

Interacting with a gaze-aware virtual character

Nikolaus Bee, Johannes Wagner, Elisabeth André, Fred Charles, David Pizzi,
Marc Cavazza

Angaben zur Veröffentlichung / Publication details:

Bee, Nikolaus, Johannes Wagner, Elisabeth André, Fred Charles, David Pizzi, and Marc Cavazza. 2010. "Interacting with a gaze-aware virtual character." In *EGI/HMI '10: Proceedings of the 2010 workshop on Eye gaze in intelligent human machine interaction, Hong Kong, China, February 07 - 07, 2010*, edited by Elisabeth André and Joyce Y. Chai, 71–77. New York, NY: ACM Press. <https://doi.org/10.1145/2002333.2002345>.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

Deutsches Urheberrecht

Weitere Informationen finden Sie unter: / For more information see:

<https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/>



Interacting with a Gaze-Aware Virtual Character

Nikolaus Bee, Johannes Wagner
and Elisabeth André

Institute of Computer Science
Augsburg University
Universitätsstr. 6a, 86135 Augsburg,
Germany
{bee, wagner,
andre}@informatik.uni-augsburg.de

Fred Charles, David Pizzi
and Marc Cavazza

University of Teesside, School of Computing
Borough Road, Middlesbrough
TS1 3BA, United Kingdom
{f.charles, d.pizzi, m.o.cavazza}@tees.ac.uk

ABSTRACT

In this paper, we present the user's attentive state interpreted through eye gaze while interacting with a virtual character. The underlying narrative in which the approach was tested is based on a classical XIXth century psychological novel: *Madame Bovary*, by Flaubert. We connected a remote eye tracker with a dynamic 3D world. An empirical study revealed individual user experiences and behavioral patterns. In particular, we identified two different groups of users: one that was showing natural eye gaze behaviors with rhythmic eye gaze shifts between the characters' eyes, face and the scene and another one permanently staring at the character. Interestingly, the group with more natural behaviors towards the character also rated the experience with the system more positively.

INTRODUCTION

A frequent metaphor for interactive storytelling is that of the Holodeck [3, 7, 12, 14], the science-fiction ultimate entertainment system, where narratives take the form of virtual reality world in which the user is totally immersed, interacting with other characters and the environment in a way which drives the evolution of the narrative. Implementing the interactive storytelling concept involves many computing technologies: virtual or mixed reality for creating the artificial world, and artificial intelligence techniques and formalisms for generating the narrative and characters in real time. As a character in the narrative, the user communicates with virtual characters much like an actor communicates with other actors. This requirement introduces a novel context for multimodal communication as well as several technical challenges. Acting involves attitudes and body gestures that are highly significant for both dramatic presentation and communication.

In this paper, we track user's eye gaze to detect their level of attention during an interaction with the agent. On the one

hand, this information is used to enable more natural eye gaze behaviors from the agent. For example, the agent would avert her gaze when the users are staring at her. On the other hand, the users' eye gaze behavior is employed as an assessment for social presence. In the context of human-agent communication, the degree of social presence characterizes to what extent a human is aware of a virtual agent. There is evidence of a high degree of social presence if humans behave towards virtual agents in a similar way as towards other humans. We therefore analyze the users' eye gaze behavior during an interaction with the agent in order to find out to what extent they are experiencing a feeling of social presence.

RELATED WORK

A number of studies informed by human-human conversation that investigate the role of eye gaze in human-agent communication provide evidence that natural eye gaze behaviors of an agent are not only more positively perceived, but elicit more natural responses in human users (see, for example, [4, 9, 13, 23]).

Colburn and colleagues [4] investigated whether natural eye gaze behaviors of an avatar elicit more natural eye gaze behaviors in users communicating with it. When an avatar was present, subjects spent more time looking at the screen. Even more attention was directed to the avatar when the agent relied on an eye gaze model that was informed by psychological studies on human-human conversation. Colburn and colleagues hypothesize that humans feel less shy when talking to a monitor than when talking to a real human. The effect occurred, however, only in the user-as-speaker condition which Colburn and colleagues attribute to the bad quality of the employed lip-synch mechanism. While Colburn and colleagues concentrate on the behavioral response to avatars employing an informed eye gaze model, Garau and colleagues [9] as well as Lee and colleagues [13] investigate the effect of informed gaze models on the perceived quality of communication by means of questionnaires. Both research teams observed a superiority of informed eye gaze behaviors over randomized eye gaze behaviors. A follow-up study by Vinayagamoorthy and colleagues [23] focused on the correlation between visual realism and behavioral realism. They found that the model-based eye gaze model improved the quality of communication when a realistic avator

was used. For cartoonish avatars, no such effect was observed. While all these approaches modify parameters, such as the timing of changes in eye gaze, depending on whether the agent is speaking or listening, they do not track the users' eye gaze behaviors.

Step toe and colleagues [21] used mobile eye trackers in order to drive the eye gaze behaviors of a user's avatar in a multiparty CAVE-based system. They found that eye gaze behaviors known from human-human communication also occurred in their 3D environment. For example, participants looked at the speaker when being asked a question or looked away when thinking of an appropriate response. The avatars in their 3D environment just mimicked, however, the eye gaze behavior of the human users and did not generate eye gaze behaviors autonomously.

Rehm and André [17] described an experiment where they investigated the user's level of attention in a multi-party scenario consisting of two human and one synthetic interlocutors. Their agent was not able to perceive the users. However, since the conversation followed a pre-defined sequence of turns, the agent knew whether the user to her left or to her right was speaking and could move her head into that direction. Similar to Step toe and colleagues, they found that certain eye gaze practices known from human-human conversation were followed. However, the users looked significantly more often to the agent when she was talking to them than when a human user was talking to them. The experiment left open whether this difference was caused by the novelty effect of the agent or by difficulties of the users to understand the agent.

Many systems investigating interactive models of visual attention make use of head trackers. They are able to roughly assess in which direction the user is looking, but do not have more detailed information on the user's eye gaze direction. Another application using an virtual agent is the MACK system [15]. The authors use a head tracker to determine a user's gaze in a direction giving task. The animated agent explains directions on a map and monitors the user's head. In that application, lack of negative feedback indicates successful grounding. If grounding fails, the agent will perform a repair action to help the user. Based on an analysis of human-human conversation, Sidner and colleagues [19] developed a model of engagement for a conversational robot that was able to track the user's face and adjusted its gaze accordingly. Even though the set of communicative behaviors of the robot was strongly limited, an empirical study revealed that users indeed seem to be sensitive to a robot's conversational gestures and establish mutual gaze with it.

One of the earliest work of using eye trackers for agent-based human interaction comes from Starker and Bolt [20]. They adapt "The Little Prince" to the users' current interest in a virtual scene that shows one planet from the story by Antoine de Saint-Exupéry. Dependent on the duration and focus of the user's gaze further details of the scene are described via a text-to-speech system. Another example is the FRED system [22] that makes use of 3D animated facial

agents in a multi-agent setting that are controlled by a conversational gaze model. The agents have the capability to notice if the user (or another agent) is looking at them. Together with the speech data they can determine if they have to listen to someone else or if they can talk. The focus of that work is the regulation of conversational flow in a multi-agent environment. Unlike Vertegaal and colleagues, Eichner and colleagues [5] who made use of an eye tracker to detect interest and attentiveness in a virtual presentation. In an experiment, they showed that agents that adapted the content of their presentation to a user's eye gaze were perceived as more natural and responsive than agents that did not have that capability.

VIRTUAL CHARACTER

The background narrative for our IS system is an adaptation of three chapters of the XIXth century classic *Madame Bovary* by Gustave Flaubert [6]. Emma Bovary is married to a country doctor, Charles Bovary, but boredom in her married life has drawn her towards Rodolphe Boulanger. The user plays the role of Rodolphe who may address Emma or respond to her complaints and love declarations.

Facial Expression and Speech

Emma (see Fig. 1), was enhanced to use the FACS to synthesize a huge set of different facial expressions. The action units were designed using morph targets and thus give the designer the full power in defining the facial expression outlook. The system includes a tool to control the single action units [2]. The tool allows us to store the result in an XML file for later usage in our agent system. For speech, the system interfaces the Microsoft Speech API to synchronize the audio output with the lip movements. As FACS defines several action units involving mouth muscles (e.g. lip funneler, lip tightener, mouth stretch), we utilize the FACS system for lip movements. The approach is similar to displaying facial expressions. The output from the editor to modify the single action units is stored in an XML file. Reusing the FACS approach for visemes enables Emma to display facial expressions and lip movements in parallel. For rendering the character and its animations the Horde3D GameEngine [1] is used.

Gaze Model

In our work, we start from the gaze model developed by Fukayama and colleagues [8] which allows us to specify a number of gaze parameters that influence the impression a character conveys. Their model includes two states: looking at the user and averting the gaze from the user. Three parameters define how often, how long (500 to 2000 ms) and where the virtual agent looks. The gaze targets consist of a set of random points from either all over the scene, above, below or close to the user. They found that a medium amount of gaze and a mean duration between 500 to 1000 ms conveys a *friendly* gaze behavior.

Our gaze model was extended with further parameters as our virtual agent is capable to react to the user's current gaze using an eye tracker. The maximal and minimal duration of mutual gaze can now be set as well. Furthermore, we

may indicate the maximal duration the virtual agent gazes around. Finally, we may specify how long the virtual agent waits until the user responds with mutual gaze. We modeled three different gaze modes for our agent. In the *non-interactive normal* mode, the character looks for about 2 s (between 1 and 3 s) at the user before she averts her gaze again. The agent’s gaze model in the *interactive* mode is parameterized as in the non-interactive condition, but the agent notices whether the user is looking at her or not and responds accordingly, for example, tries to establish mutual gaze or looks away when the user starts staring. In particular, the virtual character looks on average 2.5 s (2-3 s) at the user until she meets the user’s eyes. She then tries to hold the mutual gaze on average for about 2 s (1-3 s) before she looks away to repeat the same behavior again. This protocol is inspired by studies on human eye gaze behaviors, see, for example, [10]. In the *non-interactive staring* mode the agent’s gaze model is parameterized in such a way that the agent seems to stare at the user and the virtual character gazes on average for a period of 6.5 s (5-8 s) directly at the user. In each mode, the character looks away on average for about 4.5 s (3-6 s) before she looks back to the user.

System for Interaction

Many systems investigating interactive models of visual attention make use of head trackers [15, 19]. They are able to roughly assess in which direction the user is looking, but do not have more detailed information on the user’s eye gaze direction. In our work, we make use of the SMI iView X RED eye tracker. It operates with a sampling rate of 50 Hz, the latency for a gaze point is less than 35 ms and the tracking accuracy is less than 0.5° . The distance between the eye tracker and the user should be about 60 - 80 cm. The advantages of an unobtrusive, contact-less eye tracker include that users do not have to wear a sometimes bulky apparatus and thus are not steadily reminded that their gaze is tracked. Further, the SMI iView X RED eye tracker allows head movements horizontally and vertically up to 20 cm in each direction.

To find fixations we use the I-DT algorithm described by Salvucci et al. [18]. According to I-DT, a fixation is detected when the eye coordinates of a frame lie within the distribution *disp*. For each frame *disp* is calculated with the following formula:

$$disp = (max_x - min_x) + (max_y - min_y)$$

where min_x , max_x , min_y and max_y are the minimum and maximum coordinate values of all points inside the frame. If *disp* is beyond a certain threshold the current frame is detected as the beginning of a fixation and then expanded by following points until the threshold is exceeded. This marks the end of a fixation. The samples in the final window are averaged to a single fixation point. For our purpose a minimum length of 120 ms and threshold of 15 pixels have been found to give reasonable results.

We connect the eye tracker with the Horde3D GameEngine to detect where the user is looking at in the dynamic 3D scene. We use invisible hull objects around the virtual char-

acters’ eyes and head to detect, where the user is looking at. Ray casting allows us to map the screen coordinates obtained from the eye tracker to the objects in the virtual world. In this vein, we are able to detect whether the user looks at the virtual agent, the left eye or the right eye or something else in the virtual scene. Further, it allows us to detect the focused object in the 3D world in real-time. This was necessary for the eye gaze based interaction on a level of mutual gaze and to see if the user is looking at the virtual character’s eyes, face or away. The hull objects cover a bigger space as the eyes are usually jittering around a fixation point and thus it is sometimes not obvious what the user is exactly focusing [16]. With a bigger hull object around the eyes we make sure the jittering is taken into account.

Setting

The optimal dimensions for a video projector based eye tracking setting are limited. The user is placed in front of a table on which the eye tracker was placed. The eye tracker with an incline of 23° is placed 80 cm above ground and 140 cm away from the projection surface. The user is seated 60 - 80 cm in front of the eye tracker. In total the user is about 2 m away from the virtual agent, which is within the *social space* according to [11]. The projection surface sizes 120×90 cm, which displays the virtual agent in life-size (see Fig. 1).



Figure 1. Set-up for the interaction with Emma.

EVALUATION

In the following, we present the results of a study we conducted in order to find out how users respond to attentive characters in an interactive storytelling environment. As a first step, we started on a very early phase of human-agent conversation, namely the initiation of contact, and compared the three different gaze behavior models introduced in Section : *non-interactive standard*, *interactive* and *staring*.

We conducted a study with eleven male subjects to investigate their gaze behavior in relation to their social presence perception while interacting with a virtual character. The order of the three gaze models (i.e. *interactive*, *non-interactive normal* and *non-interactive staring*) was randomized for each subject to avoid any bias due to ordering effects. The procedure was as follows: First, the subjects were placed in front of the eye tracker. Then a calibration, which took less than 2 minutes, was carried out. The subjects were first informed about the background of the story. Then,

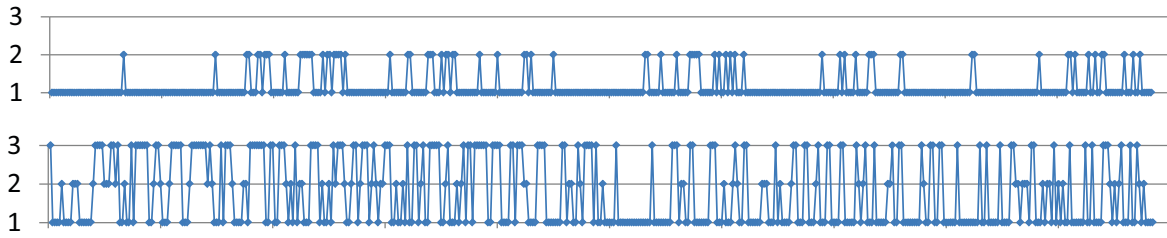


Figure 2. Comparison of two typical eye gaze behaviors for a person that stares (upper chart) and a person that gazes natural (lower chart) while interacting with the virtual character (1: eyes, 2: head, 3: scene).

they were told that they would enter the story in the role of Rodolphe who finds Emma alone in the salon and should try to engage her in a conversation by initiating eye contact with her, which means to look at her and away to establish and break mutual gaze. For instance, to engage Emma with your eye gaze the user could look at her and whenever she is looking at the user, the user could look away, like it is often described in flirt behavior to increase the opposites attraction. They were informed that they would have to run the eye gaze interaction sequence three times, but they did not know that there were different modes of gaze behavior.



Figure 3. Eye gaze heat map of fixations for the *staring* group.

Investigation of the Subjects' Eye Gaze Behaviors

First of all, we investigated to what extent the subjects' eye gaze behavior can be considered as natural. We calculated the fixation points out of the raw eye gaze data. Further, we divided the scene into three dynamic areas. The first area covers the eyes of the virtual character, the second area the rest of the character's head and the third area the rest of the scene. Figure 2 shows two examples of gaze behavior. The upper chart shows a subject that solely stares at the virtual character's eyes (1) and sometimes at the head (2) during the whole interaction. It is apparent from the chart that the subject never fixated anything in the scene, but was only concentrating on the virtual character's face, where the focus lay on the virtual character's eyes. The lower chart shows a subject that not only stared at the virtual character's eyes (1) or head (2) but also averted the gaze from the character (3). These two examples were prototypical for the users' gaze behavior in this interaction.

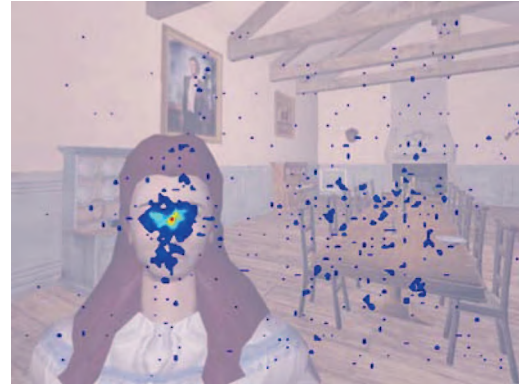


Figure 4. Eye gaze heat map of fixations for the *non-staring* group.

Measurement of Social Presence, Engagement and Interactional Rapport

Since the subjects' eye gaze behavior was extremely different, we decided to group the subjects into *starkers*, who mainly fixated the virtual character's eyes and head (see Fig. 3) and *non-starkers*, who rhythmically averted their gaze between the character and the scene during the interaction (see Fig. 4), and to evaluate their sense of social presence separately. To group the subjects into *starkers* and *non-starkers*, we took the average over the total fixations time of the subjects looking at the scene during the whole interaction within all three interaction modes and split them into those fixating the scene above average and below average. Out of the eleven subjects, eight were grouped as *starkers* and three were grouped as *non-starkers*. The objective of the study was to find out whether the different modes had any impact on the subjects' experience ratings. In particular, we used a post-questionnaire which used 22 attitude statements with a 9-ary rating scale (from strongly disagree to strongly agree) to assess the subjects' sense of social presence (P), their level of rapport with the character (B), their engagement (E and I), the social attraction of the character (T) and the subjective perception of the story (S). In addition, we assessed the subjects' feeling of social presence by behavioral eye gaze measurements. In particular, we investigated to what extent the subjects' eye gaze behaviors could be considered as natural.

Results of Questionnaire

The significance analyses were conducted using a two-tailed *t*-test. The mean over all social presence questions was 2.9

for the *staring* group and 5.1 for the *non-staring* group. A two-tailed *t*-test reveals significant differences between those two groups in presence perception throughout all three interaction modes ($t(196) = 6.8, p < 0.001$). A look at Figure 5 reveals that all attitude statements got more positive ratings from the *non-starers* than from the *starers*. For social presence, all but one difference in the ratings were significant. The *non-starer* rather had the feeling that Emma was flirting with them (P1: $t(31) = 2.9, p < 0.01$), that she was aware of them (P2: $t(31) = 3.8, p < 0.001$), that she had personal contact with them (P3: $t(31) = 3.7, p < 0.001$), that she was more warm-hearted P5: $t(31) = 2.7, p < 0.05$) and that she was more open (P6: $t(31) = 2.8, p < 0.01$). Both groups did not give significantly different ratings for the statement "Emma was personal." (P4). The *non-starers* also built better rapport with the character than the *non-starers*. They found Emma's behavior more natural (B8: $t(31) = 2.8, p < 0.01$) and thought her behavior was more synchronized with their own behavior (B10: $t(31) = 2.9, p < 0.01$). The differences for the other three attitude statements "I would have liked to continue the interaction with Emma." (B7), "I had the impression that Emma responded to me" (B9) and "It was fun to interact with Emma." (E11) were not significant. Furthermore, their engagement was higher. They showed a greater interest in continuing the interaction with Emma (I14: $t(31) = 2.4, p < 0.05$). However, they did not find it significantly easier to flirt with Emma (I12) and they did not feel significantly more immersed in the scenario (I13). The *non-starers* did not find Emma significantly more socially attractive than the *starers*. The differences for the attitude statements "I did not feel stared at by Emma." (T15), "I had the feeling that Emma was interested in me." (T16), "I trusted Emma." (T17) and "I found Emma sympathetic." (T18) were not significant. Finally, the *non-starers* rated their experience with the story more positively than the *starers*. They were significantly more curious how the story would go on (S19: $t(31) = 2.7, p < 0.05$) and they had more than the other group the feeling that they were able to influence the course of the story with their gaze behaviors (S22: $t(31) = 3.0, p < 0.01$). The differences for the attitude statements "Emma behaved appropriately according to her role" (S20) and "It was easy for me to feel in the role of Rodolphe" (S21) were not significant.

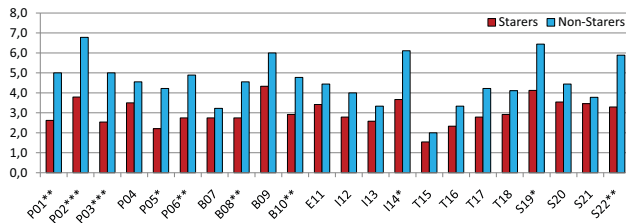


Figure 5. Results for the questions compared with the *staring* and *non-staring* group (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Analysis of the Subjects' Eye Gaze Behaviors

For the analysis of the subjects' eye gaze behaviors, we compare the fixation times over the different interaction modes (i.e. interactive, non-interactive normal and non-interactive staring) and at which area (i.e. eyes or rest of the face and

anywhere else) of the 3D scene the users are looking at. Comparing the fixation times of the *starer* group focusing on the eyes of the virtual character does not reveal significant differences with ANOVA ($F(2, 4861) = 2.25, p = 0.105$) between the three different interaction modes (see Fig. 6 left). Also, the fixation times of this group looking anywhere else does not reveal significant differences ($F(2, 3043) = 1.12, p = 0.327$). But looking at the *non-starer* group, ANOVA reveals significant differences between the fixation times of the different interaction modes ($F(2, 1426) = 5.97, p < 0.01$). The post-hoc Tukey-HSD analysis reveals that the interactive mode differs significantly from the normal ($p < 0.01$) and from the staring mode ($p < 0.05$) (see Fig. 6 right). In contrast, the fixation times looking at somewhere else in the face or the scene does not reveal any significant differences within this group during the three interaction modes ($F(2, 2307) = 1.86, p = 0.16$).

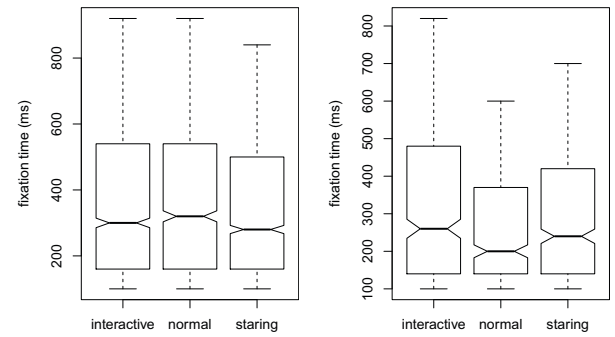


Figure 6. Comparison of all three gaze modes where the users are gazing at the virtual character's eyes (left: *starer*, right: *non-starer*).

Discussion

Overall, the evaluation led to interesting results regarding individual user experiences. First, we identified two groups of users interacting with the virtual character. One group, which we named *starer* was continuously focusing on the virtual character's face. Here, the average interaction with Emma lasted 166.8 seconds. Such a behavior could be considered as unnatural since in human-human interaction interlocutors do not stare at the counterpart's face permanently. Furthermore, we found that the *starer* group experienced a reduced feeling of social presence. The correlation between the *non-starer* group and the significant variations in their gaze behavior shows that this group does not only give more positive ratings for social presence, but also behaves more naturally when interacting with the character. Finally, the *non-starer* group showed a different eye gaze behavior depending on the interaction mode.

CONCLUSION

In this paper, we integrated the user's eye gaze into an existing storytelling system to enable richer interactions with the characters. We were interested in interactions between humans and characters. When conducting an empirical study we identified two different groups of users: one that was permanently staring at the character and one that showed more natural eye gaze behaviors with rhythmic eye gaze shifts between the characters' eyes, face and the scene. We further

found that the first group did not only rate their experience of social presence more negatively, but also did not change their eye gaze behavior when conversing with the different character versions. No matter whether the character was in the interactive, non-interactive normal or non-interactive staring mode, their fixation behavior was more or less the same while the second group showed a different fixation behavior depending on whether the character was in the interactive mode or not. We found, however, no significant differences in the ratings for social presence for the different interaction modes - neither for the first nor for the second group.

Acknowledgments

This work has been funded in part by the European Commission under the grant agreement IRIS (FP7-ICT-231824).

REFERENCES

1. Augsburg University. Horde3D GameEngine. <http://mm-werkstatt.informatik.uni-augsburg.de/projects/GameEngine/>.
2. N. Bee, B. Falk, and E. André. Simplified facial animation control utilizing novel input devices: A comparative study. In *International Conference on Intelligent User Interfaces (IUI '09)*, pages 197–206, 2009.
3. M. Cavazza, J. L. Lugin, D. Pizzi, and F. Charles. Madame bovary on the holodeck: immersive interactive storytelling. In *MULTIMEDIA '07: Proc. of the 15th international conference on Multimedia*, pages 651–660. ACM, 2007.
4. A. Colburn, M. Cohen, and S. Drucker. The role of eye gaze in avatar mediated conversational interfaces. Technical report, Microsoft Research, 2000.
5. T. Eichner, H. Prendinger, E. André, and M. Ishizuka. Attentive presentation agents. In *Intelligent Virtual Agents (IVA 2007)*, pages 283–295, 2007.
6. G. Flaubert. *La revue de Paris*. France, 1856.
7. J. Foote. Where's my holodeck? the new frontier of media display. *IEEE MultiMedia*, 13(3):104+, 2006.
8. A. Fukayama, T. Ohno, N. Mukawa, M. Sawaki, and N. Hagita. Messages embedded in gaze of interface agents — impression management with agent's gaze. In *CHI '02: Proc. of the SIGCHI conference on Human factors in computing systems*, pages 41–48, New York, NY, USA, 2002. ACM Press.
9. M. Garau, M. Slater, S. Bee, and M. A. Sasse. The impact of eye gaze on communication using humanoid avatars. In *CHI '01: Proc. of the SIGCHI conference on Human factors in computing systems*, pages 309–316. ACM Press, 2001.
10. D. B. Givens. The nonverbal basis of attraction: flirtation, courtship, and seduction. *Psychiatry*, 41(4):346–359, 1978.
11. E. T. Hall. A system for notation of proxemic behavior. *American Anthropologist*, 65:1003–1026, 1963.
12. R. Hill, J. Gratch, W. L. Johnson, C. Kyriakakis, C. LaBore, R. Lindheim, S. Marsella, D. Miraglia, B. Moore, J. Morie, J. Rickel, M. Thiébaux, L. Tuch, R. Whitney, J. Douglas, and W. Swartout. Toward the holodeck: integrating graphics, sound, character and story. In *AGENTS '01: Proc. of the fifth international conference on Autonomous agents*, pages 409–416. ACM, 2001.
13. S. P. Lee, J. B. Badler, and N. I. Badler. Eyes alive. *ACM Transactions on Graphics*, 21(3):637–644, 2002.
14. J. H. Murray. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. The MIT Press, 1998.
15. Y. I. Nakano, G. Reinstein, T. Stocky, and J. Cassell. Towards a model of face-to-face grounding. In *ACL '03: Proc. of the 41st Annual Meeting on Association for Computational Linguistics*, pages 553–561. Association for Computational Linguistics, 2003.
16. Z. Prasov and J. Y. Chai. What's in a gaze? the role of eye-gaze in reference resolution in multimodal conversational interfaces. In *ACM 12th International Conference on Intelligent User Interfaces (IUI)*, 2008.
17. M. Rehm and E. André. From chatterbots to natural interaction - face to face communication with embodied conversational agents. *IEICE Transactions on Information and Systems, Special Issue on Life-Like Agents and Communication*, 88-D(11):2445–2452, 2005.
18. D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *ETRA '00: Proc. of the symposium on Eye tracking research & applications*, pages 71–78. ACM Press, 2000.
19. C. L. Sidner, C. D. Kidd, C. Lee, and N. Lesh. Where to look: a study of human-robot engagement. In *IUI '04: Proc. of the 9th international conference on Intelligent user interfaces*, pages 78–84. ACM Press, 2004.
20. I. Starker and R. A. Bolt. A gaze-responsive self-disclosing display. In *CHI '90: Proc. of the SIGCHI conference on Human factors in computing systems*, pages 3–10. ACM, 1990.
21. W. Steptoe, R. Wolff, A. Murgia, E. Guimaraes, J. Rae, P. Sharkey, D. Roberts, and A. Steed. Eye-tracking for avatar eye-gaze and interactional analysis in immersive collaborative virtual environments. In *CSCW '08: Proc. of the ACM 2008 conference on Computer supported cooperative work*, pages 197–200. ACM, 2008.
22. 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 *CHI '01: Proc. of the SIGCHI conference on Human factors in computing systems*, pages 301–308. ACM Press, 2001.

23. V. Vinayagamoorthy, M. Garau, A. Steed, and M. Slater. An Eye Gaze Model for Dyadic Interaction in an Immersive Virtual Environment: Practice and Experience. *Computer Graphics Forum*, 23(1):1–11, 2004.