

Establishing Remote Conversations Through Eye Contact With Physical Awareness Proxies

Baha Jabarin, James Wu, Roel Vertegaal and Lenko Grigorov

Human Media Lab
Queen's University
Kingston, ON K7L 3N6
Canada

{jabarin,wuj,roel,grigorov}@cs.queensu.ca

ABSTRACT

We present a mechanism for initiating mediated conversations through eye contact. An eyePHONE is a physical proxy of a remote individual that senses and conveys attention using an eye tracking device and a pair of actuated eyeballs. Users may initiate calls by jointly looking at each other's eyePHONE. We discuss how this allows participants to implement some of the basic social rules of face-to-face conversations in mediated conversations.

Keywords Attentive User Interfaces, Mediated Communication, Eye Tracking, Alternative Input.

INTRODUCTION

This paper discusses how we might initiate remote conversations using nonverbal regulatory cues commonly found in face-to-face conversation. Mediated systems typically require callers to interrupt remote individuals before engaging into conversation. While previous research focused on solving this problem by providing awareness cues about the other person's availability for communication, there has been little work on supporting the negotiation of availability that typically precedes communication in face-to-face situations. Face-to-face interactions provide a rich selection of verbal and non-verbal cues that allow potential interlocutors to negotiate the availability of their attention with great subtlety. For example, when person A is in his office cubicle working behind his computer, there are a number of ways in which person B may request a conversation without interfering with A's activities. Firstly, person B may position himself such that person A becomes peripherally aware of his presence. Secondly, person B may use proximity, movement, gaze or touch to capture A's attention without using verbal interruption. The use of nonverbal visual cues by person B allows person A to finish his task before acknowledging B's request for attention,

e.g., by seeking eye contact. When A does not provide acknowledgement, B may choose to withdraw his request by moving out of A's visual field. Indeed, Frolich [1] found that initiators of conversations often wait for visual cues of attention, in particular, the establishment of eye contact, before launching into their conversation during unplanned face-to-face encounters. In this paper, we describe how we might augment a telephone with the above capabilities. We discuss a prototype phone augmented with an eye tracking device and a pair of actuated eyeballs that allow the sensing and display of a remote individual's attention before initiating conversations.

PREVIOUS WORK

The negotiation of interlocutor attention through nonverbal cues is typically not supported by mediated systems [6]. Media spaces [5] address the problem of providing casual awareness of the activities and presence of remote individuals. For example, the availability of continuous virtual windows into a remote environment encourages spontaneous communication, albeit at a cost of privacy [5]. In Montage [3], such privacy issues were addressed by allowing only momentary and reciprocal video glances into remote offices. Peepholes [2] further mitigated privacy concerns by using only iconic presence indicators. Similarly, physical awareness proxies [4] convey a remote user's availability using the rich but abstract representation of a tangible interface. The present work extends work on physical awareness proxies by allowing proxies to not just convey the availability of a remote user for communications, but sense eye contact.

eyePHONE: AN ATTENTIVE TELEPHONE

Figure 1 shows a prototype eyePHONE. Instead of a telephone, there is a physical surrogate of each participant at each location. This proxy consists of a pair of Styrofoam eyes, actuated by a motorized Sony EVI-D30 camera. The eyes can rotate 180° horizontally and 80° vertically around their base. When a user looks at the eyeballs on the proxy, this is detected using a small eye tracking device mounted on top. This eye

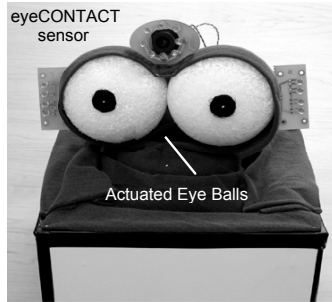


Figure 1. eyePHONE proxy with eye contact sensor.

contact sensor consists of a video camera with a set of infrared LEDs mounted on-axis with the camera lens. Another set of LEDs is mounted off-axis. By syncing the LEDs with the camera clock, a bright and dark pupil effect is produced in alternate fields of each video frame [7]. An algorithm finds any pupils in front of the camera by subtracting the even and odd fields of each video frame. The LEDs also produce a reflection from the cornea of the eyes. These glints appear near the center of the detected pupils when the user is looking at the proxy, allowing the sensor to detect “eye contact” without calibration. Once it has located the pupils of a user, the proxy can maintain eye contact by adjusting the orientation of the eyeballs such that pupils stay centered within the eye tracker image. Audio communication between eyePHONES is established through a Netmeeting Voice-over-IP connection.

Using Eye Gestures to Establish Communication

To communicate the negotiation of mutual attention we developed a set of gestures for eyePHONES, shown in Figure 2. The following scenario illustrates how users may gradually negotiate connections through these eye gestures. Connor wishes to place a call to Alex. He looks at Alex’s proxy, which begins setting up a voice connection after a user-configurable threshold of 1.5 s of prolonged eye contact. The proxy communicates that it is busy by iteratively glancing up – and looking back at Connor (see Figure 2b). On the other side of the line, Connor’s proxy starts moving its virtual eyeballs, and uses the eye contact sensor to find the pupils of Alex (see Figure 2a). Alex observes the activity of Connor’s proxy on his desk, and starts looking at the proxy’s eye balls. When Connor’s proxy detects eye contact with Alex, the eyePHONES establish a voice connection (see Figure 2c). If Alex does not want to take the call, he either ignores the proxy or looks away after having made brief eye contact. Alex’s proxy on Connor’s desk conveys Alex’s unavailability by shaking its eyes, breaking eye contact, and not establishing a voice connection (see Figure 2d). If Connor decides his call is too urgent, he may choose to press a button that produces an audible ring. Optionally, calls may be set to complete automatically when proxies determine

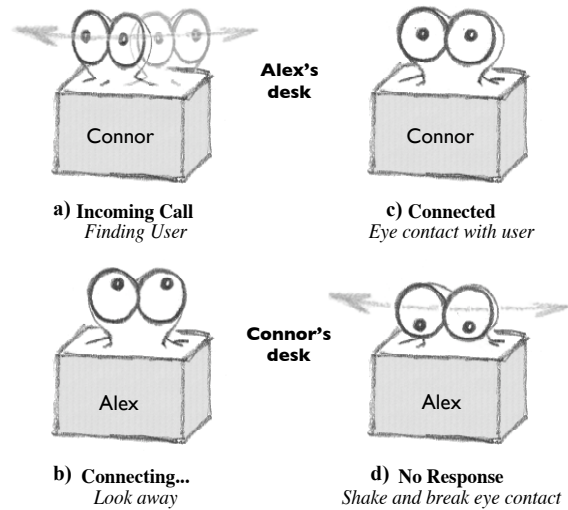


Figure 2. eyePHONE gestures.

a lack of eye contact over a user-configurable time period. We are currently experimenting with the use of eyePHONES for representing multiple participants during conference calls. Unlike regular conference calls, the negotiation of connections using nonverbal cues allows group members to enter at different times without interrupting the meeting. Furthermore, we implemented a “Cocktail Party” feature that may facilitate the establishment of side conversations. When this is active, the speaker volume of a person’s proxy depends on the amount of eye contact received from that person.

CONCLUSIONS

We presented eyePHONE, an attentive telephone that allows users to negotiate the initiation of communications through simulated eye contact. The phone acts as a physical proxy of a remote individual. Using an eye tracking device and a pair of actuated eyeballs it senses and conveys availability for communications. Users may initiate calls by jointly looking at each other’s eyePHONE. This facilitates the use of basic social rules for initiating conversations.

REFERENCES

1. Frolich, D., et al. (1994). Informal Workplace Communication: What is It Like and How Might We Support It? HP Tech. Report.
2. Greenberg, S. (1996). Peepholes: Low Cost Awareness of One’s Community, In Extended Abstracts of CHI’96.
3. J. C. Tang, et al. (1994). Supporting Distributed Groups With a Montage of Lightweight Interactions. In Proceedings of CSCW’94.
4. Kuzuoka, H. and Greenberg, S. (1999). Mediating Awareness and Communication through Digital but Physical Surrogates. In Video Proceedings of CHI’99.
5. Mackay, W.E. (1999). Media Spaces: Environments for Multimedia Interaction. In M. Beaudouin-Lafon (Ed.), Computer-Supported Cooperative Work, Trends in Software Series.
6. Vertegaal R. (1999). The GAZE Groupware System: Mediating Joint Attention in M-P Communication and Collaboration. In Proceedings of CHI’99.
7. Vertegaal, R., et al. (2002). Designing Attentive Cell Phones Using Wearable EyeContact Sensors. In Extended Abstracts of CHI’02.