Manuscript

Communications of the Association for Information Systems

CAIS

Mutual Gaze Support in Videoconferencing Reviewed

Holger Regenbrecht

Information Science, University of Otago, New Zealand

holger@infoscience.otago.ac.nz

Tobias Langlotz

Information Science, University of Otago

Abstract:

Videoconferencing allows geographically dispersed parties to communicate by simultaneous audio- and video transmissions. It is used in a variety of application scenarios with a wide range of coordination needs and efforts, e.g. private chat, discussion meetings, and negotiation tasks. In particular in scenarios requiring certain levels of trust and judgement non-verbal communication cues are of high importance for effective communication. Mutual gaze support plays a central role in those high coordination need scenarios, but generally lacks adequate technical support from videoconferencing systems. Here we review technical concepts and implementations for mutual gaze support in videoconferencing, classify them, evaluate them according to a defined set of criteria and give recommendations for future developments. Our review gives decision makers, researchers, and developers a tool to systematically apply and further develop videoconferencing systems in "serious" settings requiring mutual gaze. This should lead to well-informed decisions regarding the use and development of this technology and to a more widespread exploitation of the benefits of videoconferencing in general. For example, if videoconferencing systems would support high quality mutual gaze in an easy to set-up and easy to use way more and more effective and efficient recruitment interviews, court hearings, or contract negotiations could be held.

Keywords: video conferencing, telepresence, eye gaze, eye contact, eye-to-eye contact

Volume xx, Article x, pp. xx-xx, Month Year

The manuscript was received mm/dd/yyyy and was with the authors x months for y revisions. [Note: this box is used only for peer-reviewed papers.]

Communications of the Association for Information Systems

Communications of the Association for Information Systems

CAIS

Communications of the Association for Information Systems

Abstract:

I. INTRODUCTION

In an increasingly globalized world real-time communication and collaboration become more and more important. The on-going upgrade and extension of network infrastructure allows us to virtually connect any 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 (V/C) is now possible at high quality standards. This includes high-resolution video and audio recording, streaming and playback, filtering of environment noise, and often focus-and-context cameras. Business videoconferencing solutions are nowadays offered by many manufacturers – for instance Cisco[1], Polycom[2] and Lifesize[3] – and with different prices and feature sets ranging from off-the-shelf solutions to custom business solutions exceeding a 100.000 Dollar price-tag.

Videoconferencing has the potential to provide a number of significant benefits over traditional face-to-face meetings. For instance Davis & Weinstein (2005) list: (a) Faster decision making and shorter time to market for products and services enable dispersed teams to collaborate easily, solving problems and speeding coordination (ultimately delivering faster time-to-consensus and hence a shorter time-to-market for new products and services). (b) Higher productivity / efficiency from a scheduled environment to ad-hoc, unscheduled work style. (c) Higher impact and focus shorter, more effective meetings with minimal workflow disruption (videoconferencing meetings tend to be shorter than in-person meetings). (d) Competitive advantage example recruitment: interviews with more people, from more locations, in less time, and with less cost and disruption; also better hiring decisions. (e) Enhanced quality of life / decreased stress - business travel negatively impacts life, sleep, and general welfare. (f) Increased reach by personal touch between company and client, (g) Improved management of dispersed teams by allowing impromptu, face-to-face meetings. And, finally (g) the reduction of travel costs in the form of direct expenses for airfare, hotels, meals, taxis and car service, etc. and extended expenses for hours of downtime and days away from the office. Given all those advantages and benefits one might wonder why videoconferencing isn't used more widely in all kinds of contexts. Two main factors might come into play here: (a) the nature of the meeting purpose to be supported and (b) the support for communication aspects beyond audio and head-and-shoulders video transmissions.

The nature of a (videoconferencing) meeting is determined by the coordination need and effort, heavily depending on the task and group situation (Cornelius and Boss, 2003). Informal communication, like a chat, does not require a high communication effort, idea generation tasks require at least some protocol, problem-discussing tasks do have a defined goal but require high efforts and massive communication, judgement tasks (decision making, problem solving) require the highest coordination effort and finally negotiations require the highest coordination effort in combination with a trust enabling environment (face-to-face negotiations as the „gold standard“). Hence, the videoconferencing system needs to address the coordination effort of the meeting task. For instance, while a Skype-like system might be appropriate for a private chat or for a well-defined and brief decision making task, a legal contract negotiation meeting amongst new business partners will afford support for non-verbal communication cues and for the integration of collaboration tools (e.g. interactive document sharing).

In this paper we are focusing on those "serious" conferencing settings requiring a certain degree of communication effort and trust support. Such settings are also characterised by a need for the integration of meeting artefacts and/or collaboration tools and by the need to support gestural communication cues (body language availability; Teoh et al., 2011) and with this to provide a certain interaction space in front of the screen. Of particular interest in settings relying on levels of trust is the provision of eye-to-eye contact and gaze awareness. Gaze awareness tells the communication partners where a person is looking, e.g. at the documents discussed or addressing another partner. Eye-to-eye contact, also known as mutual gaze or simply eye contact, is a special form of gaze awareness to detect whether a person is directly looking at the partner. This forms the basis for forming empathy and trust-building in a lot of situations (Teoh et al. 2010) – would you agree to a risky million dollar deal without looking into the business partner’s eyes? Of particular interest are so called telepresence systems, which give the participants the feeling of being in one, shared room or space expressed with the concepts of co-presence and social presence.

Basic issue

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 (see Figure 1).

[1] www.cisco.com/web/telepresence/index.html
[2] www.polycom.com
[3] www.lifesize.com

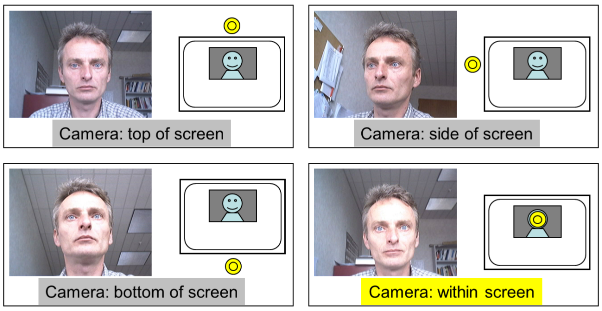

Figure 1. Eye-to-eye contact: separation of camera and screen

Unfortunately, placing a camera at this position would normally block one's view of the partner, which makes the solution unsuitable. As a best practice approach most non-consumer videoconferencing systems try to place the camera as close as possible to the displayed partner video, as for instance illustrated in figure 1 in the top left image. This can be implemented in desktop videoconferencing and in room-like systems. The size of the screen and video image, the distance from the user to the camera and screen and the position of the video image on the screen are the parameters to be considered here. Because of the practical spatial limitations in most environments, true eye-to-eye contact cannot be achieved with this approach. Other technical solutions have to be applied to achieve a real sense of mutual eye contact.
How should a videoconferencing system be set up to allow for maximising empathy- and trust-building needed in many business and other relevant communication situations? The size of the displayed face of the partner and the provision of mutual eye gaze are important factors to build trust and deliver the basis for high communication quality that is combined with non-verbal communication queues.

In the remainder of the paper we are reviewing the literature on the characteristics of gaze and eye contact within videoconferencing and video-based telepresence and we are surveying different approaches for the implementation of eye contact within videoconference systems and discuss their pros and cons. While videoconferencing technology is also available and extensively used for private communication between relatives and friends using applications such as Skype[4] or Google Hangout[5], we are mainly focussing on professional and business oriented videoconferencing solutions, while considering affordability and feasibility at the same time.

II. THE VIDEOCONFERENCING EXPERIENCE

In most cases one would hold a videoconferencing meeting because you can't come together with one or more people in one place physically. This videoconferencing meeting can be improved by enhancing the video or audio quality and the interface to an extent that the virtual coming-together is done in a seamless way, with a disappearing interface and in a quality hardly distinguishable from a physical meeting or even providing advantages over a physical meeting. Furthermore, in particular in business environments, the shared access to files and virtual artefacts becomes increasingly important. The ultimate technology for videoconferencing would deliver an immersive experience leading to the concept of *social presence* - the perceived sense of being together in one (virtual or real) place. In that context the literature also often refers to *telepresence videoconferencing*. Social presence can be considered as the central defining feature for a (professional) videoconferencing system. It requires a certain technical and setup fidelity comprising factors of space requirements, quality of the video, complexity of the system, flexibility in posturing oneself in front of the screen/camera, hardware and software requirements, price, availability and the ease of integration into existing environments.
Biocca et al.'s (2001) developed a framework with three theoretical dimensions to describe and measure what determines social presence: co-presence, psychological involvement, and behavioural engagement. They associated a number of factors to those dimensions, namely the feelings and perceptions of Isolation / Inclusion, Mutual Awareness; Mutual Attention, Empathy, Mutual Understanding; Behavioural Interaction, Mutual Assistance, and Dependent Action, respectively. When looking at those factors, even if only intuitively, one gets quite a good idea on what makes videoconferencing an artificial, sometimes frustrating or the opposite, an immersive experience. While those dimensions are often naturally supported in face-to-face situations, videoconferencing solutions don't support the same experience (yet). Adopting Bondareva & Bouwhuis's (2004, 2006) criteria to create social presence for the participants in a videoconferencing meeting main factors influencing the experience can be expressed as: Mutual Gaze (direct eye contact is preserved), a wide field of view (FOV), the display of a life-sized upper body, a high video and audio quality (including a high quality image and correct colour reproduction, audio

[4] www.skype.com
[5] www.google.com/tools/dlpage/res/talkvideo/hangouts/

Journal of the Association for Information Systems

with high signal-to-noise ratio, a directional sound field and the minimization of the video and audio signal asynchrony), the availability of a shared working space and the way the videoconferencing partners are positioned in relation to the equipment and their partners (see Figure 2).

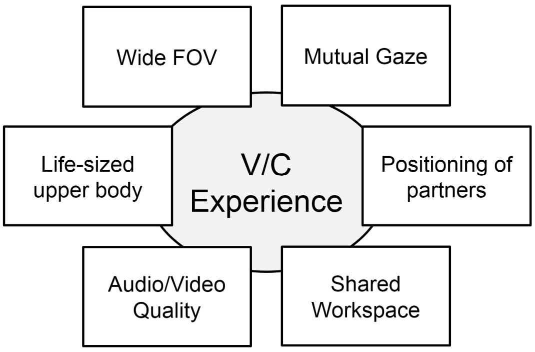

Figure 2. Factors influencing the videoconferencing experience

The relative importance of those criteria depends on the communication task and the nature of the videoconferencing meeting. For instance, a directional sound field might be very important in a situation with a high number of participants, but is of less importance for an ad-hoc, face-to-face meeting. Nevertheless, direct eye contact plays an important role in a wide range of videoconferencing situations and tops the list not without reason. This is supported by earlier work investigating the use of different multi-party videoconferencing systems (Buxton, Sellen, & Sheasby, 1997) who list “establish eye contact with other participants" also top on their list. However, we should also consider influencing factors like-life size appearance or the availability of a shared working space when it comes to non-verbal, gestural communication. Meeting situations with higher levels of collaboration effort require the integration of the communication with the collaboration space. For instance, the meeting table space found in physical meetings is used to define the physical and social positioning of the partners, to support gestural and postural communication and to share meeting artefacts like documents or physical prototypes. The displayed and perceived size of the communication partner and the collaboration environment and the way people are positioned in that (virtual) environment determine the way how social presence develops. Also of interest her are gender and cultural factors as well as the familiarity with videoconferencing meetings and setups in general.

Life-sized Upper Body and Positioning of Partners

Related to the issue of lacking eye-to-eye contact is the perceived scale of the communication partner. Am I talking to a miniature representation of my partner? Is the partner presented poster-sized? How convincing is the scaled presentation of the communication partner and can we actually maintain eye-to-eye contact within the scaled representation of my communication partner (e.g. with much too small or much too big video faces)? On one hand, Okada et al. (1994) found that the size of the communication partner on screen is an important factor for achieving a sense of reality. If the partner is presented smaller than life-size he/she might be perceived as far away. Also, it is difficult to read facial expressions or gestures. On the other hand, a larger than life-sized communication partner in videoconferencing implies dominance. Buxton (1992) suggests that social relationships, such as power, may be more balanced and natural in life-size video conferencing. Detenber and Reeves (1996) found that the display size has an effect on people’s arousal, and Lombard (cited in Detenber and Reeves) found that people evaluated others more positively when presented on large screens (in the right size). There are commercial systems supporting life-sized videoconferencing, e.g. business solution such as LifeSize[6] or the Cisco Telepresence[7] series.

Related to the perceived size of the videoconferencing partners is their seating arrangement in group videoconferencing sessions. While in 1:1 settings the partners would normally facing each other directly, potentially with some interaction space (virtual and/or real) between them, in group settings the positioning of the partners influences the support for non-verbal communication, partner and workspace awareness, possibilities for interaction and collaboration, and social-psychological balance of dominance. Yamashita et al. (2008) investigated two different versions of seating positions in 2 (local) x 2 (remote) table-centred, life-sized video communication. They found, amongst other aspects that a side-by-side seating (remote-remote, local-local) is preferable regarding balance of turn-taking and sense of unity.

Fish, Kraut, & Chalfonte (1990) present an early room-sized V/C system built with the aim of providing an ad-hoc, informal, as-if face-to-face communication channel. A technically detailed description of the implementation of gaze

[6] www.lifesize.com
[7] www.cisco.com/en/US/products/ps7060/index.html

and eye contact is missing, and it is doubtful that this was achieved at all. However, a life-sized video image was achieved with a wide NTSC camera and view. Also, Gibbs, Arapis, & Breiteneder (1999) present their early TELEPORT concept and partial implementation of a room-to-room videoconferencing system allowing for eye contact and gaze awareness in a combination of a virtual environment with life captured video feeds, compositing, tracking, and 3D projection. More recently other research groups try to achieve or support eye contact through collaborative virtual environments. Wolff, Roberts, Murgia, Murray, Rae, Steptoe, Steed, & Sharkey (2008) provide gestural communication and gaze-contact in an avatar-based, collaborative CAVE environment. Users equipped with head- and eye-trackers are communicating with each other in that CAVE system.
Gamer & Hecht (2007) emphasize the importance of observer distance, head orientation, visibility of the eyes, and the presence of a 2nd head on the perceived direction and width of the gaze cone in videoconferencing. Also, we are less sensitive to eye contact when people look below our eyes than when they look to the left, right, or above our eyes. Additional experiments support a theory that people are prone to perceive eye contact. That is, we will think that someone is making eye contact with us, unless we are certain that the person is not looking into our eyes (Chen, 2002). Those aspects help to mitigate the effects of the lack of (well configured) videoconferencing.
Vertegaal (1999) or Regenbrecht et al. (2004) use video planes in three-dimensional space to indicate gaze direction, but do not allow for actual eye contact. Almost eye contact could be achieved by the “Reciprocal Video Tunnel” (Buxton & Moran, 1990) using a half-silvered mirror and a Miniature User Node.
Ishii, Konayashi, & Grudin (1993) support eye-contact between two parties with their Clearboard system, where users are standing in front of an acrylic glass drawing board with cameras and projections placed behind. Here, the emphasius is on the collaborative work on the clear board as such and less on the quality of the video communication, even significant video artefacts can be tolerated.

Other confounding aspects

The implementation and perception of eye contact is complex. Some aspects can be controlled technologically or organisationally while others have to be considered but can hardly be controlled.
Heaton, L. (1998) for instance explores cultural aspects in Japan of eye-gaze allowing CSCW systems, namely ClearBoard and MAJIC (unfortunately, MAJIC was never actually used, not even in laboratory). “Given the constantly fluctuations and redefinitions involved in any activity which is out of the ordinary, they view the task of trying to support “delicate” communication, such as negotiation, as an impossible one. Eye-contact might also be considered as rude.
Swaab & Swaab (2008) investigated gender aspects and eye contact and found in their study that unacquainted females have a better agreement with eye-contact, in contrast to males where this leads to the opposite effect (no eye-contact leads to better agreement). Also, Teoh et al. (2011, 2012) and Hauber et al. (2012) found significant gender effects in videoconferncing sessions.
In general, if people are using videoconferencing frequently, they might learn to interpret gaze direction to a very high degree of accuracy if the equipment is configured optimally (Grayson & Monk, 2003). This is helpful when addressing objects in the environment, but does not necessarily provide the perception of eye-to-eye contact.

III. THE IMPACT OF MUTUAL GAZE

The last section demonstrated the importance of mutual gaze for the quality of the videoconferencing experience. But, which influence has mutual gaze on different dimensions of interest in video-mediated communication and collaboration?
The importance of gaze was already identified in psychology in the second half of the 20th century. “How long – and when – we look “in the eye” is one of the main signals in non-verbal communications” (Cook, 1977). Thereby, *gaze* is defined as looking at another person’s upper body, or sometimes more specifically between the eyes. Consequently, *mutual gaze* is defined as looking into each other’s eyes (Cook, 1977). In the context of this paper we use the term mutual gaze and eye-to-eye contact with the same meaning. We further consider gaze in a wider sense of looking at the others upper body but also including the importance of gaze awareness, which sometimes can be close to eye contact. In the following we present research investigating the effect and importance of gaze in telepresence applications (see Figure 3).

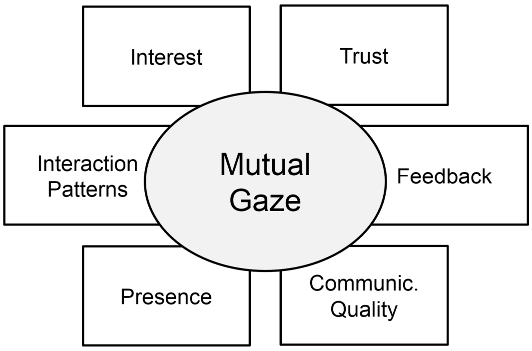

Figure 3. Factors impacted by mutual gaze

Interest and Trust

Gaver, in his seminal 1992 paper pointed out that "Video is anisotropic, interfering with the design of communicative gesture and with gaze awareness." Apart from the lack of a shared media space in V/C a compensation for the lack of eye contact (to facilitate turn-taking, indicate interest, and reflect social relations) is needed.

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 induced behaviour similar to face-to-face communication (Mukawa et al., 2005). In interview situations for instance, perceived eye contact and mental workload were identified issues when using videoconferencing (Ferrán-Urdaneta and Storck, 1997).

In experiments with half-silvered mirrors Quante & Muehlbach (1999) show that users have significantly more often the feeling of being addressed, i.e. of being looked at and recognise that they are addressed when eye-contact was provided. We look at the partners' faces to continuously make sure that they are still with us in the meeting and that we can trust them. Predictive valid faces (targeting the partner) are appearing more trustworthy (Bayliss & Tipper, 2006). Bekkering & Shim (2006) found that the absence of eye-to-eye contact in videoconferencing systems is the main factor for the lack of trust: "People associate poor eye contact with deception" (p.103). Furthermore they argue that this is a main reason for the missing large-scale adoption of the technology. Also, Teoh et al. (2014) show that eye contact and the availability of body language effect perceptions of trust, social presence, dominance, and impression management.

Feedback and Communication Quality

In a study on rural psychotherapy in Norway Sorlie et al. (1999) compared different types of distant delivery: "Most participants reported reduced eye contact, less nuances in mimics and other nonverbal cues, and a corresponding increase in dependency on verbal cues under V/C conditions. Some trainees felt that this reduced their ability to monitor how the supervisor reacted towards their presentations." (p. 458)

However, as O'Malley et al. (1996) point out: in comparison to audio only, video can lead to interruptions in communication flow. Co-presence and high bandwidth are important. A lack of or low degree of social presence might lead to an increase in verbal and non-verbal communication as a form of overcompensation for missing confidence in mutual understanding. Gaze is serving a feedback function, i.e. communication partners try to elicit feedback from their listeners. The interplay between audio and video, with or without eye-contact is much more complex though. For instance, people sometimes over-use the visual channel, which leads to an increase in cognitive load resulting in redundant verbal communication.

For a remote, collaborative design situation Olsen, Olsen, & Meader (1995) found that video with gaze awareness versus video and audio only, gaze video was superior in terms of discussion quality and time (less clarifying discussions). Also Vertegaal & Ding (2002) investigated experimentally the effect of different types of gaze and tasks on the (positive) quantity of turn-taking with positive results.

Interaction Patterns and Presence

Joiner et al. (2002) conducted three experiments in remote learning and discussion situations (physics and statistics) as part of a larger project investigating technologically mediated collaboration in learning. They found that eye contact influenced problem solving and interaction patterns and that eye contact facilitated conceptual understanding.

Mukawa et al. (2005) present a study comparing eye contact versus non eye contact and found that eye contact induced a similar behavior to face-to-face communication.

Even if we cannot see a real or video-mediated representation of our partners, for instance in Second-Life-like conferencing, gaze awareness with avatars (here 3D, cartoon-like representations of humans) is important (Garau et al., 2000). In 2001 Garau et al. conducted an experiment testing the influence on avatar gaze on different measures. Informed (inferred, synchronized, conversational) gaze (male and female avatars) outperformed non-gaze and random gaze in many measures, like involvement and co-presence. Another experiment investigating the importance of gaze in avatar-based CVE's is presented by Bailenson, Beall, & Blascovich (2002). The authors found that avatar head-movements play an important role in terms of higher levels of reported co-presence and positively changed patterns of interacting with each other.
Mutual gaze (as close as this can be achieved with standard desktop videoconferencing) has a significant effect on social presence and communication experience for children in a game playing situation. The absence of mutual gaze dramatically decreases the interaction quality (Shahid, Krahmer, & Swerts, 2012).
Fullwood & Doherty-Sneddon (2006) conducted two experimental studies using a sales pitch for a cosmetic product it was found that the absence of gaze has a negative impact on recalling what was said. This finding might have consequences in many areas, such as remote teacher-student learning environments and might lend itself to the conclusion that it is better to have an audio channel only in comparison to a video-audio system without eye-contact (or workarounds like artificially speaking into the camera have to be used).
Carville & Mitchell (2000) investigated videoconferencing used in teaching and learning in early childhood studies with early years' tutors. Students and tutors developed skills and strategies to deal with shortcomings of V/C. "Many of the tutors identify the need to remember to keep looking into the camera to make eye contact with the … [other side]" A similar workaround was reported by Birden & Page (2005) in situations of remote health education where they instructed the participants to look into the camera when speaking: "The lecturer must remember that to give students at the far site the impression that s/he is making eye contact; they must look into the lens of the camera, not at the screen.". In the same application realm, as part of their "Twelve tips for teaching using videoconferencing" Gill, Parker, & Richardson (2005) advise "Eye contact with the distant site is important; this can be simulated by talking to the lecturer camera … we have found a bright balloon tied to the lecturer camera encourages the lecturer to remember the camera is there!" In a V/C environment to teach and learn business French students were told to look into the camera while speaking to compensate for the lack of eye-contact. This workaround was seeing as being mainly successful, but in a few cases this also lead to distraction (McAndrew, Foubister, & Mayes, 1996).

In summary, videoconference systems offer a high quality in terms of audio and video performance, but still lack to transport non-verbal communication queues such as eye-to-eye contact and have perception issues such as the 1:1 scale representations of the communication partners. All these factors add to the artificial experience that can arise from existing videoconference solutions.

IV. IMPLEMENTATION APPROACHES

In the following, we present seven different categories that represent different approaches of how to implement videoconferencing solutions supporting eye-to-eye contact. To our knowledge, those categories virtually account for all systems supporting mutual gaze that are reported on in research and market today. For each category we describe the core idea and technology in brief, give representative references where appropriate and discuss advantages and disadvantages including guidelines for their application.
We discuss the systems with respect to several criteria's that are crucial when setting up such a system. These are foremost the space requirements of the system, the quality of videoconferencing that can be achieved in particular focussing on mutual gaze and the complexity of the overall system. We further discuss the hardware and software requirements, the support of a flexible posture, setup price, integration effort, and the availability of the components.
One of the most obvious criteria for deciding on a particular videoconferencing technology is the space available and need for it. Those *space requirements* range from a desktop space in an office or simply videoconferencing on a mobile or portable device (e.g. laptop computer) to full room-sized systems. We consider systems in our comparison as positive if they do not require a lot of space even if under some circumstances this might be seen as negative and the other way around. Delivering an often desirable life-sized video representation cannot be achieved on a small screen, for instance. However, usually the less space is required the better.
The achievable *quality* depends on many factors and is mainly influenced by the ability of the system to deliver high resolution, high frame rate, and correct colour video, preferably life-sized, with no or only few visual artefacts, and a high quality audio channel optimised for speech frequencies and speaker awareness.
A higher *complexity* of the technology is considered more negative than a lower complexity. The simpler the better. More complex systems tend to be more expensive, less mature, more error-prone, harder to maintain, and harder to set up, configure and to use.
The *hardware and software requirements* are linked to the complexity of the system and describe how demanding the particular solution is. For instance, on one hand a PC-based videoconferencing system delivering eye contact by a complex algorithmic software solution would have very little hardware requirements (a PC will do), but a rather specialised software is required to achieve this. On the other hand, a hardware solution for eye contact, like a beam-

splitter would require very special silvered glass arranged in a very specific way, but would not require any special software, any standard videoconferencing application would do.
Some solutions require the users to sit in a very well defined position in front of the camera or screen. This is considered as negative. Other systems would allow for a wide range of flexible movements of the users. This is considered as positive - as an option for *flexible posturing*.
A low investment *price*, market *availability*, and the ease of *integration* into existing videoconferencing solutions and infrastructures are considered as positive.

The approaches to implement mutual gaze can be divided into three groups: (1) Custom hardware dependent setups for videoconferencing (Hole in Screen, Long Distance / Small Angle, Half-silvered Mirror), (2) Custom software dependent setups for videoconferencing (2D Video-based techniques, 3D Video-based techniques) and finally (3) Videoconferencing setups dependent on both; Custom hardware and custom software (Shuttered screen, Unshuttered screen).

Hole in Screen

The naïve, obvious solution is to drill a hole in the screen exactly at the desired position, place a camera there and with this allow for eye-to-eye contact. Apparently this is not a suitable technique for CRT or LCD monitors, but can be implemented with a screen canvas and a projector (Figure 3). Back projection is generally not possible because of the size and position of the camera: this would cast shadows on the back projection screen resulting from the back-mounted camera.
The opening and the camera should be as small in diameter as technically possible and the rim of the camera should be painted in the canvas screen colour to minimize the visibility of the camera – the hole will be visible in the very centre of the focus of attention between the eyes of the videoconferencing (V/C) partner. Even with very small (i.e. in the order of 5 mm in diameter) cameras the user has to be positioned decently far from the screen to mitigate the disturbing effect of the hole-between-the-eyes effect, which directly affects the overall size of the setup.

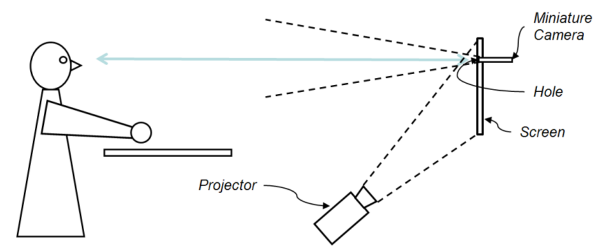

Figure 4. Schematic of hole in screen approach

Despite the need for a frontal projection the hole-in-the-screen approach allows for an unrestricted interaction space in front of the screen (see Figure 4). It can easily be implemented: all it needs is a projector, an inexpensive canvas (should be inexpensive because one has to drill a hole in it) and a small camera. The main disadvantages with respect to the quality are (a) the visible artefact of seeing a (black) spot between the eyes and (b) to minimize the spot effect the user has to be rather far away from the screen which requires much more space and also determines the achievable size of the video face display.

Table 1: Overview of hole in the screen approach

	Size	Quality	Flexible posturing	Complexity	Hardware requirements	Software requirements	Price	Availability	Integration	Examples
Hole in screen	-	-	+	++	+	++	+	+	+	No example in literature

Long Distance / Small Angle

If the room size permits, eye-to-eye contact can also be implemented by viewing the screen from a far distance and by placing the camera as close as possible (i.e. at the edge of the screen) to the displayed video stream. If the angle between the viewing axis and the eye axis (ß in Figure 5) is small enough then the offset between eyes and camera isn't noticeable.

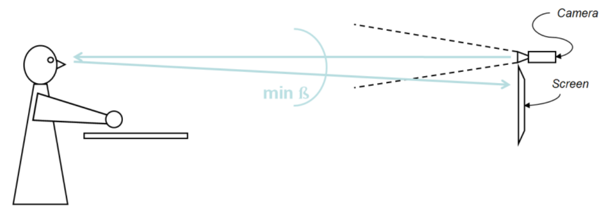

Figure 5. Schematic of long distance approach

Tam et al. (2007) found that in medical tele-consultation doctors sit rather close to the monitor (around 1m). The authors investigated the influence of eye-gaze angle on perceived eye contact. Comparing scenarios with ß = 7deg vs. ß = 15deg and different V/C environments (ranging from mobile devices to room sized environments) they concluded that smaller angles/greater distances are better for perceived eye contact. A more thorough discussion of the influence of eye-gaze angle on perceived eye contact can be found in the work of Chen (2002). He summarized that the eye-gaze angle is an asymmetric issue in a way that it depends on the general head position with respect to the conference partner. Consequently, different angles have been reported but are often stating that the acuity of eye-contact is as good as the visual acuity (1minute of arc) which means that the eyeballs are rotated by ~2.8°. These results match experiments within our laboratory showing that ß should be not much greater than 3 degrees and with this a rather long distance is needed to achieve the desired effect. Hence, if one wants to present a life-sized head (and only the head) and the camera is placed as closest as possible to the rim of the display a distance of at least 3 meters is required (atan (160mm/3000mm) = 3.05°), assuming an average adult head size. Even if one can free that much space the face of the videoconferencing partner appears rather small. Increasing the screen size to display a bigger face or a bigger portion of the partners' body would increase the angle ß and consequently requires even a longer distance. However, because of the asymmetric character it can be demonstrated that the vertical threshold is up to 5° before a deviation can be noticed (Chen, 2000). Also, Gale & Monk (2000) could show that users are quite accurate about guessing gaze from videoconferencing streams with an error of estimation of only a couple of degrees.

A special form a long distance approach is presented by Nguyen and Canny (2007) in their work on multiview group videoconferencing. Here the screen and camera have a rather large distance to the viewer. However, the main focus of this work is on generating a personal view for each user even though they see on the same projection surface. This is achieved by using a specific retro-reflective material reflecting the projection only back in the horizontal direction of the source while using a vertical diffuser to allow a varying vertical height.

In summary, the long-distance technique is the most affordable one and does not require specialized equipment and calibration. Similar to the Hole in screen approach there is enough room for interactions in front of the display and there are no visual artefacts. However, this technique requires a lot of space. The distance needed between the user and the display does not allow for close communication between the partners, this can only be achieved by much bigger than life-size displays, which is undesirable in most cases.

Table 2: Overview of long distance/small angle approach for establishing eye-to-eye contact.

	Size	Quality	Flexible posturing	Complexity	Hardware requirements	Software requirements	Price	Availability	Integration	Examples
Small angle	--	+	-	++	++	++	++	++	++	Chen 2002

Half-silvered Mirror

Using a half silvered-mirror, or any other kind of an optical beam-splitter, is probably the most commonly used solution in research and on the market to achieve eye-to-eye contact. In the literature there are several similar systems that use a half-silvered mirror for creating a video conferencing experience with eye-to-eye contact that can be implemented by either placing the camera above or below the mirror and the screen behind the mirror or the other way round by placing the screen above or below and the camera behind the mirror. Figure 6 shows one of the possible but most common configurations.

Kannes (1990, 1995) presented an early system using a half silvered mirror when introducing his courtroom conferencing system. This system allowed interviewing remote defendants and witnesses while maintaining eye-to-eye contact. The basic idea is that a user can see through half transparent mirror while being observed by a well-positioned camera at the same time. Nelson and Smoot (1992) describe a similar system but added polarizers that mitigate the effect of light entering the camera from the screen used to display the conferencing partner. Large et al. (2009) also used polarizers to avoid light passing from the display into the camera but contrary to the existing solutions they do not use one large half silvered mirror but many smaller ones that (similar to a lens array) are attached to the display and reflect a portion of the incoming light to the camera mounted in front of the screen. This

greatly reduces the required space but at the cost of visual quality (discontinuities caused by not properly aligned mirrors) and production costs for the mirror array.
McNelley and Machtig presented a series of improvements for setups not using polarizers. First improvements aimed for improving the size of the overall setup by introducing more optical elements (beam splitter, mirrors) to the system together with a projector and consequently allowed to reduce the distance of the projector to the mirror (McNelley, & Machtig, 1999). Later setups used a screen that was in front of the mirror. This setup made again the system smaller but required a beam splitter avoiding direct view into the screen (McNelley, & Machtig, 2001). Another approach aimed for integrating the conferencing partner into the environment by also capturing the environment and blending it with the displayed image of the conferencing partner for the cost of an increased system size (McNelley, & Machtig, 2004). The previous systems were static setups optimized for eye-to-eye contact. McNelley (2001) also introduced a dual mode system that permits normal use of the display as the eye-to-eye component (half silvered mirror and camera) could been conveniently folded away when not used. Similarly, Libbey (2004) created a system integrating a half-silvered mirror and a normal mirror into a box that can be attached to a normal desktop. Assuming that the camera is placed on top of the screen it aligns a portion of the screen showing the conference partner with the view of the camera. The advantages of this system were the size and the fact that the system could be easily added to normal hardware and be removed when not used while the main disadvantage is the limited size of the video.

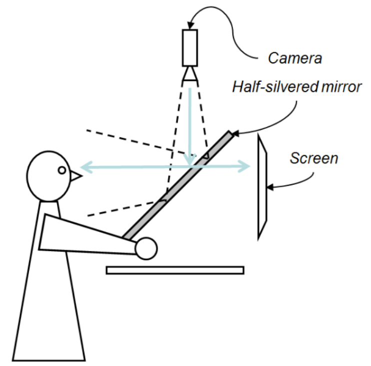

Figure 6 Schematic of half-silvered mirror approach as described by Kannes (1990, 1995).

The space in front of the user, where usually the desk is, provides only limited access because of the half-silvered mirror placement. However careful positioning, akin to ReachIn[8] setups can produce an interesting interaction space, where virtual objects can be blended with manual interaction (augmented reality interaction space).
The main advantage of half-silvered mirror systems lies in the simplicity of the setup: only the mirror and a standard monitor and camera are needed. It needs careful calibration though and usually produces optical artefacts due to the fact that the camera captures only half (or a certain percentage) of the true image of the user. The use of beam splitters can reduce the effect but not avoid it while also further reducing the brightness of the camera image. In fact, besides unwanted reflections, the maximum of achievable brightness and contrast levels might be problematic. The other main disadvantages are the high price and limited availability of large enough half-silvered mirrors needed for life-size displays that is shared by all approaches but those who only use small screens. Furthermore the space in front of the user is occupied with the display setup. While this was of lesser concern with the previously introduced approaches, here and with the following systems, the user should be placed in a way that the camera directly captures the eyes without too much deviation from the ideal spot.
Nowadays there are also commercial solutions available, normally with smaller screen sizes (e.g. iris2iris[9]) that make use of half-silvered mirrors.

Table 3: Overview of key techniques using a half silvered mirror approach for establishing eye-to-eye contact. Note: We focus on systems supporting life-sized videoconferencing. While reducing the size decreases the costs mainly driven by the large half-silvered mirror but als heavily affects the video conferencing experience.

	Size	Quality	Flexible posturing	Complexity	Hardware requirements	Software requirements	Price	Availability	Integration	Examples
Half-silvered	+	+	--	+	-	++	-	-	+	Kannes 1990,

[8] en.souvr.com
[9] www.iris2iris.com

mirror										1995, Libbey 2004, Mukawa et al. 2005, Quante & Muehlbach (1999)
Half-silvered mirror and polarizers	+	+	--	-	--	++	-	--	-	Nelson and Smoot 1992
Half silvered mirror and beam splitter	+	+	--	-	--	++	--	--	-	McNelley, & Machtig 1999, 2001, 2004

2D Video-based techniques

So far, most of the presented approaches to support eye-to-eye contact were mostly hardware solutions that required software for calibration. Contrary to this, 2D video-based approaches apply a software-based approach to support eye-to-eye contact. In the following we will discuss 2D approaches that work without 3D scene knowledge while we later also present 3D video-based techniques.

The literature shows different approaches for implementing eye-to-eye support using 2D video approaches. The first category of existing works uses a single camera. They compute the position of the pupil using computer vision techniques and apply image operations to remodel the pupil in such a way that it appears they look straight into the camera. Example images, if provided, illustrate this approach, but are not convincing, because of the amount of visual artefacts. Andersson et al. were the first claiming a conceptual solution for this kind of approach (1996 Andersson, Chen, & Haskell, 1996). Their idea was to use light reflections to detect the iris. Once detected, a three dimensional ellipsoid facial model is textured using the camera feed and finally re-oriented to simulate eye-to-eye contact. Andersson presented a simplified idea of the previous work that reorients the eyes by only shifting iris and eyelids within the 2D camera image (Andersson, 1997). Jerald & Daily also shift the iris but apply a non-linear warp (Jerald & Daily 2002). However, their work requires beforehand calibration to determine the maximum iris offset. This calibration creates a sweet spot in which the user has to stay and therefore greatly reduces the ability of having a flexible posture in front of the screen.

An image-based approach using epipolar geometry was presented by Cham et al. (Cham, Krishnamoorthy, & Jones, 2002). However, their approach was limited to small corrections as larger corrections introduced artefacts due to the lack of knowledge about the face’s shape.

Some approaches use multiple cameras while still relying on image-based operations. Lewis conceptually introduced this basic idea in 1994 (Lewis, 1994). Here, two or more cameras are positioned on the sides or corners of the screen (Figure 7). The closer the cameras can be positioned to the targeted screen position the better usually the results.

Criminisi et al, (Criminisi, Shotton, Blake, & Torr, 2003) use two cameras and applied a smart blending without 3D reconstruction. All pure image based approaches have in common that they require carefully calibrated cameras and are very sensitive to lighting effects within the image of the user. Furthermore, it is difficult to obtain convincing results for larger deviations in the viewers gaze. The results are also not backed up by studies giving evidence of the results and applicability.

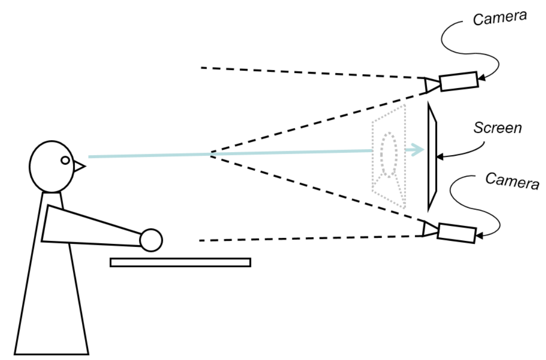

Figure 7. Schematic of 2D video interpolation approach. In some prototype systems a computer-vision approach is used to synthesize/interpolate the views of multiple cameras around the screen to compute an image from a virtual camera place in front of the user's eyes, which is then transmitted to the conferencing partner

To overcome these problems some groups worked with proxy geometry and model based approaches for image synthesis. Gemmell et al. (Gemmell, Toyama, Zitnick, Kang, & Seitz, 2000; Zitnick, Gemmell, & Toyama,1999) applied image warps by using a computer-vision and image synthesis based approach to redirect eyeballs and reorientate the displayed head (figure 8). They are using a model-based approach and are presenting early results (with artifacts). "For small angles of rotation (<5 degrees) we were successful in warping a person's head to a new orientation. For larger changes in rotation, realism was lost due to distortions." Similarly, Yip & Jin (2003) presented another approach applying image-warping techniques for warping the iris into the final camera image. Later Yip (2005) extended this system by also training an artificial neural network and gives additional input to remodel the users head and apply image warping. While it makes spontaneous videoconferencing impossible it also shows problems with rapid head movements leading to an artificial experience due to poor visual results.

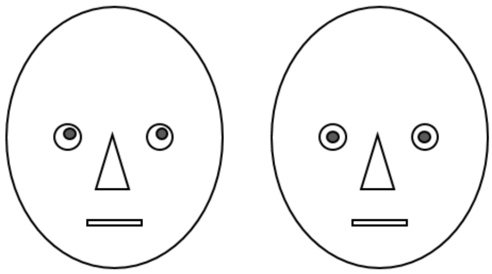

Figure 8. Eyeball correction - Left: eyeballs as seen by a camera from below/left, Right: corrected eyeballs

Synthesizing videoconferencing images is a vivid area of research (Ott, Lewis, & Cox, 1993) ranging from one or multiple camera systems which artificially replace the eye gaze in the video stream (Jerald & Daily, 2002; Gemmel et al., 2000; Tsai, Kao, Hung, & Shih, 2004; Schreer, Feldmann, Atzpadin, Eisert, Kauff, & Belt, 2008; Vertegaal, Weevers, Sohn, & Cheung, 2003) to systems which combine a half-silvered mirror with multiple cameras for multi-party videoconferencing (Vertegaal et al., 2003).
Test implementations and studies in our laboratory have shown that one should not expect perfect interpolation results (Wetzstein, 2005). Humans are very sensitive when it comes to realize subtle artificial elements in other faces. Even the slightest artefacts will be noticed and eventually destroy the eye-to-eye illusion.
Videoconference systems relying on video-interpolation to compute a synthetically frontal video allow for life-sized display and close proximity eye-to-eye contact. Due to their compact size (usually a screen with two or more cameras) they do not occupy the space in front of the screen and therefore are quite flexible in terms of positioning. The visual display parameters like brightness, contrast and colour can be very well controlled making it also a suitable approach for complex environments (e.g. lighting).
Contrary to many other systems using a fixed positioned camera and therefore require a fixed posture, systems using video-interpolations for computing the view of a virtual camera can tolerate to some extent movements of the communication partners head. This requires tracking the head of the communication partner and repositioning of the virtual camera.
However, the movement has to be within certain limits to still be able to compute an error-minimized image of the communication partner. Furthermore solutions using video interpolation require specialized and often expensive

hardware and software. Due to the complexity of the vision-based computation of the virtual camera these approaches also introduce perceivable artefacts while realizing eye-to-eye contact, which can be very distracting.

Table 4: Overview of 2D video-based techniques for establishing eye-to-eye contact.										
	Size	Quality	Flexible posturing	Complexity	Hardware requirements	Software requirements	Price	Availability	Integration	Examples
Image-based techniques (single camera)	+	--	-	--	+	--	+	+	++	Andersson, 1997 Jerald & Daily 2002
Image-based techniques (multiple cameras)	+	-	+	--	-	--	-	+	-	Criminisi, A., Shotton, J., Blake, A., Torr, P. H. S., 2003
Model-based techniques	+	-	+	--	+	--	+	+	++	Gemmell, Toyama, Zitnick, Kang, & Seitz, 2000

3D Video-based techniques

The previously presented approaches relied exclusively on 2D image operations working on one or more camera feeds to synthesize a new view. A logical step forward is to use several camera images to reconstruct a 3D model of the users head for the synthesis. Some systems even go a step further by also integrating a stereoscopic display into the system allowing for a 3D telepresence experience with eye support. In the following we will present several approaches using 3D reconstruction, some of them also with a stereoscopic 3D display.

Xu et al., (1999) presented a system using two cameras for stereo tracking and stereo analysis of image features representing the user's head. This allows to build a 3D model of the head and consequently to create a synthesized view supporting eye contact but requires several calibrated cameras and is prone to visible artifacts resulting from falsely matched image features. Yang R., & Zhang, Z. (2002) presented a similar system but using different image operations and also applying a Delaunay triangulation between the matched image points to compute the 3D model. However, the system shows similar remaining artifacts in particular between the background and the synthesized view of the users' head.

To improve the visual quality of reconstruction (e.g., closing holes and increasing the resolution of the underlying depth information) Zhu et al. proposed to use stereo cameras and fuse their results with a time-of-flight depth camera (Zhu, J., Yang, R., & Xiang, X., 2011).

Kuster et al. presented an approach making use of commodity 3D depth camera (MS Kinect). Using a 3D depth map and relying on a single depth camera introduces the problem that the resulting view has holes due to occlusions and errors in the measured depth. This is in particular noticeable in the reconstruction of the background and distance objects. In their work, Kuster et al. only create a synthesized view of the face and blend this in the original image to overcome these problems in image integrity (Kuster, Popa, Bazin, Gotsman, & Gross, 2012).

Some existing works do not focus on creating eye-to-eye contact but to allow for a full telepresence experience, this usually includes the reconstruction of the environment in addition to the reconstruction of the communication partner's face. One example for this approach is the work by Petit et al. that reconstructs the full human body of the communication partners (Petit, Lesage, Menier, Allard, Franco, Raffin, Boye, & Faure, 2010). As a full 3D model exists these systems can usually also create a view allowing for eye-contact. The quality thereby largely depends on the used hardware (e.g. camera resolution), system configuration (e.g. size of the overall system and distance to the camera), and used algorithms. Some approaches (e.g. visual hull computing) tend to blur the details, such as eyes, making it hard to detect gaze in the final reconstruction of the user's body.

Prince et al. (2002) created live 3D reconstructions of videoconference partners, using a camera array and visual hull algorithms. The dynamic 3D model is visualized at the partners' location and can be used within a mixed reality-based videoconferencing setup. Mutual gaze was not demonstrated but could possibly be achieved.

Many systems for supporting telepresence rely on expensive hardware such as hardware-synced cameras running in specifically prepared environments, and often even having multiple computers for the reconstruction. Maimone and Fuchs presented a system for telepresence that uses off-the shelf- hardware such as multiple Microsoft Kinects (Maimone & Fuchs, 2011). The presented system uses several Kinects simultaneously and merges their depth information into one model. While the results show visible artefacts resulting from error in the reconstruction enough details are preserved to identify gaze.

A more technically complex approach is the combination of a real-time 3D face scanner with a sophisticated projection system based on a rotating mirror which gives a "true" 3D impression of the remote user's head in a hologram-like manner (correct for the nearest (visually tracked) viewer). Very convincing results are printed in the article, but highly specialized equipment and setup knowledge is required here (Jones, Lang, Fyffe, Yu, Busch, McDowall, Bolas, & Debevec, 2009). The same is true for telepresence robot systems, like the BiReality robot surrogate with four displays arranged around a cube presented by Jouppi, Iyer, Thomas, & Slayden (2004).

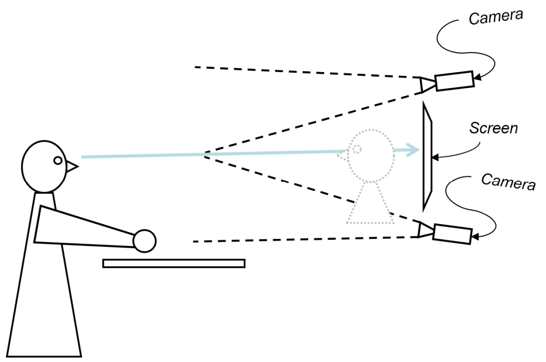

Figure 9. Schematic of 3D interpolation approach

As part of the VIRTUE teleconferencing research prototype, Schreer & Kauff (2002) describe an approach and implementation of a model-based approach to capture, transmit and display participants, environment and artefacts in a table-based conferencing situation. Because the users' heads (and eyes) are captured in a 3D model, a view-independent presentation and with this eye-contact can be made possible. The presented pictures are impressive, but still show significant signs of artifacts in the representation of the user.

Most of the systems relying on 3D techniques for computing a synthesized view, and with this allowing for eye contact, have similar drawbacks as systems using 2D. They are very sensitive to light conditions and good results have often only been reported in controlled environments. Furthermore, many need dedicated expensive hardware such as cameras or significant computational resources. Fortunately, the rise of consumer level depth cameras such as MS Kinect gave new research opportunities in terms of affordable systems.

Table 5: Overview of 3D video-based techniques for establishing eye-to-eye contact.

	Size	Quality	Flexible posturing	Complexity	Hardware requirements	Software requirements	Price	Availability	Integration	Examples
3D Video-based techniques relying on video cameras only	+	-	+	-	+	--	+	+	-	Xu et al., 1999
3D Video-based techniques using professional depth cameras	+	+	+	-	-	--	--	-	-	Zhu, J., Yang, R., & Xiang, X., 2011
Kinect-based 3D techniques	+	+	+	+	+	--	+	+	+	Kuster, C., Popa, T., Bazin, J.C., Gotsman, C., Gross, M., 2012

Shutter Techniques

As first shown by the blue-C system (Gross et al., 2003) and later by the HoloPort system (Kuechler & Kunz, 2006) 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 (Figure 10).

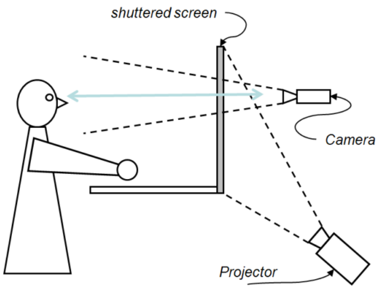

Figure 10. Schematic of shutter screen approach

Between update cycles of the system a black projection image is rendered and a camera image is captured. During that period of time the screen is transparent because a shutter glass is used and triggered to transparent mode synchronously. Hartkop (2007) placed cameras behind an OLED screen synchronizing the video camera(s) with the OLEDs illumination levels for the same purpose.

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 some artefacts in the displayed videoconferencing video though. Furthermore, the shuttered screen needs some time to fully switch its state and depending on the screen quality the direction and progress of switching the surface can cause visible artefacts.

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 solution might be too expensive, require too much environmental space in front of the display, might require too much distance, occupy the interaction space in front of the display, require specialized hardware and software or produce poor visual quality (e.g. flicker, visible artefacts, and brightness).

Table 6: Overview of techniques for eye-to-eye contact relying on shuttered screens.

	Size	Quality	Flexible posturing	Complexity	Hardware requirements	Software requirements	Price	Availability	Integration	Examples
Shuttered screens	--	+	+	-	--	-	--	--	--	Gross et al., 2003

Unshuttered Screen

Approaches in this category are usually based on the idea of the shuttered screen, but try to minimize the visible artefacts caused by shuttering the screen and by the limited opacity of the screen. One approach was to replace the shuttered screen with a screen based on standard *Holographic Optical Element* (HOE) screen (Tedesco, 1999). From certain predefined angles the HOE will be opaque and can be used for projecting onto it while for other viewing angles the HOE is transparent and a camera can be placed to record the user in front of the HOE screen. Similarly Nelson & Vaning (1997) presented the conceptual idea of using a transmissive layer in front of the screen with a micro-louver assembly. A similar, but also not shuttered approach uses a “light transmissible screen”. A special film/material on the screen is used that allows only passing a certain percentage of light (e.g. MAJIC system by Okada et al. 1994).

However, for all those systems using HOEs, the micro-louver assembly, or the ones using a film on the screen, we can still expect a considerable amount of unintended reflections and diffusions visible on the back of the screen resulting from the back projection, which will be captured by the camera.

Regenbrecht et al. (2014) have produced a conferencing system that uses a HOE-based screen but eliminates remaining reflections by using polarizing filters in front of the projector and the camera.

Table 7: Overview of techniques for eye-to-eye contact relying on unshuttered screens.

	Size	Quality	Flexible posturing	Complexity	Hardware requirements	Software requirements	Price	Availability	Integration	Examples
HOE or use of screen with layers/films attached	--	+	+	-	--	++	--	--	-	Tedesco, 1999, Okada et al. 1994
HOE with filters	--	++	+	-	--	+	--	--	-	Regenbrecht et al. 2014

V. SUMMARY AND DISCUSSION

To date, there don't seem to be an ideal solution which maximizes all desired qualities for a certain setting and scenario. Half-silvered mirror techniques are well researched and studied and therefore lend themselves to be used in a variety of scenarios. 2D and 3D techniques are less well studied but offer good potential for future research and development. New algorithms for 2D/3D interpolation as well as first studies are in progress. System designers would need metrics to develop and optimize their systems though.

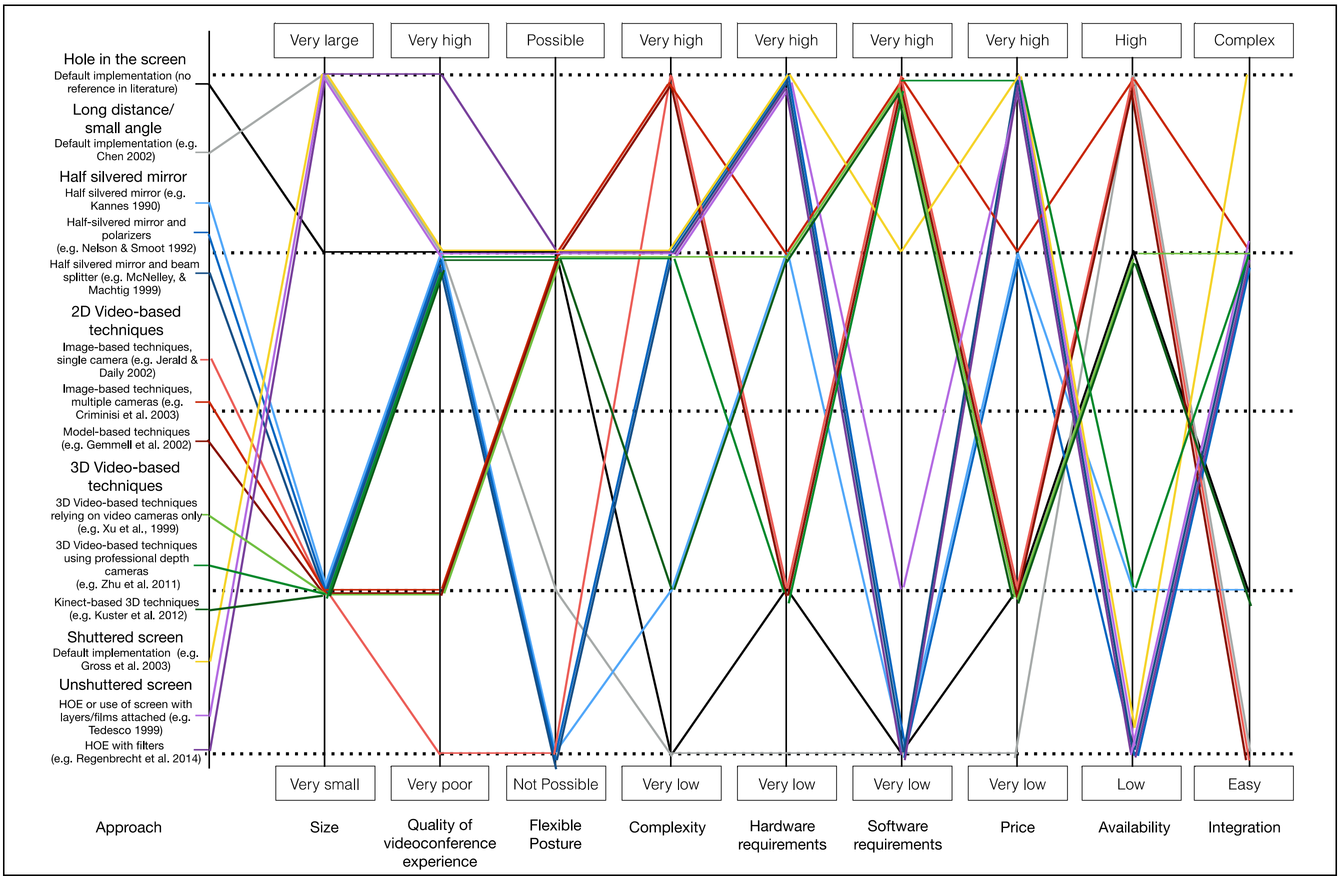

Figure 11. Parallel line plot visualising the main approaches for eye-to-eye videoconferencing and their performance according to our criteria's. As one can see there are a wide variety of different approaches with different strengths. However, all existing approaches have also major drawbacks. Usually either the hardware or the software requirements are very high but still don't guarantee perfect quality.

With figure 11 we are presenting a compiled and condensed overview of our findings to give developers and decision makers a tool at hand for current and future technologies to implement mutual gaze in videoconferencing.

We showed that there are different technologies for implementing eye-to-eye contact in video conferencing systems that differentiate themselves through the complexity of hardware and software. Overall, there are naïve approaches (hole in the screen and small angle) that are, despite their advantages in terms of system complexity and costs, rarely used. This might be mostly due to the general size and in particular to the screen distance which prohibits several use-cases. Still, in prototypical applications where the large screen distance can be compensated by a bigger screen or where the screen distance is not critical these setups might be worth a try in particular due to their affordability.

Video conferencing setups supporting eye-to-eye contact using half silvered mirrors are well researched and studied. They have been technically described as well as been used to research eye-to-eye contact and other perceptual aspects. Still, systems using half-silvered mirrors have downsides in terms of quality (usually ghosting and lower contrast) as well price due to the need for the half-silvered mirror.

Video conferencing setups using video-based gaze control techniques such 2D interpolations or 3D reconstructions while technically described in some detail are still having problems in terms of quality due to interpolation artifacts. We argue, that these techniques probably have the biggest potential in terms of future research and possible improvements. Both of these approaches, 2D and 3D, heavily rely on real-time computer vision techniques and consequently benefit from new and faster hardware allowing more complex computations as well as they benefit from new approaches in the area of computer vision. We therefore think that the quality of these approaches can still be significantly improved while keeping minimal hardware requirements (of usually having additional cameras), the support for flexible postures and freedom of movement in front of the screen and a general small setup space size. However, even if the quality was improving over the last couple of years all these interpolation and reconstruction techniques are challenged mainly by two aspects: (1) even the slightest artefacts in a person's face, and in particular in and around the eyes, are easily noticeable and therefore a very high and reliable synthesis quality has to be provided and (2) related to this first aspect, developers have to overcome the "uncanny valley" known from anthropomorphic robotics research: if the eye contact looks and feels almost exactly like in human-human communication, but not exactly, this might lead to revulsion.

To date, the best quality is offered by approaches using shuttered or unshuttered screens, with several different systems described in the literature. Still we argue that these approaches do not offer as much potential for future improvements as for example 2D video and 3D video techniques while at the same time have the drawback of a complex hardware setup.

In the past, most research was spent on half-silvered mirror setups while many of the later systems (2D video techniques, 3D video techniques, shuttered and unshuttered screens) have not been studied or evaluated in terms of perception of eye-to-eye contact. There is especially no work comparing these existing systems in terms of user perception with respect to eye-to-eye and mutual gaze support. However, these evaluations and studies also require well-defined metrics for evaluating mutual gaze in videoconferencing, which need to be defined beforehand. To date, there are no common metrics used and mutual gaze was only investigated targeting specific aspects (e.g., trust).

VI. CONCLUSION AND FUTURE WORK

In this paper we presented an overview on videoconferencing solutions supporting eye-to-eye contact between participants. We motivated the advantages of gaze and mutual gaze for supporting communication in business-type videoconferences by reviewing main works in the field. We could show that gaze and eye-to-eye contact is fundamental for the quality of the experienced interest and feedback but also for trust between communication partners; an important parameter that is especially important in business meetings. Further results demonstrated that lack of eye-to-eye contact affects the interaction quality, making it a more artificial experience, but if eye contact or gaze are supported the quality of discussion can be increased leading to shorter discussion times or improve turn-taking.

We further surveyed and discussed existing solutions for video-conferencing supporting eye-to-eye contact, ranging from affordable solutions to complex systems. We developed a set of criteria to be used as a guide for the selection of research directions, developments, and investment decisions and evaluated all commonly known technical approaches to achieve mutual gaze according to those criteria.

Mutual gaze support in videoconferencing is still an open research topic. For instance, for future setups one could think of a constellation where the reflections on a reflective LCD screen surface itself, normally an annoyance for users, is used as a mirror. The screen is inherently polarizing the light and a camera with a polarizing filter properly positioned near the user at a certain angle to a slightly titled screen could deliver the desired effect. This would be an affordable solution, needing not much space and leaving the space in front of the screen unoccupied. The main challenge is the reduction of unwanted artefacts though. 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. In particular software solutions using one or more 2D or depth cameras offer huge potential.

ACKNOWLEDGMENTS

We want to thank Simon Hoermann, Lavell Mueller and Andreas Dünser for their valuable input to this paper as well as the reviewers of an earlier version of this paper for their feedback helping us to improve this paper. This work was partly funded by a University of Otago CALT – RUT research grant (JDLJ17096).

REFERENCES

Andersson, R.L., Chen, T., & Haskell, B.G. (1996). Video Conference System and Method of Providing Parallax Correction and Sense of Presence. United States Patent 5,500,671, Mar 19, 1996.

Andersson, R.L. (1997). Method for achieving eye-to-eye contact in a videoconferencing system. United States Patent 5,675,376, Oct 7, 1997.

Bailenson, J.N., Beall, A.C., & Blascovich, J. (2002). Gaze and taskperformance in shared virtual environments. J. Visual. Comput. Animat. 2002; 13: (DOI: 10.1002/vis.297), 313–320.

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? Psychological Science, 17(6), 514-520.

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

Biocca, F., Harms, C. and Gregg, J.: 2001, The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity, Presence 2001, 4th international workshop, Philadelphia, PA.

Birden, H. & Page, S. (2005). Teaching by videoconference: a commentary on best practice for rural education in health professions. Rural and Remote Health 5: 356. (Online) 2005. Available: http://www.rrh.org.au

Bondareva, Y. & Bouwhuis, D. (2004). Determinants of Social Presence in Videoconferencing. Proceedings of the Workshop on Environments for Personalized Information Access Working Conference on Advanced Visual Interfaces AVI 2004, Gallipoli, Italy, May 25, 2004 (eds.: Liliana Ardissono and Giovanni Semeraro), 1-9.

Bondareva, Y., Meesters, L.M.J. & Bouwhuis, D.G. (2006). Eye contact as a determinant of social presence in video communication. In 20th international symposium Human factors in telecommunication, Sophia-Antipolis, France, 20-23 March 2006.

Buxton, W. & Moran, T. (1990). EuroPARC's Integrated Interactive Intermedia Facility (IIIF): Early Experience, In S. Gibbs & A.A. Verrijn-Stuart (Eds.). Multiuser interfaces and applications, Proceedings of the IFIP WG 8.4 Conference on Multi-user Interfaces and Applications, Heraklion, Crete. Amsterdam: Elsevier Science Publishers B.V. (North-Holland), 11-34.

Buxton, W. (1992). Telepresence: integrating shared task and person spaces. Proceedings of Graphics Interface '92, 123-129.

Buxton, W., Sellen, A. & Sheasby, M. (1997). Interfaces for multiparty videoconferencing. In K. Finn, A. Sellen & S. Wilber (Eds.). Video Mediated Communication. Hillsdale, N.J.: Erlbaum, 385-400. Review (back then) of different multi-party V/C systems.

Carville, S. & Mitchell, D. R. (2000). 'It's A Bit Like Star Trek': The Effectiveness of Video Conferencing. Innovations in Education & Training International, 37:1, 42-49.

Cham, T.-J., Krishnamoorthy, S., and Jones, M. (2002). Analogous view transfer for gaze correction in video sequences. InICARCV, vol. 3, 1415–1420.

Chen, M. (2002). Leveraging the Asymmetric Sensitivity of Eye Contact for Videoconference. In proceedings of CHI 2002, April 20-25, 2002, Minneapolis, Minnesota, USA.

Cook, M. Gaze and Mutual Gaze in Social Encounters. American Scientist 65, (1977), 328–333.

Cornelius, C., & Boss, M. (2003). Enhancing Mutual Understanding in Synchronous Computer-Mediated Communication by Training. Communication Research 30(2), April 2003, 147-177

Criminisi, A., Shotton, J., Blake, A., Torr, P. H. S. (2003). Gaze Manipulation for One-to-one Teleconferencing, Proceedings of the Ninth IEEE International Conference on Computer Vision, p.191, October 13-16, 2003

Davis, A.W. & Weinstein, I.M. (2005). The Business Case for Videoconferencing - Achieving a Competitive Edge. White Paper. Wainhouse Research. March 2005.

Detenber, B. and Reeves, B. (1996). A bio-informational theory of emotion: Motion and image size effects on viewers. Journal of Communication, 46(3), 66-84.

Eisert, P. (2003). Immersive 3-D Video Conferencing: Challenges, Concepts, and Implementations. In Proceedings of SPIE Visual Communications and Image Processing (VCIP), Lugano, Switzerland, July 2003.

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

Fish, R.S., Kraut, R.E., & Chalfonte, B.L. (1990). The VideoWindow System in Informal Communications. In Proceedings of CSCW 1990, October 1990.

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

Fullwood, C. & Doherty-Sneddon, G. (2006). Effect of gazing at the camera during a video link on recall. Applied Ergonomics 37 (2006), 167–175.

Gale, C. & Monk, A.F. (2000). Where am I looking? The accuracy of video-mediated gaze awareness. Perception & Psychophysics 62 (3), 586-595.

Gamer, M., & Hecht, H. (2007). Are You Looking at Me? Measuring the Cone of Gaze. Journal of Experimental Psychology: Human Perception and Performance, 33(3), 705-715.

Garau, M., Slater, M., Bee, S., & Sasse, M. A. (2001). The Impact of Eye Gaze on Communication using Humanoid Avatars. Proceedings of the SIGCHI conference on Human factors in computing systems. Seattle, Washington, United States, 3(1), 309-316.

Gaver, W.W. (1992).The Affordances of Media Spaces for Collaboration. Proceedings of the 1992 ACM conference on Computer-supported cooperative work (CSCW '92), 17-24.

Gemmell, J., Toyama, K., Zitnick, C. L., Kang, T., & Seitz, S. (2000). Gaze Awareness for Video-Conferencing: A Software Approach. IEEE Multimedia Magazine, 7(4), 26-35.

Gibbs, S.J., Arapis, C., & Breiteneder, C.J. (1999). TELEPORT – Towards immersive copresence. Springer: Multimedia Systems 7: 214–221.

Gill, D., Parker, C., & Richardson, J. (2005). Twelve tips for teaching using videoconferencing. Medical Teacher 27 (7), 573–577.

Grayson, D. M., & Monk, A. F. (2003). Are you looking at me? Eye contact and desktop video conferencing. ACM Transactions on Computer-Human Interaction, 10(3), 221-243.

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.

Hartkop, D. (2007). Method and Apparatus for maintaining eye contact during person-to-person video telecommunication.

Hauber, J., Regenbrecht, H., Cockburn, A. & Billinghurst, M. (2012). The Impact of Collaborative Style on the Perception of 2D and 3D Videoconferencing Interfaces. The Open Software Engineering Journal 6(1), [DOI: 10.2174/1874107X01206010001], Bentham Science, 1-20.

Heaton, L. (1998). Preserving Communication Context. In C. Ess and F. Sudweeks (eds). Proceedings Cultural Attitudes Towards Communication and Technology '98, University of Sydney, Australia, 207-230.

Ishii, H., Konayashi, M., & Grudin, J. (1993). Integration of Interpersonal Space and Shared Workspace: ClearBoard Design and Experiments. ACM Transactions on Information Systems 11 (4), 349-375.

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

Jerald, J., & Daily, M. (2002). Eye gaze correction for videoconferencing. Proceedings of the 2002 symposium on Eye tracking research & applications. New Orleans, Louisiana, USA, 77-81.

Joiner, R., Scanlon, E., O'Shea, T., Smith, R.B., & Blake, C. (2002). Evidence from a series of experiments on videomediated Collaboration: Does Eye Contact Matter?. Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community (CSCL '02), 371-378

Jones, A., Lang, M., Fyffe, G., Yu, X., Busch, J., McDowall, I., Bolas, M., Debevec, P. (2009). Achieving Eye Contact in a One-to-Many 3D Video Teleconferencing System. ACM Trans. Graph. 28, 3, Article 64 (August 2009).

Jouppi, N.P., Iyer, S., Thomas, S., & Slayden, A. (2004). BiReality: MutuallyImmersive Telepresence. In Proceedings of MM'04, October 1016, 2004, New York, New York, USA.

Kannes, D. (1990). Video Conferencing System for Courtroom and other applications. United States Patent 4,965,819, Oct 23, 1990.

Kannes, D. (1995). Video Conferencing System for Courtroom and other applications. United States Patent 5,382,972, Jan 17, 1995.

Kuechler, M. and Kunz, A. (2006). HoloPort – A Device for Simultaneous Video and Data Conferencing Featuring Gaze Awareness. Proceedings of IEEE Virtual Reality 2006, March 25 - 29, Alexandria, Virginia, USA

Kuster, C., Popa, T., Bazin, J.C., Gotsman, C., Gross, M., Gaze (2012). Correction for Home Video Conferencing, Proceedings of ACM SIGGRAPH Asia (Singapore, November 28 - December 1, 2012), ACM Transactions on Graphics, vol. 31, no. 6, pp. 174:1-174:6

Large, T.A. and Rosenfeld, D. & Emerton, N. (2009). Establishing eye contact in video conferencing, US Patent App. 12/474,044

Levine, T. R., Asada, K. J. K., and Park, H. S. (2006). The Lying Chicken and the Gaze Avoidant Egg: Eye Contact, Deception, and Casual Order. Southern Communication Journal, 71(4), 401-411.

Lewis, J.P. (1994). Videoconferencing system using a virtual camera image. United States Patent 5,359,362, Oct 25, 1994.

Libbey, K.A. (2004). System for Maintaining Eye Contract[sic!] during Videoconferencing. United States Patent US2004/0155956 A1, Aug 12, 2004.

Maimone, A. and H. Fuchs. (2011). Encumbrance-free Telepresence System with Real-time 3D Capture and Display using Commodity Depth Cameras. Proceedings IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2011 (Basel, Switzerland, October 26-29, 2011)

McAndrew, P., Foubister, S.P., & Mayes, T. (1996). Videoconferencing in a language learning application. Interacting with Computers 8(2), 207-217.

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.

McNelley, S.H., & Machtig, J.S. (1999). Desktop Large Image and Eye-Contact Projection Display. United States Patent 5,890, 787, Apr 6, 1999.

McNelley, S.H., & Machtig, J.S. (1999). Reflected Display Teleconferencing Eye Contact Terminal. United States Patent 5,953, 052, Sep 14, 1999.

McNelley, S.H. (2000). Foldable Eye Contact Components for a Dual Mode Display. United States Patent 6,104424, Aug 15, 2000.

McNelley, S.H., & Machtig, J.S. (2001). Integrated Reflected Display Teleconferencing Eye Contact Terminal. United States Patent 6,243,130 B1, Jun 5, 2001.

McNelley, S.H., & Machtig, J.S. (2004). Adaptable Teleconferencing Eye Contact Terminal. United States Patent 6,710,797 B1, Mar 23, 2004.

Mukawa, N., Oka, T., Arai, K. & Yuasa, M. (2005). What is Connected by Mutual Gaze? — User's Behavior in Video-mediated Communication —. In proceedings of CHI 2005, April 2–7, 2005, Portland, Oregon, USA, 1677-1680.

Nelson, T.J. & Smoot, L.S. (1992). Eye Contact Apparatus for Video Conferencing. United States Patent 5,117,285, may 26, 1992.

Nelson, T.J. & Vaning, B.R. (1997). Eye contact apparatus employing a directionally transmissive Layer for Video Conferencing. United States Patent 5,612,734, Mar 18, 1997.

Nguyen, D., & Canny, J. (2007). MultiView: improving trust in group video conferencing through spatial faithfulness. Proceedings of ACM CHI'07, 1465-1474.

Olsen, J.S., Olsen, G.M., & Meader, D.K. (1995). What Mix of Video and Audio is Useful for Small Groups Doing Remote Real-time Design Work? In proceedings of CHI' 95, Denver, Colorado, USA, 362-368.

Okada, K., Maeda, F., Ichikawa, Y., & Matsushita, Y. (1994). Multiparty videoconferencing at virtual social distance: MAJIC design, Proceedings of the 1994 ACM conference on Computer supported cooperative work (CSCW '94), 385-393.

O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G., & Bruce, V. (1996). Comparison of face-to-face and video-mediated interaction. Interacting with Computers 8(2), 177-192.

Ott, M., Lewis, J. P., & Cox, I. (1993). Teleconferencing eye contact using a virtual camera. Paper presented at INTERACT '93 and CHI '93 conference companion on Human factors in computing systems in Amsterdam, The Netherlands, 109-110.

Petit, B., Lesage, J.-D., Menier, C., Allard, J., Franco, J.-S., Raffin, B., Boyer, E., and Faure, F. (2010). Multi-camera real-time 3D modeling for telepresence and remote collaboration. Intern. Journ. of Digital Multi. Broadcasting.

Prince, S., Cheok, A.D., Farbiz, F., et al. 3D Live: Real Time Captured Content for Mixed Reality. ISMAR '02 Proceedings of the 1st International Symposium on Mixed and Augmented Reality, (2002), 7.

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

Regenbrecht, H., Lum, T., Kohler, P., Ott, C., Wagner, M., Wilke, W., Mueller, E. (2004). Using Augmented Virtuality for Remote Collaboration. Presence: Teleoperators and virtual environments, 13(3), MIT Press, Cambridge/MA, USA. 338-354.

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

Regenbrecht, H., Müller, L., Hoermann, S., Langlotz, T., Wagner, M. & Billinghurst, M. (2014). An Eye-to-Eye Contact System for Life-Sized Videoconferencing. submitted.

Schreer, O., Feldmann, I., Atzpadin, N., Eisert, P., Kauff, P., & Belt, H.J.W. (2008). 3Dpresence - A System Concept for Multi-User and Multi-Party Immersive 3D Videoconferencing. In Proceedings of 5th European Conference on Visual Media Production (CVMP 2008), 1 - 8.

Shahid, S., Krahmer, E., & Swerts, M. (2012). Video-mediated and co-present gameplay: Effects of mutual gaze on game experience, expressiveness and perceived social presence. Interacting with Computers 24 (2012), 292–305.

Sorlie, T., Gammon, D., Bergvik, S., & and Sexton, H. (1999). Psychotherapy Supervision Face-To-Face and by Videoconferencing: A Comparative Study. British Journal of Psychotherapy 15(4), 452-562.

Swaab, R.I., Swaab, D.F. (2008). Sex Differences in the Effects of Visual Contact and Eye Contact in Negotiations, Journal of Experimental Social Psychology (2008), doi: 10.1016/j.jesp.2008.06.009.

Tam, T., Cafazzo, J.A., Seto, E., Salenieks, M.E., & Rossos, P.G. (2007). Perception of eye contact in video teleconsultation. Journal of Telemedicine and Telecare 13 (2007), 35–39.

Tedesco, J.M. (1999). Apparatus facilitating eye-contact video communications. United States Patent 5,856,842, Jan 5, 1999.

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, Brisbane, Australia, 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, Canberra, Australia, 273-280.

Teoh, C., Regenbrecht, H., & O'Hare, D. (2012). How the Other Sees Us: Perceptions and Control in Videoconferencing. Proceedings of ACM OZCHI 2012, Nov 26 - 30, 2012, Melbourne, Australia.

Teoh, C., Regenbrecht, H., & O'Hare, D. (2014). Body language availability in video-conferencing: Perceptions of others and self-transmission. Manuscript submitted (under revision). Information Technology & People.

Tsai, Y, Kao, C., Hung, Y., & Shih, Z. (2004). Real-Time Software Method for Preserving Eye Contact in Video Conferencing. Journal of Information Science and Engineering 20 (2004), 1001-1017.

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

Vertegaal, R. (1999). The GAZE Groupware System: Mediating Joint Attention in Multiparty Communication and Collaboration. In Proceedings of ACM CHI'99 Conference on Human Factors in Computing Systems. Pittsburgh, PA USA: ACM, 1999.

Vertegaal, R., van der Veer, G., & Vons, H. (2000). Effects of Gaze on Multiparty Mediated Communication. In Proceedings of Graphics Interface 2000, Montreal, Canada.

Vertegaal, R. & Ding, Y. (2002). Explaining Effects of Eye Gaze on Mediated Group Conversations: Amount or Synchronization? In proceedings of CSCW'02, November 16–20, 2002, New Orleans, Louisiana, USA, 41-48.

Yamashita, N., Hirata, K., Aoyagi, S., Kuzuoka, H., & Harada, Y. (2008). Impact of Seating Positions on Group Video Communication. In proceedings of CSCW'08, November 8-12, 2008, San Diego, California, USA, 177-186.

Vertegaal, R., Ding, Y. (2002). Explaining effects of eye gaze on mediated group conversations: amount or synchronization? In: CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work, New York, NY, USA, ACM (2002) 41–48.

Vertegaal, R., Weevers, I., Sohn, C., & Cheung, C. (2003). GAZE-2: Conveying Eye Contact in Group Video Conferencing Using Eye-Controlled Camera Direction. In proceedings of CHI 2003, April 5–10, 2003, Ft. Lauderdale, Florida, USA, 521-528.

Wainhouse Research. (2002). The business case for videoconferencing: Understanding the benefits, costs, and risks of videoconferencing over ISDN and IP. Retrieved from:http://www.radvision.com/NR/rdonlyres/0BA46C30-9F92-4136-A457-A04AE0133064/0/RADVISIONBusinessCaseforVideoconferencing.pdf

Webster, J. & Hackley, P. (1997). Teaching Effectiveness in Technology-Mediated Distance Learning. The Academy of Management Journal 40(6) (Dec., 1997), 1282-1309.

Wetzstein, G. (May 2005). Image-Based View Morphing for Teleconferencing Applications. Project Report, University of Otago.

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

Wolff, R., Roberts, D. Murgia, A., Murray, N., Rae, J., Steptoe, W., Steed, A., & Sharkey, P. (2008). Communicating Eye Gaze across a Distance without Rooting Participants to the Spot. In Proceedings of the 12th 2008 IEEE/ACM International Symposium on Distributed Simulation and Real-Time Applications, 111-118.

Xu, L-Q, Loffler, A., Sheppard, P.J., and Machin, D. (1999). True-view Videoconferencing System Through 3-D Impression of Telepresence. BT Technology Journal, 1999, Volume 17, Number 1, 59-68.

Yang, R., & Zhang, Z. (2002). Eye gaze correction with stereovision for video-teleconferencing. In ECCV, 479–494.

Yip, B. & Jin, J.S. (2003). An effective eye gaze correction operation for video conference using antirotation Formulas. In Proceedings of IEEE ICICS-PCM 2003, 15-18 December 2003, Singapore.

Yip, B. (2005). Face and Eye Rectification in Video Conference using Artificial Neural Network. Proceedings IEEE International Conference on Multimedia and Expo (ICME 2005), 2005, 690 – 693..

Zitnick, C.L., Gemmell, J., & Toyama, K. (1999). Manipulation of Video Eye Gaze and Head Orientation for Video Teleconferencing. Technical Report MSR-TR-99-46. Microsoft Research.

Zhu, J., Yang, R., & Xiang, X. (2011). Eye contact in video conference via fusion of time-of-flight depth sensor and stereo. 3D Research 2, 1–10.

ABOUT THE AUTHORS

Holger Regenbrecht is an Associate Professor at the department of Information Science at Otago University and leads the Computer-Mediated Realities Lab. His research interests include Human-Computer Interaction (HCI), Applied Computer Science and Information Technology, (collaborative) Augmented Reality, Videoconferencing and Collaboration, Psychological aspects of Mixed Reality, Three-dimensional user interfaces (3DU) and Computer-aided therapy and rehabilitation.

Tobias Langlotz is a Lecturer at the department of Information Science at Otago University. His main research interest is in location-based mobile interfaces, where he works at the intersection of HCI, Computer Graphics, Computer Vision and Pervasive Computing.