Judging People's Availability for Interaction from Video Snapshots

Brad Johnson
Dept of Educational Psychology
The University of Calgary
Calgary, Alberta CANADA T2N 1N4
Tel: +1-403-220-5910
email: bfjohnso@acs.ucalgary.ca

Saul Greenberg
Dept of Computer Science
The University of Calgary
Calgary, Alberta CANADA T2N 1N4
+1-403-220-6087
email: saul@cpsc.ucalgary.ca

ABSTRACT
Several groupware systems support casual real time interaction over distance by providing periodically updated snapshots of other people's offices. People then monitor these snapshots to determine how available others are for communication. In this research, we try to isolate what information people use from these snapshots to help them infer another's availability. Research participants examined video snapshots of people posed in typical office situations, and judged how available those people were for interaction. Our first result suggests that people have difficulty extracting information from these images unless their resolution was at least 128x128 pixels. Our second result indicates that people interpret stereotypic situations as indicating varying degrees of availability. In general, people are judged as less available when they are seen to be absent from their office, or in conversation with others. People are judged more available when they are in transition (e.g., entering or leaving a room), and when they are not working. People at work seem to portray a more ambiguous situation. However, all situations had a minority of people who interpreted the image quite differently.

KEYWORDS
Real time groupware, casual interaction, awareness, contact facilitation.

INTRODUCTION
Casual real time interaction is an essential ingredient of group cohesiveness. Yet the bottleneck to rich spontaneous interaction is distance (Kraut, Egido and Galegher 1988), and users of wide area networks will be at a disadvantage unless a prosthesis that overcomes distance barriers is available. Consequently, many groupware researchers are designing interfaces that facilitate how distance-separated people can establish real-time contact with one another across computers. A crucial component of these systems is that they provide information that helps people stay aware of who is around in their community, whether those people are available for conversation, and whether it is socially acceptable to initiate a conversation with them (Cockburn and Greenberg 1993).

Media spaces, for example, let people observe both offices and public areas through continuous audio/video channels. Through these, people can see who is around and what others are doing at remote sites (Abel 1990). Alternatively, *video glimpses* give short video-only views into one's offices (Tang, Isaacs and Rua 1994), while *video snapshots* provide periodically updated still images of other people's offices (Dourish and Bly 1992). *Iconic indicators* are a different approach, where hints of availability are provided through an icon that indicates a person's recent activity on their computers (Greenberg 1996).

While demonstrably useful for supporting casual interaction, these techniques are also fraught with problems. These include privacy violations, excessive requirements on technology (such as video cameras), bandwidth costs, scaling to large communities, and so on. We believe that some of these issues come about because we do not know precisely what information people require for contact facilitation. Consequently, we either put too much information on the channel (resulting in problems with privacy, bandwidth, and/or scaling), or too little (resulting in inappropriate contacts and/or lost opportunities).

In this research, we try to isolate what information people use to decide whether another person is both available and interruptible. We want to articulate the types of information used, and to determine the weight that people place on these types of information.

A FIRST STEP IN STUDYING CONTACT FACILITATION
Dourish and Bly (1992) suggested that video snapshots were reasonably successful for contact facilitation. A set of small video snapshots taken from cameras in people's offices and updated every minute or so suffices to give people a sense of who is around and what they are doing. Its disadvantage is that screen real estate is consumed, privacy violations are easy, and that it relies on video cameras. In order to understand if alternate strategies can provide the same information in a more concise and secure form, we

decided to determine what information people found useful in these snapshots.

Intuitively, we expect that people's judgement of other's availability will rely on several attributes. Some of these are part of the video image (e.g., whether a person is actually present). Others are part of a culture (e.g., whether it is polite to interrupt at this particular moment); part of the history (e.g., inter-personal relationships, past responses); and part of the current needs (e.g., level of urgency). We consider the first of these in this paper: to discover the essential awareness information contained within a video image. Our specific goals follow.

1. *Determine the critical threshold of image resolution necessary to convey the cues people required to make availability decisions.* The visual cues people use to make availability decisions may be difficult to extract if image resolution is reduced beyond some threshold (see Figure 1). Consequently, decisions or determinations regarding availability will deteriorate to guesswork.
2. *Determine how people judge availability from a set of typical poses.* The typical ways people work in their offices can be loosely categorized e.g., whether people are present or absent in their offices, whether they are entering or leaving the office, whether or not they appear to be hard at work, whether they are talking to others. We captured these situations as image poses (see Figure 2), and we hypothesize that people's judgement of availability will vary with these poses.

METHODS

Participants

Research participants were sixty-five university people, all with some computer experience.

Materials

Ten poses each of two male and two female actors were photographed for a total of 40 photographs. Each actor was photographed in their natural workplace. Photo resolution was 640 x 480 pixels and 16 million colors.

Figure 2 illustrates the ten categories of poses of one actor. Other actors had similar poses, and Figure 3 shows an example of how each actor appeared in a single pose. Photos were then digitally reduced to lesser resolutions: 16x16, 32x32, 64x64, 128x128, 256x256, and 512x512 pixels (Figure 1).

Custom software presented screen sequences to participants. Each screen contained a particular image (which depended on the research condition). Participants used a 7 point Likert scale to indicate how available the person in the image was, from 1 (least available) to 7 (most available).

Figure 3. The four actors in the eating pose.

Procedure

Participants were randomly placed into one of six conditions corresponding to the six photo sizes. Participants then completed the two tasks below.

Task 1. Using the software mentioned above, participants looked at and rated individual images according to how available they believed the person in the picture was. A short scenario was given to provide context for the task.

Task 2. Afterwards, participants were given the four sets of high-resolution printed pictures corresponding to the four actors, and asked to sort each set into a sequential order from most available to least available.

DESIGN AND RESULTS

Critical threshold of image resolution

Design. We measured the *basis of agreement* in task 1 between any two people regarding whether or not the person in a particular image is available. Using a basis of agreement assumes that people will generally base their decisions on similar cues or cue sets, and that different people would sort similar cue sets into similar positions or rankings. To determine the critical threshold of image resolution necessary for people to interpret awareness cues, we expect that agreement approaches chance below a particular threshold, with more agreement as image quality improves above it. The key is not necessarily high levels of agreement, because people may interpret similar images differently. Rather, we expect a statistically significant difference between agreement levels across this threshold.The design was one-way ANOVA, where the independent variable was image resolution and the dependent variable was agreement, as scored by the Kappa coefficient (K). The Kappa statistic, designed to measure agreement between a number of raters on a set of criteria, corrects for chance agreement. Kappa values may take ranges from 0 (no agreement) to 1 (perfect agreement). Thus any amount greater than 0 is a measure of agreement, however weak that agreement may be. Participant scores were randomly paired in each condition and a Kappa coefficient generated. There were 5 pairs in each cell.

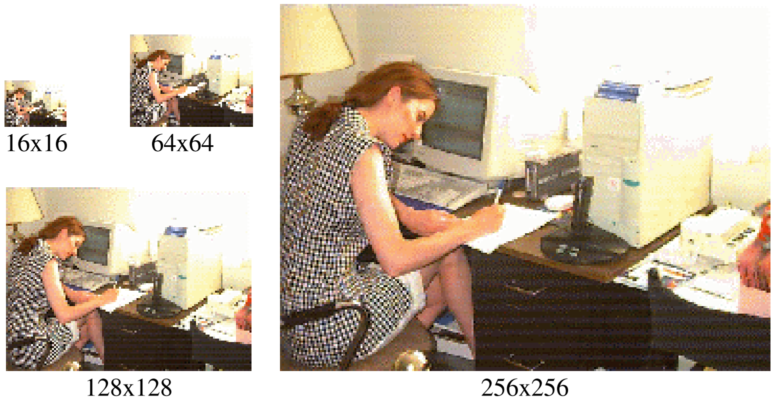

Figure 1. Various image resolutions of one actor working at their desk. The 32x32 and 512x512 images are not show. This printed version likely differs in quality from the ones used in the experiment.

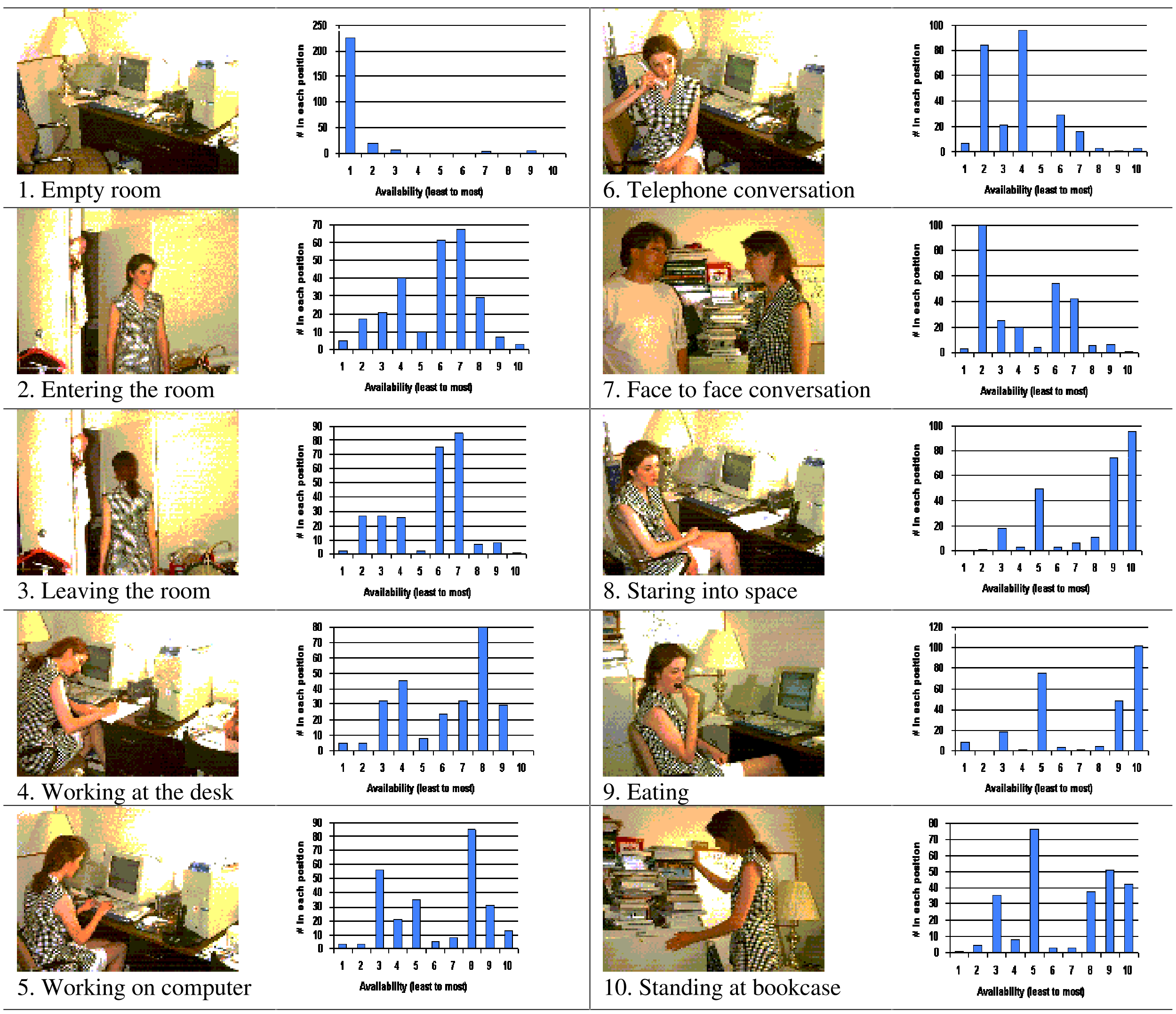

Figure 2. The ten poses, photographed over one actor (printed here at low resolution). Accompanying histograms show pooled data for that pose by all actors for the sort performed in task 2. The X axis is the sorted position, from least available (1) to most available (10). The Y axis plots the actual number of images found in each position.

Results. Mean Kappa scores are plotted in Figure 4 for each pose. Scores indicated only modest agreement. A significant difference was found between groups $F(5,24)=6.389$, $p<.001$. Post-hoc test comparisons showed that the difference occurred between the 64x64 and 128x128 conditions.

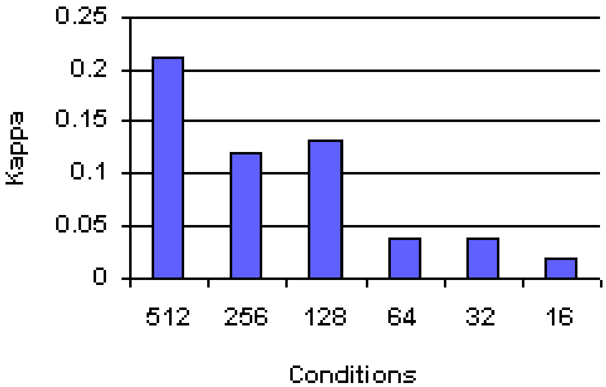

Figure 4: Mean Kappa scores. There is a significant difference between the 64x64 and 128x128 condition.

Differences between poses

Design. In the second task, each person sorted the sets of poses (grouped by actor) from most to least available. Every image thus had a position in the sorted pile, from 1 (least available) to 10 (most available). We plotted the pooled data for each pose as a frequency histogram, and we counted how frequently a pose appeared in particular position. We then examined the histograms qualitatively for trends.[1]

Results. The histograms for the pooled data by pose are illustrated in Figure 2 next to its pose type (although the figure shows only one actor in that pose). The X axis shows the sorted position, from least available (1) to most available (10). The actual number of images found in each position is shown on the Y axis.

DISCUSSION

Critical threshold of image resolution.

When picture resolution dropped past some threshold, people found it difficult to make determinations regarding availability. From our results, it appears as if the threshold lies somewhere above 64x64 pixels and below or at 128x128. As we would expect, people's availability cues depend upon sufficient visual information being presented. These cues may be obscured when image quality is degraded. The practical implication is that systems using a snapshot approach for displaying awareness information should use an image size of at least 128x128 pixels. We should note, however, that the expected image quality in a real situation should be considered as well, as poor lighting conditions and camera placement may compromise the cues that should be visible in an image of this resolution.

Differences between poses

Several poses indicated definite trends on how people ranked their availability. The vast majority (87%) rated the "empty room" (Figure 2, pose 1) as 1 (least available), with 97% scoring it 3 or less. Images involving telephone conversations (pose 6) were also rated as generally unavailable, where 80% judged them as 4 or less. On the other extreme, people generally rated the "staring into space" image (pose 8) as most available, where 65% scored it as a 9 or above. Surprisingly, 27% scored it as modestly unavailable (between 3-5), indicating a dichotomy of how people rated that pose.

This dichotomy in judgement rankings appears in varying degrees throughout the rest of the poses. For example, the way people interpret face to face conversation (pose 7) is similar to how they see telephone conversations: about half thought it indicated unavailability (a score of 3 or less), while 36% thought it indicated modest availability (a score of 6-7). The act of entering or leaving a room (poses 2 and 3) also tended towards availability (60% scored entering between 6-8, while 61% scored leaving between 6-7), although as seen in Figure 2 about 40% of the people scored these poses as four or less. People's judgements on each of the other poses was split almost equally, again in a bimodal distribution. About half thought the "working" images (poses 4 and 5: at desk and at computer) indicated modest unavailability, with the other half interpreting it as quite available. Many thought that "eating" and "standing at bookcase" (poses 9 and 10) indicated high availability, although again a few though this reflected unavailability.

These results suggest that although there may be trends on how people judge availability in these images, different people may interpreting the same pose differently depending upon their social understandings. For example, while a majority thought that it was not appropriate to interrupt people in conversation, a few thought that this was acceptable (e.g., some could view the conversation in pose 7 as an informal one and thus interruptible). Similarly, activity transitions, such as when a person enters or leaves a room, were flagged as an indication of availability by most, but not all people. The same applies to idle activities, such as standing by the bookcase, staring into space, and eating. However, work activities, whether behind a desk or computer, generates a mixed response. Some thought it appropriate to interrupt that person, but others felt they should not disturb that person.

[1] We also produced histograms of data pooled by pose and by actor, to see if the actual actor had any effect on how people judged poses for availability. Differences, while present, were slight enough to allow us to consider the data pooled across all actors.

CONCLUSIONS AND IMPLICATIONS

This study provided some initial data points in terms of how people tend to view availability.

First, resolution is a factor when images are used to present awareness information. As a preliminary guideline, the image size should be at least 128x128 pixels. This assumes that the image was taken in reasonable conditions.

Second, people do interpret stereotypic situations as indicating varying degrees of availability. In general, people are less available when they are seen to be absent from their offices, and in conversation with others. They are more available when they are in transition (e.g., entering or leaving a room), and when they appear to be not working. People at work seem to portray a more ambiguous situation. However, these statements are not absolutes, as all situations had a minority of people who interpreted the image quite differently. As a guideline, other interface devices, such as iconic indicators (Greenberg 1996), must be careful on how they capture and display these stereotypic situations. A straight representation of a particular activity may provide enough information to allow others to estimate awareness (e.g., that a person is present and active on their computer). Alternatively, certain situations can be detected and the degree of availability inferred by the system and portrayed on the interface. However, this is risky as different people could have interpreted the same baseline information differently from the system's interpretation.

REFERENCES

Abel M. (1990). Experiences in an exploratory distributed organization. In J. Galegher, R. E. Kraut, and C. Egido (Eds.), Intellectual Teamwork: Social Foundations of Cooperative Work, pp. 489-510. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Cockburn, A. and Greenberg, S. (1993). Making contact: Getting the group communicating with groupware. Proc ACM Conference on Organizational Computing Systems, ACM Press.

Dourish P. and Bly S. (1992). Portholes: Supporting awareness in a distributed work group. Proc ACM Conference on Human Factors in Computing Systems, pp. 541-547, ACM Press.

Greenberg, S.(1996) Peepholes: Low Cost Awareness of One's Community. ACM SIGCHI'96 Conference on Human Factors in Computing System, Companion Proceedings, p206-207, ACM Press.

Kraut, R., Egido, C. and Galegher, J. (1988) "Patterns of contact and communication in scientific collaboration." In Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW '88), pp. 1-12, ACM Press.

Tang, J.C., Isaacs, E., and Rua, M. (1994). Supporting Distributed Groups with a Montage of Lightweight Interactions. In Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW '94), p23-34, ACM Press.