Eye-contact in Multipoint Videoconferencing

Birgit Quante and Lothar Mühlbach
Heinrich-Hertz-Institut für Nachrichtentechnik Berlin GmbH (HHI)
Einsteinufer 37, D-15087 Berlin, Germany, http://www.hhi.de/
Phone: +49 30 31002 237, Fax: +49 31002 213, E-mail: quante@hhi.de
Theme: Collaborative Work and Education

Abstract
This paper presents the major outcomes of a human factors experiment which investigated the benefits of a new multipoint desktop videoconferencing feature. The videoconference system enables interlocutors to establish individual eye-contact, a nonverbal cue which has not been transmitted by standard videoconferencing systems until now.
Eye-contact is seen as an important component of interpersonal interaction particularly with regard to the individual addressability of interlocutors. It was assumed that the transmission of this nonverbal signal produces, among other things, a smoother conversation flow and therefore enhances the feeling of telepresence as well as affects other relevant dependent variables in a positive direction (e.g. overall user satisfaction).
Two different videoconferencing systems were compared against each other, one supported individual eye contact, the other did not. An audio conference system as a reference condition was included in the experimental design. The scenario for the experiment was a group discussion on a given topic. The results show that users have significantly more often the feeling of being addressed i.e. of being looked at and recognise if they are addressed or not. The expected advantages in terms of a higher degree of telepresence or a higher overall satisfaction did not occur.

Keywords
Multipoint videoconferencing, telepresence, eye-contact, individual addressability

Introduction

The experiment described in the following was part of the research project „Telepresence in Desktop Multimedia Conferencing“ aiming at investigating means that can improve the feeling of telepresence which was seen as being particularly important for enhancing the effectiveness and attractiveness of tele-cooperation and telework as former findings suggested.

This article focuses on the effects of the reproduction of eye contact onto telepresence by eliminating eye contact angles in desk top multipoint videoconferencing.

The Effect of Eye-Contact in Multipoint Videoconferencing

The uptake of videoconferencing services for studio as well as for desktop systems continues to be lower than experts previously expected. One reason might be the limited degree of telepresence that can be achieved with conventional videoconferencing technology. The term telepresence embraces two concepts: Telepresence refers to the capacity of a medium for transmitting cues that affect the impression of sharing space with a geographically separated interlocutor (spatial presence) as well as to the capacity

of the medium for transmitting verbal and non-verbal communicative signals (communicative presence) [1] [2].

One reason that reduces the feeling of telepresence in videoconference meetings is, among other things, the limitations in using non-verbal communicative cues, particularly the inability of participants to establish eye-contact with each other like in face-to-face meetings. This in turn reminds the conferees of the fact that they are separated at different sites and therefore hampers the feeling of sharing space with their interlocutors.

Eye contact performs several key functions in communication: It is an important turn taking cue which regulates the flow of conversation and provides feedback from the interlocutors. Moreover studies of nonverbal communication have demonstrated that eye contact is, among other things, a cue for dominance, for friendliness, for approval, for romantic love, for status relationship, and expression of emotions [3] [4]. Video systems which fail to support eye contact may affect any of these functions.

Individual addressability by eye-contact

In conventional videoconferencing the use of nonverbal signals is limited. As outlined in the following, this is due to eye contact angles and the fact that each conferee sees the same image, captured by one camera at the other site. Eye contact angles are caused by the fact that the camera is not at the same position as the eyes of a displayed conferee. This leads to the effect that when a user looks at an interlocutor who is displayed on the screen, the latter has the feeling that the user is looking down rather than into his/her eyes.

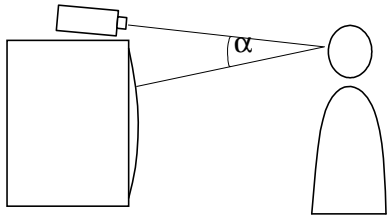


Figure 1. Eye- contact angle in videocommunications

One technical means to overcome this deficiency is to install half silvered mirrors at the conferencing terminals in a way that allows a person to simultaneously look at the eyes of a displayed interlocutor and into the camera.

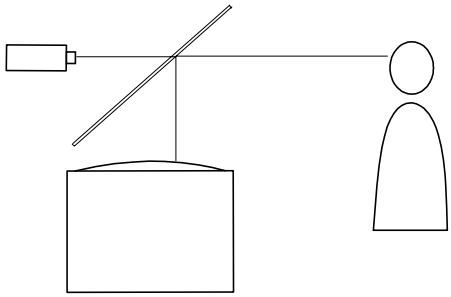


Figure 2. Reducing the eye-contact angle by means of a half-silvered mirror

However, in the case of multipoint videoconferencing, in which more than two participants are involved, the problem of individual addressability by means of eye-contact remains unsolved. Due to the fact that only one camera per site is used, the following two errors may occur: a) If the user looks into the camera *all* participants have the feeling of eye-contact, and b) if the user does not look into the camera *none* of the participants feels being looked at.

In order to allow individually addressable eye-contact, additional (real or virtual) cameras - one for each remote participant - can be installed. This would lead to the effect that a conferee perceived eye-contact if and only if he/she is actually being looked at. Thus, all participants are individually addressable.

This feature can be implemented by means of various display techniques, such as a time division multiplexing as was the case in the experiment reported below.

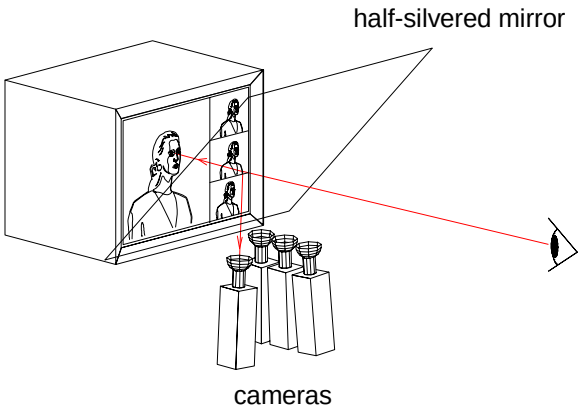


Figure 3. Individual eye-contact

Hypotheses

We expected that the possibility of individual addressable eye contact could improve the impression of telepresence and moreover would lead to better ratings in terms of user satisfaction and acceptance in comparison to conventional video conferencing systems.
In conventional video conferencing systems vertical and horizontal eye contact angles reduce the individual addressability by nonverbal signals and therefore hamper the impression of individual eye contact.

Method

The experimental conditions:

In a pilot study the three versions described above were tested with regard to the extent of eye-contact. The results of this study show that in the individual eye-contact condition subjects felt significantly more often been looked at compared to subjects in the standard- and ‚half-silvered mirror'- condition. The comparison of the two latter did not produce any significant differences concerning the feeling of being individually addressed by means of eye- contact.
Consequently we dropped the ‚Half silvered Mirror' version which is in comparison to the Standard variant more bulky and expensive in the following communication experiment.
From our pilot experiment an additional finding derived. With regard to the recognizability of eye-contact the sizes and arrangement of the windows, in which the conferees are displayed seems to play a role. We assumed that an asymmetric splitting(see Table 1) causes very small vertical angles between the small windows and this could hamper the visibility of nonverbal cues. Furthermore, we assume that in a quarter split mode (the screen is divided in four quadrants) the vertical angles are more obvious than in the asymmetrical mode and therefore could support the transmission of visual cues. Considering this we tested the individual eye-contact version in two different screen splitting modes: asymmetric and quarter split.

The following conditions were included in the design of the main experiment:

Condition 1: The Standard Videoconferencing system
The camera was placed above the monitor, this led to a horizontal eye contact angle as well as a vertical eye contact angle. Therefore eye contact was not possible.
Condition 2 and 3: The Videoconferencing System with individual eye contact asymmetric/ quarter split
A half silvered mirror and four cameras were installed each at the virtual position of the conferee. To each subject, only the image from his/her camera was displayed. Eye contact was possible.

Condition 4: The Audio conferencing system
As a reference condition an audio-only condition was implemented in the experimental design. The connection between the interlocutors was restricted to the transmission of audio signals.

Nine 5-person teams were tested in each video condition, seven teams in the audio condition giving a total of 34 trials. Each subject used only one of the four mentioned conference (a between-subject design).

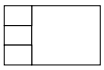
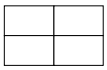
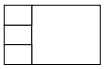
Variant	Screen splitting	subjects
Standard (without eye-contact)	asymmetric 	27 (9 groups)
View-per-person (eye-contact)	quarter-split 	27 (9 groups)
	asymmetric 	27 (9 groups)
Audio (no visual cues)		21 (7 groups)

Table 1: Experimental design

The laboratory set up
The communication laboratory system consisted of five work-places each in a separate room. The work places were provided with video and audio equipment described in the following and with equipment for data collection. All technical set ups were multipoint desktop-conferencing systems:

The audio equipment
The audio equipment consists of a microphone and two loudspeakers per work place. The audio representation followed a pseudo stereophonic sound reproduction principle. The voices of the conferees were reproduced at constant positions within the stereo base, independent from the video setting. Each of the remote conference members could be heard either from a left, half left, right or half right position. The audio equipment was identical in all four experimental conditions.

The video equipment
The video equipment consisted of a person camera or four person-cameras at each site in the eye contact version, respectively. The image separation was accomplished by means of a time division multiplex technique. The video images were displayed on 17‘‘ monitor with 100 Hz image frequencies.

The task
During each experimental session a group of three subjects and two confederates of the experimenter discussed a controversial issue, namely the German legislation on shop-opening hours, which was regarded as a representative communicative task for using a multipoint videoconference system. The two confederates were instructed to hold a counter perspective of each other in order to stimulate the discussion.

Procedure
Subjects were randomly assigned to the discussion teams without taking into account prior acquaintance or to control for any other social-psychological factors.
The subjects were introduced to the purpose of the experiment. Subjects were then familiarised to the communication system. Each subject was then led to a workplace in a separate room. After answering some preliminary questions they were advised to

prepare for the discussion for 10 min. The conference sessions lasted about 20 minutes. After that data were collected by means of a questionnaire. Finally the subjects and the experimenters met again in one of the laboratory rooms to summarise and reflect the experiment. Overall, a session lasted up to two hours.

Data collection

Data were collected by means of a questionnaire on relevant dependent variables and control variables such as recognizability of eye contact, communicative presence subdivided in several aspects (conversation flow, individual addressability, visibility of mimics and gesture, feedback), telepresence, overall satisfaction, appeal, and acceptance.

Usually, the dependent variables were measured with a 5-point agreement scale. Furthermore, a Kunin rating scale was used to measure overall satisfaction.

In order to get a profile of the subjects the questionnaire comprised items like sex, age, education, number of meetings per year, business trips per year, experience of computer use, communication media use, (videoconference, picture phone), etc..

Subjects

A total of 102 subjects took part in the experiment, 55 were male, 47 were female. The majority were between 20 and 30 years. 25% took part in conference meetings regularly, 20% went on more than one business trip per year. The majority was not familiar with videoconferencing, less than 10% indicated prior experience with telecommunication devices such as a picturephone or a video conferencing system.

Results

The comparison of the eye-contact- and the standard condition showed that in the eye-contact condition subjects became clearly aware of eye-contact. Under that condition subjects stated significantly more often that they had the feeling of eye-contact than did subjects under the standard condition ($p = 0.026$). Additionally, subjects became more aware of who is looking at them ($p = 0.004$) and who is not looking at them ($p = 0.019$). The latter applied particularly to the quarter split condition.

Secondly, our data showed that the individual addressability by means of non verbal signals in general was rated superior in the eye-contact condition in comparison to the standard condition ($p = 0.003$). The horizontal and vertical eye contact angles in the standard condition hampered the recognizability of mimics ($p = 0.021$), especially in the case when mimics were directed not to the own person but to other conferees. However, this does not hold for the recognizability of gestures.

Moreover, our results revealed that the eye-contact conferencing system supported the transmission of nonverbal signals in general. Besides the feeling of being looked at subjects in the eye-contact condition indicated that they had significantly more often the feeling to communicate by means of nonverbal cues than did subjects in the standard condition ($p = 0.003$).

Due to the fact that eye-contact plays an important role with regard to floor control, turn taking and keeping the track we expected that in the eye-contact condition the flow of conversation would be rated to be smoother than in the standard and audio condition. Contrary to this hypothesis the possibility of eye-contact did not lead to any significant

media difference. In all four experimental conditions subjects rated the regulation of conversation to be good and indicated that they have had no difficulties in co-ordinating the turn taking. However, subjects in the audio condition judged the conversation flow to be more stagnated than subjects did in the video conditions ($p = 0.001$).

Although the possibility for establishing eye-contact was realised by the users of the eye-contact version this had no impact on dependent variables such as telepresence, efficiency, overall satisfaction, or acceptability (that is operationalized by the willingness to use the videoconferencing systems for own purposes). Nevertheless, subjects rated the video conferencing versions as to be more fascinating than the audio system, especially the eye-contact version with the quarter split screen ($p = 0.048$).

Concerning the audio system employed in the experimental set up, subjects in all four conditions judged the quality of transmission of audio signals to be good to excellent. This is valid particularly in the audio-condition ($p = 0.025$). A spatial impression due to the feature of pseudo stereophonic attracted more attention in the audio condition and was judged as more helpfully in this condition compared to the video conditions ($p = 0.032$). One explanation for this result could be that subjects in the audio condition were more dependent on the transmission of audio signals and consequently appreciated the audio transmission as being good whereas subjects in the video conditions took this probably more obviously.

Discussion

This experiment—a comparison of four different versions of a multipoint conferencing system—was carried out in order to examine the possible advantages of an eye-contact videoconferencing system. In the social psychological literature eye-contact is usually seen as an important nonverbal cue in interpersonal communication and, hence, eye-contact may enhance the feeling of telepresence.

The most striking result of the study was that the hypothesis on the superiority of the eye-contact feature could not be supported by the results. Although the eye-contact condition allowed the conferees of individually addressing one of their remote interlocutors and although it enhanced the recognizability of nonverbal cues in general the comparison revealed surprisingly no significant effect on dependent variables such as telepresence, flow of conversation, overall user satisfaction, and acceptance.

One explanation for the low media differences in our study might stem from the task applied in our experiment. An analysis of control variables regarding the task indicates that the group discussion was perceived less conflictive and emotionally involving than we had anticipated. Ratings on the five-point-agreement scale indicate an indifferent position of the item “The discussion was controversial” (overall mean: 3,19, $p = 0.104$). Additionally, the ratings of the item “The discussion was exhausting” were rather low (overall mean: 1,85, $p = 0.417$). This may indicate that the communicative scenario would be better characterised as an information exchange than a conflictive group discussion. Therefore it was more necessary to express cognitive concepts transmitted by verbal signals than to express emotions and feelings conveyed by nonverbal cues. Accordingly, nonverbal cues were of minor importance.

There is also evidence from research that nonverbal cues and the associated verbal

communication are highly redundant. This could be a reason for the fact that users in the standard condition and in the audio condition did not miss any nonverbal cues and rated the variables such as telepresence or satisfaction as well as or worse than subjects did in the eye-contact condition.
Moreover, findings from Human Factors in telecommunication research suggest that people in a mediated communication situation adapt their communicative behaviour to the situation [3]. Conferees use additional verbal or visual signals in order to compensate for media shortcomings and are perhaps more tolerant, particularly, in a laboratory environment. Therefore actual differences between the experimental conditions could not become apparent.

Besides the common explanations of low media differences described above, an additional interpretation of the results might be valid: We presume that subjects used different anchors for their judgements in the experimental conditions. In the eye-contact condition subjects perceived the communication situation as being similar to a face-to-face situation and accordingly rated in reference to that. By doing so they were not aware of additional non verbal cues such as eye-contact but focused on other shortcomings in the mediated communication situation. Their ratings in terms of overall satisfaction, acceptance and others became low. In the standard and the audio condition subjects were in contrast aware of the limited use of nonverbal cues. Therefore their judgements were based on a low level of expectations. If the communication was perceived as to be quite pleasant, the subjects' ratings became high and—as a consequence—actual differences were covered or biased, respectively. Adopting an alternative experimental design i.e. a within design where subjects have to use all of the three conferencing systems might probably reduce this bias.

Another interpretation of the results obtained in the study is the following: Although it conveyed additional nonverbal cues the eye-contact version was still not as perfect as a face-to-face situation. In non-mediated communication the interlocutor can always assess, in the case that he or she has not been addressed, who of the remaining persons is addressed. In the eye-contact condition, conferees perceived eye-contact, when they were addressed, and noticed, if they were not addressed, but in the latter case, they had no nonverbal information about who of the other conferees was addressed. It might be the case that interlocutors experienced this as a lack of certainty. This possibly may have decreased the feeling of telepresence.

Taking into account the high technical effort and costs in constructing a multipoint videoconference system with an eye contact feature, it must be judged as being pleasant and appealing but not necessary.

Acknowledgements

The work reported here has been supported by the Ministry of Education, Science, Research and Technology (BMBF) of the Federal Republic of Germany and by the Federal State of Berlin. The authors alone are responsible for the contents of this paper.

References

[1] Prussog, A., Mühlbach, L. and Böcker, M.. Telepresence in Videocommunications, Proc. Human Factors in Telecommunication, 1993, pp. 201-211.

[2] Mühlbach L., Böcker, M. and Prussog, A.. Communicative Presence in Video-communications, Proc. Human Factors in Telecommunications, 1993, pp 213-222.

[3].Williams, E.. Experimental Comparisons of Face-to-Face and Mediated Communication. A Review. Psychological Bulletin, 1977,Vol. 84, No. 5, S. 936-976.

[4] Sellen, A.. Speech Patterns In Video Mediated Conversations. Proc. ACM Conference on Human Factors in Computing Systems, Monterey, California May 3 –7, 1992, pp 49-59.