in *Proceedings of the International Conference on Multimedia Computing and Systems (ICMCS)*, May 1994

Casual Collaboration

Judith S. Donath
MIT Media Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139
donath@media.mit.edu

Abstract

As computers are increasingly used to mediate social interaction, tools are needed not only to support direct communication, but also to create a richer social environment for the networked group. These include tools that provide information about the presence of others and that give the user a sense of the ongoing activities. This paper describes two ongoing research projects in this area. Collaboration-at-a-Glance is a visual interface to an electronic group. It creates an animated image of a remote meeting by compositing pictures of the participants so as to represent their actions. Visual Who is a window onto a larger community. Like a window onto a street, it shows the comings and goings of the community members.

1: Visualizing electronic presence

William Whyte, describing what makes for a successful urban environment, said: "What attracts people most, in sum, is other people. If I labor the point, it is because many urban spaces are designed as though the opposite were true and as though what people liked best are the places they stay away from. "

Whyte went on to describe a number of places that attract people to them: they are places in which there are other people – sometimes familiar, often not. The key is the ongoing presence of others, of an environment that is vigorous and populated, one that has a constant flow of human activity. Whyte's book is an argument against the design of spaces that do not draw people to them, against the bland corporate landscapes and the sterile plazas, against places that may perform their stated function efficiently, but that fail in their community role.

The same principle – that the presence of other people is the key to a vital urban environment – is likely to hold true for electronic communities as well. Yet relatively little attention has been given to building the infrastructure to support this type of environment. Unlike the direct, task-oriented activities which are easy to describe (such as shared editing windows, or group calendars), the presence-oriented activities are less obvious and less well understood. Yet to a large extent, the future success of online communities depends on how well the tools for this activity are designed. If they are poorly designed, the online world may feel like a vast concrete corporate plaza, with a few sterile benches: a place people hurry through on their way to work or home. If the tools are well designed, the online world will not only be inhabited, but will be able to support a wide range of interactions and relationships, from close collaboration to casual people watching.

Casual collaboration. Computers are increasingly used for social tasks, such as reading email and working cooperatively. For collaborative work, a variety of shared workspaces have been created, most of which emphasize task performance. Yet a social environment does not consist solely of direct communication and purposeful collaboration. Knowing who else is around, sensing the level of activity, and generally being aware of the presence of others is an important part of participating in a social milieu.

Casual collaboration – the chance discussion at the coffee maker, the suggestions of colleagues who happen to pass by – is an essential and highly productive part of the work experience [9][3]. In a well designed work environment, these chance encounters occur frequently. They are facilitated by easy access to one's co-workers – not formal meetings, but proximity and awareness of presence. The two projects described in this paper bring this awareness and the opportunity for chance encounters to people for whom proximity is sharing a network connection, rather than a physical space.

2: Collaboration-at-a-Glance

Lindsey is working at her computer – editing a screenplay, reading the news. Among the windows on her screen are several which show groups of faces, turned toward each other as if in conversation. Occasionally, there is movement in one of these windows – a head turns to face a different person. At one point, one of the window grows quite active. Many of the faces in it turn first towards one person, then to another; below the heads, a text window

The work reported herein is supported by the Movies of the Future consortium,
including Bellcore, Intel Corporation and Viacom International.

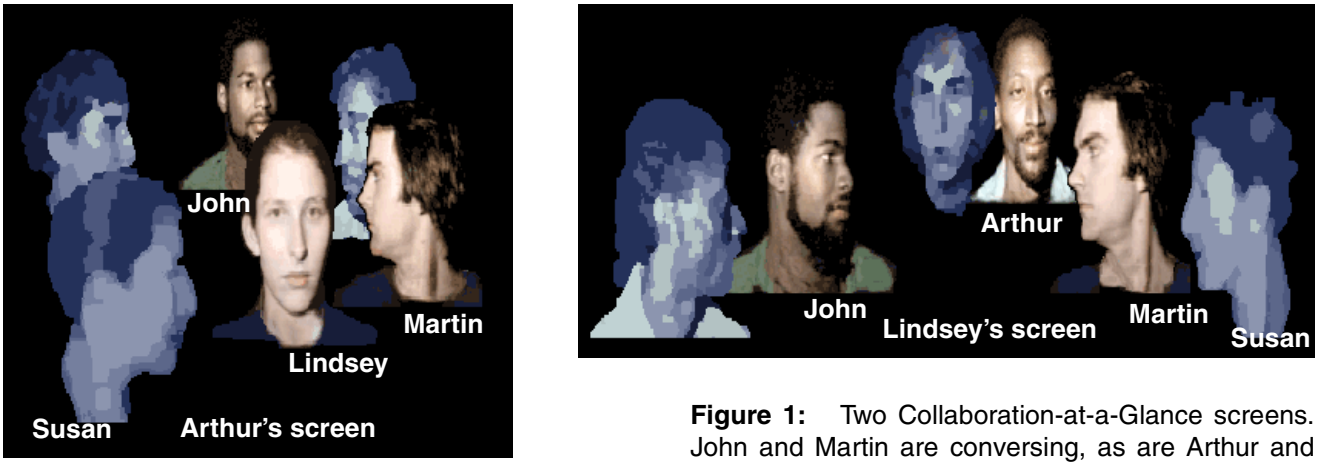

Figure 1: Two Collaboration-at-a-Glance screens. John and Martin are conversing, as are Arthur and Lindsey.

fills with messages. Lindsey is curious: this window is the marketing group with whom she works closely. The discussion, it turns out, is about a proposal to stage surprise bicycle stunts in shopping malls to promote their new feature – an idea she thinks is ludicrous. She clicks on the face of the idea's main proponent, Arthur, and types her objections.

On forty screens scattered across the country, in the window showing the marketing group, Lindsey's head turns to face Arthur's and the forty other participants in the argument read her remarks.

There is no picture of Arthur on Arthur's screen. Instead, he sees the picture of Lindsey looking straight out at him. He responds to her comments – and on all the screens, Arthur turns to face Lindsey (and on her screen, this means he looks directly out at her).

Meanwhile, the discussion continues, sometimes in public, sometimes privately. Martin, who is new to the company, asks for his friend John's opinion before he ventures a suggestion. Their conversation is a private aside within the group – they are seen conversing, but the contents of their notes are not included in the general text, appearing only on each other's screens.

Lindsey goes back to her editing. She's curious to hear what Susan, the producer of the film, will say about the proposed stunts. But Susan is not around. Her image appears as a stylized drawing, which means that she has her window set to record, and may review the discussion later, but is not actually present to participate.

2.1: Social visualization

The Collaboration-at-a-Glance window on Lindsey's screen provides her with a casual connection to her coworkers. Not only can she quickly see who is around, she can also see when and where an interesting conversation is taking place. The participants in the conversation know to whom they are talking – in contrast with many email and news reading systems, those who are just listening are also visible.

The bandwidth requirements of Collaboration-at-a-Glance are extremely low: no images are sent across the network, only data about the state of the group. The pictures that the participants see are synthesized locally; they are a visualization of the data about the group's interactions. Yet Collaboration-at-a-Glance is not simply a low-bandwidth substitute for video conferencing. If limited bandwidth were not an issue – if one could have live video images of all of one's co-workers running simultaneously – Collaboration-at-a-Glance would not be redundant. Collaboration-at-a-Glance creates a simple movie of an unfilmable event: a meeting among widely separated people. The coherent 3D space inhabited by these images shows the interactions between participants in a way that individual video windows cannot.

Representation	Meaning
Gaze direction	Attention & communication
Image style	Presence / absence
Location	Viewer preference
Image features	Physical appearance

Table 1: Collaboration-at-a-Glance representations

Collaboration-at-a-Glance maps abstract relationships and states of being to concrete visual representations. Some mappings are intuitively obvious, others are arbitrary; some show a range of values along a continuum,

others the presence or absence of a particular quality. The choice of mappings is the fundamental issue in designing the visualization.

Representing presence The mapping from presence to image style is open to wider interpretation. Photographic images appear much more present and life-like than stylized drawings: using the photographs to indicate an active user is not an arbitrary choice. The meaning of the drawings, however, is application dependent. They can be simple placeholders for absent group members, or they can indicate passive listeners in an otherwise active group (auditors in a virtual classroom, audience members in an electronic panel discussion, software agents sent out to record online conversations).

Many collaborative interfaces create a viewer-centric universe, placing images of the participants in individual windows, each shown facing the user. Collaboration-at-a-Glance uses gaze to provide information. The screen usually shows people facing in a variety of directions. Only if you actually are the center of attention do you see everyone facing you.

2.2: Controlling the representation

The real-world pedestrian controls his own appearance. I may think that his outfit poorly chosen, but I cannot improve it through force of will. The virtual window through which I see my online co-workers is, in theory, capable of letting me make such adjustments to other people's apparel. Beyond reasons of ethics and etiquette (important, but outside the scope of this paper) there are communication reasons why this should not be done: it trades customization for information. Instead of hearing what others have to say, I have chosen to talk to myself.

Yet customization – being able to filter information according to one's interests, to adjust the virtual window to show particular highlights of the view – is quite useful. The goal is to find the right balance, to adjust the view of the scene and not its contents.

The use of location. As it is currently implemented, the user determines the location of people in the Collaboration-at-a-Glance space. The layout thus represents the user's view of the group – whose image he wishes to see easily, what subdivisions he makes in the larger group.

For some applications it might be more useful to have the participant's actions rather than the viewer's preference control the participant's location on the viewer's screen. For example, if a single Collaboration window is used to represent several sub-discussions (threads in a newsgroup, virtual rooms in an audio conference) a participant could move from discussion to discussion. As he moved in "topic space" his image on all screens would change location to indicate his current involvement.

2.3: Applications

Collaboration-at-a-Glance currently includes the simple note sending system described in the scenario above. It also includes live-video windows: one can choose to see, in the place of a participant's image, live video of that person (given that the subject has the necessary video capture hardware).

Future work on the project will include integrating it with other communication applications. One potential project is to use it to visualize message patterns within a large project that generates a lot of email. It would be used as a real-time window onto the group and as a way of reviewing a large backlog of mail, allowing one to watch patterns of activity while replaying the message actions. Collaboration-at-a-Glance could also serve as a front end for a MUD – shared communities, usually text-based, which users log into and then move around in a common space, encountering other users. Another possible application is to use Collaboration-at-a-Glance as an interface to a remote audio conference. Such conferences suffer from voice confusion: it can be hard to tell who is speaking and difficult to address remarks to a particular individual. Collaboration-at-a-Glance provides a visual interface that can indicate who is speaking (perhaps here linking image style with source sound level) and that would allow users to face particular people to indicate direct address.

3: Visual Who

Collaboration-at-a-Glance shows the interactions between a small, known group of people. Its analog in the real world is a meeting, or the casual commentary that passes between people working in the same room. Visual Who is a tool for visualizing a larger group of electronically connected people. It provides a screen window that, like a real window onto a busy street, lets the viewer watch the comings and goings of the populace.

Patterns of activity. Media-Lab is the Media Laboratory's main "social" computer – the one used for checking email, finding phone numbers, etc. There are often well over a hundred people logged into this machine. However, the usual terminal screen connected to this bustling hub looks like any other screen: a prompt, some text and a cursor. It looks like that at four in the morning when only a few late-working students are active; it looks exactly the same at eleven a.m. when many people are busy sending mail, checking announcements, and reading the news. The Visual Who window, on the other hand, reflects the ongoing activities. It displays the names of the current logins, fading them with increasing idle times. At four a.m. the Visual Who window is dark, with only an occasion bright spot. At mid-morning, the window is bright and animated.

Figure 2: The Visual Who window. Names appear brighter to show recent activity.

Visual Who displays a lot of brief, casual activity – people active for a minute or two as they read their mail, or look for a phone number. It is to a large extent patterns such as these that make it seem so life-like. "The city is still the prime place. It is so because of the great likelihood of unplanned, informal encounters or the staging of them."[15] The key idea is that participation in Visual Who needs to be incidental – it must be a by-product of other, more deliberate activities.

Conceptual map. Visual Who provides a conceptual map of the community, where the different areas of the screen are neighborhoods of interest. In the current version, these areas indicate research group affiliation. The color of a name indicates community role – staff, faculty, graduate student, etc. (At four a.m., the few names that are still lit are almost invariably green, the color of graduate researchers, and an occasional blue faculty name, logging in from a visit to Japan. Morning comes with a flurry of magenta names – the arrival of the administrative staff.)

Work is currently in progress to make this map adaptive, letting the viewer define the neighborhoods, and the program determine who would be found where. To do this, Visual Who uses the alias file – the place where people sign up for mailing lists, ranging from the administrative lists enumerating all the faculty, or all the students, to the self-chosen lists, such as the ones for hardware designers, vision researchers, or bicycling enthusiasts. People are motivated to add themselves to various lists because they wish to be involved in the discussions; the effect is that everyone creates a simple public profile of their role in the community and their current interests. By choosing which mailing lists are to be the focus of a neighborhood, Visual Who user can create a view of the community that is adapted to his interests.

The color of names is another dimension for indicating community role. In general, the spatial layout is left constant: the viewer grows accustomed to the layout of the community, able to quickly determine whether a particular person is present. Color is useful for highlighting temporarily desired information. As a default it can indicate, for example, job title, and be easily set to show in bright red letters particular features of the community, such as who may be able to fix a broken file server or who can answer a question about signal processing.

Future work. Visual Who is still in the development stage – the near future work is a full implementation of the ideas described in this paper. After that, the work will move in two directions. One is to use faces rather than names – to explore the use of graphical images to visualize a large community. The second is to add sound, using sampled voices to create an auditory representation of activity.

4: A final word

Like the design of physical spaces for incarnate populations, the design of interfaces for electronic communities requires technical knowledge, an understanding of how people interact with each other and with their surroundings, and the design mastery to create a space that is inviting, intriguing and functional. Our understanding of how electronic communities will function is still very limited, as is our understanding of which representation of people will make their presence truly felt and what information one needs to know about a person one meets online in order to gain a sense of their personality. Yet, far more than architecture shapes the character of a city, the design of online communities will shape the way they function – for the design determines the actions and appearance of the inhabitants.

5: Implementation details

Collaboration-at-a-Glance is a networked client-server system with extremely low bandwidth requirements. No images are sent across the network, just data about the status of the group – who is logged in, who is communicating with whom. Each client synthesizes an appropriate group picture based on the status data and its own layout design.

Server-client communication. The server maintains information about the state of the group. All information is routed through the server, which forwards private messages to individuals and broadcasts status updates to all the participants. Whenever a participant does something that is to be shown on the other users' screens (logs out, sends a message to another client, etc.) the server broadcasts a message to all the clients, which then update their screen image.

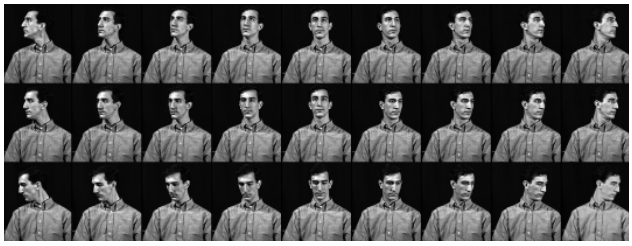

Figure 3: A set of user images.

The set of messages that need to be supported is quite small. The minimal set of messages is: login/logout <name>, look-at <name>, and request-download. The user does not send these commands directly; they are invoked by the application that is running Collaboration-at-a-Glance as a visual interface.

In the current implementation, a simple note sending capability is integrated into the system. Sending a note automatically invokes the look-at message. A note sent from a to b is routed through the server. The contents of the note are sent to b, and a message saying that a is communicating with b is sent to all the currently logged-in group members. The individual clients then update their screens to show *a* looking at *b*.

The image database. A set of images looking in predetermined directions is needed for each group member. In the first implementation of the project, 5 images were used per person, sufficient for small groups with simple layouts. A more recent version uses 28 images per set: the finer angular resolution makes the focus of attention clearer, which is important as the groups get larger.

The stylized cartoons used in the current system were made by altering the photographs using a paint program. Automatic cartoon generation would be quite useful, especially for larger groups. We have been doing some work in this area, expanding the technique described in [8] to create color images.

The client uses three pieces of information to synthesize a view: the 3D location of all the participants (according to that client's layout); the 3D location of the viewer; and the current state of the group – who is present and where are they looking. Knowing these things, the client determines from what angle the viewer would see each person, chooses the frame whose gaze angle matches most closely for each person and composits the final image.

Very low bandwidth & platform independent. The network bandwidth requirements of Collaboration-at-a-Glance are extremely low. Only the system messages and names need to be send exchanged; no images are sent over the network. Because the bandwidth is so low the group size can grow to be fairly large: size is limited by how big a group can be clearly displayed, not by network bottlenecks. Furthermore, a user can have multiple Collaboration windows running at the same time, each one a window onto a different group.

No special network hardware is required. One can participate from any machine that has access to the image database (which is itself not large: the 7-frame image sets are a little over 100K each). Since only the system messages are sent between clients, and the details of accepting input and compositing the images are left up to the clients, the clients themselves can be implemented on a wide range of platforms. Other participant's system choices are invisible to the users. Collaboration-at-a-Glance is currently implemented on several Unix machines running X-Windows; clients could be written for other platforms without difficulty.

Visual Who is implemented as an X-Windows client on a Unix platform. It obtains information about current logins and idle times by reading the utmp file.

References

See [2][3] and [4] for other work on casual electronic collaboration. See [9] and [10] for a thorough discussion of the sociological aspects of electronic communication. [7] and [13] cover issues in the creation of synthesized movies. Regarding faces: see [12] and [14] for various aspects of facial modelling, [8] on automatic cartoon generation, and [5] for an introduction to the perception of pictures of faces. [11] is an excellent exposition on issues in the visualization of information. See [6] for a study of how people make highly individual mental maps of real neighborhoods.[1] describes the client-server software that was used to implement Collaboration-at-a-Glance, including the live video windows.

[1] Abramson, Nathan. "The Dtype Library or, How to Write a Server in Less Time than it Takes to Read this Manual.". MIT Media Lab, June 1992.

[2] Ahuja, Sudhir R. and Ensor, J. Robert. "Coordination and control of multimedia conferencing". In *IEEE Communications Magazine*, Volume. 30, Number 5. May 1992. pp 38-43.

[3] Dourish, Paul and Bly, Sara. "Portholes: Supporting Awareness in a Distributed Work Group". Proceedings of ACM Conference on Human Factors in Computer Systems, CHI '92, Monterey, CA.

[4] Fish, Robert S., Kraut, Robert E. and Root, Robert W. "Video as a Technology for Informal Communication". In *Communications of the ACM,* Vol. 36. Jan. 1993, pp. 48-61.

[5] Hochberg, Julian. "The Representation of Things and People.", in *Art, Perception and Reality,* pp. 47-94. Baltimore: The Johns Hopkins University Press, 1972.

[6] Milgram, Stanley. *The Individual in a Social World*. Reading, MA: Addison-Wesley, 1977.

[7] McLean, Patrick Campbell."Structured video coding." Master's thesis, MIT, 1991.

[8] Pearson, Don E.and Robinson, John A. "Visual Communication at Very Low Data Rates," in *Proceedings of the IEEE,* vol... 73., no. 4, pp.795-812, 1985.

[9] Sproull, Lee and Kiesler, Sara. *Connections: New Ways of Working in the Networked Organization.* Cambridge, MA: MIT Press, 1991.
[10] Sproull, Lee and Faraj, Samer. "Atheism, sex, and databases: The net as a social technology". 1993. In *Public Access to the Internet* (in press). Brian Kahin and James Keller (eds.)
[11] Tufte, Edward R. *Envisioning Information*. Cheshire, CT: Graphics Press, 1990.
[12] Waters, Keith and Terzopoulos, Demetri. "The Computer Synthesis of Expressive Faces", in *Philosophical Transactions of the Royal Society.* Volume 335, B, 1992, pp 87-93.
[13] Watlington, John A. "Synthetic Movies." Master's Thesis, MIT, 1989.
[14] Welsh, W.J., Searby, S., and Brigant, E. "Model-based coding of videophone images." (1990).
[15] Whyte, William. *City: Rediscovering the Center.* New York: Doubleday, 1988.