Gaze Communication between Human and Anthropomorphic Agent
--Its concept and examples--

Naoki Mukawa, Atsushi Fukayama, Takehiko Ohno, Minako Sawaki and Norihiro Hagita
NTT Communication Science Laboriatories, NTT
3-1, Morinosato Wakamiya, Atsugi-shi
Kanagawa 243-0198, Japan
E-mail mukawa@eye.brl.ntt.co.jp

Abstract

Psychologists have revealed that the human gaze conveys a wealth of information about a person's mental state and its personality. Gaze contributes to human-to-human communication. In this paper we propose using gaze as a new nonverbal communication tool between humans and computers. This paper presents the concept of gaze communication between humans and an anthropomorphic agent and shows that the gaze of an agent can convey personality when the gaze is moved using our gaze model. Finally, it presents some examples of gaze communication.

1 Introduction

Psychological studies have shown that human gaze indicates one's focus of attention, degree of attention, and one's emotion or mood. In human-to-human communication, a person generates a gaze and recognizes gaze of others. The human brain is equipped with a mechanism for controlling gaze and recognizing another's gaze. Much research on the roles of the gaze has been carried out. In his pioneering work, Kendon [1] suggested : a monitoring function that extracts the partners intentions and affections, an expression function that expresses one's own intention and affections, and a regulatory function that regulates turn-taking of conversation.

In this paper, we categorize these functions more precisely as:

F-1: intention monitoring,
F-2: intention expressing,
F-3: affective information monitoring ,
F-4: affective information expressing ,
F-5: communication regulatory .

Affective information in this paper includes emotion, mood and personality. This section briefly reviews the research works of gaze-based communication.

Kendon [1] focused on F-5 and clarified that gaze can regulate the change and maintenance of the speaker role. Baron-Cohen [2] focused on F-1 and F-2 and proposed a cognitive model for gaze. His model consists of four modules: the "intentionality detector," which reads mental states in behavior, the "eye direction detector," the "shared-attention mechanism," which builds interesting things, and the "theory-of-mind mechanism," which estimates mental states. For intention sharing between humans he emphasized the importance of *focus of attention*. Vertegaal [3] measured the *focus of attention* for multiparty conversation using gaze. He found that there is a high probability that a person being looked at is the person who is listened or spoken to. Cook et al. [4] investigated F-3. They reported that the gaze deeply effects the impression of personality, such as likability and capability. Physiological research has also begun. Using fMRI (functional Magnetic Resonance Imaging), Hoffman [5] investigated the perception of face identity and eye gaze in the human brain. She revealed that face identification and gaze reading are processed in a different area in the brain.

Application-oriented research efforts concerning gaze have also been made, aiming at richer and smoother human-to-human and human-to-computer communication. Dating back to the 70's, it has been pointed out in visual telephony and teleconferencing, that eye contact is critical for better communication when using such equipment. Many systems that realize eye contact between users have been proposed (an example is shown in [6]). In recent years, to support intuitive communication between human and computer, anthropomorphic agents and avatars with an animated gaze for eye contact have been studied.

As for F-5, Novick et al. [7] revealed that the turn of speaker and listener can be detected by observing gaze in human conversation. Base on psychological findings, Colburn et al. [8] proposed a gaze model in dialogue where the amount of gaze during speaking is less than that of during listening. They applied the model to a human–computer interface using an avatar. Garau et al. [9] clarified that the impact of communication was intensified by gaze control where gaze is moved synchronously by the turn of speaking and listening rather than gaze is randomly moved.

As for F-1, research has been performed on *focus of attention*, that is, who is looking at whom by detecting gaze direction in a multiperson conversation. Stiefelhagen et al. [10] modelled the *focus of attention* for multiple

IEEE International Worksop on Robot
and Human Interactive Communication
0-7803-7222-0/01/$10.00©2001 IEEE

participants. The system they developed codes the *focus of attention* based on the gaze and also codes the location of participants. It then tracks the attention along the conversation using an neural network model and a hidden Markov model.

Existing application-oriented research on human gaze are mainly concerned with F-1, F-2, and F-5. Little research has been done on F-3 and F-4, i.e., affective communication by gaze of emotions, moods, and personality. We must admit that facial expressions and gestures are considered as main channels in affective communications and have been intensively studied. By intuitive considersation, however, the expressions of face and gesture without gaze is not lively nor attractive. Therefore the purpose of the paper is to clarify the importance of affective communication using gaze. Focusing on the F-3 and F-4, the paper first describes the concept of the communication of affection and a model of communication using gaze. Then as one of the usefull affective informations we show that personality of anthropomorhic agent can be conveyed by its gaze. Finally, we conclude by mentioning future work.

2 Concept of Gaze Communication

This paper focuses on F-3 and F-4. So, the information to be transmitted is not intention but affection such as emotion, mood, and personality. Affection is detected and recogized in F-3 by observing the gaze, and affection is generated by controlling the gaze in F-4. In this model, F-3 infers affection using inherited knowledge and knowledge obtained by one's expriences. In this paper, we assume an affective communication system using human gaze is represented in Figure 1. We propose using this model as a new nonverbal communication tool between humans and computers.

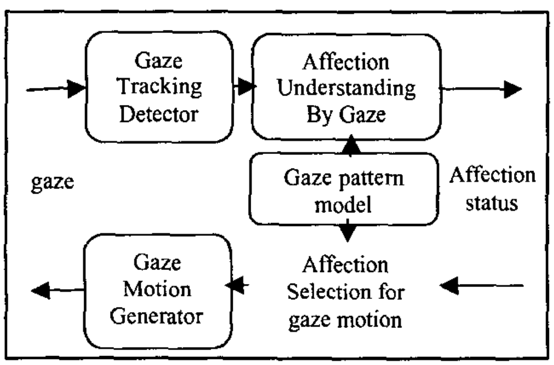


Figure 1. Affective communication model by gaze

Our research target is to develop a system consisting of the components in Figure 1 to enable richer and smoother communications between humans and computers. For this pupose, we are developing an eye tracking system for gaze direction detection and investigating a human-like agent (anthropomorphic agent) having gaze that conveys affection by gaze. As for the recognition system we must develop the gaze database before we start studying the gaze recognition that understands human affection by gaze. Research results for the recogntion system may therefore appear in the future.

The rest of the paper discusses our anthropomorphic agent that can convey its affective status or personality by gaze. It shows agents can use gaze to convey personality when the gaze is moved according to our gaze model. We then present a newly developed eye tracking system called *FreeGaze* that is applicable for human-computer communication. Finally, we describe a demonstration system as a concrete example of our concept.

3 Anthropomorphic Agent with Gaze

Various kinds of affective information: emotions, moods and personalities can be conveyed from agents to users through nonverbal means. Among them personality is important because it can induce some crucial responses from users. For example, a user can become attached to an attractive agent. Besides, an agent that looks trustworthy makes it easier for users to accept its suggestions. However, expressing personality through an agent has not been widely discussed. Thus we investigate the gaze of agent for expressing its personality. Whether nor not agents can express emotion and mood by gaze need further study, although it may give strong impression to users and be an interesting research issue.

As mentioned in the Introduction, Cook et al. suggested that gaze of a human deeply effects the impression of his/her personality [4] in human communications. We apply their findings to interface agents, and propose to clarify how to control the personalities that users perceive from the gaze of interface agents.

As the first step, we conducted experiments to examine the feasibility of such an interface agent that can control the impression of its personality by its gaze.

3.1 Preparations

Hypotheses

H-1. *The gaze of agents can express human-like personalities.*

Supported if personalities perceived from some types of gaze of an interface agent are evaluated on the basis of personality measures similar with those for humans.

H-2. *The personality perceived from the gaze of agents is reproducible.*

Supported if two different agents induce a similar impression of personality when they move their eyes in the same way.

3.2 Preparations

To verify these hypotheses, we prepared three gazers as shown in Figure 2, i.e., human (G_H), agent1 (G_{A1}), agent2 (G_{A2}) [11], and nine video-recorded movies for each gazer as it moves its eyes based on one of nine gaze types (Table 1).

G_{A1} and G_{A2} move their eyes so as to appear to look at a gaze point calculated at each time-step. The gaze points obey the gaze-point distributions of the specified gaze type. The distributions of all types of gaze were calculated beforehand from measurements of the gaze points of G_H.

Figure 2. Three gazers.

Table 1. Nine types of gaze

Type of gaze	Description
1. Normal	Nomal gaze pattern of G_H
2. Stare	Always staring at user's eyes
3.-5. Avoid {L,U,S}	Always looking at Lower /Upper/Sides of user's face
6. Unstable	Unstable eye movement
7.-9. Glance {L,U,S}	Sometimes looking at user's eyes

3.3 Procedure

Participants: University students (G_H: 7 males and 11 females / G_{A1}, G_{A2}:4 males and 13 females)

Stimuli: For each gazer, the Normal recording was presented first as the standard of the gazer. Then, the other eight recordings were presented and the personality of the gazer in each movie was evaluated by questionnaire.

Questionnaire: Based on the *Semantic Differential Method.* Here, twenty pairs of bipolar adjectives that describe human personalities, e.g., *warm-cold, extrovert-introvert,* were used. For each pair, the participants were asked to indicate which adjective described the personality of the presented gazer better and rate it on one of seven levels.

3.4 Results

Measures of Personlity : In order to extract some basic measures on which personality of each gazer could be evaluated, the rated points on the questionnaire were analyzed by a factor analysis. Table 2 shows the measures, i.e., common factors, extracted from the evaluations of each gazer. Figure 3 shows personality scores obtained by the analysis. The number in the figure corresponds to the type of gaze in Table 1.

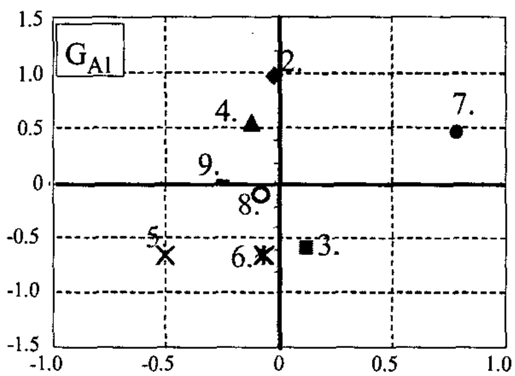
For each gazer, two measures were extracted. Judging from the related adjectives, we can interpret what the measures mean. First, all of the primary measures can be interpreted as ones that represent concepts like the "likability" of gazers. On the other hand, all of the secondary measures can be interpreted as "capability" measures.

Likewise, the measures on which the personalities of both G_{A1} and G_{A2} were evaluated show similarity with the measures of G_H to the extent of the comparison between the interpretations. If the gazes of G_{A1} and G_{A2} didn't have the ability to express human-like personalities, it is likely that the measures with high contribution wouldn't be extracted, or that the measures wouldn't have related adjectives consistent with one another as those of G_H . Therefore, this result supports hypothesis H-1.

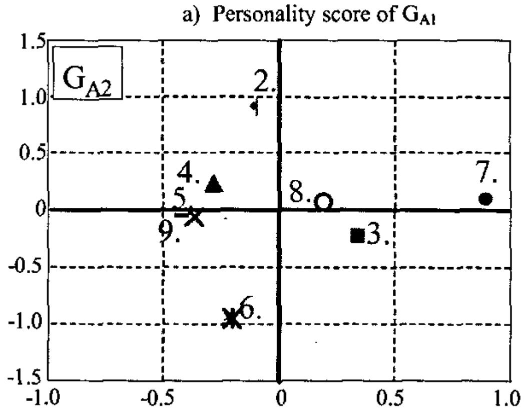
Personalities Perceived from Gaze : Based on the measures of "likability" and "capability", we can score the eight types of gaze of G_{A1} and G_{A2}. In order to compare G_{A1} and G_{A2}, we conducted t-tests to find the significant differences between all possible pairs of the eight. On the basis of both primary and secondary measures, we confirmed that if a certain pair of types of gaze $i, j (2 \leq i, j \leq 8)$ satisfy the relation $i \succeq j$ (the

Table 2. Extracted measures by factor analysis

Gazers	Measures	Related adjectives
G_H	primary (53.1%)	warm, friendly, flexible, sympathetic, intimate, attractive
	secondary (11.0%)	confident, strong, successful, extroverted
G_{A1}	primary (48.6%)	warm, tolerant, sympathetic, polite, intimate
	secondary (22.5%)	strong, confident, successfule, extroverted
G_{A2}	primary (50.5%)	tender, warm, cooperative, honest
	secondary (12.8%)	confident, strong, successful, with strong sense of responsibility



a) Personality score of G_{A1}



b) Personality score of G_{A2}

Figure 3. Personality score of G_{A1} and G_{A2}. The numbers corresponds to the gaze types in Table 1. Horizontal coordinates represents the "likability" and vertical coordinates represent the "capability" of the agents.

personality score of gaze type i is bigger than j or there is no significant difference) in either G_{A1} or G_{A2}, the same relation $i \succeq j$ is also observed in the other gazer. In other words, there is no conflict in the order of the personality scores between G_{A1} and G_{A2}.

Though this is a qualitative comparison, G_{A1} and G_{A2} induced similar impressions of their personalities, which supports hypothesis H-2.

Hypotheses H-1 and H-2 are supported by the results of the experiments. These results indicate the feasibility of an interface agent that can control impressions of its personality by its gaze.

4 FreeGaze – A New Eye Tracking System

An eye tracking system is a key device in our human-computer communication. However, existing eye tracking systems have been developed mainly for research of human cognitive processes. Thus, not many are suitable for interfaces, interactions, and communications between humans and computers. Some of them are head-mounted types and too heavy for communication purposes. Furthermore, most of them need to be carefully calibrated before each measurement session. Calibration needs many calibration points, as a result, time consuming. Aiming for human-computer communication, we have developed a new eye tracking system called *FreeGaze,* which drastically simplifies the calibration procedure.

4.1 Advantages of the Eye Tracking System

The advantages of the system are ease of calibration and compactness. Table 3 shows the characteristics of *FreeGaze*. Figure 4 is a photograph of the system hardware.

Table 3. Advantages of *FreeGaze*

Frequency of calibration	Once for every user
Number of calibration points	Two (at least)
Portability	Portable
Constraint to users	No goods to wear

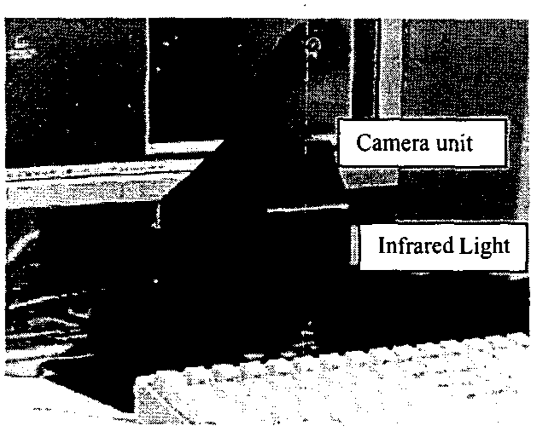


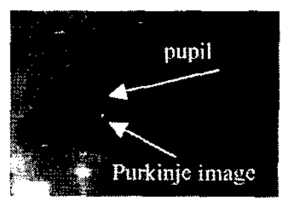
Figure 4. *FreeGaze*- Eye tracking system.

4.2 Gaze detection Algorithm

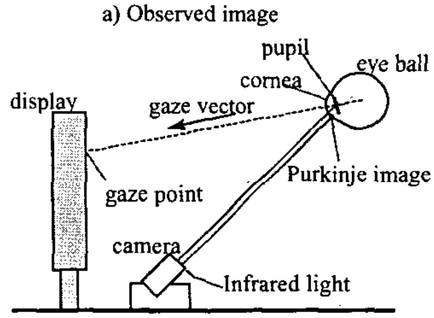
The configuration is shown in Figure 5. *FreeGaze* consists of an NTSC camera and an infrared light source mounted on the same box. The infrared light is first incidented on the user's eye, then is reflected on the cornea surface. The reflection (called *first Purkinje image*) is detected from the image taken by the camera. The center of the pupil is also detected from the image. Gaze direction is calculated from these two points. To detect the center of the pupil, refraction phenomenon was taken into account [12].

4.3 Calibration

The most sophisticated operation in *FreeGaze* is calibration. Intrinsic and geometrical parameters of both the camera and infrared light are carefully calibrated and refraction is compensated. The remaining parameter to be calibrated explicitly is the radius of cornea curvature and the distance between the center of the curvature and center of the pupil which may differ user to user. Consequently, in theory, only two points are neccesary for calibration which involves the adjusting translation and resizing parameters. For calibrating the system a small spot is displayed at the upper left corner of the display screen intructing the user to press down a key while gazing the point. The next point appears at the lower right, and the user is requested to do the same. Additional calibration points will promise more accurate and stable measurement. A preliminary evaluation test showed that the accuracy was between 0.23 degree and 0.46 degree in the view angle.



a) Observed image



b) System configuration

Figure 5. Configuration of eye tracking system

5 Demonstration

Based on the concept of the communication between human and anthropomorphic agent proposed in Sec.2, we plan to develop the system shown in Figure 6. At present, we have accomplished a demonstration system shown in Figure 7. The agent reads the user's gaze through the eye tracker and controls its gaze in response to the user's gaze. The proposed gaze model is used for agent gaze control, although the agent has no gaze model for reading the user's gaze.

In the demonstration scenario, the agent asks five questions to the user about his/her everyday behaviour. For example, "In a train, you overhear a boy sitting next to you explain a computer operation proudly but wrongly. Do you correct?" The user must answer 'Yes' or 'No'. During the question and answer the agent will look at the user's eyes through *FreeGaze*. Based on the user's gaze motion, the agent selects one gaze motion pattern among the nine motion patterns given in the Sec. 3. For example, if the user gazes into the agent's eyes continuously, the agent selects the gaze pattern 3-5 (Avoid) and 7-9 (Glance) from Table 1 to avoid continuous mutual gaze. At the end of the demonstration session, the agent analyzes the user's character (just for fun). The evaluation of the system needs further study.

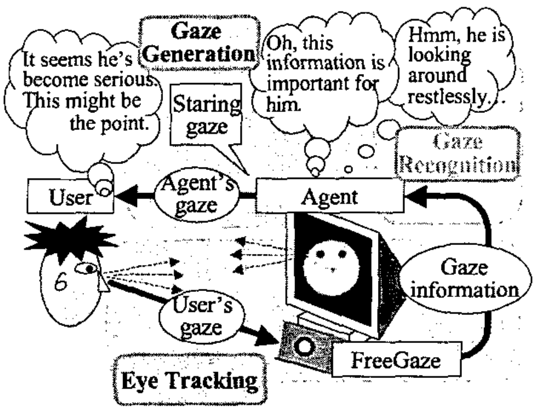


Figure 6. Concept of gaze communication between user and agent

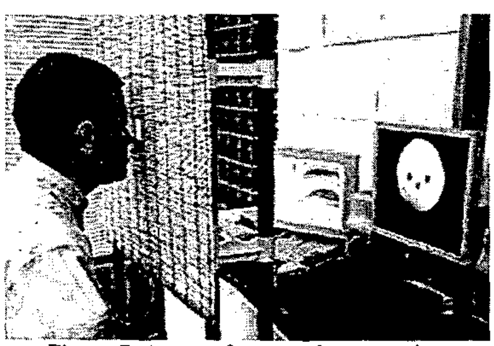


Figure 7. A scene from our demonstration.

6 Conclusion

In this paper we proposed a gaze-based affective communication model between humans and anthropomorphic agents and to use the model as a new nonverbal communication tool. Through the research activities conducted at our Laboratories, we showed that gaze can aid in fluent communication. We also have confirmed that impression is conveyed by the gaze of the agent and that the eye tracking system can be applied for this purpose. At present we have only limited results. We are now collecting human gaze data for constructing database. Recognition of the human gaze remains for further study, but will be studied using the database. As for next research direction we are going to investigate the interaction between face, gesture and gaze motion for aiming comprehensive nonverbal communication.

Acknowledgements

The authors are indepted to Dr. Ken-ichiro Ishii, the Director of NTT Communication Science Labs for his support and encouragement. The authors appreciate the free distribution of the CG agent tool at Electrotechnical Lab.

References

[1]A. Kendon "Some functions of gaze-direction in social interaction",*Acta Psychologia*, 26, pp.22-63. (1967)

[2]S. Baron-Cohen "Mindblindness – An Essay on Autism and Theory of Mind," *MIT Press*, 1995.

[3] R. Vertegaal, R. Slagter, G. van der Veer and A. Nijholt "Eye Gaze Patterns in Conversations: There is More to Conversational Agents than Meets the Eyes," *Proc. of Graphics Interface*, pp.95-102 (2000)

[4] M. Cook and J.M.C. Smith "The Role of Gaze in Impression Formation," *J. Soc. Clin. Psychol.* 1975 14, pp.19-25

[5] E.A. Hoffman, J.V. Haxby "Distinct representation of eye gaze and identity in the distributed human neural system for face perception," *Nature Neuroscience*, Vol.3, No.1, pp.80-84, January 2000.

[6]T. Aoki, K. Widoyo, N. Sakamoto, K. Suzuki, T. Saburi, H. Yasuda, "MONJUnoCHIE System : Videoconference System with Eye Contact for Decision Making", *International Workshop on Advanced Image Technology (IWAIT) '99*, 1999.

[7] D.G. Novick, B. Hansen and K. Ward "Coordinating Turn-taking with Gaze," *Int. Conf. On Spoken Language*, Vol.3. pp.1888-1891, 1996.

[8] R. A. Colburn, M.F. Cohen and S. M. Drucker "The Role of Eye Gaze in Avatar Mediated Converstaional Interface," *MSR-TR-2000-81*, July, 2000

[9] M. Garau, M. Slater, S. Bee, and M.A. Sasse "The Impact of Eye Gaze on Communication using Humanoid Avatoars," *Proc. of CHI 2001*, ACM Press, pp.309-316 (2001)

[10] R. Stiefelhagen, Jie Yang and Alex Waibel " Modeling focus of attention for meeting indexing", *ACM Multimedia*, pp. 3-10, Vol. 1, No. 1, 1999.

[11] O. Hasegawa, K. Sakaue, "CG Tool for Constructing Anthropomorphic Interface Agents", *Proc. IJCAI-97 WS (W5)*, pp.23—26, (1997)