The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Improving Video Communication for Elderly or Disabled by Coordination of Robot's Active Listening Behaviors and Media Controls

Tomoko Yonezawa†, Yuichi Koyama†,‡, Hirotake Yamazoe†, Shinji Abe† and Kenji Mase‡

Abstract— **In this paper, we propose and evaluate a video communication system that compensates for user's uncongenial attitudes by coordinating the robot's behaviors and media control of the video. The system facilitates comfortable video communications between elderly or disabled people by an assistant robot for each user that expresses *a*) active listening behaviors to compensate for the listener's attitude when he/she is not really listening to another user's talking and *b*) a cover-up behavior (gaze turned to the user) to divert attention from the other user's uncongenial attitude when that person is not looking at the talking user but toward the robot at her/his side; this behavior is performed by coordinating the automatic switching of cameras to give the impression that the congenial person is still looking at the user. The results obtained in the system evaluation show the significant effectiveness of this design approach using the robot's behavior and media control of the video to compensate for the problems in video communication that we aimed to overcome.**

I. INTRODUCTION

Our aging societies suffer from a serious deficiency in the communication abilities of elderly people. People who have disabilities sometimes confront a similar problem due to the labor shortage in welfare services. Such citizens are not well served by the system interfaces of most IT solutions, which focus on new technologies for industry and commerce On the other hand, the conventional telephone is a long-established communication method. For the elderly and disabled who live with somebody or feel not so lonely, the telephone is nearly sufficient for communication with others. However, there are many elderly and disabled people who live alone and feel lonely.

A previous trial of videophone communication, between an elderly person and an active-listening volunteer sharing the contents of the elderly user's memories [1], showed the effectiveness of using a videophone to simulate a real conversation. However, the human-resource problem was not resolved because this system needed the volunteer and an assistant at the user's side. Consequently, the need for elderly-elderly or disabled-disabled communication arises, and it is likely that the two groups share common problems in practical communications.

Meanwhile, many pet robots and communication robots have been designed as partners for reducing their loneliness. To an extent they fulfill this purpose, but human-human communication is still the most effective type of communication to satisfy the social appetite of elderly or disabled people living alone. In such cases, the partner robot should become a kind of bystander that assists its owner.

People sometimes lose their concentration, especially the elderly suffering from dementia. Normally, if a participant's attention changes toward other things during the conversation, the communication cannot be sustained for a long period. To compensate for such extreme cases in communication, it is important to prepare a comfortable communication environment with appropriate assistance. In this research, we aim to develop an anthropomorphic assistance using a stuffed-toy robot for real-time video communication on the Internet. Accordingly, we integrate the remote communication support system for elderly people using not only behavioral controls of the robot but also media controls to display the participant's appearances in such a way that compensates for the other person's possible lack of attention.

In order to make the participant feel comfortable even when another person is not concentrating on the video communication, this system incorporates a paradigm of changing the direction of interests in communications that adopts i) *active listening behaviors* (nodding and back-channel feedback) of the robot, ii) *automatic switching* of two cameras' images, and iii) behaviors of the robot for "*covering-up*" the other user's inattention by gazing at the user during switching of the camera images. The automatic switching chooses the camera that views the image of the user's front face so that the system and the robot give an illusion that makes the primary user feel as if she/he is gazed at by the other user even when the other's attention is not actually attracted to the TV conversation. Therefore, each user can freely talk to the other person or to the robot sharing her/his location.

Although there are actually four presences (2 users and 2 robots), each user is not aware of the other user's robot, so we can say that this configuration forms an atypical triadic relation among the user, the other person in the 2D monitor, and the robot in the 3D real world. Consequently, it is important to verify the effectiveness of the controls of the robot and the remote user's image. For verification of our proposed system including the triadic relation, we discuss whether the proposed components of the system and their coordination provide the user's positive impressions through empirical evaluations for each element.

II. RELATED RESEARCH

Based on the effectiveness of the real-world anthropomorphism, we have developed gaze-communication based robot systems for activating the elderly communication [2], [3]. In this paper, we propose a video communication system with

† ATR IRC Labs. 2-2-2 Hikaridai, Seika, Soraku, 6190288 Kyoto, JAPAN. {yone,yamazoe,sabe}@atr.jp
‡ Nagoya University Yoro-cho, Chikusa, Nagoya, 4648601 Aichi, JAPAN. {koyama,mase}@itc.nagoya-u.ac.jp

a stuffed-toy robot using its gazing behaviors in order to support and improve these communication and relationship between the users.

The trial of videophone network for one year by an elderly community achieved satisfaction of members who used the videophone everyday [4]. The videophone conversation with a volunteer using contents of user's original memory was effective as same as the volunteer's real visit [1]. Thus there are demands and effectiveness of the elderly communication through videophone, however, there are still the problem of labor shortage.

On the other hand, there are many works using robot for media and information support [5], [3, etc.]. These systems need the coordination of the robot's behavior and the media; the latter system uses a projection corresponding to the user's gazing position with a direction of the robot. The service robots can collaborate with media controls like them to bring the user immersive experience. There are also many researches about facing and gazing directions of the robot [6], [7, etc.] and virtual agents [8]. Not only anthropomorphic behaviors but also the physical relationship of human and robot are discussed [9], [10, etc.]. Their results indicate that the various aspects of the user's impression of communication can be changed by controls of these factors. Therefore, it is important to build the video communication system with considering the effective use of these factors.

In the video communication with a robot, the users feel a triadic relation from the viewpoint of each other. There are many discussions on triadic relation among human and robots [11], [12], [13, etc.]. In order to make our proposed system more effective, the balance between the remote-side person in the monitor and the robot's presence should be considered to improve the relation between the users in the video communication.

III. SYSTEM IMPLEMENTATION

A. Concept of Proposed System

In the proposed system, we focus on communication assistance. People's communications are not necessarily good. In video communication, since available channels are limited, the reality/quality of the communication deteriorates even further.

On the other hand, an increasing number of solitary elderly people have few chances for communication. Even if video communication is a limited form of communication, it is very useful for such elderly and mildly demented patients and offers real potential to help their communications. Therefore, we propose an improved video communication system that utilizes an anthropomorphic medium to help the user's communication and coordination/interaction among other media. In communications, a user's feelings can be hurt due to misunderstandings and a decrease in the communication partner's attention. To deal with this problem, we implement a stuffed-toy robot to compensate for the shortcomings of video communication. In addition, re-connectivity of the communication is supported by the controlling medium.

In the proposed system, a robot exists in the 3D space

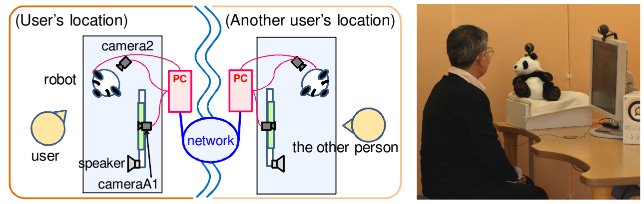

Fig. 1. The system implementation and view of system use

Fig. 2. Usual performance

Fig. 3. Video without image switching

Fig. 4. Image switching and cover-up behavior

of each user while the companion (other user) is displayed in 2D images, and under this scenario we assume that the robot gains greater influence. Furthermore, since the relation of the users and each robot forms a triadic relation from the viewpoint of each user, we also assume we can variously control the user's impressions of the other-side person and the conversation itself by changing the coordination of robot behaviors and media controls. Such a configuration is necessary to develop a continuous communication system.

B. Overview and Implementation of the System

Figure 1 shows the configuration of our proposed system. Our system consists of a TV conference system (a monitor, a USB camera (monitor camera), and a microphone), a stuffed-toy robot for conversation assistance, and another USB camera (robot camera). The stuffed-toy robot is placed equidistantly from the user and the monitor. The robot camera is placed just above the stuffed-toy robot. The monitor camera and robot camera are used for image switching and detecting the user's face direction. Since both cameras capture only a user's face region, users cannot see the remote-side robot.

We implemented the following two functions of the proposed system to assist the user's conversation.

i) Active Listening Behavior of Robot: Unlike active listening volunteers, it is difficult for elderly people and especially dementia patients to concentrate on conversation for a long time. Therefore, we introduced a stuffed-toy robot in our video communication system. The robot listens to the user's speaking along with the remote user. By employing the robot, we assume that the user can easily speak even when the remote user doesn't concentrate on the conversation.

To show the robot's interest in the user's speaking and sympathetic responses, we implemented two types of robot behavior as active listening behavior. *(i-a) Robot gazing behavior based on user's gaze:* the robot looks in the same direction as the user (Joint attention). When the user looks at the robot, the robot make eye-contact with the user. *(i-b) Robot nodding/back-channel feedback based on the user's utterance:* when the user says something (not talking to the robot), the robot only nods towards the user so as not to disturb the user's conversation. When the user talks to the robot, the robot gives back-channel feedback to the user.

Gazing/facing directions of the user are estimated as

Fig. 5. Views of the demonstration experiment

follows. First, we perform face-detection with both monitor camera and robot camera based on Haar-like features [14]. Comparing face-detection scores, we assume that the user is looking in the direction of the camera that detects the frontal face. Utterances of the user are detected by threshold processing based on audio signals. Based on these results, the behaviors of the stuffed-toy robot are determined. Thus the robot listens to the user while the user talks to the remote user, listens to the remote user while the remote user talks, and replies to the user when the user talks to the robot.

ii) Image switching and Cover-up Behavior of Robot: When the user loses concentration and doesn't listen to the conversation, the remote-side user's attention might move to the robot and she/he may start talking to the robot because the robot is performing active-listening behaviors. In such situation, since the user feels that the remote-side user is not talking to her/him (Figure 3), there is the possibility that the user will further lose the interest in the video communication.

To solve these problems, we introduce *(ii-a) image-switches:* the remote-side monitor image is switched between the two cameras based on the primary user's face direction in order to make the remote user feel that the primary user is always looking at her/him (Figure 4). When the image switching occurs, the background of the remote-side user might change and become discontinuous. The appearance and direction of the remote-side user's face might also change. These problems can cause a feeling of strangeness in the user. Therefore, we also introduce *(ii-b) cover-up behavior of the robot:* the robot's behaviors are controlled to reduce the feeling of strangeness. The head direction of the user-side robot is changed in conjunction with the image switching. As shown in Figure 4, the user-side robot looks at the user (makes eye-contact with the user) for one second to get the user's attention before the image switch. It is assumed that these behaviors make it difficult for the user to notice the image switching.

C. Demonstration Experiments in Day-care Center

We performed demonstration experiments of the proposed system in a day-care center for one week. A set of our system was installed in the center. Twelve elderly people participated in the experiment. They talked remotely for about fifteen minutes per session using our system. The total number of experimental session is twelve. Figure 5 shows the appearance of the demonstration experiments.

The following events were observed during experiments. In a situation in which a user was not looking at the other user, the user returned to the conversation while the other user was talking toward the robot, and the conversation between the users continued. This is an example of the effect of the robot's active listening behavior. When a user said to the robot "Thank you for listening to my story," the other user, watching the switched image, thought that the user was talking to him, and answered, "You're welcome" (Figure 5(A)). This example shows the effectiveness of the image switching. Some users nodded in response to the robot's nodding (Figure 5(B)), others petted the robot. We assume they made these actions because the robot's active listening behavior gave the users a sense of affection toward the robot.

As described above, the proposed system has the effect of continuing conversation between elderly people, and giving the users a sense of affection toward the robot, among other possible effects.

IV. EVALUATIONS

In this section, we verify the effectiveness of our system design with A) the active listening behaviors of the robot in order to build smooth communication that compensates for the other person's behaviors, and B) the automatic switching of camera images coordinated to the robot's cover-up behavior designed to distract the user's attention from the image switching.

Settings in both experiments: A common experimental co-agent participated in the system evaluation as the other person talking with the subject in video communication. The co-agent was instructed to behave corresponding to the conditions. In these experiments, the setting adopts extension cables for two-way voice and video signals between the room of the co-agent and that of the subject, allowing the experimental system to avoid any delay of the signals which occurs in a real network system. The system works according to the assigned conditions.

A. Active Listening Behaviors of Robot

Purpose of experiment A: In this evaluation, we verify whether there are positive effects of the robot's active listening behaviors on the video communication. The evaluation includes another factor, i.e., the attitude (direction of gaze) of the other person, to investigate *the triadic exchange relations* among the user, the other person, and the robot.

Hypotheses: *A-H1*) The attitude of the active listening behaviors of the robot do not affect the impressions of the video communication (neither toward the robot, the other person, nor the conversation itself). *A-H2*) The attitude of the other person does not affect the impressions of the video communication (neither the robot, the other person, nor the conversation itself).

Subjects: 34 people from 18 to 26 years old (13 females and 21 males, average age is 21.29 years).

Conditions: There are two factors in this experiment. The first factor is the other's attitude, i.e., whether she/he shows continuous gaze at the participant (`Op`) or looks off in the side direction toward the camera (`On`). The second factor is the robot's attitude, i.e., whether it shows active listening behaviors with nods and gaze at the user (`Rp`) or nods and gazes in randomized timing and directions (`Rn`). Therefore, the conditions combining these factors are four types, for example `RpOp`.

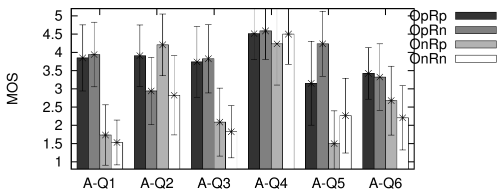

Fig. 6. Results of MOS in user evaluations (Ex.A)

TABLE I
TWO-FACTOR ANOVA RESULTS FOR MOS IN EX.A

factor item	the other person $F_{(1,32)}$	p	robot $F_{(1,32)}$	p	interaction $F_{(1,32)}$	p
A-Q1	171.0	<.001	1	0.325	3.347	0.077*
A-Q2	0.368	0.549	49.3	<.001	3.516	0.070*
A-Q3	102.0	<.001	1.27	0.269	1.763	0.194
A-Q4	1.19	0.284	0.074	0.788	0.725	0.401
A-Q5	72.07	<.001	34.98	<.001	3.99	0.054*
A-Q6	36.0	<.001	3.256	0.081*	3.347	0.077*

Procedures: The co-agent gives a general topic such as transportation conditions around the subject's house (5[sec]), and the subject answers to the topic in about 25[sec]. During the conversation, the co-agent shows his attitude (`Op`/ `On`), and the robot shows its attitude (`Rp`/`Rn`) corresponding to the condition. The image of the other person disappears from the monitor after 30[sec] from the beginning of the conversation. The experiments in different conditions were held in repeated measurements for each subject with mixed orders of conditions. The prepared topics are used in the same order, so the interaction between the conditions and the topics are counter-balanced.

After each experiment, the subject uses a five-point rating scale to evaluate the relevance (5: very relevant, 4: somewhat relevant, 3: neutral, 2: somewhat irrelevant, 1: irrelevant) of the statements below.

A-Q1 The other person seemed to be conversable.
A-Q2 The robot showed desirable behaviors.
A-Q3 It was easy to talk to the other person.
A-Q4 You felt an undisturbed conversation with the robot.
A-Q5 The other person's interest in you seemed to be greater than the robot's interest in you.
A-Q6 You would like to have communication in this condition.

Instructions: The participant was instructed i) to talk about the topic that the other person brings up at the beginning of the conversation until the other's image disappears (after about 30 seconds), and ii) to answer the evaluation items for each session.

Results: There are missing data of one subject. Figure 6 shows MOS (means opinion score) averages and standard deviations for each evaluation item, and Table I shows two-factor ANOVA results with the significance level (α) = 0.05 and DOF (degree of freedom, ϕ = (1,32)). The underlined p values mean significance, and $\star$ marks indicate a significant tendency.

First, the statement A-Q6 implies the possibility of repeated use of the system. The results show a significance between `Op`/`On` and significant tendencies between `Rp`/`Rn` with both factors' interaction. The effect of interaction is especially noticeable in the case of `On`. The results for A-Q4 do not show any significance, which indicates that the active listening behaviors of the robot do not disturb the video communication. The results of A-Q3 simply show the significance of the difference between `Op` and `On`. A-Q5 is a comparison between the active listening behaviors

TABLE II
SUMMARY FOR NUMBER OF TIMES LOOKING AT ROBOT

OpRp ave.	(stdev.)	OpRn ave.	(stdev.)	OnRp ave.	(stdev.)	OnRn ave.	(stdev.)
1.088	(1.288)	0.588	(1.158)	1.647	(1.721)	0.971	(1.487)

TABLE III
ANOVA RESULTS FOR NUMBER OF TIMES LOOKING AT THE ROBOT

factor item	the other person $F_{(1,33)}$	p	robot $F_{(1,33)}$	p	interaction $F_{(1,33)}$	p
number	5.651	.0234	15.1	<.001	.360	.552

of the other person and the robot. The results clearly show significant differences by both factors, as well as a significant tendency in their interaction. `OnRn` and `OpRp` show different results: When there is no active listening behavior, neither by the other nor by the robot, the robot's presence was interpreted as showing stronger interest than that of the other user in the monitor.

In A-Q1, not only the significance of the factor `Op`/`On` but also a significant tendency in the interaction between the factors (`Op`/`On` and `Rp`/`Rn`) was confirmed. On the other hand, the results in A-Q2 show not only the significant difference between `Rp`/`Rn` but also a significant tendency in the interactions of the two factors. The results of A-Q1 indicate a positive effect of the robot's active listening behavior on the impression toward the other person, and the results of A-Q2 suggest a positive effect of the robot's active behaviors especially when the other person is not looking at the subject.

Next, we observed the number of times each subject's gaze fell on the robot. Table II summarizes the results in each condition and Table III shows the two-factor ANOVA results. These tables show significant results of both factors (`Op`/`On` and `Rp`/`Rn`) and a significant tendency between them. The results directly contradict the results of A-Q5 on the other's interest compared to the robot's interest.

From these results, both hypotheses *A-H1*) and *A-H2*) were rejected, and the effectiveness of each factor in this experiment was confirmed.

B. Media Control and Coordination with Robot's Behavior

Purpose of experiment B: In this evaluation, we verify *1)* whether the image switching between the cameras, when the other person looks at the remote-side robot, has positive effectiveness on the video communication, and *2)* whether the robot's gaze coordinated with the timing of the image switch has positive effects on the user's attention to the switch.

Hypotheses: *B-H1*) The switching of the two camera images does not affect the impression of the video communication (neither toward the robot, the other person, nor the conversation itself). *B-H2*) The robot's gaze behavior, which tries to draw the user's attention from the switching of the camera images, does not affect the impression of the video communication (neither toward the robot, the other person, nor the conversation itself).

Subjects: Same participants as used in Experiment A.

Conditions: There are two factors in this experiment. Factor 1 is the switching between the two camera images, which show either the other's gaze at the participant regardless of the other's real gazing direction (`Sp`) or the other person just looking off into the side direction (`Sn`). Factor 2 is the

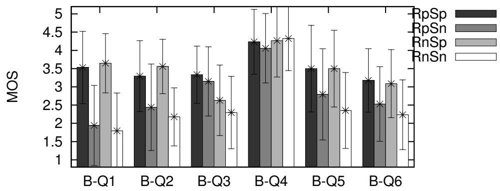

Fig. 7. Results of MOS in user evaluations (ex.B)

TABLE IV
TWO-FACTOR ANOVA RESULTS FOR MOS IN EX.B

factor	robot		camera switch		interaction	
item	$F_{(1,33)}$	p	$F_{(1,33)}$	p	$F_{(1,33)}$	p
B-Q1	0.024	0.879	93.3	<.001	3.476	0.071*
B-Q2	0.016	0.899	40.89	<.001	10.82	0.002
B-Q3	22.65	<.001	3.225	0.084*	.026	.873
B-Q4	1.539	0.223	0.314	0.579	1.291	0.264
B-Q5	3.741	0.062*	21.85	<.001	3.103	0.087*
B-Q6	3.279	0.079*	19.96	<.001	1.417	0.242

∘ the item B-Q3 have 7 missing data, so we use $F_{(1,26)}$ for 27 samples.

robot's behavior, either turning its gaze to the user during the other's gaze turn toward the side direction (Rp) or the motionless behavior during the timing of this action (Rn).

Procedures: The subject asked the other person about an assigned general topic, which was given in advance for each experiment such as their favorite TV program (about 5[sec]). The co-agent (other person) replies to the question for about 25[sec]. Next the co-agent asks the question back to the subject (about 3[sec]), and the subject replies to the question until the other's image disappears (about 27[sec]). The co-agent turns his face to the left side (60[deg]) from the front in order to simulate the other's decreased concentration. When the co-agent looks away from the subject, the system changes its performance of the switching between the two camera images (Sp/Sn), and the robot changes its behavior (Rp/Rn, i.e., with or without its cover-up behavior) corresponding to each condition. Experiments under different conditions were conducted in repeated measurements for each subject with a mixed order of conditions. The prepared topics are used in the same order, so the interaction between the conditions and the topics are counter-balanced. The subject uses a five-point rating scale (the same as for experiment A) for each statement below.

B-Q1 The other person seemed to be interested in your talk.
B-Q2 It was easy to talk to the other person.
B-Q3 The robot showed desirable behaviors.
B-Q4 You felt an undisturbed conversation with the robot.
B-Q5 There was no unreasonable feeling in the conversation.
B-Q6 You would like to have communication in this condition.

Instructions: The participant was instructed i) to ask an assigned topic to the other person, ii) to answer when the other asks the question back until the other's image disappears (for about 30 seconds), and iii) to answer the evaluation items for each statement.

Results: The data of 7 persons were missing in B-Q3. Figure 7 shows MOS averages and standard deviations for each evaluation item, and Table IV shows the two-factor ANOVA results, with the significance level (α)=0.05 and DOF (ϕ=1,33) except B-Q3 (ϕ=1,26).

The results in B-Q5 show a significance in Sp/Sn with significant tendencies of Rp/Rn and the interaction of both factors. The robot's cover-up behavior was especially effective in Sn when the co-agent looked away. The statement B-Q6 indicates the possibility of repeated use of the system. The results show a significance between Sp and Sn and significant tendencies between Rp/Rn. The switching of the two camera images affects the positive impression of the video communication with a slight effect of the robot's cover-up behavior. The results in B-Q4 do not show any significance, which indicates the cover-up behavior of the robot does not disturb the video communication.

In B-Q1 and B-Q2, there were significances in the switching factor (Sp/Sn), indicating that the camera switch for the front face of the other person is effective when other's positive attitude appears, in line with our aim. There were also a significant tendency (B-Q1) and a significance (B-Q2) for the interaction of the two factors but without any significance in the robot's factor (Rp/Rn). On the other hand, the results of B-Q3 show a significance of the robot's cover-up behavior and a significant tendency of the switch factor, although the statement mentions only the robot's attitude.

From these results, both hypotheses *B-H1*) and *B-H2*) were rejected, and the effectiveness of each factor in this experiment was verified.

V. DISCUSSION

The robot's behaviors and the media controls in our proposed system are designed expecting smooth and comfortable video communications. As a basis of the system evaluation, the results of A-Q4 and B-Q4 clarify that the designed behaviors of the robot do not disturb the subjects.

A. Active listening behaviors of the robot

In experiment A, we verified the effectiveness of a-i) the robot's active listening behavior through different situations of a-ii) the other person's attitudes.

First, A-Q2 shows the effectiveness of robot's attitude on the impression of the robot itself, especially when the other does not show a congenial attitude. A-Q1 and A-Q3 show natural results on the importance of the other person's attitude for the impression of her/him when the user talks to her/him. A-Q1 shows the possibility of the robot giving assistance while the other person does not concentrate on the conversation.

Statement A-Q6 denotes the potential of the continuous use in this configuration. The results in the statement show a strong effect of the other's attitude, while the robot's active listening behaviors differ from its random behaviors with a significant tendency. It is conjectured that the active listening behaviors of both the other user and the robot affect the user's impressions. Furthermore, the robot's behavior is especially effective when the other user does not look at the subject. Statement A-Q5 and the number of times the subjects look at the robot (Table II and III) show not only significant results by the factors of the robot's and the other's active listening behaviors; in addition, the user's interest toward the robot and gaze at the robot are affected by the balance of the factors, as we expected. From these results, OnRn and OpRp show different tendencies: When there is no active listening behavior, by neither the other user nor the robot, the robot's presence was interpreted as eliciting stronger interest than the presence of the other user in the monitor, contrary to the results when they both showed active listening behaviors. Both A-Q5 and A-Q6 results imply that the effectiveness of the robot's active listening behaviors grows when the other's attitude is negative or not so appropriate. This tendency is what we aimed toward in this system implementation.

From these results, the effectiveness of a-i) and a-ii) is confirmed. In this triadic relation among the user, the other person, and the robot, the positive attitude of the other person is a strong factor in achieving better communication, but the robot's attitude can still recover the user's positive impression to make a comfortable communication setting.

B. Switching of cameras and robot's cover-up behavior

In experiment B, we verified the effectiveness of b-i) the switching between two camera images corresponding to the other's face direction, and b-ii) the robot's cover-up behavior to camouflage the change in the other's face direction or the switching of the camera images when the other person does not look at the subject when she/he talks.

From the results of B-Q1 and B-Q2, there is an apparent advantage of the switching between the two camera images which always shows the other person's frontal face, although this other person may be looking away from the front view.

B-Q3 provides some intriguing results; not only the robot's cover-up behavior but also the image switching can affect the subject's impression toward the robot. On the other hand, the robot's cover-up behavior affects not only the impression of the robot but also that of the communication itself as seen in the results of B-Q5 and B-Q6. Moreover, the significant results of the interactions in experiment B suggest that the robot's cover-up behavior properly compensates for the other's inappropriate attitude. This tendency is very similar to the results in experiment A.

From these results, the effectiveness of b-i) and b-ii) is confirmed. The robot's cover-up behavior was designed for diverting the user's attention from the switch and the change in the other's interest. This function can help avoid the user's hurt feelings by diverting her/him from the other's uncongenial attitude or appearance.

C. Summary of Discussions

First, our proposed controls of the robot's behavior are effective in improving comfortable communications by expressing its active listening attitude. Next, the media controls (switching of two camera images), which simulate the other person's active listening attitude, sufficiently generate the user's impression of receiving adequate attention.

The system configuration, which integrates the coordination of the robot's behavior and the media control, provides an atypical triadic relation among the user, the other person on the monitor, and the robot. From the results of the experiments, the robot's role in the triadic relation explicitly appeared in the effectiveness of both the robot's behaviors and the media controls. The robot is not a competitive presence with the other person on the monitor. For this reason, in the case of the other's uncongenial attitude, the appropriate coordination of the robot's behavior and the media control is considered a strong method for providing comfortable communication using videophones.

VI. CONCLUSION

In this paper, we proposed and evaluated a video communication system that compensates for the participants' uncongenial attitude by coordinating the robot's behaviors and the media control of video images. The system provides comfortable communications between elderly people by using a robot that appropriately behaves corresponding to the user's situation: a) when the other person is not really listening to the user's talk, the robot expresses its active listening behaviors in order to compensate for the listener's attitude; and b) when the other person is not looking at the user but at the remote-side robot, the image on the monitor switches to the remote robot-view camera to show the user as if the other person is still looking at her/him, and at the same time, the user-side robot expresses a cover-up behavior (gaze turned to the user) in order to divert attention from the other's uncongenial attitude.

The results in the system evaluation show that these designs have positive effects on the video communication, as we expected. We conclude that the effectiveness of the robot's behavior can shape and improve the video communication between elderly or disabled people in this atypical triadic relation. As future work, it is necessary to adopt an automatic recognition system for both users' attitudes to handle more delicate situations in video communication.

VII. ACKNOWLEDGMENTS

This research is supported by the Ministry of Internal Affairs and Communications of Japan. We would like to thank Dr. Norihiro Hagita for the opportunity to conduct this research and the staff and participants at the Seika-cho Social Welfare Council for allowing us to carry out the demonstration experiments.

REFERENCES

[1] N. Kuwahara, S. Abe, K. Yasuda and K. Kuwabara, Networked reminiscence therapy for individuals with dementia by using photo and video sharing, *Proc. ACM SIGACCESS*, pp.125–132, 2006.

[2] 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. ICMI*, pp.140–145, 2007.

[3] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe, GazeRoboard: Gaze-communicative Guide System in Daily Life on Stuffed-toy Robot with Interactive Display Board, *IEEE IROS2008*, pp. 1204–1209, 2008.

[4] H. Ezumi, N. Ochiai, M. Oda, S. Saito, M. Ago, N. Fukuma and S. Takenami, Peer support via video-telephony among frail elderly people living at home, *Journal of Telemedicine and Telecare*, Vol.9, No.1, pp. 30–34, 2003.

[5] W. Burgard, A.B. Cremers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, W. Steiner, and Sebastian Thrun, Experiences with an interactive museum tour-guide robot, *Artificial Intelligence*, Vol. 114, Issue 1–2, pp. 3–55, 1999.

[6] C.L. Sidner, C. Lee, C.D. Kidd, N. Lesh, and C. Rich, Explorations in engagement for humans and robots, *Artificial Intelligence*, Vol. 166, Issue 1–2, pp.140–164, 2005.

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

[8] Y. Nakano and R Ishii, Estimating User's Engagement from Eye-gaze Behaviors in Human-Agent Conversations, *Proc. IUI*, pp.139–148, 2010.

[9] H. Huttenrauch, K.S. Eklundh, A. Green, and E.A. Topp, Investigating Spatial Relationships in Human-Robot Interaction, *Proc. IEEE IROS*, pp.5052–5059, 2006.

[10] H. Tochigi, K. Shinozawa, and N. Hagita. User impressions of a stuffed doll robot's facing direction in animation systems. *Proc. of ICMI*, pp.279–284, 2007.

[11] S. Baron-Cohen, Mindblindness, *MIT Press*, 1996.

[12] T. Ono, M. Imai, and R. Nakatsu, Reading a Robot's Mind: A Model of Utterance Understanding based on the Theory of Mind Mechanism, *Advanced Robotics*, Vol.13, No.4, pp.311–326, 2000.

[13] J. Terken, I. Joris, and L.D. Valk, Multimodal Cues for Addresseehood in Triadic Communication with a Human Information Retrieval Agent, *ICMI2007*, pp.94–101, 2007.

[14] P. Viola and M. Jones. Rapid Object Detection Using a Boosted Cascade of Simple Features, *Proc. CVPR*, vol.1, pp.511–518, 2001.