Look Who's Talking: the GAZE Groupware System

Roel Vertegaal
Ergonomics Department
Twente University
The Netherlands
roel@acm.org

Harro Vons
Baan Development
Barneveld
The Netherlands
hvons@baan.nl

Robert Slagter
Ergonomics Department
Twente University
The Netherlands
r.j.slagter@student.utwente.nl

ABSTRACT
The GAZE Groupware System is a multiparty mediated system which provides support for gaze awareness in communication and collaboration. The system uses an advanced, desk-mounted eyetracker to metaphorically convey gaze awareness in a 3D virtual meeting room and within shared documents.

KEYWORDS: CSCW, awareness, eyetracking, VRML2.

INTRODUCTION

With recent advances in network infrastructure and computing power, desktop video conferencing and groupware systems are rapidly evolving into viable solutions for remote communication and collaboration. The central premise for the use of video mediated communication over traditional telephony has been that video images improve the quality of communication between individuals by increasing the available sensory bandwidth. In a face-to-face situation, auditory, visual and haptic expressions are freely combined to convey messages and regulate interaction. It has been presumed that by adding video to an audio-only communication link, mediated communication would bear a significantly closer resemblance to face-to-face communication. However, usability studies [2] reveal a problem in multiparty video mediated communication: knowing who is talking or listening to whom. We believe multiparty communication and collaboration tools should provide simple, yet effective means of capturing and representing the attention participants have for one another and their work [7]. Conveying gaze direction is a good way of providing such information [1, 5]. In this paper, we will discuss the eye-controlled GAZE Groupware System (GGS), a virtual meeting room which supplements multiparty audio conferencing with gaze awareness, allowing users to see where other participants look, be it at each other or within documents.

THE GAZE GROUPWARE SYSTEM

Our prototype provides awareness about the participants' gaze position without some of the drawbacks of earlier systems such as Hydra [4]. Instead of conveying gaze direction by means of multiple streams of video, the GAZE Groupware System (GGS) measures directly where each participant looks using an advanced desk-mounted eyetracking system. The system represents this information metaphorically in a 3D virtual meeting room and within shared documents. The system does this using the Sony Community Place [6] plug-in, which allows interactive 3D scenes to be shared on a web page using a standard multiplatform browser such as Netscape. The GAZE Groupware System can be used in conjunction with any multiparty speech communication facility such as an internet-based audio conferencing tool, or standard telephony.

The Virtual Meeting Room

Fig. 1 shows how GAZE simulates a four-way round-table meeting by placing a 2D image (or *personification*) of each participant around a table in a virtual room, at a position that would otherwise be held by that remote participant. Using this technique, each person is presented with a unique view of each remote participant, and that view emanates from a distinct location in space. Each personification rotates around its own x and y axes in 3D space, according to where the corresponding participant looks. When person A looks at person B, B sees A's personification turn to face her. When A looks at person C, B sees A's personification turn towards C. This metaphorically conveys whom each participant is listening or speaking to [5].

When a participant looks at the shared table, a lightspot is projected onto the surface of the table, in line with her personification's orientation. The color of this lightspot is identical to the color of her personification. This "miner's helmet" metaphor enables a participant to see exactly where the others are looking within the shared workspace. With their mouse, participants can put document icons on the table representing shared files. Whenever a participant looks at a document icon or within the associated file, her lightspot will be projected onto that document icon. This allows people to use deictic references for referring to documents (e.g., "Here, look at these notes"). Shared documents are opened by double clicking their icon on the table. When a document is opened, the associated file contents appears in a separate frame of the web page (see fig 1). In this frame, an editor associated with the file runs as an applet. When a participant looks within a file, all participants looking inside that file can see a lightspot with her color projected over the contents. This lightspot shows exactly what this person is reading. Again, this allows people to use deictic references for referring to objects within files (e.g., "I cannot figure this out"). We realize, that providing such information may invade the privacy of individual users. By (annoyingly) projecting their own gaze position whenever it is shared, we hope to ensure that individuals are aware their gaze position is transferred to others.

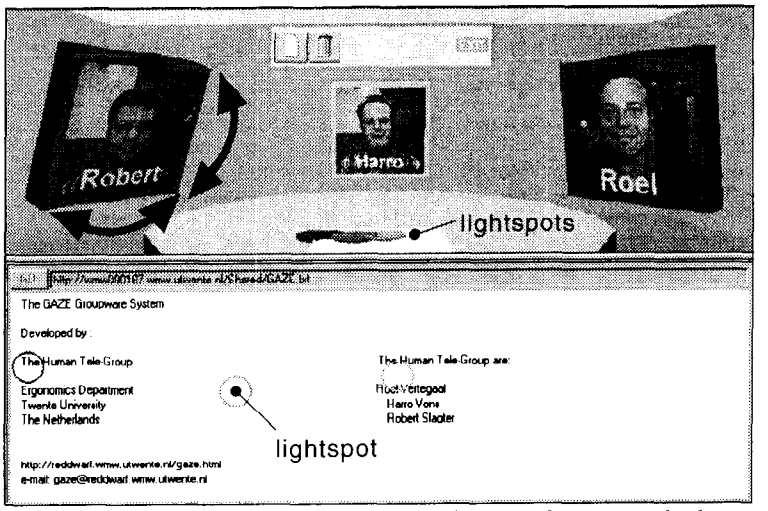

Fig 1. Personifications rotate according to where users look.

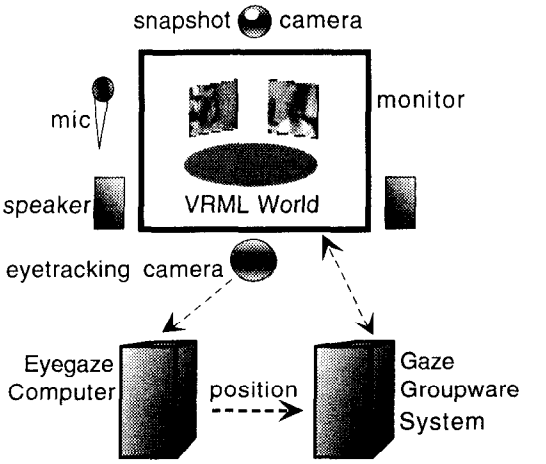

Fig 2. The GAZE Hardware Setup.

HARDWARE SETUP

Each user, seated in front of a computer screen, has a hardware setup as shown in fig 2. The GAZE Groupware System consists of two key components: the LC Technologies Eyegaze system [3], which determines where the participant looks; and the GGS computer, a Windows '95 Pentium running Netscape, the GAZE Groupware System, a web server, framegrabber software and an audio-conferencing tool. The Eyegaze system, which tracks the user's eye movement by means of a camera underneath her monitor, reports the on-screen position of the user's gaze over a serial link to the GGS computer. The GGS computer determines where the user looks, manipulates her personification and lightspot, and conveys this information through a TCP/IP connection via a server to the other GGS computers. The Eyegaze system is not required. Users can also use their mouse to indicate point of interest. The video conferencing camera on top of the monitor is currently used to make snapshots for the personification (future versions will also incorporate motion video). When making a snapshot, it is important that users look into the video conferencing camera lens, as this will allow them to achieve a sense of eye contact during meetings.

SOFTWARE IMPLEMENTATION

The GAZE Groupware System was implemented using the *Virtual Reality Modeling Language 2.0* [6]. This cross-platform standard separates 3D graphic descriptions (rendered natively) from their dynamic behaviour (running on a JAVA Virtual Machine). Sony Community Place [6] is a plug-in for Netscape which implements the VRML 2 standard and adds a client-server architecture for sharing 3D graphics and behaviour over TCP/IP. For each dynamic object a user sees in the virtual meeting room, there is a corresponding JAVA object. Whenever such an object does something, its behaviour is broadcast via the Community Place Server by means of messages to the other systems. This way, all participants' copies of the meeting room are kept in sync. Eyetracker input is obtained from a small native driver application polling the serial port or the mouse. Document editors are JAVA applets running separately from the VRML world, although they do communicate with it to obtain eyetracking data and URLs.

EVALUATION OF THE SYSTEM

Informal sessions with several hundred novice users at ACM Expo'97 indicated our approach to be a promising one. Most participants seemed to easily interpret the underlying metaphors. The eyetracking technology was, in many cases, *completely* transparent. The prototype has not yet been tested for usability. Potential usability issues include: limited to four users; conferencing software not integrated: no spatial separation of audio or visual encoding of speech activity; no option for motion video; users cannot control point-of-view; and privacy. Although the eyetracker works fine with most glasses and contact lenses, a small percentage of users has problems with calibration. The eyetracker works well while talking, but head motion is still limited to about 5-10 cm in each direction.

CONCLUSIONS

Conveying gaze direction may solve a usability problem of multiparty video mediated systems: knowing who is talking or listening to whom. The GAZE Groupware System is a multiparty mediated system which conveys the gaze direction of participants. The system measures directly where each participant looks using an advanced, desk-mounted eyetracker. It represents this information metaphorically in a 3D virtual meeting room and within shared documents.

ACKNOWLEDGEMENTS

We thank Nancy and Dixon Cleveland for their support.

REFERENCES

1. Argyle, M. and Cook, M. Gaze and Mutual Gaze. London: Cambridge University Press, 1976.
2. Isaacs, E. and Tang, J. What video can and can't do for collaboration: a case study. In *Proceedings of Multimedia'93.* Anaheim, CA: ACM, 1993.
3. LC Technologies. *The Eyegaze Communication System.* Fairfax, VA, 1997. *http://www.lctinc.com*
4. Sellen, A., Buxton, B., and Arnott, J. Using Spatial Cues to Improve Desktop Videoconferencing. *Proceedings of CHI'92.* Monterey, CA: ACM, 1992.
5. Slagter, R. *Visual Attention in Multiparty Communication.* Master Thesis. Twente University, The Netherlands, 1997.
6. *Sony Community Place*, 1997. *http://vs.spiw.com/vs/*
7. Vertegaal, R. Conversational Awareness in Multiparty VMC. In *Extended Abstracts of CHI'97.* Atlanta, GA: ACM, 1997, pp. 6-7.