Eye-to-Eye Contact for Life-Sized Videoconferencing

H. Regenbrecht[1], L. Müller[1], S. Hoermann[1], T. Langlotz[1], M. Wagner[2], M. Billinghurst[3]
[1]University of Otago, Department of Information Science
P.O. Box 56, 9054 Dunedin, New Zealand
{holger|lmuller|shoermann|tlanglotz}@infoscience.otago.ac.nz
[2]IPO.Plan GmbH, Germany
info@ipoplan.de
[3]HITLabNZ, University of Canterbury, New Zealand
mark.billinghurst@canterbury.ac.nz

Figure 1: i2i system setup (left), see-through view shows the videoconferencing camera when no image is projected (centre), video conferencing result of our system showing eye-to-eye-contact (right).

ABSTRACT
Videoconferencing systems available for end users do not allow for eye-to-eye contact between participants. The different locations of video camera and video display make it impossible to directly look into each others eyes. This issue is known as the lack of mutual gaze. Combined with a lack of a life-sized video image of the communication partner videoconferencing becomes an artificial experience leading to decreased communication quality, empathy and trust. In this work, we present life-sized videoconferencing solution supporting mutual gaze and report on the experiences made with our system in empirical evaluations.

Author Keywords
Videoconferencing, eye contact, trust, mutual gaze

ACM Classification Keywords
H.4.3 Information Interfaces and Presentation (e.g., HCI) - Communications Applications: *Computer conferencing, teleconferencing, and videoconferencing;* - H.5.3 Group and Organization Interfaces: *Computer-supported cooperative work*

INTRODUCTION
In an increasingly globalized world real-time communication such as text-based chat but especially audio chat and videoconferencing becomes more and more important. It reduces travel times as well as travel costs and leads to more and faster decision-making amongst the communication partners. The ongoing upgrade and extension of the network infrastructure allows us to virtually connect any two partners in the world. Together with the increased performance of video encoding and decoding algorithms needed to compress the high resolution signals of the integrated cameras and microphones, video conferencing is now possible at high quality standards. This technology is also extensively used for private communication such as in Skype[1].

However, in certain other situations, for instance business meetings with a strong negotiation character, people are often reluctant to use videoconferencing and often prefer face-to-face meetings. Bekkering & Shim (2006) found that the absence of eye-to-eye contact in videoconferencing systems is the main factor for the lack of trust and argue that this is a main reason for the missing large-scale adoption of the technology. And so do we in this paper: We are interested in the influence of mutual eye gaze on trust. Would you agree to a risky million dollar deal without looking into the business partner's eyes? Fox (2005) stresses the importance of eye gaze to indicate another's person intentions, interest in conversation etc. In business meetings (and other "non-chat" situations) this is of high importance. Relationships involving complex tasks can be maintained by increasing the frequency and flow of communication (McKinney &Whiteside, 2006) – this requires the indication of gaze and mutual eye contact. It was shown that systems using communication with eye contact induce behaviour similar to face-to-face communication (Mukawa et al., 2005). In interview situations for instance, perceived eye contact

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
OzCHI '14 , December 02 - 05 2014, Sydney, NSW, Australia
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-0653-9/14/12...$15.00 http://dx.doi.org/10.1145/2686612.2686632

[1] www.skype.com

and mental workload were identified issues when using videoconferencing (Ferrán-Urdaneta and Storck, 1997).

In the remainder of the paper we present two contributions: (1) We present our own system, which delivers the required characteristics of size and eye contact and finally, (2) we report on experiences made with our system in empirical investigations.

BACKGROUND

The principal problem for the lack of eye-to-eye contact is the positional offset between the capturing camera and the display of the partner's video image. Ideally, the camera should sit between the displayed eyes of the videoconferencing partner. Using small screens will reduce the effect of lacking eye-to-eye contact but also reduce the size of the displayed face of the partner. Okada et al. (1994) found that the size of the communication partner on screen is an important factor for achieving a sense of reality and argued for life-sized videoconferencing solutions.

Using a half silvered-mirror is probably the most common solution in research and on the market to achieve eye-to-eye contact in life-sized videoconferencing. The user can see through a mirror while being observed by a well-positioned camera at the same time. Half-silvered mirror solutions in videoconferencing setups have successfully been used by Mukawa et al. (2005), Quante & Muehlbach (1999) and others. However, it needs careful calibration though and usually produces optical artifacts due to the fact that the camera captures only half (or a certain percentage) of the true image of the user. Besides unwanted reflections, the maximum of achievable brightness and contrast levels are problematic.

Other researchers applied computer-vision approaches to achieved eye-to-eye contact by synthesizing the views of multiple cameras to compute an image from a virtual camera place in front of the user's eyes. These approaches require high quality cameras, careful calibration and a real-time and error-free computation (Schreer & Kauff, 2002) as humans are very sensitive when it comes to realize subtle artificial elements in other faces.

Other approaches using shuttered screens and cameras or a Holographic Optical Element (HOE). As first shown by the blue-C (Gross et al., 2003) a camera can be placed behind a back projection screen to virtually see-through the screen if the screen itself and/or the projection and cameras are shuttered. Similarly can be achieved with a HOE where the images are not separated over time (shutter) but dependent on the viewing angle. This type of installation allows for 1:1 scale videoconferencing but requires (expensive) instrumentation (i.e. shuttered glass). The shuttering (flicker) and the limited achievable transparency of the used screens introduce additional artifacts in the displayed videoconferencing video though.

In summary, all the approaches presented above are able to produce eye-to-eye contact in videoconferencing and a life-sized display of the communication partner to some degree. However there is no optimal solution: The solutions might be too expensive, require specialized hardware and software or produce poor visual quality (e.g. flicker, visible artifacts, and decreased brightness).

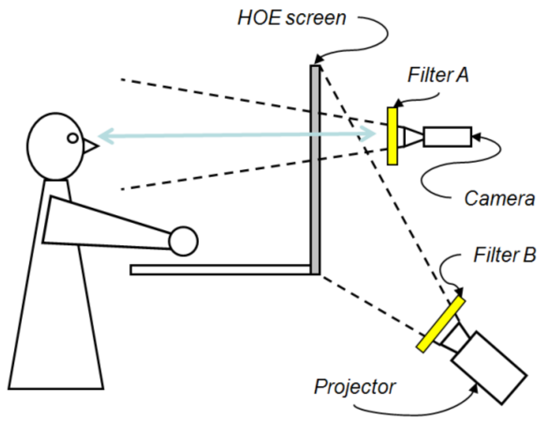

Figure 2. Schematic of "i2i" approach. A Holographic Optical Element (HOE) is used to separate the camera view from the projected video conferencing image. Additional polarization filters are added in the light path.

In the following we present the principal setup and report on the evaluation of a system we have developed addressing those issues. The requirements derived from previous studies targeting eye-to-eye contact, body language availability and trust in videoconferencing in non-chat situations (business-type scenarios) (Teoh et al. 2010, Teoh et al., 2011).

"I2I" SYSTEM

Our own approach is based on the idea of the shuttered screen, but minimizes the artifacts introduced by shuttering the screen and the limited opacity of the screen. When using a standard HOE-based screen in combination with a back projection there are still considerable, unintended reflections and diffusions visible on the back of the screen, which will be captured by the camera. Instead of shuttering we are minimizing this effect by using polarizing filters in front of the projector and the camera (Figure 2). These are oriented perpendicularly to each other so that the camera does not capture portions of the projected image and indirect light from the projector. Because the filters as well as the screen do not allow for 100% of the light to pass through we apply simple post process filtering to increase contrast and brightness to the desired level. Together with carefully positioned lights we achieve a good conferencing quality surpassing techniques relying on shuttered screens.

We have produced two identical systems based on our HOE and filters approach. The systems (called "i2i") were tested in symmetrical (i2i connected to i2i) and in asymmetrical constellations (i2i connected to standard desktop videoconferencing systems). As depicted in figures 1 and 3 we designed a special metal frame construction to hold the holographic optical element (HOE) screen. The 40" HOE screen can be used in landscape and portrait mode and requires a back projection from a certain, manufacturer specified angle (38 degrees in our case). We can adjust the height of our

Figure 3. The camera behind the HOE can't be seen by the users and only the image of the conferencing partner is visible providing eye-to-eye contact.

HOE screen within the frame, either to be used at desks (lower positions) or for standing users (higher positions). The entire aluminum frame construction is portable on wheels.

Behind the frame/screen a standard PAL camera is positioned on a tripod at eye level and equipped with a polarizing filter. Under the tripod (between the legs) a DLP projector with XGA resolution projects the video stream onto the HOE screen at an angle. Here too, a polarizing filter is used in front of the lens omitting the projected light to be visible for the camera. The projection is sized and keystone-corrected appropriately. Projector and camera are connected to a PC running the videoconferencing software. An echo cancelling microphone-speaker system is used for audio communication (black device on desk in figure 3). With this, we developed a novel solution while being affordable and delivering an almost artifact-free, 1:1 scale, eye-to-eye contact videoconferencing solution.

EVALUATION

In a series of informal and formal, qualitative and quantitative studies we investigated whether the effect of eye-to-eye contact can be achieved with our system and what possible factors might be affected by this. Our system proved to be technically sound and reliable and to provide life-sized eye-to-eye contact. On the basis of those earlier studies and observations we designed an experiment. We were particularly interested to investigate our main dimension of interest: the influence of mutual eye gaze on perceived trust. Is our system able to affect trust ratings and therefore might be suitable for "serious", business-type videoconferencing situations? To test this we installed our two i2i systems in different locations: one system in Dunedin and the other in Christchurch (360 km apart).

Task

The task was adapted from Regenbrecht and Hoermann (2008) and Bekkering and Shim (2006): The participants were required to discuss current events (e.g. championships in rugby or soccer) in (a) an honest and (b) a dishonest way. I.e. they had to lie or to tell the truth. This gave each participant something to judge the other participant on and also gave the users enough time to evaluate the communication quality during the session. Even if a current event task might influence internal validity it is strengthening external validity - so we chose that one over an artificial task (like prisoners' dilemma) By making the participant lie or tell the truth, it was assumed that this would contribute to whether the one participant felt trustful or not of the other participant.

Experiment Design

Twenty people participated in a within subjects experimental design. Each of the participants performed an eye-to-eye and a non-eye-to-eye videoconferencing trial in each experimental session. The eye-to-eye or non-eye-to-eye setup was randomized for both trials of the experiment with the setup (i2i or non-i2i) being switched to the opposite setup for the second trial. Whether the participant lied or told the truth was also randomized with and overall equal number of participants telling either lies or the truth. We investigated the influence of eye-to-eye contact and deception on trust as measured by the ITS (Wheeles & Grotz, 1977) and perceived communication quality as measured by (modified, see below) ITU scales (ITU, 1999).

Procedure

The participants were recruited using Facebook, word of mouth and by sending bulk text messages. The experimental sessions were scheduled to be carried out over two weeks; each session taking about 30 minutes. Each participant was positioned by adjusting the position and height of the chair at the table; this was relevant to guarantee the best position within regards to the camera and consequently to guarantee the best eye-to-eye experience for the participants. Once the facilitator determined (initially by flipping a coin) whether the participant was to tell lies or truth, the participants were asked to begin the discussion. The participants were not made aware whether their communication partner was instructed to lie or to tell the truth. The participants were told that they would have at least 5 minutes to discuss their chosen topic, thereafter they were asked by the facilitator to stop their conversations which would end the first trial of the session.

Each participant was then asked to fill out a perceived communication quality and perceived trust questionnaires for that trial, while the facilitators would change the setups of the systems on both ends. Once the participant had completed the questionnaires, they were told by the facilitator whether they would be lying or telling the truth for the next trial. The participants would then perform another discussion of about 5 minutes and once they had reached 5 minutes they were asked to stop the discussion, which would end the trial. They were then given the same two questionnaires as after the first trial, with an additional questionnaire about the entire experiment.

Results and Discussion

The reported trust between the conditions (eye-to-eye; non-eye-to-eye) was not statistically significant ($p = .096$). However, this result gives a slight indication in support of the hypothesis that the implementation of direct eye-to-eye contact in videoconferencing might improve perceived trust. Since we only found a trend, the first assumption made was that due to the small number of participants used in this study the results did not reach statistical significance. Therefore the results have to be corroborated with a larger sample.

The second assumption that we made was that asking the participant to lie or tell the truth would have an effect on the level of perceived trust between participants. To test this assumption we compared the eye-to-eye setup against the non-eye-to-eye setup separately for being lied to by and for truth being told by the communication partner. On one hand, there was no significant difference between eye-to-eye and non-eye-to-eye when the participants were being lied to ($n=11$, $p = .109$). On the other hand, when participants were told to tell the truth there was a marginal significant difference ($n=9$, $p = .052$). There was no significant difference for perceived communication quality ($p = .396$) between eye-to-eye and non-eye-to-eye. These results do not support the hypothesis that the implementation of direct eye-to-eye contact in videoconferencing will improve the perceived communication quality.

We found some initial support for the notion that the i2i implementation of direct eye-to-eye contact can improve perceived trust. Our results indicate trends for increased overall perceived trust and perceived trust when participants were being told the truth. However, an improvement of perceived communication quality was not supported by the results of this study. However, the questionnaire might not have been suitable to measure the perceived communication quality of this system.

CONCLUSION AND FUTURE WORK

We introduced our own i2i system approach using a holographic optical element screen in combination with polarizing filters implementing 1:1 scale videoconferencing with mutual gaze in high quality by reducing artifacts that were visible in other eye-to-eye videoconference systems. We evaluated our presented system in a real remote communication scenario.

We could show that our implementation does provide eye contact amongst other optimized factors. Supporting earlier work, we could illustrate that perceived trust will increase with eye-to-eye contact in “truthful situations”. To date we examined two-party situations only, but it would be very interesting to test it with three or more parties.

In a not too distant future, technical solutions might arise which implement eye-to-eye contact in a more elegant way (e.g. by placing light sensors in-between light emitting elements in computer displays; Uy, 2009). In the meantime we can and should apply one of the techniques described here and elsewhere to improve the quality of our videoconferencing experiences. We hope that our i2i approach presents a valuable alternative here.

ACKNOWLEDGMENTS

We would like to thank all participants in our studies and Raphael Grasset, Andreas Duenser, Magbagbeola Adejumo, Mohammed Alghamdi and Cameron Teoh for help with the i2i system. This work was partly funded by a University of Otago research grant (JDLJ17096).

REFERENCES

Bekkering, E., and Shim, J. P. (2006). i2i Trust in Videoconferencing. Communications of the ACM, 49(7), 103-107.

Ferrán-Urdaneta, C., & Storck, J. (1997). Truth or Deception: The Impact of Videoconferencing for Job Interviews. International Conference on Information Systems. 183-196.

Fox, E. (2005). The role of visual processes in modulating social interactions. VisCogn, 12(1), 1-11.

Gross, M. et al. (2003). blue-c: a spatially immersive display and 3D video portal for telepresence ACM Transactions on Graphics 22(3), 819-827.

ITU (1999). International Telecommunication Union (ITU-T P.920). Subjective video quality assessment methods for multimedia applications.

McKinney, V. R., & Whiteside, M. M. (2006). Maintaining Distributed Relationships: Electronic communication works best when it increases interaction and collaboration through a variety of media. Communications of the ACM, 9(3), 82-86.

Mukawa, N., Oka, T., Arai, K., & Yuasa, M. (2005). What Is Connected By Mutual Gaze?: User's Behaviour in Video-Mediated Communication. CHI '05 extended abstracts. Portland, Oregon, USA, 1677 -1680.

Okada, K., Maeda, F., Ichikawa, Y., & Matsushita, Y. (1994). Multiparty Videoconferencing. In Proceedings of CSCW.

Quante, B. & Muehlbach, L. (1999). Eye-contact in Multipoint Videoconferencing. Heinrich-Hertz-Institut für Nachrichtentechnik Berlin GmbH (HHI).

Regenbrecht, H. & Hoermann, S. (2008). Eye-to-eye contact in videoconferencing-mediated course advising situations - technology and evaluation. 2nd Spotlight Colloquium, Dunedin, New Zealand, Nov 19/20, 2008.

Teoh, C., Regenbrecht, H., & O'Hare, D. (2010). Investigating Factors influencing Trust in Video-Mediated Communication. Proceedings of ACM OZCHI 2010, November 22-26, 2010, 312-319.

Teoh, C., Regenbrecht, H., & O'Hare, D. (2011). The Transmission of Self: Body Language Availability and Gender in Videoconferencing. Proceedings of ACM OZCHI 2011, Nov 28 - Dec 2, 2011, 273-280.

Uy. M. (2009). Integrated sensing display. United States Patent, 7,535,468 , May 19, 2009.

Wheeless, L.R., & Grotz, J. (1977). The measurement of trust and its relationship to self-disclosure. Human Communication Research, 3(3), 250-257.