Supporting Coherence with a 3D Instant Messenger Visualization

Carman Neustaedter & Saul Greenberg
University of Calgary
Department of Computer Science
Calgary, AB, T2N 1N4 Canada
+1 403 220-6087
[carman *or* saul]@cpsc.ucalgary.ca

ABSTRACT
Instant messengers have become a popular medium for providing awareness of others and supporting casual interaction. To smoothly move into and out of computer-mediated conversation, coherence is necessary not only as a means to represent conversations, but also to afford an awareness of who is around and if they are available for interaction. We have developed a peripheral visualization for an instant messenger designed to utilize people's natural cognitive abilities. Each contact is represented by pictures for each availability state (e.g. online, offline) or video snapshots embedded within a 3D environment using a space metaphor. Contacts that are more available—determined as a function of availability state and a viewer-settable interest level—are placed in the foreground and contacts less available are placed closer to a single focal point in distant space. The viewer is able to move contacts throughout the space to create a spatial mapping. Contacts that are *of interest* display conversation bubbles containing incoming messages.

Keywords. Casual interaction, computer mediated communication, information visualization, instant messaging, peripheral awareness, space metaphor.

INTRODUCTION
Supporting coherence in computer-mediated conversation (CMC) involves creating a logical connection that holds people together during casual interaction and conversation. In this sense, we typically think of coherence in terms of gaining a sense of how people attend an existing conversation, or how people view as a collective the threads of past conversations.

Just as crucial, however, is providing coherence in how people view their opportunities to move into and out of communication with others [6,7]. Here, coherence refers to how *informal awareness information* about others is presented and acted upon, that is, how people efficiently gain an understanding of who is around, what others are doing, and whether or not they are available for interaction or collaboration. Informal awareness provides cues used to decide if and when to initiate a conversation with a colleague [3,6].

Neustaedter, C. & Greenberg, S. and Edwards, C. (2001) **Supporting Coherence with a 3D Instant Messenger Visualization**. *Report 2002-967-01, Department of Computer Science, University of Calgary, Alberta, Canada, November.*

http://www.cpsc.ucalgary.ca/grouplab/papers/index.html.

Our premise is that our graphical user interface must provide coherence in how it presents awareness information of others. We wanted to construct a display where information is understood at a glance and is acquired with minimal distraction.

To explore our ideas, we have developed a visualization to Microsoft's MSN Messenger system, which we call IMVis (Instant Messenger Visualization). As will be described, IMVis has several features that we believe enhances coherence.

- It appears on a second monitor, which others have shown are often used in practice as a peripheral display [4,3].
- Availability state is mapped onto several redundant cues: a series of posed photos mapped onto availability state [2], and a 3D depth and blurring cue for rapid perception of who is very available vs. who is not.
- One's interest state in another person is adjustable, where contacts can be moved to less spatially prominent positions that mute how their availability state is portrayed.
- Certain events are made more salient to attract attention, while others are made less salient. For example, as the availability of a contact changes, the contact is animated through the 3D environment to attract the user's attention and provide peripheral awareness.
- Conversations are associated with the person who generated them, where incoming messages resulting from interaction are placed next to the contact in the form of a conversation bubble.

IMVis DESCRIPTION
IMVis, pictured in Figure 1, visually represents each contact with an actual picture of the person, or with live video snapshots to offer more awareness. These pictures and videos are positioned on the circular wall of a 3D tunnel, where people appearing deeper in the tunnel (and thus smaller) are less available or of less interest to the user.

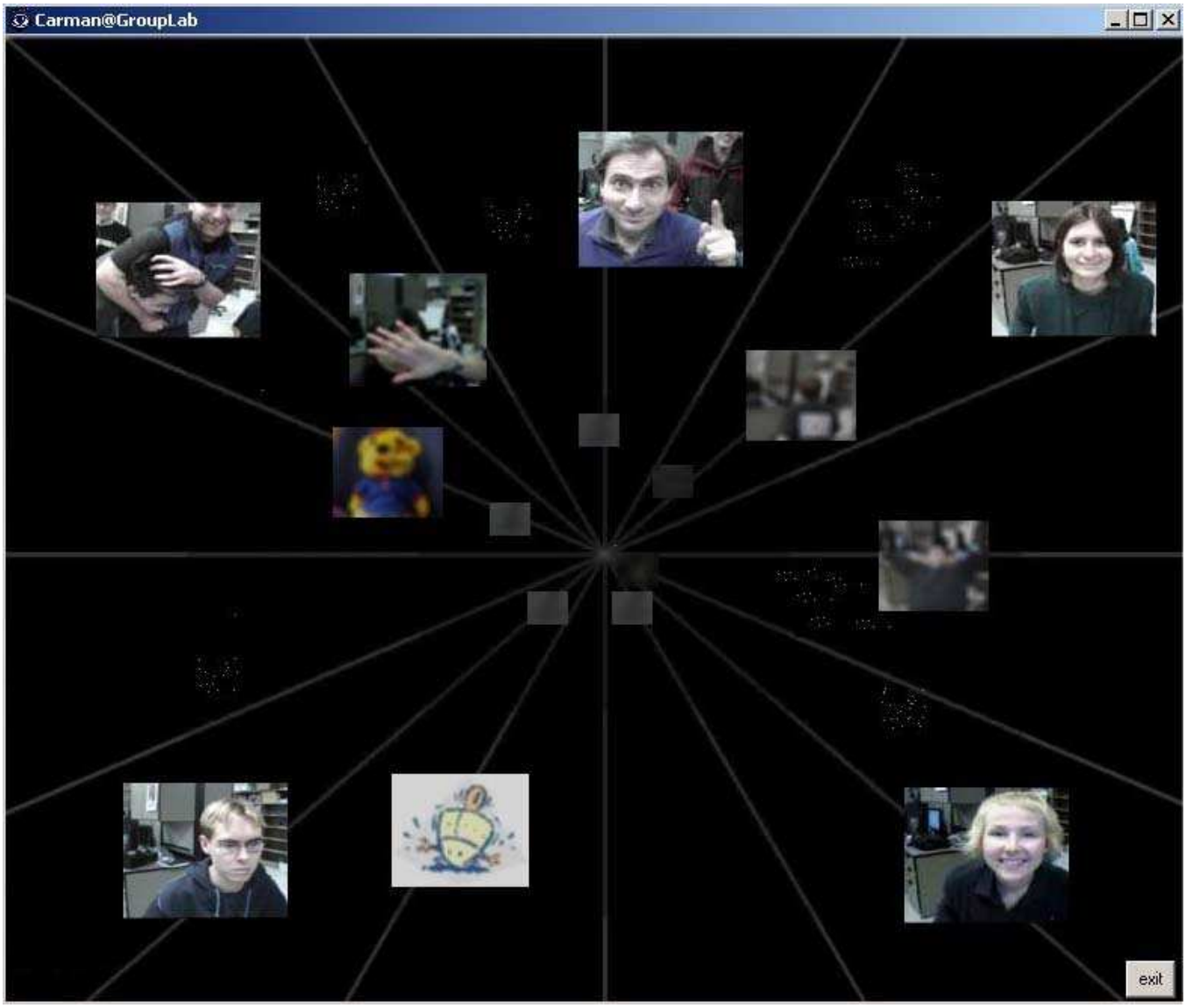

Figure 1: Instant Messenger Visualization (IMVis) showing each contact from a MSN Messenger contact list.

Contact Representation

Each person is visually displayed in IMVis as a set of photos representing their availability, or as a periodically-updated live video snapshot, or as generic clipart. Photos exploit the natural ability to recognize and remember faces and availability postures [2,5], while videos accurately capture the current state of the person. Both forms likely create a lower cognitive load than the text names typically found in instant messenger systems. However, it is the person who decides (perhaps on-the-fly) how he/she wants others to view them by right-clicking his/her own image in Figure 1 and raising the dialog shown in Figure 2 and then by selecting the appropriate radio button.

That same dialog allows people to define the set of photos that best represents their various availability states. The first row in Figure 2 shows the current snapshot used for that availability state, while the middle row shows the video image currently on their PC camera (to help them pose for a new snapshot). Unlike SideShow, people define their own poses [2][1]. For example, Carman posed himself with a 'waving hand' to indicate he is available (top left), and took the snapshot. Similarly, he posed himself as getting up from his chair (for 'Be Right Back'), with his empty chair (for 'Away'), and so on.

Alternatively, a person can opt to transmit live video snapshots—where they can adjust the update frequency—taken from their attached PC camera (Figure 2 bottom). This provides others with better and more accurate awareness information for judging availability, which in turn affords an increased level of coherence.

Next, a generic set of clipart images represent contacts who have not yet set their IMVis images, or who come through via another channel that does not make images available (such as the standard MSN Messenger system) as well as

[1] We are not yet sure if people are good at choosing their own availability pose.

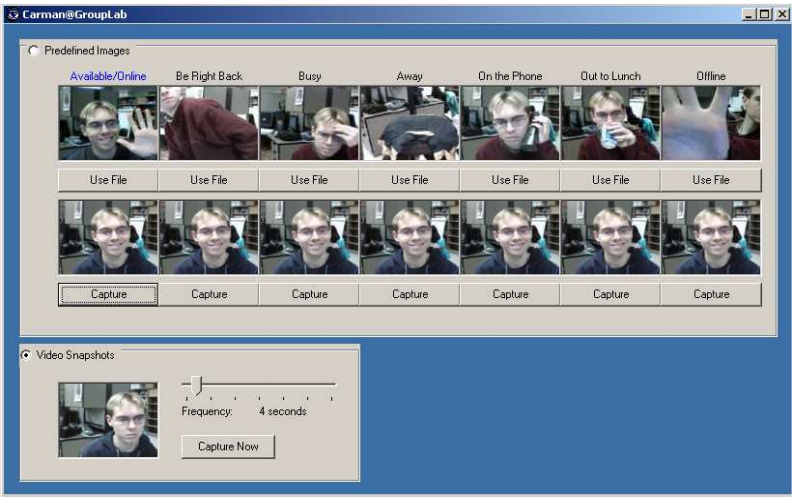

Figure 2: The pop-up dialog, where users select and adjust how they want to be represented.

those that choose to do nothing (Figure 1, lower left). This helps solve two problems common to many systems: users must spend a large amount of initial effort to gain any additional benefit, and they can only be part of a community if they are using a particular brand (and even version) of software.

Finally, people can view additional details of a contact by hovering the mouse over the image. This raises the contact's name, a textual description of their availability state, and the full-size, undistorted view of the contact's image.

3D Tunnel Metaphor

We used a 3d tunnel as a metaphor for visualizing, placing and interacting with people's image representations (Figure 1). Edges of items stick to the circular wall of the tunnel. The tunnel contains a single distant focal point positioned at the window's centre. Similar to a real tunnel, objects near the entrance (e.g, those illustrated on the outside edge of Figure 1) are clear and large. As they move back into the tunnel towards the focal point (i.e. away from the viewer), they grow smaller and fade into the background (see details in Figure 3). To emphasize the 3d nature of this tunnel, we draw perspective lines from the focal point to various points on the window's edge.

IMVis reinforces its 3d effect by using other depth cues. As contact objects move back into the tunnel, they are alpha-blended into the background, and the images are blurred somewhat (Figure 3). Both create the illusion that these objects are distant from the viewer.

There are limits to 3d. If objects are moved very far back into the tunnel, they would become quite blurry, almost solid black, and extremely small. Thus we limit how deeply objects can be moved in the tunnel, which means images cannot go below a critical size, blurring amount, and alpha blending level.

Measuring Availability

While there are many ways to measure and capture availability, we limit ourselves to the seven states currently present and captured by MSN Messenger: Online, Busy, Be Right Back, Away, On the Phone, Out to Lunch, or Offline. Indeed, our system works directly with MSN Messenger, which means that people currently on Messenger can use it immediately. We rank-ordered these availability states, following guidelines described in a previous study by Johnson and Greenberg [5]. These states are captured by MSN Messenger in the standard way, i.e., by implicitly monitoring computer activity, or by letting people explicitly set their availability states. We then intercept these states from MSN Messenger to drive our own display.

The availability is then mapped onto particular depths of objects in the tunnel. Less available contacts are positioned deeper into the tunnel, while highly available contacts are located at the tunnel's front close to the viewer (Figures 1+3). Movement between availability depths always follows a virtual linear (z-axis) track from outside to the center, which we believe aids spatial memory. For example, the image of an offline person will return back to its previous online position when they become active. This virtual track is only changed if the viewer explicitly moves a contact within the visualization (discussed shortly).

Attracting Attention

Animation of objects is used to provide a smooth cognitive

Figure 3: Using depth to represent availability.

Figure 4: A conversation bubble with an incoming MSN Messenger message.

transition as a contact changes depth. The animation of the image increases an object's visibility, where it attracts the user's peripheral attention when they are not looking directly at the window. This helps them notice changes in availability states of others. The degree of animation is asymmetric. It is relatively fast and noticeable as contacts become available, as we want people to notice these contact opportunities. However, animation is slower and much more subtle as contacts become less available to avoid excessive distraction. Finally, a sound cue is played when a contact moves online, as we feel this is an important event that should also attract the viewer's attention.

Communication

IMVis intercepts messages from a contact on MSN Messenger and visualizes them in a conversation bubble (accompanied by an audio cue) next to the contact that sent it (Figure 4). This creates a direct link between the message sender and their representation in IMVis. A standard chat dialog box also appears in a separate window (not shown). After a few seconds the bubble disappears, but the user is still able to use the chat window.

IMVis also gives a visual indication of each contact's frequency of interaction over time. As each incoming message arrives, a star is placed in a random location around the top left corner of a contact (see Figure 4, left top corner of the photo of Saul). Over time, old stars move towards the focal point, following the same linear path that the object uses. This results in a starfield that provides users with a simple and easy understanding of each contact's interaction frequency (e.g., frequent speakers have many stars) as well as conversation recency (e.g., recent conversations are closer to the edge, bursts of conversations are noticeable, etc.).

Interest

New contacts are initially placed randomly within the IMVis perimeter space, and are treated equally. However, IMVis recognizes that a viewer may not have equal interest in all contacts. To overcome this, we let viewers adjust their interest level of others, which personalizes how various people appear in the IMVis space.

First, the viewer can move contacts around the surface, which affects the linear track followed by each contact as their availability changes. Thus different people can be

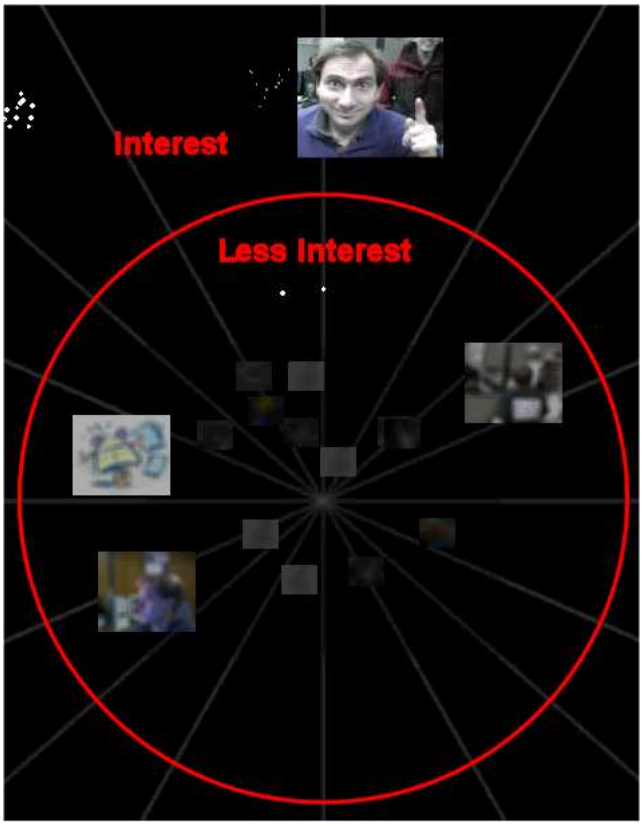

Figure 5: Contacts can be positioned within 3D to reflect the user's desired level of interest in the contact. The interest threshold is shown with the circle. Contacts within the circle do not generate conversation bubbles.

placed in different spatial tracks, and 'groups' can even be clustered close together.

Second, the viewer can set their interest level by moving a contact deeper into the tunnel e.g., if one moves a contact closer to the focal point, it will shrink, become blurred, and fade into the background. For example, contacts that are not of immediate interest may be placed at a distance further from the viewer than contacts of more interest.

This interest setting also adjusts how each contact object moves up and down the tunnel depth as its availability state changes. Specifically, the tunnel has an invisible interest threshold, which it uses to adjust an object's interest level as a person moves the object. If (say) a person moves an online contact object from the perimeter deeper into the tunnel (across the interest threshold, as marked in Figure 5), that contact is marked as being of less interest. In that case, the contact object will no longer be able to move towards the very front of the tunnel; the depth threshold becomes a barrier it cannot cross. That is, when that person goes off-line, her image will go back into the tunnel (as normal), but if her state changes to one that is more available, her image will stop before getting to the very front. This obviously makes that person less salient. These contacts also do not generate conversation bubbles for incoming messages, although their conversation is still tracked in the star field display. In the opposite case, when

the viewer moves the contact closer to the front, passing the interest threshold, he/she is marked as being of interest and thus can move all the way to the front as their status changes (Figure 5).

Occlusion

Occlusion is typically a very difficult problem to address when placing objects within an interface, whether it be two or three-dimensional. Normally there is some portion of an object that will be blocked by other objects. To try to avoid occlusion, IMVis uses a collision-detection algorithm while the user is moving contacts. During movement, if a stationary contact is in the way of the moving contact, the stationary contact will slide out of the way. Once the moving contact has passed, the stationary contact slides back to its original location, providing it is currently uninhabited. This method is similar to the active page avoidance found in Robertson et al.'s Data Mountain [8].

Occlusion can still result when objects change their availability. An object may follow a path away from or towards the focal point and can easily occlude an object along the way, or worse, stop directly in front of another object. One possible solution is to use the collision-detection algorithm as contacts change their availability. We did not include this for several reasons. First, it results in excessive animation that may distract the viewer. Second, it may force other objects off their visual track, which in turn can affect a person's spatial memory of where an object should be located.

Discussion

We began this paper by discussing the importance of coherence for how people view their opportunities to smoothly move into and out of communication with others. With this in mind, IMVis was designed 3d contact list to take advantage of several things, as listed below.

First, we wanted to exploit people's cognitive abilities to recognize people and their availability. We do this by representing people and their availability state through images and videos *vs.* text labels [2,3].

Second, we wanted to exploit people's spatial memory [8], where they can associate information with spatial location. However, we do not know how well this works for linear tracks.

Third, we wanted to exploit people's ability to attend information at their periphery. We use 3d tunnel depth as the metaphor to control this: image size, position towards the center, animation, apparent visual depth, blurring and alpha-blending to accentuate or to attenuate how noticeable a contact's image as the contact's state changes. Other 3d visualization systems exploit similar aspects of depth and spatial position [8,9].

Fourth, we wanted to give people the ability to easily adjust their interest in particular contacts. We do this by letting people move contact objects to particular positions, and to decrease/increase a contact's interest by moving them up and down the tunnel. In turn, this adjusts how noticeable a contact could become as their state changes, for it automatically restricts how close to the surface an object can move to.

IMVis is still highly experimental, and has not yet been evaluated. We are exploring it as one of several alternate ways of representing contacts and resulting conversations; one of our former systems used a collage metaphor [3], and we are currently building a version of the Data Mountain [8] for similar purposes. While we are far from any definite conclusions, we are excited by the opportunities and richness of IMVis and other metaphors to serve as a contact list visualization.

Acknowledgements. Thanks to Michael Boyle for his invaluable help and the use of his collaborative multimedia toolkit and Sheelagh Carpendale for her idea of using a space metaphor.

REFERENCES

1. Ahlberg, C., Shneiderman, B. Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. *Readings in Information Visualization: Using Vision to Think*, Morgan Kaufmann Publishers, Inc., pp. 244-250.
2. Cadiz, J.J., Venolia, G., Jancke, G. and Gupta, A.. Sideshow: Providing Peripheral Awareness of Important Information. *Microsoft Research Technical Report MSR-TR-2001-83*.
3. Greenberg, S. and Rounding, M. The Notification Collage: Posting Information to Public and Personal Displays. *Proc ACM CHI, 2001.*
4. Grudin, J. Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use. *Proc ACM CHI, 2001*
5. Johnson, B. and Greenberg, S. (1999). Judging People's Availability for Interaction from Video Snapshots. *Proc Hawaii International Conference on System Sciences, Distributed Group Support Systems Minitrack,* IEEE Press.
6. Kraut, R., Egido, C., and Galegher, J. (1988) Patterns of contact and communication in scientific observation. *Proc ACM CSCW '88.*
7. Nardi, B., Whittaker, S., Bradner, E., (2000) Interaction and Outeraction: Instant Messaging in Action. *Proc. ACM CSCW 2000.*
8. Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D., Dantzich, M., (1998) Data Mountain: Using Spatial Memory for Document Management. *Proc ACM UIST '98.*
9. Robertson, G., Dantzich, M., Robbins, D., Czerwinski, M., Hinckley, K., Risden, K., Thiel, D., Gorokhovsky, V., (2000), The Task Gallery: A 3D Window Manager. *Proc ACM CHI 2000.*
10. Whittaker, S., and O'Conaill, B. (1997), The Role of Vision in Face-to-Face and Mediated Communication, in *Video Mediated Communication*, Finn, Sellen, and Wilbur eds., LEA Press.