confirmed that the user could experience both the anthropomorphic reactions and guidance through the operation of the guide system. Figure 9-A and 9-B show examples of the system view in operation. Figure 9-A is a view of gaze-communicative operation (with eye-contact reaction) and Figure 9-B is a view of a guide operation (*LR*) combined with joint attention. Thus the system operations in real situations were verified in both the gaze-communicative function and the communicative guide function.

Figure 10. Demonstration Experiment in Lobby of Hotel

4.3. Demonstration Experiment in Semipublic Space

Our proposed guide system had a trial use in the lobby of a hotel (Hotel Kintetsu Universal City) for one week. About 170 people/groups experienced interactive uses. Many people had tried to use the system in the demonstration experiment in the lobby (see Figure 10-A). The system guided them to a restaurant (*LR*), a special room (*CR*), or a particular shop (*RR*) corresponding to their interest at that moment. Most users looked at both the stuffed-toy robot and the illuminated guide board. Not only the adult users but also children were interested in and used the interactive guide system as shown in Figure 10-B. There were several users who looked at the guide expressions provided for the other user's gaze (as Figure 10-C).

5. System Evaluations

To verify i) detailed effectiveness of anthropomorphic gaze behaviors in eliciting the user's interest and ii) the total effectiveness of our proposed gaze-communicative guide system, we conducted following two evaluations.

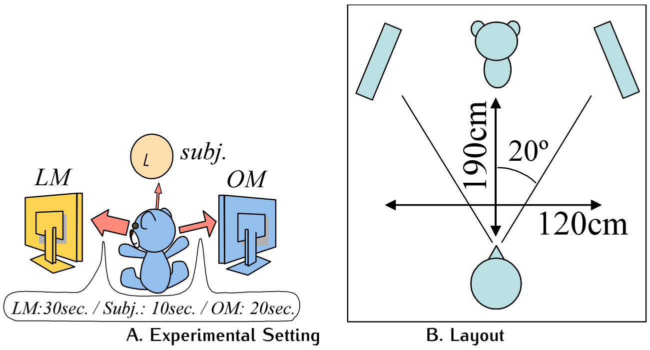


Figure 11. Setting and Layout of Experiments on Drawing Interest

5.1. Elicitation of User's Interest using Robot's Gaze Behaviors

To confirm elementary effectiveness of anthropomorphic gaze-communicative behaviors by the stuffed-toy robot, we examined whether the gazing behavior of the stuffed-toy robot effectively lead the user's gaze. In this experiment, we focused on the effect of the gazing orientation of the stuffed-toy robot in drawing the user's interest.
Hypotheses: I) Subjects gaze longer at an animation content that is also gazed at by the robot than at the other animation. II) Subjects gaze in the same direction as the robot's gaze at any given time.
Subjects for Elementary Experiments: Twenty-two subjects aged from 21 to 40 years (ten females and twelve males).
Stimuli: A stuffed-toy robot was placed between two monitors. The robot turned its head in following three directions: to each monitor and the subject. As shown in Figure 11-A, the gazing durations of the robot for each direction were set to be about i) thirty seconds for the primary-gazed monitor (*LM*), ii) twenty seconds for the less-gazed monitor (*OM*), and iii) ten seconds for the subject in a common scenario. The monitors displayed two different animations on each side as secondary stimuli for objects of joint attention.
Procedure and Instructions: The environment of this experiment is shown in Figure 11-B. Participants were just instructed only to guess what the robot had in mind in order to refrain from the case which the subject concentrates on the animations and never looks at the robot.
The participants were initially instructed to briefly describe what the stuffed-toy robot had in mind after each one-minute experiment. A stuffed-toy robot was placed in front of the subjects as shown in Figure 11-B. Two monitors were placed at left and right side of the robot. A hidden camera captured images of the subject's face. The gazing behaviors of the robot were generated corresponding to or diverging from the detected gazing position.
We also prepared simple animation pairs changing every six seconds. For the object of joint attention, the animation pairs were displayed on 17-inch monitors at half-size.
Results of Effectiveness in Elicitation: Table 3 shows the averages of total gaze durations for *LM*, *OM*, and ***robot***. Analysis of variance (ANOVA) with repeated measures ($\alpha = .05, \phi = 20, 2$) resulted in a significance with $F_{(20,2)} = 8.34$ and $p < \underline{.01}$. The results of the post-hoc test using Scheffé showed significant differences between *LM* and ***robot*** with $p < \underline{.01}$ and between *LM* and *OM* with $p = \underline{.01}$. Accordingly, Hypothesis I was supported by these results (Figure 12-A).
Next we observed the relation between the gaze direction of the robot and the subjects (Figure 13). It was observed that many subjects pursued the robot's gaze (Figure 13-subject A), while some subjects looked at monitors regardless of the robot's gaze (Figure 13-subject B). Then we focused on the transitions of participant's gazing point (gaze-switches) from ***robot*** to *LM* ( ) or to *OM* ( ) as shown in Figure 13. Table 4 shows the results tabulated by following two situations: i) the relation between the gaze targets of the robot and the participant (gazing at the ***same monitor***, " " or ***anywhere else***, " ") and ii) the switching of the user's gaze target. We calculated the predicted values (prd.) from the rates of the robot's gazing durations. The χ^2 tests showed significant results. From the results, it is conjectured that the gazing behavior of the stuffed-toy robot drew the subject's gaze at the common object (Figure 12-B). Thus Hypothesis II was also confirmed.

Table 3. Total Gazing Duration of Subjects in Each Direction

	LM	*OM*	*robot*
average (stdev)	25.45 (7.255)	16.05 (7.06)	17.49 (5.41)
(comparative analyses)			
preliminarily test:	$F_{(20,2)} = 8.34$,	$p < \underline{.01}$	
post-hoc test (Scheffé):	(LM,OM) $p = \underline{.01}$,	(LM,robot) $p < \underline{.01}$	

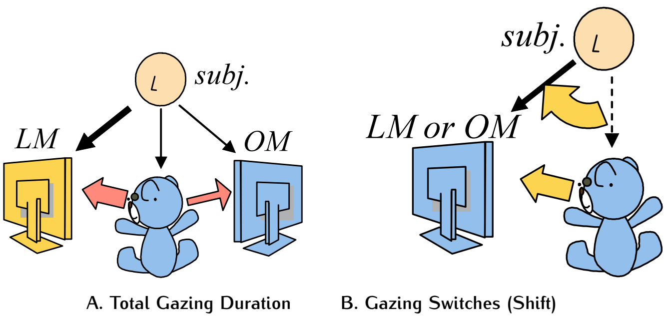


Figure 12. Observation of User's Gaze

5.2. Evaluation of Guide-communicative Guide

We show our three experiments focusing on the effectiveness of 1) the presence of the robot, 2) the motion of the robot's head direction, and 3) gaze corresponding guide.
Participants through Experiments of Guide System: Twenty-five subjects (thirteen females and twelve males) aged from twenty-one to forty years.
Common Settings of Experimental Guide Plate: In order to avoid the effect of the pictures on the guide plate and to verify simple effectiveness of multiple factors, we prepared simplified contents of the guide plate as shown in Figure 14. The simple figures are also applicable for various guide contents by conditions. A triangle pole was drawn in the left region of the guide plate (*LR*), a cube was drawn in the right region (*RR*), and no figure was drawn in the center region (*CR*).
Common Procedures: In each condition of the experiments, the subjects evaluate each stimulus using a five-point rating scale on relevance (5: very relevant, 4: somewhat relevant, 3: even, 2: somewhat irrelevant, 1: irrelevant) of the following statements:
(a) certainty factor (credibility of the guidance),
(b) attractive factor,
(c) naturalness, and

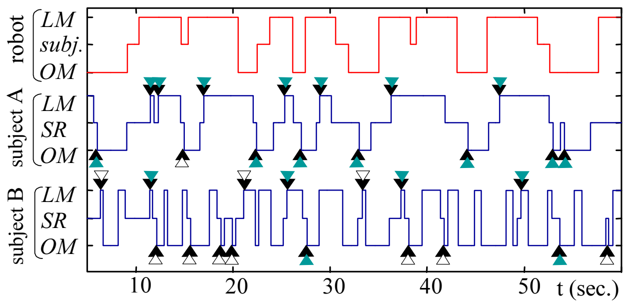


Figure 13. Relation of Robot's and Subject's Gazes

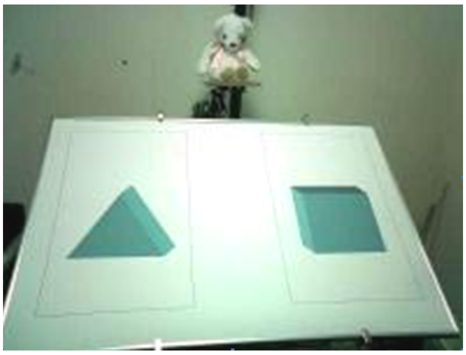


Figure 14. Experimental Environment for Gaze-reactive Robot Guide

Table 4. Summaries of Gaze-switching and χ^2 Test

switch to	total	*same* ▼▲ (prd.)	*else* ▽△ (prd.)	χ^2	p
LM "▼"	191	127 (95.5)	64 (95.5)	20.8	< .01
OM "▲"	157	87 (52.3)	70 (104.7)	34.4	< .01

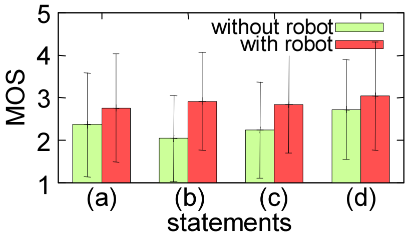


Figure 15. Evaluations for Robot's Existance

(d) change of the subject's interest after the guidance.
The subjects stood in front of the guideboard. The distance from the guideboard to the subjects was about four feet. To avoid the effect of the experimental order, all the conditions were counter-balanced.

5.2.1. Effectiveness of Presence of the Stuffed-toy Robot

In this experiment, we verified the effectiveness of the anthropomorphic presence in the guide system. To confirm the effectiveness as the following hypothesis, we compared the following two conditions.
Hypothesis: The presence of the robot affects on the user's impresson for the guide itself while the user is passive attitude to the system.
Conditions: We prepared the condition "without robot (R-)" as the automatic vocal guidance without the robot and the condition "with robot (R+)" as the automatic vocal guidance while showing the robot. In R-, the stuffed-toy robot was simply hidden by a large piece of paper to make the sound sources the same (from a speaker on the robot's belly) in both conditions.
Stimuli: The preceding vocal guidance, "I would like to talk about these cookies," was made before the stimuli. The next vocal guidance was the explanation of the contents. In this experiment, two guidances assuming each region were vocalized; 1) "this is a cinnamon cookie, " and 2) "this is a coconut cookie." The robot in both conditions did not make any motion.
Experimental Results: Figure 15 shows the summary of MOS (means opinion scores) for each statements. Table 5 shows the results of ANOVA with repeated measurements (degree of freedom (Dof,ϕ) = 24, rate of rejection (α) = 0.05). In all the statements, we could confirm that the significant differences between "without robot (R-)" and "with robot (R+)". The hypothesis was confirmed and it is conjectured that the presence of anthropomorphic robot can draw the user's positive impressions for the guide contents. Thus, the anthropomorphic presence show the effectiveness when it is fixed to the guide plate.

5.2.2. Effectiveness of Gazing Behavior of the Stuffed-toy Robot

In this experiment, we verify the effectiveness of the anthropomorphic gazing behaviors of the system's stuffed-toy robot. To confirm the following hypothesis, we compared the following two conditions.

Hypothesis: The facing and gazing behaviors of the robot affects on the user's impresson for the guide itself while the user is passive attitude to the system.
Conditions: We prepared the condition "without gaze (G-)" as a motionless robot with vocal guidance and and the condition "with gaze (G+)" as the robot's vocal guidance with its gazing behavior corresponding to the guided content. In G-, the stuffed-toy robot was simply sitting on the guide board without gazing behaviors.
Stimuli: The preceding vocal guidance, "I would like to talk about these housings," was made before the stimuli. The next vocal guidance was the explanation of the contents. In this experiment, two guidances assuming each region were vocalized; 1) "this is a tent for one person, " and 2) "this is a concrete building."
Experimental Results: Figure 16 shows the summary of MOS (means opinion scores) for each statements. Table 6 shows the results of ANOVA with repeated measurements (degree of freedom (Dof,ϕ) = 24, rate of rejection (α) = 0.05). In all the statements, we could confirm that the significant differences between "without gaze (G-)" and "with gaze (G+)". Thus, the anthropomorphic gaze behaviors show the effectiveness when it is explaining the contents of the guide plate and the hypothesis was confirmed. From the results, it is conjectured that the gazing behaviors of anthropomorphic robot toward the guide plate can draw the user's positive impressions for the guide contents.

5.2.3. Effectiveness of Correspondence to the User's Gaze in Guidance

In this experiment, to observe the user's feelings about the correspondence of the guidance to the her/his gaze, which the system sets based on the user's unconscious initiative, we verified the effectiveness of both

Table 5. ANOVA of Evaluations for Robot's Existance

	(a)	(b)	(c)	(d)
F-value	9.6	27.9	12.0	5.37
p-value	<.01	<.01	<.01	.029

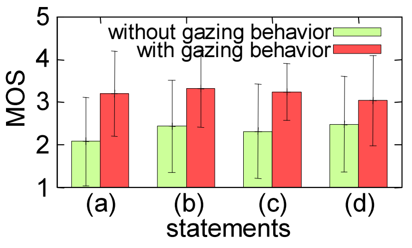


Figure 16. Evaluations for Motion of Joint Attention

Table 6. ANOVA of Evaluations for Motion

	(a)	(b)	(c)	(d)
F-value	23.1	27.9	23.3	7.19
p-value	<.01	<.01	<.01	.013

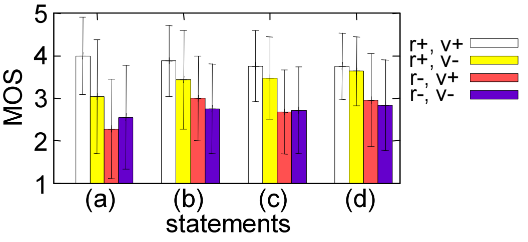


Figure 17. Evaluations for Gaze-correspoinding Guide

a) the anthropomorphic gazing direction of the stuffed-toy robot and b) the appropriate contents of vocal guidance. To confirm the effectiveness of both, we compared the following four conditions based on the following hypotheses.

Hypotheses: I) The correspondance of the robot's gazing direction to the participant's gaze affects on the user's impresson for the guide itself. II) The correspondence of the vocal guidance to the direction to the participant's gaze affects on the user's impression for the guide itself.

Conditions: In order to verify two different criteria, gaze-corresponding vocal guidance and gaze-corresponding direction of the robot's head, we combined the factors and prepared four conditions. We made a comparative analysis among the variables of correspondence (+) and noncorrespondence (-) as well as guidance with joint attention (looking at the same place with the participant, **r**) and using vocal guidance (appropriate content of the user's gaze, **v**). Here, **r+** means the robot's gaze corresponding to the direction of the user's gaze, and **r-** means the robot gaze behavior in the opposite direction to the user's gaze. In addition, **v+** means the vocal guide gives the appropriate contents for the gazed figure, and **v-** means the voice guide of inappropriate contents. Accordingly, our prepared condition **r+ v+** denotes that the robot gazes at the same figure as the subject looks at and talks about the figure by vocal guidance. Condition **r+ v-** denotes the robot gazing at the same figure as the subject while speaking explanation of the content of the opposite figure. In condition **r- v+**, the robot gazes at the opposite figure but speaks guidance corresponds to the subject's gazed figure. In condition **r- v-**, both the robot's gazing direction and the content of vocal guidance correspond to the opposite figure of the subject's gaze.

Stimuli: The preceding vocal guidance, "I would like to talk about these foods," was made before the stimuli. The next vocal guidance was the explanation of the contents. In this experiment, four guidances assuming the left or right region were vocalized; 1) "this is a triangle shortcake," 2) "this is a square tiramisu," 3) "this is a triangle Onigiri," and 4) "this is a square lasagna," in association with the behaviors of the robot corresoponding to each condition.

Experimental Results: Figure 17 shows the summary of MOS (means opinion scores) for each statements. Table 7 shows the results of two-factor ANOVA with repeated measurements (degree of freedom (Dof,ϕ) = 24, rate of rejection (α) = 0.05). The results are marked with underlines for each significance, with * for significant tendencies. Factor 1, the difference in the correspondence of the robot's gazing direction, shows significance in each statement. On the other hand, Factor 2, the difference in the correspondence of the vocal guidance, shows weak significant tendencies only in statements (a) and (b). Interaction between factors found in statement (a) shows that vocal guidance was not effective while the robot is not looking at the same area with the subject. These results show that the quality of the guidance obviously increases when the robot's gazing direction is corresponding to the user's gazing position. Compared with the results, the correspondence of the vocal guide's contents show only somewhat significant tendencies. Consequently, Hypotheses I) and II) were confirmed. It is conjectured that the gaze-corresponding guide provides positive impressions of the user and that there is a strong effectiveness of the gaze-corresponding behavior of the robot rather than the guide contents.

6. Discussions

Now we discuss the system evaluations. First, we could confirm the system operation in total with both the remote and ambient gaze-tracking method and the controls of the robot corresponding to the estimated target of the gaze. As a basis of the user interface, our gaze estimation method handles various and unrestricted behaviors of the user and realizes accurate gaze tracking in a real space, contrary to traditional gaze-tracking systems. The proposed system facilitates natural and gradual guidance by employing anthropomorphic gaze-communications between the user and the robot. The demonstration experiments were performed for different scenes from that of the prototype system, and we could confirm that the configuration of the guide system can perform attractive experiences for the general public.

Next, we evaluated a) the effectiveness of communicative guide robot adopted for a guide plate, and b) the effectiveness of the robot's gaze behaviors on eliciting user's interest based on our gaze-communicative model for human-robot interaction.

From the experiments described in Section 5.1, the elemental factors of the gazing behaviors were revealed. The objective evaluations, observing the participant's gazing durations and directions of gaze-switches, revealed that the subject's gaze are significantly affected to be drawn by the direction of the robot's gaze. It is conjectured that the interests of

Table 7. Two-factor ANOVA of Gaze-correspondence

comparison		(a)	(b)	(c)	(d)
rp vs op	$F_{(1,24)}$	11.6	14.3	23.8	14.2
(direction)	p	<.01	<.01	<.01	<.01
rv vs ov	$F_{(1,24)}$	3.68	4.11	0.61	1.13
(voice)	p	0.067	0.054	0.442	0.298
interaction	$F_{(1,24)}$	9.17	0.342	1.30	0
	p	<.01	0.563	0.266	1

the subjects were subconsciously drawn by the robot's gazing behavior while the user's gaze orientation is elicited.

From the experiments described in Section 5.2, it is clear that there are positive effects on the impression of the quality of guidance according to the evaluated factors. It is conjectured that there is a beneficial effect of gaze-communicative guidance for increasing the attractiveness of a guide system and reliability of the guide, based on positive impressions such as affection and naturalness.

Consequently, we could confirm the basic operation of the guide system, the effectiveness of gaze-corresponding guide by the anthropomorphic robot in the guide plate system, and the effectiveness of gaze behaviors in eliciting users' interests. Although the settings and instructions of the experiment were different from the real situations, individual effectiveness should be considered as one of the criteria for the contents design of applied scene. The experiments by various ages and situations of the user should be helpful for the adjustments of detailed guidelines.

The verified effectiveness of the gazing behaviors is expected to be employed for building an enriched interaction model of embodied robotics, since anthropomorphic representations have various multiple modalities. The robot's gazing behaviors and the conjunction with an interactive display board must be not only sophisticated but also comprehensible to the target user. In present implementation, the illumination on the board is turned on and off by a timing of the guide system that corresponds to the user's gaze. In order to improve the system, we should develop more appropriate and natural coordination between the display component and the robot's gazing behavior in a first step toward improvement.

In the past decade, neural mechanism of perceiving gaze or attention has been shown by recent studies [27, 28]. Analyses of brain activities for our studies are expected to verify the process of the gaze communication with artificial presence. More reliable evaluation would help to design appropriate gazing behaviors of the robot.

7. Conclusions

This paper proposed a gaze-communicative guide system consisting of a stuffed-toy robot and a display board to provide users detailed information of the contents in the guide plate. The stuffed-toy robot attracts her/him with natural gazing behaviors (joint attention and eye-contact reactions) corresponding to the user's gazing position, and guides the user to detailed information. The gazing orientation of the user is estimated using our remote gaze-tracking method without any fixture or attachment. The guide system is expected to be applicable not only for guide assitance in elderly care but also for information boards in amusement parks, supermarkets, or advertisement boards. We confirmed the sufficient accuracy of gaze-tracking method and the system operation through demonstration uses in semipublic spaces. The results of elementary effectiveness of gaze behaviors of the robot add the strength of the positive results with suppositions as the gazing direction of the stuffed-toy robot can draw the user's gaze. Our system evaluations showed that the anthropomorphic presence, the anthropomorphic behaviors, and gaze-corresponding interaction of the gaze-corresponding guide make a positive impression on the users of the guide system.

These results present the possibility of removing the psychological or physical burdens that are actually experienced by disabled people or individuals who are hesitant to ask for further detailed information beyond that on the guideboard. The capabilities of the anthropomoprhic media are well suited to our focus of supporting wandering people in semipublic spaces.

As future work, we are going to focus on some domain-specified users. We should consider not only actual implementations with the robot's gazing behaviors during vocal guidance and simple illumination but also other expressive modalities such as sign language by the robot and interactive information on the board; our intention here is to extend the range of target users to deaf people. Additional expressiveness of the robot's behaviors would be very effective for disabled people missing the use of various communication channels.

Acknowledgements

This research was supported in part by KAKENHI 25700021 and KAKENHI 24300047, and the basis of the research achievement was supported in part by the Ministry of Internal Affairs and Communications of Japan when the author was in ATR IRC.

References

[1] Arrington Research. Viewpoint eyetracker. http://arringtonresearch.com/index.html.

[2] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, and B. Blumberg. Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Autonomous Robots*, 11(Issues 1-2):31–62, 2005.

[3] Wolfram Burgard, Armin B. Cremers, Dieter Fox, Dirk Hahnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, 114(1-2):3–55, 1999.

[4] B. R. Duffy. Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3):177–190, 2003.

[5] A. Fukayama, T. Ohno, N. Mukawa, M. Sawaki, and N. Hagita. Messages embedded in gaze of interface agents – impression management with agent's gaze–. *Proc. ACM SIGCHI2002*, 1:41–49, 2002.

[6] M. Imai, T. Ono, and H. Ishiguro. Physical relation and expression: Joint attention for human-robot interaction. *IEEE Int. Workshop on Robot and Human Communication*, pages 512–517, 2001.

[7] Lewis Johnson, Jeff Rickel, Randy Stiles, and Allen Munro. Integrating pedagogical agents into virtual environments. *Presence: Teleoperators and Virtual Environments*, 7(Issue 6):523–546, 1998.

[8] Y. Katagiri, T. Takahashi, and Y. Takeuchi. Social persuasion in human-agent interaction. *Second IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, IJCAI-2001:64–69, 2001.

[9] A. Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22–63, 1967.

[10] H. Kojima. Infanoid: A babybot that explores the social environment. *Socially Intelligent Agents*, pages 157–164, 2002.

[11] Y. Matsumoto and A. Zelinsky. An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. *Proc. Int. Conf. Automatic Face and Gesture Recognition*, pages 499–504, 2000.

[12] C. Moore, P. J. Dunham, and P. Dunham. *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum, 1995.

[13] D. Moore. 'it's like a gold medal and it's mine' - dolls in dementia care. *Journal of Dementia Care*, 9(6):20–22, 2001.

[14] J. Pineau. Towards robotic assistants in nursing homes: challenges and results. *Robotics and Autonomous Systems*, 42(Issues 3-4):271–281, 2003.

[15] M. Pollack, S. Engberg, J.T. Matthews, Sebastian Thrun, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, J. Dunbar-Jacob, C. McCarthy, Michael Montemerlo, Joelle Pineau, and Nicholas Roy. Pearl: A mobile robotic assistant for the elderly. *Workshop on Automation as Caregiver: the Role of Intelligent Technology in Elder Care (AAAI)*, 2002.

[16] L. Goldman. Counseling With Children in Contemporary Society, *J. Mental Health Counseling*, 26(2):168–187, 2004.

[17] T. Shibata, M. Yoshida, and J. Yamato. Artificial Emotional Creature for Human-Machine Interaction. *Systems,Man,and Cybernetics*, pp. 2269–2274, 1997.

[18] C. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich. Explorations in engagement for humans and robots. *Artificial Intelligence*, 12(Issues 1-2), 2005.

[19] Y. Sumi, T. Etani, S. Fels, N. Simonet, K. Kobayashi, and K. Mase. C-map: Building a context-aware mobile assistant for exhibition tours. *Community Computing and Support Systems*, LNCS 1519:137–154, 1998.

[20] Will Taggart, Sherry Turkle, and Cory D. Kidd. An interactive robot in a nursing home: Preliminary remarks. *CogSci 2005 Android Science Workshop*, pages 56–61, 2005.

[21] A. L. Thomaz, M. Berlin, and C. Breazeal. An embodied computational model of social referencing. *ROMAN 2005*, pages 591–598, 2005.

[22] Hiroko Tochigi, Kazuhiko Shinozawa, and Norihiro Hagita. User impressions of a stuffed doll robot's facing direction in animation systems. *Proc. of ICMI07*, pages 279–284, 2007.

[23] H. Yamazoe, A. Utsumi, T. Yonezawa, and S. Abe. Remote gaze direction estimation with a single camera based on facial-feature tracking without special calibration actions. *Eye-tracking Researches and Applications*, 2008. to appear.

[24] T. Yonezawa, N. Suzuki, S. Abe, K. Mase, and K. Kogure. Cross-modal coordination of expressive strength between voice and gesture for personified media. *Proc. of ICMI06*, pages 43–50, 2006.

[25] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe. Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze-tracking. *Proc. of ICMI07*, 2007.

[26] Y. Yoshikawa, K. Shinozawa, H. Ishiguro, N. Hagita, and T. Miyamoto. The effects of responsive eye movement and blinking behavior in a communication robot. *Proc. IROS2006*, pages 4564–4569, 2006.

[27] Stephen V. Shepherd, Following gaze: gaze-following behavior as a window into social cognition, *Frontiers in Integrative Neuroscience*, Vol.4, Art.5, 2010.

[28] Lauri Nummenmaa, and Andrew J. Calder, Neural mechanisms of social attention, *Trends in cognitive sciences*, Vol.13, No.3, pp. 135 -143, 2009.