GColl: A Flexible Videoconferencing Environment for Group-to-Group Interaction

Petr Slovák, Pavel Troubil, and Petr Holub

Masaryk University and CESNET z.s.p.o., Czech Republic
{*slovak, pavel, hopet*}@ics.muni.cz

Abstract. In this paper, we present GColl, a group-to-group videoconferencing environment concept, which aims to provide a natural communication channel even for ad-hoc groups or other teams that require frequent changes in the number of participants or videoconferencing locations. GColl supports mutual gaze as well as partial gaze awareness for all participants while still retaining very modest technical requirements: a camera and an echo-canceling microphone at each site; and a notebook with two USB cameras for each user. A working prototype is available for download.

1 Introduction

In the last few years, videoconferencing environments and tools came once again into the focus of human-computer interaction researchers. While videoconferencing between individuals has become quite common (e.g., Skype), attention is now turned to videoconferencing environments that would support collaboration among remote *groups* of people. In this setting, problems known from previous research (e.g., preservation of gaze awareness and other non-verbal cues) become even more salient and creation of specialized systems is therefore necessary.

Such systems might be used by distributed teams both in academic and commercial sphere. Although some of these teams might be stable enough to settle for a fixed system in a dedicated room, others require the possibility of frequent changes in the number of attendants and videoconferencing locations. Moreover, some might even be created only for very short-term tasks, meeting just a few times altogether. Throughout the rest of this paper, we will denote groups with these needs as *ad-hoc groups.*

To support ad-hoc groups, our goal was to create a videoconferencing environment, which would be a compromise between the need for preserving the non-verbal cues and the requirement of low-cost and flexibility. Our resulting system, GColl, supports *mutual gaze* (knowledge, that someone else is looking at you) as well as *partial gaze awareness* (knowing general direction someone else is looking) for all participants. At the same time, it requires only a camera and an echo-canceling microphone at each remote site, and a common laptop with two USB cameras for each user. Users can join or leave GColl without any changes to the physical layout at any site; this can be done even during a videoconference meeting.

Related Work Concept of gaze awareness has been studied in great detail in the literature due to its importance for effective communication as well as other group activities (e.g., in [7, 3]). In most videoconferencing environments, gaze information is not conveyed easily due to the usual discrepancy between the camera position and the visualization of the other person's eyes. Several multi-person videoconferencing systems were invented to mediate some or all forms of gaze awareness (e.g., GAZE2 [8], MAJIC [5], and Hydra [6]). The most recent one, Multiview [4], even achieved a full spatial faithfulness and reported very positive user evaluations. None of these are, however, directly suitable for ad-hoc groups due to either the lack of support for group-to-group interaction ([8, 5, 6]), or problems with mobility and flexibility ([4]).

In all of these systems, a change of one site (e.g., an increase in the number of participants) calls for non-trivial changes of the physical layout at remote sites. Also, a change of site location is not easy due to the need of transferring a complex structure of cameras and viewing screens. Once set up, these designs provide the users with excellent communication environment; however, ad-hoc groups might be forced to spend a lot of energy on re-installment of the system.

GColl mediates a slightly restricted variant of mutual gaze and partial gaze awareness than, e.g., Multiview. On the other hand, we believe the restriction to be very well counter-weighted by the highest mobility and flexibility of GColl among all similar systems known to the authors.

2 Environment Description

2.1 Basic Structure

Support of mutual gaze was one of our main priorities during the design phase of GColl. To fulfill this requirement, the videoconferencing system has to supply each user with her own visualization device as otherwise those users, who are sharing a common device, are in a sense sharing the "same eyes"—either all of them experience a feeling of eye-contact with a remote participant, or none does.

We could not use designs based on several projectors or monitors as this would severely limit flexibility and mobility of GColl. Instead, we opted for assumption that each user has a personal notebook or other computer at her disposal, which, as we believe, is not too restrictive for the anticipated class of possible users.

A basic structure diagram of a GColl installment at a site with three users is depicted in Figure 1. An image of the whole group is recorded by a single camera G. An echo-canceling microphone is used to capture audio at this site. Additionally, each user uses
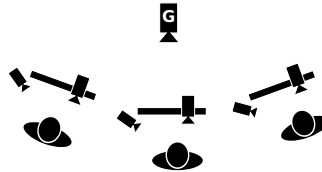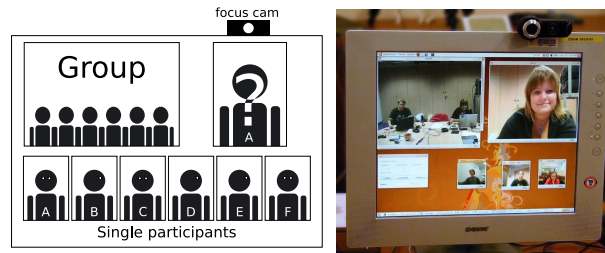

Fig. 1. A scheme of site installation.

Fig. 2. A scheme of the visualization layout and a photo of its prototype implementation (with two remote sites)

her notebook, two common web cameras and a headset with a microphone: the *focus camera* is attached to the top of the notebook screen itself while the *side camera* is placed on the desk next to the notebook. If available, a shared screen can be used for slides or any other material the group needs.

A scheme of GColl visualization screen and a photo of our prototype implementation is shown in Figure 2. The screen is divided into three parts: the top left image provides a "whole room" view of the remote group (or splits into subframes if more than two groups are communicating); images of other participants (remote as well as co-located) are shown on the bottom of the screen and after clicking on any of these small images, the chosen stream is then shown in higher quality in the top right (*focus*) frame. A click on the *whole group* frame itself unselects the chosen stream (and restores it to the original size) once it is not needed.

2.2 Functionality

The *focus* frame allows each participant to easily select some other user, on whom she wants to currently focus, and see him in more detail. The system informs other users whether or not they are "in focus" of each of the other participants by selecting the appropriate videostreams to be shown in the bottom (or top right) frames of their visualization screen: those users, who are not selected, receive the *side camera* image; the selected user receives the stream from the *focus camera*. For example, the user whose screen is depicted in Figure 2 knows that she has been selected by the participants A, C and F, and that she is herself focusing on the participant A.

This setup allows us to mediate a sensation of mutual gaze. It is based on findings by Chen [2] that show asymmetry in human sensitivity to eye-contact—that is, people would still perceive eye-contact if the other person's gaze is directed less than 5° below the camera. By having a fixed place for the *focus* window in GColl, it is easy to attach the *focus camera* appropriately close to the visualized eyes. Thus, if you are selected by another participant and that participant looks at you in his *focus* frame, you will perceive eye-contact as can be seen on Figure 2.

3 Conclusions and Future Work

In this paper, we have presented the GColl system for group-to-group collaboration with mutual gaze support. GColl is designed to be usable even for ad-hoc communicating groups due to its flexibility and modest technical requirements. Fully functional (yet still prototype) implementation is available at [1].

Currently, we are analyzing data gathered in a usability study comprising more than 90 people and the preliminary results seem to be quite promising. A real world case study is being run simultaneously, involving working teams from several universities and a national research organization.

To bring GColl closer to face-to-face communication, we are now exploring the possibilities of using USB cameras streams for eye-tracking. If we can overcome an inaccuracy of to this gaze-estimation method, it might be used for user selection instead of the currently used mouse/keyboard interface. We believe that this method could be more natural and therefore ease users' interaction with GColl. Once we have this modification ready, we plan to conduct another user study to verify our hypothesis.

GColl currently supports group-to-group communication, but has no inherent shared workplace or any other means of task support. We presume that such extensions of the design are possible (e.g., a shared tabletop at each site) and would enhance the collaboration experience greatly.

References

1. `http://www.sourceforge.net/projects/gcoll`.
2. M. Chen. Leveraging the asymmetric sensitivity of eye contact for videoconference. In *CHI '02*, pages 49–56, New York, NY, USA, 2002. ACM.
3. A. Monk and C. Gale. A Look Is Worth a Thousand Words: Full Gaze Awareness in Video-Mediated Conversation. *Discourse Processes*, 33(3):257–278, 2002.
4. D. T. Nguyen and J. Canny. Multiview: improving trust in group video conferencing through spatial faithfulness. In *CHI '07*, pages 1465–1474, New York, NY, USA, 2007. ACM.
5. K.-I. Okada, F. Maeda, Y. Ichikawaa, and Y. Matsushita. Multiparty videoconferencing at virtual social distance: Majic design. In *CSCW '94: Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pages 385–393, New York, NY, USA, 1994. ACM.
6. A. Sellen, B. Buxton, and J. Arnott. Using spatial cues to improve videoconferencing. In *CHI '92*, pages 651–652, New York, NY, USA, 1992. ACM.
7. R. Vertegaal, G. van der Veer, and H. Vons. Effects of gaze on multiparty mediated communication. In *Graphics Interface*, pages 95–102. Morgan Kaufmann, 2000.
8. R. Vertegaal, I. Weevers, C. Sohn, and C. Cheung. Gaze-2: conveying eye contact in group video conferencing using eye-controlled camera direction. In *CHI '03*, pages 521–528, New York, NY, USA, 2003. ACM.