TO SEE OR NOT TO SEE:
The role of visibility and awareness in videoconference-mediated teamwork

This paper will be published in the Proceedings of ICCCBE-VIII: 8thConference on Computing in Building and Civil Engineering, Stanford University, Stanford, CA on August 14-17, 2000

HUMBERTO CAVALLIN
College of Environmental Design, University of California Berkeley

RENATE FRUCHTER
Ph.D., Member of ASCE, Director of Project Based Learning Laboratory, Department of Civil and Environmental Engineering, Stanford University

W. MIKE MARTIN
Ph.D., AIA, College of Environmental Design, University of California Berkeley

INTRODUCTION

The recent developments in telecommunications and the spreading of the Internet culture, have made closer the possibility of having communicational resources as the once envisioned by science fiction, not too far ago. The use of video in telecommunication is already a real possibility not only for corporations and the aero-spatial industry, but also for any person with not necessarily very sophisticated equipments and a connection to the Internet. These changes facilitate collaboration and project development in geographically distributed environments to turn into common activities.

However, many aspects that are present in these new kinds of environments and interactions have not been fully studied. That is the case of the role of visibility and awareness in the computer-mediated communication process, in particular during the transmission of information and the resolution of tasks in geographically distributed collaborative environments. The role of video in interactive human communication and the possible impact of video-mediated

interaction on the social processes associated with groups working especially on informal communication was pioneered by Chapanis (1975). His study pointed out the relevance of the different communication channels (e.g., text, audio, video) as a function of the context of the teamwork and goal of the meeting. Recent studies (Heath et al., 1991, Veinott et al., 1999) have shown that video could impact areas that were not considered before by Human-Computer Interaction (HCI) research, as the social value of the images in aspects as negotiation, and non-verbal communication.

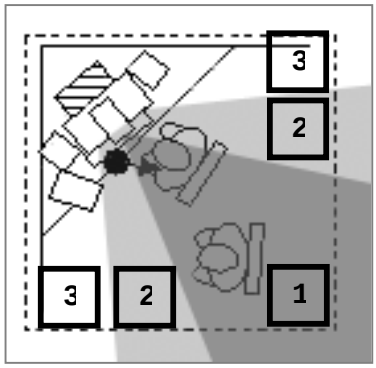
In order to know more about how this relationship between the social interaction aspect and the role of the video impacts the computer mediated interaction in the Architecture/Engineering/Construction (A/E/C) collaboration, we analyzed the protocols of interaction produced by graduate students of those three fields, enrolled in the course "Computer Integrated Architecture, Engineering and Construction (A/E/C)", organized by the Department of Civil and Environmental Engineering, at Stanford University (Fruchter 1999). This paper presents preliminary results of an ethnographic study that focuses on the definition of communication protocols and the role of visibility and awareness in videoconference-mediated teamwork, particularly in the use of computers as a communication tool by professionals in the fields of Architecture, Engineering, and Construction management (A/E/C).

METHODOLOGY

As it was mentioned before, the study was performed in the context of the "Computer Integrated A/E/C" course offered at Stanford University. The purpose of this course was to expose the students to a problem based learning (PBL) experience. The emphasis is on teamwork. Each A/E/C team works on a building project. During their design process, the participants used Internet based communication technologies that enable them to interact in geographically distributed shared environments. These students came from Stanford University, University of California at Berkeley, and Georgia Tech, in the United States, Strathclyde University in Glasgow, LjubljanaTechnical

University in Slovenia, and Aoyama Gakuin University in Japan. They were organized in five A/E/C teams, whose members interacted remotely with each other from their countries of origin. Each A/E/C team had an owner who had a budget, a site, and a program for the building. In addition, 13 professors and practitioners acted as mentors in the three different areas of expertise.

The data was captured on a Video-8 system. The camera was placed in a position that enabled to capture both the participants in the meetings, and the screen of the computer used for establishing the communication through the Internet. About 40 hours of interactions were recorded and at the end, authors reviewed the videos. The videos were transcribed using notation systems designed for keeping track of the verbal discourse, and of the non-verbal activities as gestures, movements, use of the space, and use of the tools.



CONE OF INTERACTION (COI):

Form the observations we made of the interactions, one important aspect in understanding part of the behavioral patterns of the participants was the study of the workspace used in the interactions. It was reduced in general to the area surrounding the computer, creating a restricted space for the equipments required for the interaction. The affordances of the equipments used also determine the way in which the space was used by the participants. Particularly, two aspects have shown to be relevant to the way in which people use the workspace: the locations of the Monitor and the Video camera.

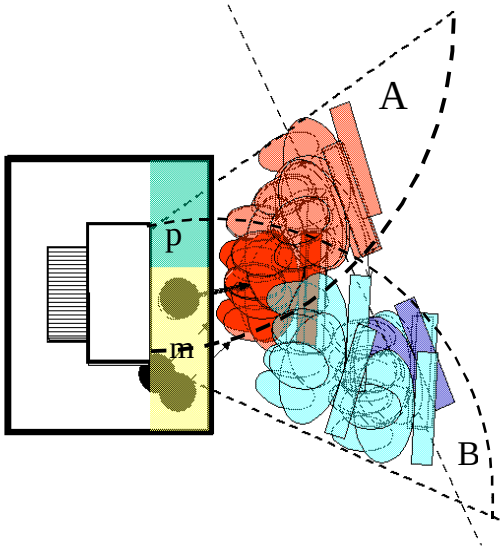
Both Monitors and Video cameras define preferred locations for participants that narrow the possibilities for using a certain area in the working space. Using the previous example, we can define a triangular area as the one shown in the picture on the right, where two different areas can be identified: the

lighter one corresponds to the area of visibility for the monitor, and the dark one to the angle of capture of the camera.

When analyzing the movements of the participants in relation to the location of the equipment, we can see that the movements are restricted to a triangular area, that we have called Cone of Interaction (COI).

The drawing on the right shows the movements of two participants during a real interaction. Four major areas can be identified in this COI:

1. Command area (A): Is the area in which the person that leads the interaction locates. The position has to do most likely with the use of the input device, and it is all the way around in the case of left-handed users.
2. Secondary area (B): this is the area occupied by default by the other person or people involved in the interaction.
3. Pointing device area (p)
4. Microphone area (m)



Some aspects have to be considered in relation to the COI. On one hand, the overlapping of functional areas created by the video camera and the monitor. This overlapping creates three zones as can be seen in the next graphic: the sector (1) defines the area in which the user of the computer can have visibility of the screen and also be captured by the lens of the video camera; the sector (3) is the area in which no visibility can happen, both for the user of the computer and for the receiver of the image captured by the video camera. However, the sector (2) is the potentially most problematic of all, because when being in this area the user of the system can be having visibility of the computer screen, but at the same time be out of the camera range without noticing it, creating a visual contact failure in the communicational process.

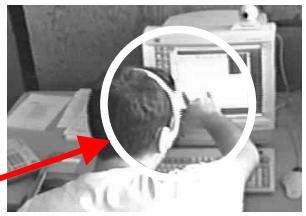
FAKED POINTING

The COI contributes in the production of a false sense of awareness in the participant of the video interaction, by creating the wrong belief that by being in the visibility range of the screen, the actions performed will be transmitted to the non collocated participant.

An example of this wrong sense of awareness is an action that happens when participants are describing information that involves graphical information is the one that we have called Faked pointing (**Fp**). The **Fp** could be considered a failure situation, because it can conduce to misunderstanding and delays in the communicational process. However, due to its particular characteristics we have decided to analyze it separately.

Consider the following segment extracted from an interaction:

Let me tell you this
Look this up
If I don't give you a scale very soon
I will give you the size of one of those rooms
So you can work out the scale off
of one of those ratios
But, I think that this is printed at a
scale of 1/32 of an inch in a foot
But yeah, the the the…long side is a 116 feet
Um…or actually, um…it's actually 116 feet
those little offices uh…
I'm going to take control for just a
second
Uh…ok, right here where the mouse is
These small offices
These are the student offices and uh…they are
10 feet by 6 feet



The pictures on the right of the text show the gestures of the speaker while describing the graphics on the screen. These gestures, that would be meaningful in a face-to-face interaction, are completely lost in the computer-mediated communication because they fall out of the range of vision of the video camera. What the speaker meant by *"long side"* and *"those little offices"* most probably was unknown by the receiver of the information. The speaker himself notices after a certain point that the information he is aiming

to transmit has not been conveyed by his gestures and explanations, and decides at the end of the segment to take control of the mouse and use it as a pointer for indicating the object of his description.

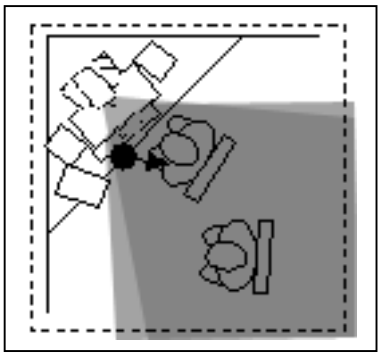
Situations like this show the misled perception in the speaker that he/she is seen by the hearer as in the case of a face-to-face interaction. The use of body gestures and its use for conveying the discourse are drastically constrained by the affordances of the video devices in use, and the lack of awareness of this fact by the speaker can lead to sometime important losses in the communicational process.

SOME FINAL REMARKS

The role of visibility and awareness in videoconference-mediated geographically distributed HCI, in particular for the case of the studied A/E/C collaboration seems to be relevant in situations in which the participants are involved in dialogues that imply the exchange of visual information or social issues, because of the social value that the visual conveyed by the use of video adds to the interaction and in the solving of tasks is highly relevant (Veinott et alt., 1999, Rocco, 1998).

A better understanding of the role of visibility and awareness in videoconference-mediated teamwork by studying the practices that are involved in the communicational process have shown to be useful in identifying situations in which the lack of that awareness can lead to miscommunication situations.

It is important to deepen into the role of this media in order to improve both the physical elements and the socials involved in the communication. Applications of the knowledge range from a better understanding of the role of the visual communication in geographically distributed interactions, to recommendations for the development of communication protocols to support efficient collaborative work when in remote meetings, and desired hardware and software functionality.



A simple example of how by introducing small changes in the work environment situations as the one observed can be by making congruent the limits of the area of visibility of the screen with that of capture of the video camera. This can be done in several ways, as by modifying the capture configuration of the capture angle of the camera (if available) or locating the camera in a position as slightly behind the monitor, so both angles can overlap as much as possible (see graphic on the right). The second factor to consider has to do with the designing of the workspace. The space has to provide space not only for the computer (CPU, monitor, etc.) but also for activities as writing or the use of other equipments as tablets or digitalizers, all inside of the COI if possible.

Using shared pointing devices instead of pointing out to the things on the screen is a good advise. However, the real solution to these situations will come by the improvement of the hardware and the expansion of the communicational devices as the Ishii's Clear Board that enables participants to see each other using the metaphor of "talking through and drawing on a big transparent glass board" (http://www.media.mit.edu/~ishii/CB.html).

However, a solution like this one does not necessarily apply to cases in which more than two non-collocated participants have to share information that requires visibility of the others. These multiple users situations should contemplate both the option of having more than one speaker on any of the sides, and more than two sides in the course of the meeting.

Requirements like these point out to the development in the near future of new creative hardware options that will enable to tailor the interfaces to the demands of the natural users' activities.

ACKNOWLEDGEMENTS

The study was partially sponsored by the Project Based Learning Laboratory, at Stanford University, and a grant from MITI through a joint project between Aoyama Gakuin University, Tokyo, Japan, and Stanford University. The authors would like to thank Christina Cho, undergraduate student at Stanford University for her invaluable contribution in processing the information. Finally, the authors thank the 6th A/E/C student generation who participated in the A/E/C course in Winter and Spring of 1999.

REFERENCES

Chapanis, A. (1975) Interactive human communication. Scientific American, 232, P. 36-42

Cooper, G., Hine, Ch., Rachel, J., and Woolgar, S. (1995) Ethnography and human-computer interaction. In Peter J. Thomas (Ed.) The social and interactional dimensions of human-computer interfaces. Cambridge: Press Syndicate and the University of Cambridge.

Finn, K.E., Sellen, A.J. and Wilbur, S.B. (Eds.) (1997) Video-mediated communication, Hillsdale, NJ: Lawrence Erlbaum Associates.

Fruchter, R. "Architecture/Engineering/Construction Teamwork: A Collaborative Design and Learning Space," Journal of Computing in Civil Engineering, October 1999, Vol 13 No.4, pp 261-270.

Gallager, J., Kraut, R.E. and Egido, C. (Eds.) (1990) Intellectual Teamwork: the social and technological foundations of cooperative work, Hillsdale, N.J.: Lawrence Erlbaum Associates

Goodwin, Ch. (1981) Conversational Organization: interaction between speakers and hearers, New York: Academic Press

Heath, Ch. And Luff, P. (1991) Disembodied conduct: communication through video in a multi-media office environment. In Human factors in computing systems conference proceedings on Reaching through technology, P. 99-103

Veinott, E. and Xiaolan, F. (1999) Video helps remote work: speakers who need to negotiate common ground benefit from seeing each other. In Proceedings of the Conference on Computer Human Interaction (CHI'99), Pittsburgh, PA, P. 302-309

Vera, H.A., Thomas, K., West, R.L. and Lai, S. (1998) Expertise, Collaboration and Bandwidth. In Proceedings of the Conference on Computer Human Interaction (CHI'98), Los Angles, CA, P. 503-510