Model-Driven Gaze Simulation for the Blind Person in Face-to-Face Communication

Shi Qiu[1], Siti Aisyah binti Anas[1], Hirotaka Osawa[2], Matthias Rauterberg[1], Jun Hu[1]

[1]Department of Industrial Design
Eindhoven University of Technology
{SQIU, S.A.B. Anas, G.W.M.Rauterberg, J.Hu}@tue.nl

[2]Faculty of Engineering, Information and Systems
University of Tsukuba
osawa@iit.tsukuba.ac.jp

ABSTRACT

In face-to-face communication, eye gaze is integral to a conversation to supplement verbal language. The sighted often uses eye gaze to convey nonverbal information in social interactions, which a blind conversation partner cannot access and react to them. In this paper, we present E-Gaze glasses (E-Gaze), an assistive device based on an eye tracking system. It simulates gaze for the blind person to react and engage the sighted in face-to-face conversations. It is designed based on a model that combines eye-contact mechanism and turn-taking strategy. We further propose an experimental design to test the E-Gaze and hypothesize that the model-driven gaze simulation can enhance the conversation quality between the sighted and the blind person in face-to-face communication.

Author Keywords

Eye contact; eye tracking; conversation quality

ACM Classification Keywords

H.5.2. [Information Interfaces and Presentation]: User Interfaces, K.4.2 [Computers and Security]: Social Issues – Assistive technologies for persons with disabilities.

INTRODUCTION

Gaze and eye contact are of importance in the development of trust and deeper relationships [1]. McNeill emphasizes that nonverbal cues such as mutual gazes are integral to a conversation and that ignoring them means ignoring part of the conversation [2]. Nonverbal cues also play an important role for the blind people in social interactions. Some findings were reported in an investigation of twenty blind people [3]. They were interviewed about the nonverbal signals for face-to-face communication with the sighted people. Some blind people in the investigation tended to wear the black glasses to hide their eyes in face-to-face communication. One possible reason was their eyes were conspicuously unattractive, and/or deformities might be present which made the person less appealing to others [4]. The other reason was some of them were not able to control the appropriate eye gestures. For example, one blind person had an illness of nystagmus and she could not control the movement of the eyeball. It was easier to cause the misleading of the sighted towards her eye gaze gestures. As a blind person, she often feels difficult to meet people, because she cannot see and establish eye contacts with the sighted people. Van Hasselt [4] pointed out, lack of eye contact might cause the sighted to feel that the blind person was not fully in communication. The impatience, discomfort, or intolerance of the sighted is an important factor in determining the possible extent of the involvement for the blind person. In this paper, we present E-Gaze glasses (E-Gaze), a reactive system based on the eye-tracking technology. It simulates gaze for the blind person to react and engage the sighted in face-to-face conversations. It is designed based on a model that combines eye-contact mechanism and turn-taking strategy. We further propose a user experiment to test the sighted person's responses to different gaze patterns of the E-Gaze, worn by the blind person, in a dyadic-conversation scenario.

HAI '16, October 04-07, 2016, Biopolis, Singapore
ACM 978-1-4503-4508-8/16/10.
http://dx.doi.org/10.1145/2974804.2980482

RELATED WORK

Several previous studies on simulating gaze behaviors concentrated on turn-taking phenomena that linked gaze behaviors with speaking and listening modes in conversations. Cassell et al. proposed a new approach to design conversational agents that exhibit appropriate gaze behavior. The exchange of looks between participants was related to both information threads and the exchange of turns during the flow of conversation[5]. Heylen et al. reported an experiment that investigated the effects of different eye gaze behaviors of a cartoon-like talking face on the quality of human-agent dialogues. The result demonstrated that the gaze strategy using a turn-taking model significantly affects the dialogue quality [6]. Garau et al. described an experiment to investigate the importance of eye gaze between humans in four mediated conditions: video, audio-only, random-gaze avatars and inferred-gaze avatars. In inferred-gaze, gaze was tied to turn taking in the conversation. The result showed that the inferred-gaze avatar significantly outperformed the random-gaze model and also outperformed audio-only model [7].

E-GAZE GLASSES

We presented our design concept of E-Gaze in a previous study [8]. It is a wearable device, worn by the blind person as glasses (Figure 1) and displays two basic gaze patterns: "look at" and "look away" from the sighted based on whether the blind person is talking. When the blind person starts talking, E-Gaze will "look away" from the sighted to concentrate on what the blind person is going to say; when she ends talking, E-Gaze will "look at" the sighted to establish the "eye contact" when giving the sighted an opportunity to take the turn. If the sighted stares at E-Gaze, E-Gaze will "look away" to avoid the long gaze.

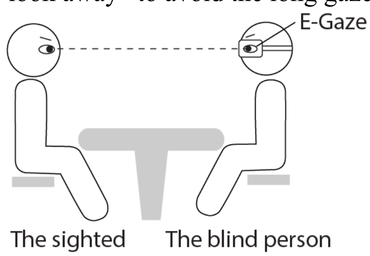

Figure 1.Design concept of the E-Gaze.

Model-Driven Gaze Simulation

A number of studies that investigated the role of gaze in human-avatar communication and provided the evidence that proper eye gaze behavior of the avatar elicited more natural responses in human users[6][7]. Most avatar systems display gaze behaviors based on turn-taking strategies rather than being reactive. In reactive systems, the sighted person's gaze behavior should trigger an instant response from the avatar side which in turn influences the user, resulting in a tightly coupled feedback loop [9]. These reactive responses can be triggered by tracking eye gaze from the sighted.

In our gaze model, E-Gaze is capable of reacting to the sighted person's current gaze using an eye tracker. The gaze model combines eye contact mechanism with turn-taking strategy which can distinguish while listening and while speaking modes in a conversation flow. We basically copy the timing of the eye gaze from research on dyadic conversations between a human and a humanoid avatar [7][10]. In this model, whenever the sighted is looking at E-Gaze, it displays "look at" eye gaze and tries to establish the eye contact with the sighted. The timing of E-Gaze "look at" and "look away" is slightly adapted depending on whether the blind person is talking or listening to account for the fact people look more at the person when listening than when speaking[1].

More specifically, we define four states of the sighted in the dyadic conversation with the blind person: **`looking at & speaking (LS)`**, **`looking at & quiet (LQ)`**, **`looking away & quiet (AQ)`** and **`looking away & speaking (AS)`**. The blind person also has the same four states since they are peer-to-peer in the conversation. The state machine diagram (Figure 2) describes four states and their transitions. The corresponding eye gaze animations of each state are presented as follows:

`LS:` when the sighted is speaking and she is looking towards E-Gaze, E-Gaze displays a "look at" eye gaze for 2.5 seconds to establish an eye contact with the sighted and then "look away" for 1.6 seconds. After that, E-Gaze displays an average frequency of 17 "look at" glances per minute if the state continues. For "look at" partner gaze (Figure 3(a)), E-Gaze eyes focus directly ahead. There are two types of the "look away" eye gaze: left look away (Figure 3(b)) and right look away (Figure 3(c)). In order to avoid repeated eye gaze animation, E-Gaze randomly displays one of the "look away" eye gaze animations.

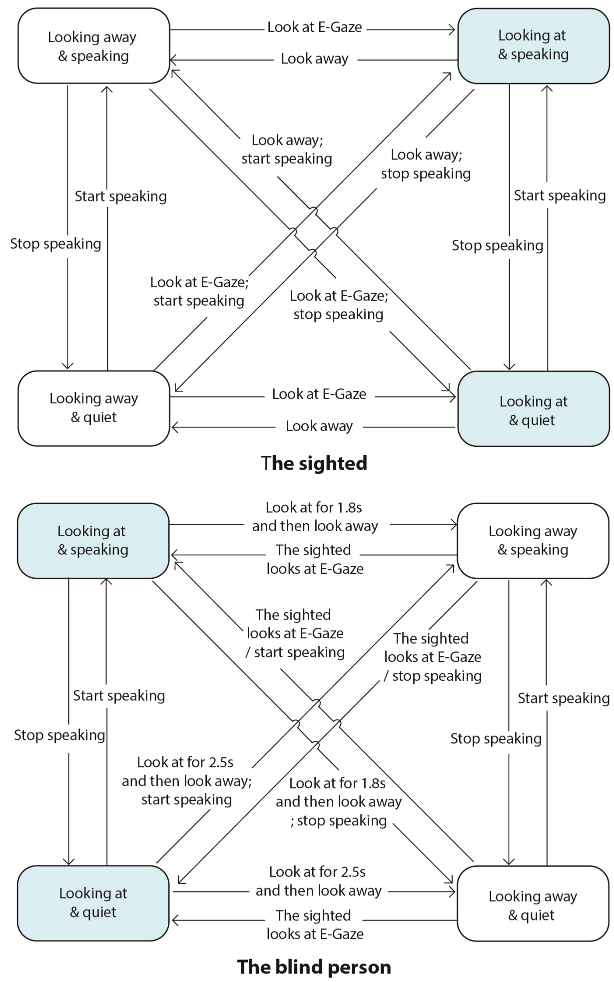

Figure 2.The state machine diagram of the sighted and the blind person with E-Gaze in a dyadic conversation. The sighted and the blind person can establish the eye contact in `LS` and `LQ` states.

`LQ:` when the sighted is not speaking and she is looking towards E-Gaze, E-Gaze displays a "look at" eye gaze for 1.8 seconds to establish an eye contact with the sighted then "look away" for 2.1 seconds. After that, E-Gaze displays an average frequency of 14 "look at" glances per minute if the state continues.

AQ: when the sighted is not speaking and she is not looking towards E-Gaze, E-Gaze displays a "look away" gaze for 2.1seconds to concentrate on what the blind person is going to say. After that E-Gaze displays an average frequency of 14 "look at" glances per minute if the state continues.

AS: when the sighted is speaking and she is not looking towards E-Gaze, E-Gaze displays an average frequency of 17 "look at" glances per minute if the state continues.

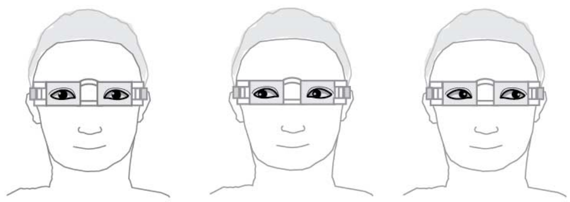

Figure 3 Face-on views of E-Gaze looking "at" and "away".

System

In the previous study[11], we used an Eye Tribe Tracker to detect the eye gaze from the sighted. The sighted used her gaze to control five gaze animations of the E-Gaze: look up, look down, look left, look right and look at. The system was programmed in Java. In this work, the system implements the gaze model. The gaze animation of the E-Gaze is driven by two sensors: the gaze signal from the Eye Tribe and the audio signal from the sound detector. The Eye Tribe is used to detect the gaze signal from the sighted to provide input for eye contact mechanism, while the sound detector is used to detect the blind person's speaking for turn-taking strategy. The E-Gaze system consists of an Eye Tribe Tracker, an Arduino microcontroller, a Bluetooth module, two uOLED-160-G2 display modules with an embedded GOLDELOX graphics processor, a sound detection sensor module and a physical glasses-shaped prototype fabricated using a 3D printer (Figure 4). Figure 5 shows the overview of the E-Gaze.

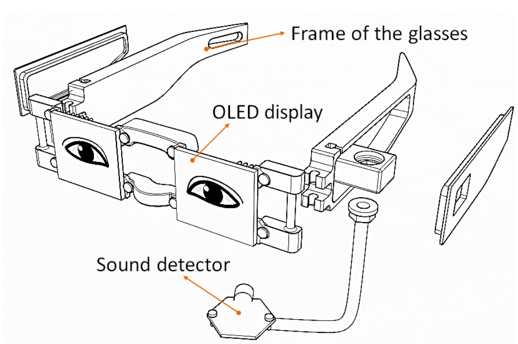

Figure 4.The exploded view of the prototype for the E-Gaze

In a dyadic-conversation setting, in order to detect the blind person's speaking clearly, the sound detector is fixed on a flexible rod that can be adjusted to near her mouth (Figure 4). We lower the sensitivity of the sound detector, to make it only detect the speaking from the blind person rather than the sighted. We use the real human's eye gaze video to display on OLED display. It needs to be saved onto the SD card in a raw format readable by GOLDELOX graphics processors.

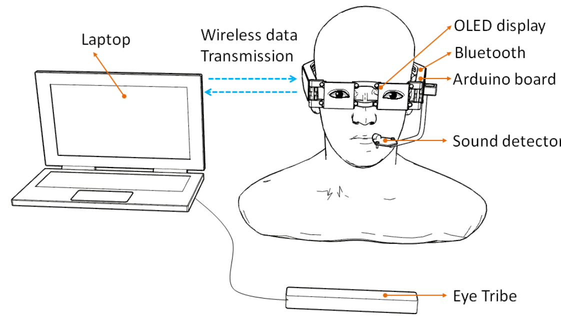

Figure 5.Overview of the E-Gaze system.

For the calibration, a graphical user interface (GUI) with 15 targeted area is created to detect the point of the interest where the sighted fixates in real-time (Figure 6(a)). When the sighted fixates on one of the target areas, the E-Gaze system activates the corresponding area to display the red points. It indicates that the sighted is now looking at the direction of the target area. The target area is parallel with the E-Gaze, so when the laptop is removed after calibration (Figure 6(b)), the sighted is equal to looking at the E-Gaze.

When the sighted looks at the E-Gaze, the system sends the command to the Arduino over a wireless Bluetooth connection. To enable the E-Gaze to interact and respond to the sighted person's eye gaze, the position of the E-Gaze is predetermined, and should be within the Eye Tribe Tracker's tracking area. The Eye Tribe Tracker detects eye gaze from the sighted, and if her gaze point corresponds to the position of the E-Gaze, a command is sent out via Bluetooth adapter from the laptop to a Bluetooth module connected to the Arduino. At the moment, E-Gaze displays a "look at" eye gaze to establish the eye contact with the sighted.

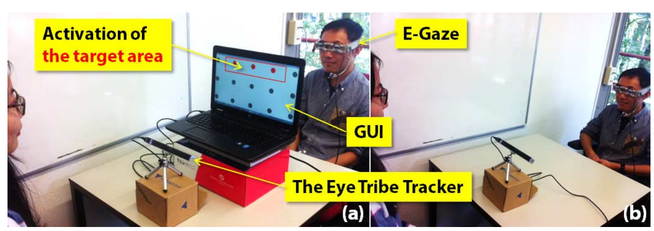

Figure 6.(a)Calibration;(b)E-Gaze.

EXPERIMENTAL DESIGN

We have planned an experiment that uses the E-Gaze system as a test bed to find out how the sighted perceive the blind person with the E-Gaze in face-to-face communication. The purpose of the experiment is two-fold: first to test whether simulating the eye appearance image for the blind person improves the conversation quality with the sighted in face-to-face communication compared to not having the eye gaze image (i.e. E-Gaze without gaze display and it is used only as the black glasses). The second, we are particularly interested in finding out which gaze patterns are appropriate for the blind person in communication. Does the model-driven gaze simulation improve the conversation quality with the sighted compared to constant gaze and random gaze? Based on the literature review, we hypothesize that the model-driven gaze simulation for the blind person can enhance the conversation quality between blind and sighted people in face-to-face communication. A between-subject experiment will be conducted and it includes four conditions (E-Gaze without gaze simulation, E-Gaze with constant gaze, E-Gaze with random gaze and E-Gaze with the model-driven gaze simulation). The level of conversation quality will be measured using questionnaires with four subjective measures: relatedness, partner closeness, engagement and partner evaluation.

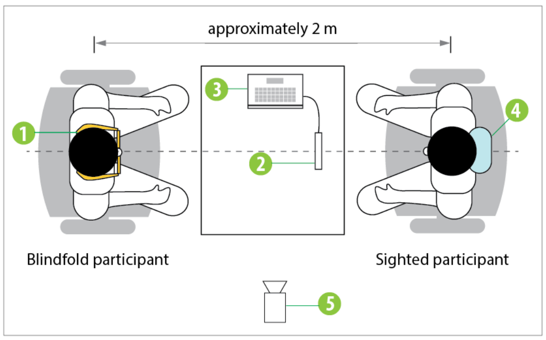

Figure 7.Overhead view of the experiment: (1) E-Gaze; (2) Eye Tribe Tracker; (3) Laptop; (4) the pillow to fix the neck of the sighted participant; (5) observation camera.

In the experiment, a specific goal is to exam the responses of the sighted towards E-Gaze. We will use the blindfolded but sighted participants as an alternative to the blind user. The participants will be 96 student volunteers from the university. They are divided into pairs to have dyadic conversations. One of each pair will wear the blindfold and the E-Gaze. Two participants sit face-to-face to have a conversation. There is a table between two participants on which the Eye Tribe Tracker connected to a laptop is placed. The Eye Tribe Tracker will be installed 40 to 60 cm away from the sighted participant. When the Eye Tribe Tracker is calibrated, the eye tracking software calculates the sighted participant's eye gaze coordinates with an average accuracy of around 0.5 to 1 degree of visual angle. In order to stabilize and track accurate eye gaze, we will place a pillow on the chair to fix the neck of the sighted participant. The observation camera captures the whole dyadic-conversation. Figure 7 shows an overview of the setup. In the next step, we will involve some real blind users to participate in the experiment.

CONCLUSION

We presented a functional prototype of the E-Gaze glasses based on the Eye Tribe tracking system to simulate the model-driven gaze for the blind person, to enhance the engagement between the sighted and the blind person in face-to-face communication. In our future work, we will implement a user experiment that uses the E-Gaze system as a test bed to find out how the sighted perceives the blind person with E-Gaze in face-to-face communication.

REFERENCES

1. Michael Argyle and Mark Cook. 1976. *Gaze and mutual gaze*. Cambridge University Press.
2. David McNeill.1992. *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
3. Shi Qiu, Jun Hu, and Matthias Rauterberg.2015. Nonverbal Signals for Face-to-Face Communication between the Blind and the Sighted. In *Proceedings of International Conference on Enabling Access for Persons with Visual Impairment*, 157–165.
4. Vincent B. Van Hasselt.1983. Social Adaptation in the Blind. *Clinical Psychology Review*, 3(1), 87–102.
5. Justine Cassell, Obed E. Torres, and Scott Prevost.1999. Turn taking versus discourse structure. In *Machine conversations*, 143-153, Springer US.
6. Dirk Heylen, Ivo van Es, Anton Nijholt, and Betsy van Dijk. 2005. Controlling the gaze of conversational agents. In *Advances in Natural Multimodal Dialogue Systems*, 245-262, Springer Netherlands.
7. Maia Garau , Mel Slater, Simon Bee, and Martina Angela Sasse. 2001. The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 309–316.
8. Shi Qiu, Hirotaka Osawa, Jun Hu, and Matthias Rauterberg. 2015. E-Gaze: Create Gaze Communication for People with Visual Disability. In *Proceedings of the 3rd International Conference on Human-Agent Interaction*, 199–202.
9. Michael Kipp and Patrick Gebhard.2008. Igaze: Studying reactive gaze behavior in semi-immersive human-avatar interactions. In *Intelligent Virtual Agents*, 191–199.
10. Nikolaus Bee, Johannes Wagner, Elisabeth André, Thurid Vogt, Fred Charles, David Pizzi, and Marc Cavazza.2010. Discovering eye gaze behavior during human-agent conversation in an interactive storytelling application. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*.
11. Shi Qiu, Siti Aisyah Anas, Hirotaka Osawa, Matthias Rauterberg, and Jun Hu.2016. E-Gaze Glasses: Simulating Natural Gazes for Blind People. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, 563–569.