SCoViA: Effectiveness of spatial communicative virtual agent based on motion parallax

Naoto YOSHIDA[1] Tomoko YONEZAWA[1]

[1] Kansai University

Abstract: We propose a spacial communication method between a user and a virtual agent that provide an illusion as if it has a physical presence. In this system, real-time 3DCG agent has been drawn with changing a viewpoint in a virtual space corresponding to the user's viewpoint detected by face-tracking. The face-tracking can perform this interaction as same as a robot in the real world. We conducted a preliminary experiment to verify whether our proposed method was effectively producing spacial interaction between human and virtual agent. The result shows the agent's talkability increase when the agent made eye-contact with the user.

1 Introduction

Various virtual agents mediate between a human and a computer in daily life. In addition, many studies are conducted about the individual gestures and expressions in the communication. Virtual agent can make various types of expression without any physical limitation, however, presence of virtual agent is not strong compared to robot which has physical presence in a real space. In relation to the presence with virtual agents, Nakatani et al.[1] proposed a Web communication system which enables to feel the existence of other participants through the appearance of multiple avatars. This study aims to maintain the friendship relations in SNS by showing multiple avatars' communication in virtual world. However, there is not many researches to make user feel a physical existence of virtual agent as though the agent were existing in real world.

The physical presence of robot has two aspects; one is physical touch, and the other is appearance in real world. Virtual agents do not have real body, so they are not able to be touched, however, the realistic appearance can be considered to make user recognize as though a virtual agent have a real body. Physical law and recognition mechanisms should help the illusion of realistic presence. Humans recognize depth and the size of real objects based on motion-parallax[2]. Thus users recognize physical existence visually by simulated stereoscopic vision based on motion-parallax. We focused on the expression of reality for the agent's presence using motion-parallax based three-dimensional computer graphics. The interaction between the agent and the user may be enriched by not only realistic view but also variation of spacial communication. We developed an interactive agent system with a simulated stereoscopic vision, that was named SCoViA (**S**patial **Co**mmunicative **Vi**rtual **A**gent system).

Main difference from conventional agent systems is a spatial interaction based on the relative positions between agent and user. In two-dimensional agent systems, users generally communicate with the agent in front of the monitor which shows the agent. These agents also look at the user regardless of the user's position. Virtual space and real space are considered to be divided by a presentation device such as a monitor. Conventional agent systems have problems in making users recognize a virtual agent as though it were a real presence. It is also difficult to express seamless connection of the virtual space in the monitor with the real space as though the agent were existing in our real world and sharing the space with us. In the SCoViA system, the user can feel as if the agent is in the real space by the motion-parallax-based system.

Our proposing method is expected to replace a part of various tasks that have been performed using a robot. Moreover, the virtual agent system can be realized by low cost and low maintenance with maintaining variety of expressions. One of robots' advantage compared to agents is an indication of a particular object by controlling its gazing orientation or gazing behaviors, for example, watching a specific object closely, and leading the interest of the person to the object. The motion-parallax-based presence of the virtual agent can overcome the limitation of the difference between robot and agent and generate natural communication related to real object. Practical scene of the system use is expected for personal space. For example, a communication support system for dementia patients in elderly care centers or a personal counselor for lifestyle guidance in the user's own room. As Yonezawa et al. have proposed a communication support system for the dementia patients by adopting the gazing behaviors of the robot, the gazing behaviors are strongly affecting on the underconscious attention of the user. Their analyses of the communication showed that the robot can draw the patient's interest. When we consider an appropriate interaction corresponding to the relative position between user and agent, it is necessary to be aware of the real world information such as user's gaze or a particular object in real space.

The second advantage of our proposed agent system is versatility and reduction in cost. The software system can work on a basic constitution in general computers, and it is easy to install. so that the system is used in a hospital, care centers, or a residence for single life household.

2 Related Works

Francone et al.[3] proposed a drawing method of three-dimensional computer graphic using face tracking based on motion-parallax in mobile devices. This system is able to show a 3D-CG object as though there were in the real world without using the specific devices. Thus the CG drawing methods with motion-parallax have indicated their effectiveness of realistic view.

Kipp et al.[4] have analyzed gaze interaction between user and avatar based on motion-parallax. In contrast to our approach, their purpose was to build an immersive communication media to draw the user into a virtual space. Results of a communication experiment based on gaze reactions are showing effectiveness of the agent drawn by motion-parallax, however, this study did not refer to communication sharing a real space such as talking a particular object in real space. Moreover, their system need to attach the glasses with two infrared LEDs and to install IR camera (Nintendo's Wii remote) behind the screen in order to detect the user's head.

Differently from their communication model, we propose various interactions in the real world where the agent and the user can talk about a real object in natural manner as if the agent physically exists. SCoViA also consists of a remote microphone, a web camera and a monitor without any fixing to the user, or any special devices. These equipments are basic components of personal computers. We apply a face-tracking software, "Face API[1]" to detect the position and direction of the user's face for assumption of the user's viewpoint. Our system is expected to be an active communication tool, and the continuous use might be led by natural behaviors and states of everyday life.

3 Spatial Communicative Virtual Agent System

3.1 Three Dimensional Rendering Based on Motion Parallax

SCoViA system performs CG rendering by synchronizing the user's viewpoint in the physical space and the position of the user's head to reflect the data on the virtual space (Figure 1). For example, the user can see the left side of the agent when she/he moves to the left, and the user can see the top side of the agent when she/he moves upward. In addition, the agent can turn the head and its eyeballs to any direction.

[1]seeing machines, Braddon, Canberra, Australia, 2008, http://www.seeingmachines.com/product/faceapi/

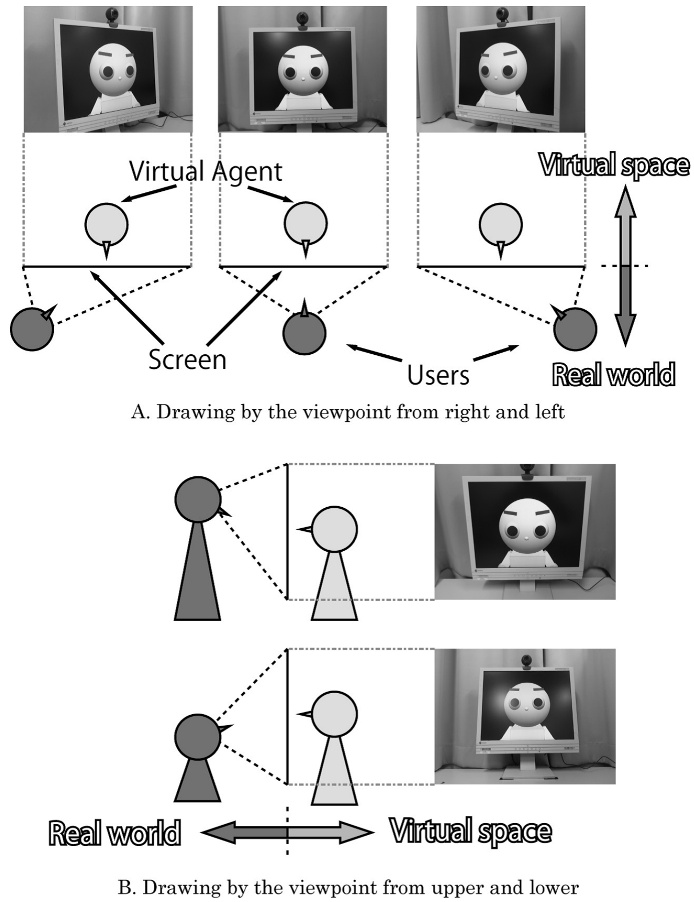

Figure 1: Drawing Method based on motion parallax

On conventional (two-dimensional) agents, it was difficult to express the eye contact with the user and the behavior of turning around by the influence of Mona Lisa effect. Our proposing agent can control its two eyes together by three dimensional drawing with considering the motion-parallax and appropriate angle adjustment.

3.2 Processing Flow

The system consists of Windows 7 PC, 17-inch liquid crystal monitor, a three mega pixels web camera attached to the upper part of the monitor, and a remote microphone. As shown in Figure 2, the interactive agent system has in the processing flow for both face tracking and utterance detection. The image analysis part extracts three dimensional face data by Face API. The coordinates and rotation angles of the user's face are sent to a graphics rendering unit. The three-dimensional computer graphics are appropriately controlled to show the agent's appearance with motion-parallax. The reactions of the agent are determined by both the spatial relationship between the user and the agent and the user's utterances.

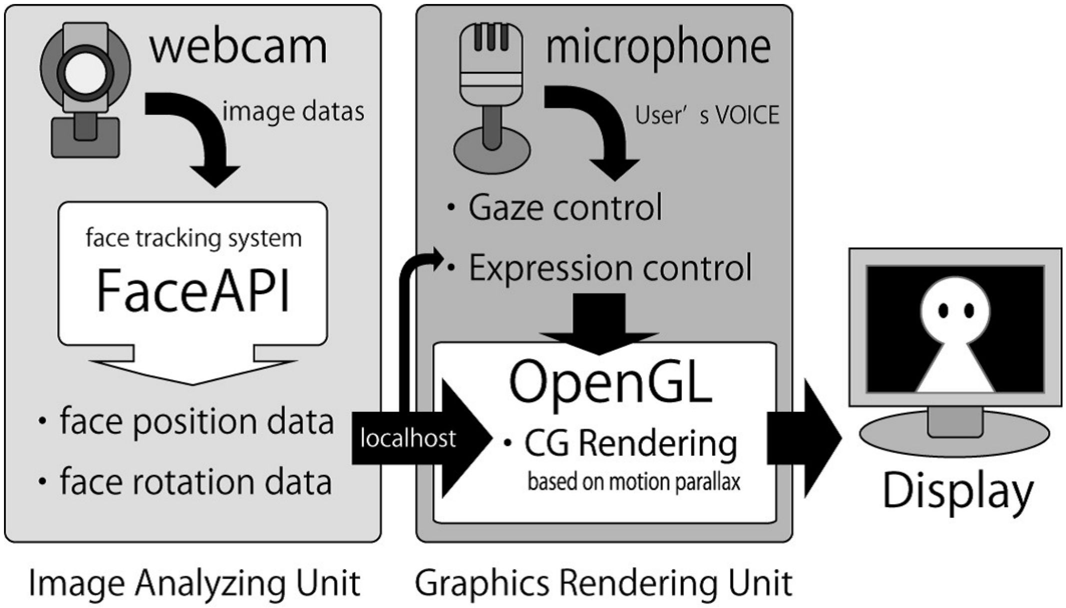

Figure 2: System Processing Flow

3.3 Design of Eyes' Movement

Humans change the gazing direction by both the head turn and eyeball movement. The movement of the head or eyeballs is automatically produced in many cases. At the same time, our eyes show convergence ocular movement depending on the focus distance. Eye-contact reaction of our virtual agent is expressed by the combination of these movements and the communication behaviors with basic eye movements as above. When the user turns her/his face to the agent, the agent watches the user's face closely in a state of the stereoscopic vision based on motion parallax. The system can reproduce human-like behavior of the agent attentive to the user's eyes, looks back towards the user, and shows joint attention with the user.

3.4 Facial Expression and Behavior in Conversation

A feature of SCoViA is a parametric control method of facial expression as the virtual agent's emotional appearance. as shown in Figure 3. The agent has eyebrows, eyes, a nose and a mouth. Size, angle and shape of these parts are controlled corresponding to the strengths of the emotional parameters. Thus the agent's expressions reflect by the parameter based on Russell's circumplex model of emotion[5]

Movement of the mouth and speaking expresses the talking state of the agent. The speech sound of the agent is made using "softalk[2]", a voice synthesis software. The size of the mouth spreads by the buffer size of an output sound.

[2]http://www35.atwiki.jp/softalk/

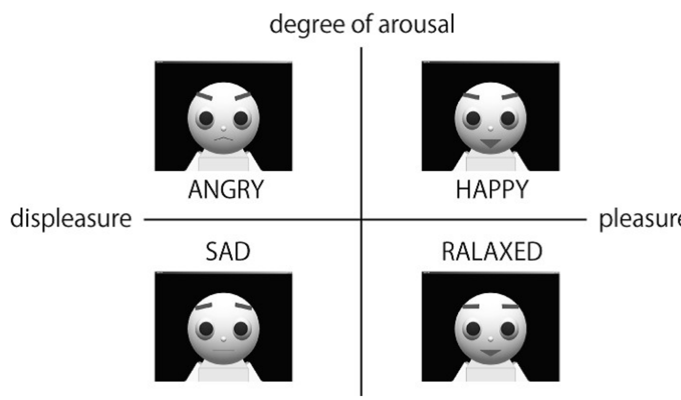

Figure 3: Emotional Expression by Facial Placements

4 Application and Verification of Spatial Interaction

4.1 Interaction Depending on Relative Height

Humans change voice tone and speaking behaviors corresponding to the emotional, external situation and social relationship with other people. Understanding of partner's attitude has an effect on user's feeling. One of the factor to recognize the partner's attitude is positional relationship, especially the difference of the vertical positions. Agent should differently react to the partner by the relation of vertical position as same as humans differently show their attitude while looking up or down. Table 1 shows our design of the agent's behavior with three parameters by the user's

Table 1: Parameterization of Viewpoints

Position of User's Viewpoint	Low	—	High
Agent's Tone of Voice	High	—	Low
Agent's Speaking Tempo	Fast	—	Slow
Agent's Level of Expression	High	—	Low

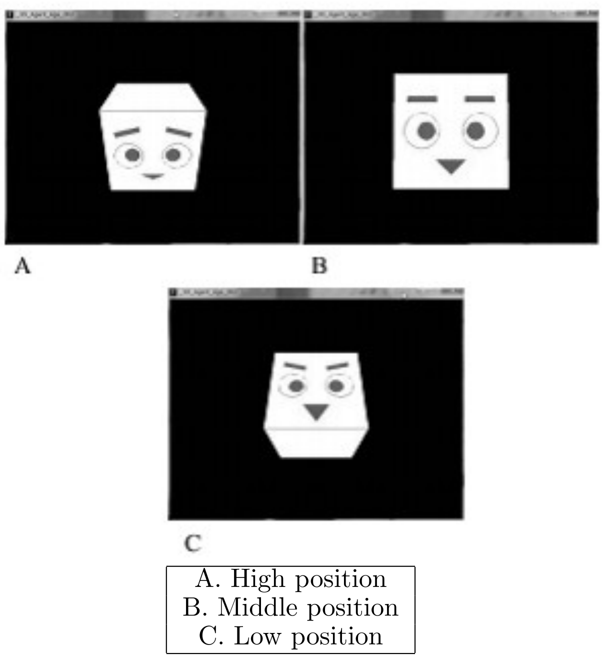

Figure 4: Example Views by Diffferent Parameter Level

viewpoints. The parameters we focused on are the tone of voice, speaking tempo and level of the expressions. The example views of the agent's expressions are shown in Figure 4. When the agent is looked down, the agent cannot show its strong expression as shown in A. When the agent is looked up, it shows strong attitude with assertive speech as shown in C.

4.2 Easiness to Talk to Agent with Gaze to User

Persons may not necessarily face each other at the beginning of when they need to talk with their partner. The intention of speech is whereat shown against some kind of signals at the beginning a conversation. The behavior of the intention of speech is called "Speech-implying Behavior[6]". The behavior was proposed to be expressed when the user is looking at some tasks, calling partner, or tapping the shoulder of them(Figure 5). It is necessary also for the partner agent to show the speech-implying behavior in order to start talking appropriately. By using this system, we can verify the difference of the impression that a user receives from a pattern of the speech-implying behavior.

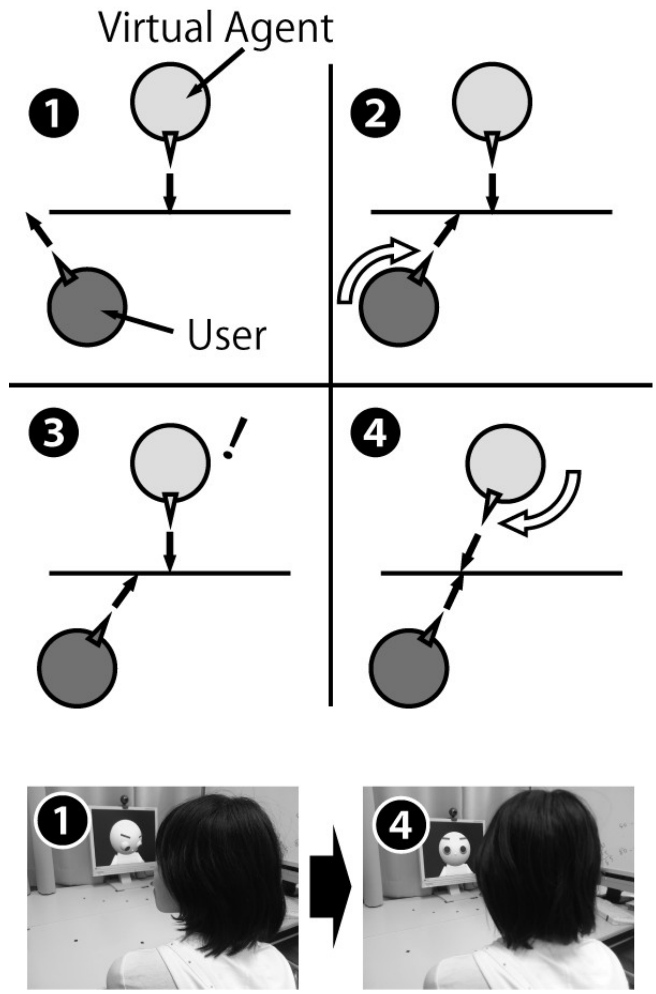

Figure 5: Example of Speech-implying Behavior

4.3 Experiment on Easiness to Talk to Agent

In this experiment, we verify whether the behaviors of the virtual agent, turning around and making eye contact, are suitable as a replying manner of the speech-implying behavior.

We asked participants to evaluate their impressions. Experiment participants are twenty-seven people (eighteen males, nine females) from 19 to 22 years old. The experimenter calibrated the participants' faces for face tracking before the experiment.

In the experimental setting, the participant stood at thirty centimeters from either right or left side from the front and one meter from a display. The participant turned their heads and looked at the agent. the agent showed three patterns of the behavioral actions as follows.

A It turned a face away from the participant more.

B It turned around and watched the face of the participant.

C It did not react to the participant at all.

The participants evaluated a five-point rating scale for the user's impression of easiness to talk to the agent.

Results Figure 6 shows the results of means opinion score (MOS). ANOVA for repeated measurements showed significance (F=18.840 and p=<0.001), and the results of the post-hoc tests are illustrated in Table 2. Based on the degree of freedom 26 and the rejection rate 0.05, there were significant differences among three conditions.

The experimental results shows the highest value of the B condition: the agent turned around and looked at the participant's face. The result shows possibility that the user feels more easily to start talking. On the other hand, the A condition showed the lowest value. That is, the occurrence of the reaction sometimes works on the negative impressions on the user even when the agent reacts appropriate timing. The agent's behavior of turning its face away from the user is suggested to avoid the communication than situation of the agent does not react at all.

From the results, it is conjectured that the agent's behavior of "turning around" or "looking at the user" leads positive impression on the easiness to talk to. From these, the direction of the agent's face and a condition of their gazes could show some kind of the acceptance state of the agent.

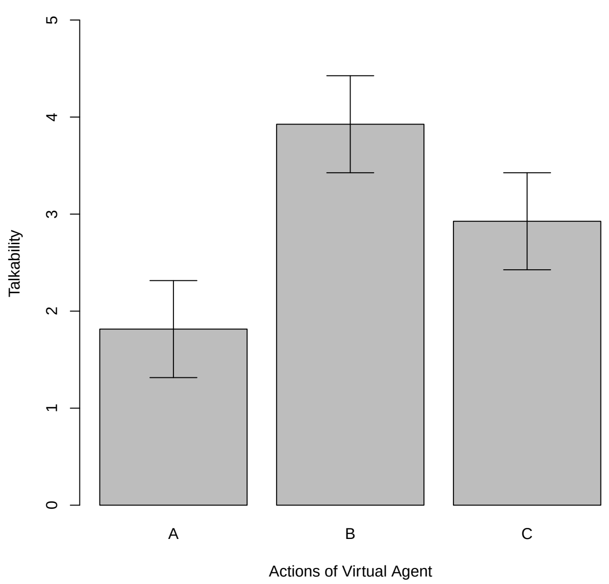

Figure 6: Result: The rate of talking easily by the ansewer of the agent

Table 2: Post-hoc Test by Ryan's method: Easiness to Talk to Agent

pair	nominal level	t	p
B - A	0.017	6.14	<0.001
B - C	0.033	2.91	0.005
C - A	0.033	3.23	0.002

5 Discussion

5.1 Discussion of SCoViA System

We proposed an interaction method as if the virtual agent is in the real world. However, the information and interaction that can be shared between virtual space and real space have limitation. The user cannot go into the virtual world, and some objects in the virtual world is not able to be brought into the real world. Therefore we should prepare the agent's graphics where the virtual camera is controlled and synchronized to a real motion of the user's face to make illusion of the presence of the agent.

The experiment was conducted to verify whether a person recognizes the easiness to talk to the agent by its turned eyes to the real user. In the human-human communication in real world, our gaze directions are reactive to the relative position to the other person. Therefore, in the virtual space, it is also required that the user can judge whether the virtual agent looks toward the user with depending on the user's viewpoint. The consideration of the results of the experiment are discussed below.

5.2 Discussion of Application

The experimental results show difference of the participants' impression by turning around motion of the agent. Even if the participants look from any viewpoint, they recognized that the agent faced towards themselves. These results show that the participant recognized spatial relation in the real world and specified the eyes direction of the agent. The results are expected to be appeared in the two-dimensional agent by Mona-Lisa effect, however, the reactions are remarkable when SCoViA is considered to the communicative agent. Also, it is considered that it may be recognized according to a physical law of the real world in the agent of the virtual space.

5.3 Summary of Discussion

The experiments in this paper were conducted as preliminary tests to verify the impression of the physical presence of the agent. Using a virtual agent, the possibility was shown that we could communicate in the real world. Our proposed system is a basis of verification of the spatial communication. The results of the impression on the agent's presence and communication feeling should be discussed for the future research of human-agent interaction. The virtual agent study may also reflect the existing study of robot. And the system should be experimented with the situations of assumed use cases, especially about the effect the gaze and speech behaviors that have been revealed in prior researches.

6 Conclusion

We proposed a virtual agent system with physical presence and interaction based on three-dimensional illusion of motion-parallax. The aim of the system is to increase the presence of the virtual agent as though it were in a real space with the user. The agent can change its behavior corresponding to the user's relative position and gaze to the agent. Furthermore, we have developed two practical applications of the agent system; the first one is the superior-inferior communication by a vertically spatial relationship between the user and the agent. The second one is the face-turn agent system corresponding to the user's position.

In order to verify whether our proposed method was effectively producing spatial interaction between human and virtual agent, we conducted a preliminary experiment. The result shows that the feeling of easiness to talk to the agent increases when the agent makes eye-contact. Consequently, we could confirm our proposed method has an effectiveness of the real-space communication with a virtual agent.

In future works, we are going to consider various types of the interactions between agents and humans in real space comparing to the interactions between robots and humans. The trials should reveal the possibility of the virtual agents. The realistically physical presence will be able to appear the agent life.

Acknowledgement

This research was supported in part by Subsidy for Supporting Young Scholars in Kansai University, KAKENHI 24300047, and KAKENHI 25700021.

References

[1] Yukari Nakatani Tomoko Yonezawa. Abotat: An expressive method of web communication using appearances of avatars attached to text message and remarks. In *Proceedings of First International Conference on Human-Agent Interaction*, to appear.

[2] Brian Rogers, Maureen Graham, et al. Motion parallax as an independent cue for depth perception. *Perception*, Vol. 8, No. 2, pp. 125–134, 1979.

[3] Jérémie Francone and Laurence Nigay. Using the user's point of view for interaction on mobile devices. In *23rd French Speaking Conference on Human-Computer Interaction*, IHM '11, pp. 4:1–4:8, New York, NY, USA, 2011. ACM.

[4] Michael Kipp and Patrick Gebhard. Igaze: Studying reactive gaze behavior in semi-immersive human-avatar interactions. In *Intelligent Virtual Agents*, pp. 191–199. Springer, 2008.

[5] Jonathan Posner, James A Russell, and Bradley S Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, Vol. 17, No. 03, pp. 715–734, 2005.

[6] Tomoko Yonezawa, Hirotake Yamazoe, Akira Utsumi, and Shinji Abe. Anthropomorphic awareness of partner robot to user's situation based on gaze and speech detection. *International Journal of Autonomous and Adaptive Communications Systems*, Vol. 5, No. 1, pp. 18–38, 2012.

[7] Minoru Hashimoto and Daisuke Morooka. Facial expression of a robot using a curved surface display. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, pp. 765–770. IEEE, 2005.

[8] Yoshihiro Sejima, Tomio Watanabe, and Michiya Yamamoto. Analysis by synthesis of embodied communication via virtualactor with a nodding response model. In *Universal Communication, 2008. ISUC'08. Second International Symposium on*, pp. 225–230. IEEE, 2008.