Impact of Simulated Gaze Gestures on Social Interaction for People with Visual Impairments

Shi QIU[a,1], Ting HAN[b], Matthias RAUTERBERG[a] and Jun HU[a]

[a]*Department of Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands*
[b]*Department of Design, Shanghai Jiao Tong University, Shanghai, China*

Abstract. Gaze and eye contact have important social meanings in our daily lives. The sighted often uses gaze gestures in communication to convey nonverbal information that a blind interlocutor cannot access and react to. In many examples, blind people's eyes are unattractive, and often with deformities, which makes the eye appearance less appealing to the sighted. All of these factors influence the smooth face-to-face communication between the blind and sighted people, which leads to blind people's poor adaptions in the communication. We implemented a working prototype, namely E-Gaze (glasses), an assistive device based on an eye tracking system. E-Gaze attempts to simulate the natural gaze for blind people, especially establishing the "eye contact" between the blind and sighted people to enhance the engagement in the face-to-face communication. The interactive gaze behaviors of the E-Gaze are based on a model that combines a turn-taking strategy and the eye-contact mechanism. In order to test the impact of the interactive gaze model in the face-to-face communication, we conducted an experiment with sixteen participants. In the user experiment, participants had a monologue with a dummy wearing the E-Gaze. Two monologues took place under two experimental conditions (i.e., Interactive Gaze and Random Gaze) with counter balancing to avoid carry-over effects. Results well support the hypothesis that the interactive gaze model of the E-Gaze can enable the sighted to feel attention from the listener, enhancing the level of engagement in the face-to-face communication. We also obtain insights and design implications from participants' comments.

Keywords. Communication quality, social interaction, eye tracking, visual impairment

Introduction

According to the information from the World Health Organization (WHO) in 2017, it is estimated that 253 million people are visually impaired, 39 million are blind and 217 million have low vision worldwide [1]. A critical problem for those people is the poor adaption in social interaction.Kemp [2] conducted a user experiment to investigate how blind people behave in conversations. Thirty blind and thirty sighted participants formed three groups: ten blind-to-blind pairs, ten sighted-to-blind pairs, and ten sighted-to-sighted pairs. Each pair participated in 15-minute discussion sessions and all the discussion sessions were videotaped. He found that fewer physical gestures were

---

[1] Corresponding Author, E-Mail: sqiu@tue.nl.

observed in blind participants, and they were less confident to report their feelings in the tests. Due to the loss of vision, they were more introverted, submissive, and less confident in social interaction.

In many instances, eyes of people with visual impairments are unattractive, and often with deformities, which makes the eye appearance less appealing to the sighted. The impatience, discomfort, or intolerance of the sighted is a crucial reason that affects blind people's involvement in the communication or activities [3]. One typical example could be found in our previous study [4]. We investigated nonverbal signals between the blind and sighted people in social interaction. In interviews, one participant stated that due to an illness of nystagmus, she could not control the movement of the eyeball. Her eye gestures easily misled the sighted during conversations, which greatly troubled her in social interaction. She often felt socially isolated because she could not see and establish the eye contact with the sighted.

To improve the communication quality between the blind and sighted people, we designed and implemented an interactive gaze model displayed on a wearable glasses device, to simulate appropriate gaze behaviors for the blind people. Afterwards, we evaluated the impact of the interactive gaze model, and reported both quantitative and qualitative findings from the user experiment.

1. Related work

We summarize related work from two major aspects: gaze interaction between humans, and gaze interaction between humans and the virtual agents (avatars). Because we cannot directly find the related work of designing gaze behaviors for the blind people, we attempt to borrow the practical approaches on how to model gaze behaviors between humans and the virtual agents (avatars).

In psychology field, researchers explored the gaze behaviors between humans in social interaction. Argyle and Ingham [5] measured the amount of gaze in the dyadic (two-person) conversations. In the experiments, participant pairs had conversations at a normal interaction distance of 1.5-1.8m. About 75% of the time that participants were looking at the speaker while they were listening. Participants that were talking looked less of the time at the listener (41%). The eye-contact in the conversation was about 31% of the time. Kendon [6] conducted an exploratory study to investigate the relationship between the direction of gaze and the occurrence of utterances in dyadic social interaction. He claimed that gaze has an important role of regulating the flow of conversation and communicating emotions and relationships. Vertegaal et al.[7] extended the research scenario from the dyadic conversations to the multiparty conversations. They used an eye tracker to measure the participant's gaze at the faces of conversation partners during four-person conversations. The results demonstrated that gaze was an excellent predictor of conversational attention in multiparty conversations.

Informed by gaze interaction between humans, some studies further investigated the turn-taking to design gaze behaviors for virtual agents. Cassell et al. [8] investigated how to display the appropriate gaze behaviors of conversation agents during the dialogues with a human. They examined the relationship between the gaze behaviors and the turn-taking phenomenon through an empirical analysis for the dyadic conversations. Based on the empirical findings, an algorithm was proposed to design the gaze behaviors of the conversation agents. Garau et al. [9] conducted an experiment

to investigate the importance of gaze gestures between humans and humanoid avatars. They compared participants' responses in four conditions: video, audio-only, random-gaze avatars and inferred-gaze avatars. In the inferred-gaze, gaze behaviors were linked to the turn-taking strategy in the conversations. The results indicated that the inferred-gaze provided remarked improvements of communication quality between humans and humanoid avatars.

Most virtual agent systems display gaze behaviors based on the turn-taking strategy rather than being reactive. In the reactive systems, a user's gaze behaviors trigger a momentary response from the virtual agent, which in turn influences the user and results in a feedback loop[10]. Kipp et al.[10] used head tracking to implement a system named IGaze, aiming at exploring reactive gaze behaviors between human and virtual agents. They implemented a virtual agent with three gaze strategies: Mona Lisa (continuous gaze following), dominant and submissive. An experiment was conducted to test how the participants could perceive three gaze strategies. The results demonstrated that the dominant and submissive strategies conveyed the intended impressions, and the Mona Lisa strategy was positively received by the participants. Bee et al.[11] proposed a gaze model for embodied conversation agents to improve the user experiences in Interactive Storytelling. An eye tracker was used to enable the interactive gaze of the conversation agent to respond to the user's gaze. A user experiment compared the participants' responses towards the gaze model and the non-interactive gaze. The results showed that the gaze model significantly outperformed the non-interactive gaze and provided a better user experience.

2. Design and system implementation

In our prior work [12], we presented a working prototype, namely E-Gaze (glasses), an assistive device based on an eye tracking system. E-Gaze attempted to imitate the appropriate eye gestures for the blind people, especially establishing the "eye contact" between the blind and sighted people, to improve the level of engagement in the face-to-face communication.

Based on the overview of related work, we identified the specific direction of designing gaze behaviors. A interactive gaze model was proposed, combining the eye-contact mechanism (i.e., the reactive gaze) and the turn-taking strategy in the conversations. When turn-taking occurs, the sound sensor of the E-Gaze system will detect the listening and the speaking modes in the conversation flows. We set the detailed timing of the interactive gaze model, which was based on the research on the two-person conversations between a human and a virtual agent [6,11]. In the interactive gaze model, whenever the sighted is looking at the E-Gaze, it reacts to the sighted with a "look at" eye gaze, and holds it for about 2 seconds, attempting to establish the "eye contact". Then it looks away for about 4 seconds, to avoid a dominance being gazing too long. Four eye gaze gestures (i.e., look up, down, left and right) are randomly chosen to display "look away". The timing of the E-Gaze "look at" and "look away" is slightly varied according to whether the blind person is talking or listening. This strategy is based on the experimental studies of Argyle and Cook [13], in which they found that people tend to look more at the interlocutor while listening than while speaking.

Overall, E-Gaze system consists of an Arduino microcontroller board, a Bluetooth module, two OLED display modules with an embedded graphics processor, a sound

detector module, and a physical glasses-shaped prototype. Figure 1 shows the overview of the systems.
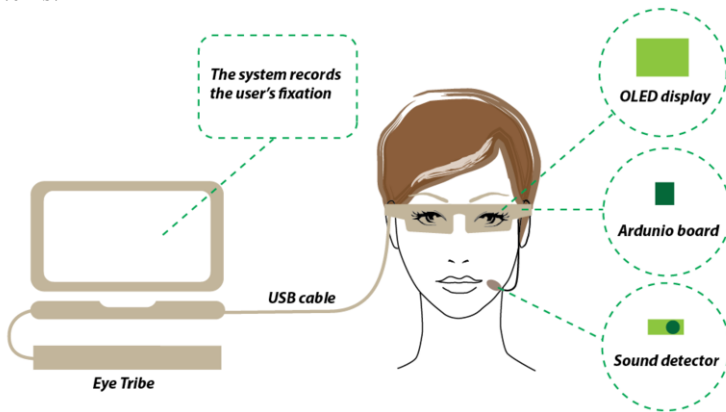

Figure 1. System overview.

In the system implementation, we used the Eye Tribe Tracker[2] to detect the eye gaze from the sighted. An average tracking accuracy was from 0.5º to 1º in visual angle, and the distance between eye tracker and the user should be approximately 60cm. The advantage of using an unobtrusive eye tracker is that users do not need to wear bulky equipment, which makes them aware that their gaze is tracked. The Eye Tribe Tracker provided an input of users' gaze for the eye contact mechanism. It estimated the user's gaze behaviors, and sent the data to the microcontroller. To enable the E-Gaze to interact and respond to the user's gaze, the position of the E-Gaze was predetermined within the tracking area. When the user looked at the area of the E-Gaze, the microcontroller processed the gaze data and animated the corresponding gaze gestures. A sound detector detected whether the blind interlocutor was talking or listening, to provide input for the turn-taking strategy in conversations. The animations of eye gestures were taken from the research prototype "Agency Glasses" [14].

3. Evaluation

We designed a study to test sighted participants' perception towards the E-Gaze, and specifically examine whether the interactive gaze model could affect the communication quality between the sighted and blind people. In our early stage of user experiments, we used a dummy as an alternative to a blind listener in the test, and collected preliminary results to improve the system. The reason of using a dummy rather than a human was: we focused on testing the E-Gaze system, and attempted to avoid the potential interferences from other nonverbal signals of a human (e.g., smiling, nodding). Figure 2 presents a dummy wearing the E-Gaze (glasses).

In the experimental design, we hypothesized that the E-Gaze with the interactive gaze model could enhance the quality of communication between the speaker and the listener. More specifically, two gaze models of the E-Gaze (interactive vs. random) were compared while participants had a monologue with the dummy. A within-subject experiment were designed, and it included two conditions in the following:

---

[2] http://theeyetribe.com/theeyetribe.com/about/index.html

- **Interactive Gaze**: E-Gaze displays gaze gestures based on the interactive gaze model introduced in Section 2. Besides, because the dummy does not speak in this experiment, and it is only a listener, the sound detector was not used in the system. The gaze gestures of the E-Gaze follows the timing of listening mode.
- **Random Gaze**: E-Gaze randomly displays five gaze gestures (i.e., look at, up, down, left, and right). The average duration of each state is around 2 seconds.
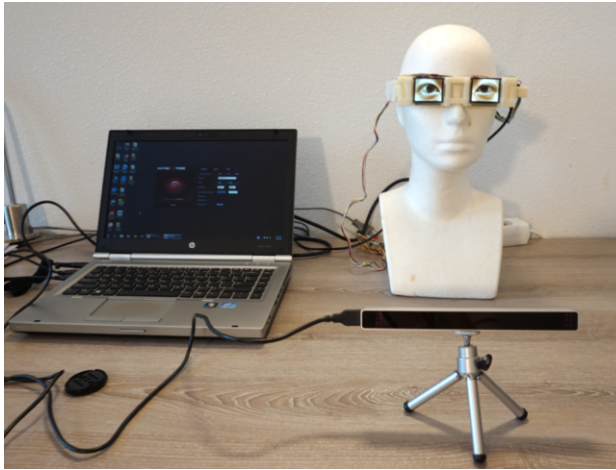

Figure 2. Dummy wearing E-Gaze (glasses).

3.1. Participants

The participants were sixteen students from the university (M_{age} = 22.44, SD = 1.59, 7 females and 9 males) with ages ranging from 19 - 25. None of the participants was visually impaired. We posted the recruitment information on the website (tongqu.me) to find students willing and able to take part in our study. Compensation for each participant was 50 CNY for approximately 1 hour.

3.2. Experimental setup and procedure

In the experiment, the participant sit in front of a round table where a dummy with the E-Gaze were placed. We aligned the Eye Tribe tracker and adjusted it towards the participant's face for the maximum trackability. The tracker connected to a laptop was installed around 50cm away from the participant. When the Eye Tribe Tracker was calibrated, the eye tracking software calculated the participant's eye gaze coordinates with an average accuracy of around 0.5 to 1 degree in visual angle. In order to stabilize and track accurate eye gaze, we used a chair with comfortable support to well fix the neck of the participant. The observation camera captured the whole scene. An overview of the experimental setup is shown in Figure 3.

In the experiment, the participant first signed the consent form and completed the pre-experimental questionnaire. We randomly picked one topic in fourteen sample topics from IELTS oral exams [15]. These topics were all about daily lives and easy for the participant to start a monologue (e.g., "Describe one of your family members").

Five minutes were given to the participant to prepare for the topic. Afterwards, the Eye Tribe Tracker was calibrated for the participant, which took less than 2 minutes. In the test, the participant had a monologue with the dummy. Two monologues took place under the experimental conditions (Interactive Gaze vs. Random Gaze) for the participant with a counter balanced order to avoid carry-over effects. The monologue lasted approximately 10 minutes. After each monologue, the participant was asked to complete the post-experimental questionnaire. Finally, we did a short interview for the participant's comments and suggestions towards the E-Gaze. All the monologues were video-taped, and the overall experiment lasted approximately 60-90 minutes.
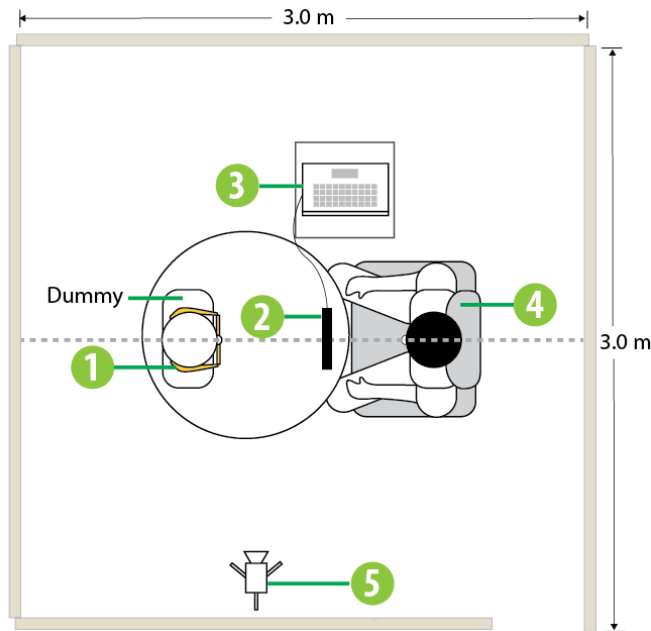

Figure 3. Overhead view of the experimental setup: (1) E-Gaze(glasses), (2) Eye Tribe Tracker, (3) Laptop, (4) support of the chair to stabilize the neck of the participant, and (5) observation camera.

3.3. Measurements

We measured the quality of the face-to-face communication with two subjective measures: social presence, and closeness.

The brief description of "social presence" was "sense of being with another". This other refers to either a human or an artificial agent [16]. In this experiment, we used an adapted version of the Networked Minds Social Presence Inventory (NMSPI) developed by Harms and Biocca [17]. NMSPI included 36 items with 7-point Likert-type scales. It is composed of 6 sub-dimensions: (1) *co-presence,* (2) *attentional allocation,* (3) *perceived message understanding,* (4) *perceived affective understanding,* (5)*perceived emotional interdependence,* and (6) *perceived behavioral interdependence.*

Inclusion of Other in the Self (IOS) Scale [18] was used to measure the closeness between the speaker and listener. It included seven increasingly overlapping circle pairs, indicating the distance of the relationship between themselves and their partners.

We collected participants' comments about the E-Gaze from an open-ended questionnaire including the item "Do you have suggestions for improving the E-Gaze?"

4. Results

Below we report both quantitative and qualitative results from the experiment.

Quantitative results are presented in Table 1 and Table 2. A paired-samples t-test demonstrated that participants experienced significantly greater co-presence to the dummy with Interactive Gaze (M = 5.23, SD =.74) than it with Random Gaze (M = 4.76, SD =.75), t (15) = 3.56, p = .003, r = .68. The effect size of r = .68 represents a large effect (above .5) [19]. As well as being statistically significant, this effect is very large, and represents a substantive finding. There was also a significantly greater perceived message understanding for participants to the dummy with Interactive Gaze (M = 3.62, SD =.66) than it with Random Gaze (M = 3.14, SD = 1.03), t (15) = 2.15, p = .049, r = .49. All other comparisons were non-significant ($p \geq .05$).

Table 1. The means for each dimension of communication quality between Random Gaze and Interactive Gaze.

	Random Gaze Mean	Std. Deviation	Interactive Gaze Mean	Std. Deviation
Co-presence	4.76	.75	5.23	.74
Attention allocation	3.21	1.11	3.48	.93
Perceived message understanding	3.14	1.03	3.62	.66
Perceived affective understanding	3.23	.88	3.49	.70
Perceived emotional interdependence	4.59	.42	4.66	.64
Perceived behavioral interdependence	3.40	.95	3.78	.75
Partner closeness	2.88	1.54	2.94	1.24

Table 2. Paired t-test summary for each dimension of communication quality comparing the average Likert scale answer results between Interactive Gaze and Random Gaze.

	Mean difference	Std. Deviation	t value	df	Sig.
Co-presence	.47	.53	3.56	15	.003[**]
Attention allocation	.27	.84	1.29	15	.216
Perceived message understanding	.49	.91	2.15	15	.049[*]
Perceived affective understanding	.26	.68	1.54	15	.145
Perceived emotional interdependence	.06	.50	.50	15	.622
Perceived behavioral interdependence	.38	1.04	1.47	15	.162
Partner closeness	.06	1.18	.21	15	.835

Significant group difference; [*] p<.05, [**] p<.01

To analyze the qualitative data, open questions in interviews were transcribed and notes were taken. To gain structured insights from the transcripts, we conducted data analysis based on the qualitative content analysis [20], aiming at interpreting the meaning from the context of text data based on a naturalistic paradigm. It consists of three primary approaches: *conventional content analysis*, *directed content analysis*, and *summative content analysis*. In analyzing, we followed the approach *conventional content analysis, in which* the coding categories derived directly from the text data. We identified four major categories from the transcripts: interests, attitudes, design suggestions, and extended potential scenarios.

- Interests

We investigated whether participants had an interest towards the E-Gaze by asking them the question "Do you have an interest towards this system? If yes, why you are interested in this E-Gaze system?" Twelve participants showed a great interest towards E-Gaze while the other four held the opposite ideas. The example reasons for interests

were: "E-Gaze can influence our attitudes in the conversation" (P6); "I am interested in the operating mechanism of the E-Gaze (e.g., how to send us appropriate gaze gestures)" (P9); "I think the E-Gaze is very cool, and it can be used in a game" (P10). However, P16 stated that "It is very pitiful if a person feels lonely, she is only able to communicate with the machine".

- Attitudes

We collected participants' positive and negative attitudes towards the E-Gaze with the question "Which aspects make you like/dislike the E-Gaze?" Example keywords and phrases are presented in Table 3. The majority of participants express the positive attitudes towards the E-Gaze. We found some example descriptions:

"E-Gaze has a similar eye appearance as we could see in our daily livings. Eyes looks vivid, and conveys a warm feeling in communication" (P2).

"Looking downwards of the E-Gaze indicates an engagement. Besides, it can stimulate the enthusiasm of conversation, if we establish the eye contact with the E-Gaze" (P4).

"The eye appearance of the E-Gaze is very realistic, and it looks much better if it has more personalized options (e.g., the eyes with the single eyelid or the double eyelid)" (P16).

In contrast, some participants held negative attitudes towards the E-Gaze:

"Sometimes the E-Gaze looks around, and makes me feel that the dummy has lost patience in listening, therefore it disturbs me in communication"(P9).

"Require a more beautiful design for the eye appearance" (P15).

"If I am visually impaired and wear the E-Gaze glasses, I am afraid of being looked strange in other people's eyes […]" (P10).

Table 3. Exemplar keywords and phrases towards the E-Gaze.

Positive	Negative
Diverse gaze gestures, looking downwards indicates an engagement, acquire one's emotional state through the eye gaze, innovative idea, high-tech, stylish, encourage the communication, realistic eye appearance, etc.	Not very smart, the eye appearance does not match the shape of the face, some eye gestures are not very friendly (e.g., sometimes look around, improperly move the eyes upwards), eye gestures inconsistent with conversation contents, etc.

- Design suggestions

Overall, design suggestions include ideas of improving aesthetics of the eye appearance, displaying appropriate eye gaze gestures in conversations, and linking other nonverbal signals (e.g., smiling, nodding) to express emotions. Example statements are given below:

"E-Gaze is suggested to display proper eye gestures based on the manner of speaking. When the sighted feels sad, the eye gestures of the E-Gaze will reveal a feeling of sympathy to comfort her" (P9).

"E-Gaze will look more realistic by synchronizing the facial expressions" (P9). The similar description was "I think it is not capable of expressing dummy's feelings well only using eye gestures. Synchronizing facial expressions can enhance the effect of feelings, for example, a smile with curved eyes often indicates the happiness" (P16).

"Designing gaze gestures is suggested to be based on the comprehensive sentiment analysis (e.g. the analysis of a person's facial expressions, language, and the speed and manner of speaking)" (P3).

- Potential extended scenarios

The participant majority claimed the recent scenario was the most reasonable (i.e., improve the communication quality between the blind and sighted people), yet several participants pictured other potential scenarios of using the E-Gaze. Example statements are:

"E-Gaze can be installed on a medical robot for the psychological counseling. Sometimes depression people feel safer to interact with a robot rather than a human doctor, especially in the talking treatments. Besides, it is not easy for human doctors to keep mental health if they always hears about the dark side of the world from patients. This field might derive the benefits from the technology of the E-Gaze" (P16).

"E-Gaze is helpful for people with autism, who are not willing to interact with others, especially establishing the eye contact. E-Gaze can assist them to be familiar with gaze gestures before interacting with other people." (P4).

"Some people feel anxious in communication, and interacting with the E-Gaze makes them feel relaxed in social activities" (P5).

5. Conclusion

In general, this pilot study contributes an understanding of how the interactive gaze model of the E-Gaze glasses could impact the quality in the face-to-face communication, how to further design and develop the future smart glasses system based on users' requirements. To design the interactive gaze model, we gained insights from approaches regarding how to model gaze behaviors between humans and virtual agents (avatars). In quantitative findings, the interactive gaze model significantly outperformed the random gaze on *co-presence* and *perceived message understanding*. This suggests that the interactive gaze model of the E-Gaze can indeed make a contribution to increase the quality of the face-to-face communication. These were aligned with the qualitative findings: the participants majority expressed interests and positive attitudes towards the E-Gaze.

Nevertheless, our study has some limitations. In the experiment, ten-minute monologue seems too long for participants. Some of them felt exhausted to present a long monologue, which might trigger negative emotions. Our final goal is to test the efficiency of the E-Gaze in conversations. The results may be various between the conversation and monologue scenarios. In this experiment, we used a dummy as an alternative to a real blind listener. Participants' perception may be somewhat different if they face a real blind person wearing the E-Gaze.

Despite the limitations discussed, design suggestions from participants identify our future smart glasses system: simulating gaze behaviors based on the multiple sensory inputs in a social scenario. For example, eye gaze behaviors in sync with facial expressions, speaking pace, tone etc. E-Gaze may be applied in the other special user groups (e.g., people with depression or autism). In our future studies, we will recruit real blind participants, and test the communication quality between the blind and sighted people in a conversation scenario.

Acknowledgement

References

[1]   World Health Organization, *Visual impairment and blindness*, Oct-2017, Accessed: Dec, 27 2017, http://www.who.int/mediacentre/factsheets/fs282/en/

[2]   N. J. Kemp, Social interaction in the blind, *International Journal of Rehabilitation Research,* Vol. 3, 1980, No. 1, pp. 87–88.

[3]   V. B. Van Hasselt, Social Adaptation in the Blind, *Clinical Psychology Review,* Vol. 3, 1983, No. 1, pp. 87–102.

[4]   S. Qiu, J. Hu and M. Rauterberg, 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*, 2015, pp. 157–165.

[5]   M. Argyle and R. Ingham, Gaze, mutual gaze, and proximity, *Semiotica*, Vol. 6, 1972, No. 1, pp. 32–49.

[6]   A. Kendon, Some functions of gaze-direction in social interaction, *Acta psychologica*, Vol. 26, 1967, pp. 22–63.

[7]   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, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2001, pp. 301–308.

[8]   J. Cassell, O. E. Torres and S. Prevost, Turn taking versus discourse structure, in *Machine conversations*, Springer, 1999, pp. 143–153.

[9]   M. Garau, M. Slater, S. Bee and M. A. Sasse, The impact of eye gaze on communication using humanoid avatars, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2001, pp. 309–316.

[10]   M. Kipp and P. Gebhard, Igaze: Studying reactive gaze behavior in semi-immersive human-avatar interactions, in *Intelligent Virtual Agents*, 2008, pp. 191–199.

[11]   N. Bee *et al.*, 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*, 2010.

[12]   S. Qiu, S. A. Anas, H. Osawa, M. Rauterberg and J. Hu, E-Gaze Glasses: Simulating Natural Gazes for Blind People, in *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, 2016, pp. 563–569.

[13]   M. Argyle and M. Cook, *Gaze and mutual gaze*, Cambridge University Press, Boston, 1976.

[14]   H. Osawa, Emotional cyborg: human extension with agency for emotional labor, in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 108–108.

[15]   IELTS Speaking Module - Part 2 - Sample Topics, *Good Luck IELTS*, 2012, Accessed: Dec,08 2017, http://www.goodluckielts.com/IELTS-speaking-topics-2.html.

[16]   F. Biocca, C. Harms and J.K. Burgoon, Toward a more robust theory and measure of social presence: Review and suggested criteria, *Presence: Teleoperators and Virtual Environments,* Vol. 12, 2003, No. 5, pp. 456–480.

[17]   C. Harms and F. Biocca, Internal consistency and reliability of the networked minds measure of social presence, In *Seventh Annual International Workshop: Presence 2004*, Universidad Politecnica de Valencia, Valencia, 2004.

[18]   A. Aron, E. N. Aron and D. Smollan, Inclusion of Other in the Self Scale and the structure of interpersonal closeness, *Journal of Personality and Social Psychology,* Vol. 63, 1992, No. 4, p. 596-612.

[19]   A. Field, *Discovering statistics using SPSS*, Sage publications, Los Angeles, 2009.

[20]   H.-F. Hsieh and S. E. Shannon, Three approaches to qualitative content analysis, *Qualitative health research,* Vol. 15, 2005, No. 9, pp. 1277–1288.