Chit Chat Club: Bridging Virtual and Physical Space for Social Interaction

Karrie G. Karahalios
University of Illinois
201 N. Goodwin Ave. 3110
Urbana, IL 61801
+1 217 265 6841
kkarahal@cs.uiuc.edu

Kelly Dobson
MIT Media Lab
20 Ames St. E15-020c
Cambridge, MA 02139
+1 617 253 5182
monster@media.mit.edu

ABSTRACT

In this work, we create an audio-video link via an interactive sculpture to facilitate casual, sociable communication between two remote spaces. This communication installation was designed to blend the benefits of online interaction such as low risk interaction, lower barriers to entry, and minimized geographical constraints with the ease and the affordances of interacting and signalling in physical space. We describe the creation and the iterative design process for creating a social virtual-physical hybrid space-interface we call the *Chit Chat Club*. In describing our design decisions, we note the advantages and disadvantages of two *Chit Chat Club* installations and their effect on interaction.

Categories & Subject Descriptors: H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces --- synchronous interaction, collaborative computing; H.4.3 [Information Systems Applications]: Communications Applications --- computer conferencing, teleconferencing, and videoconferencing

General Terms: Design, Human Factors

Keywords: computer mediated communication, media spaces, social catalysts, social computing

INTRODUCTION

The term media space refers to any environment created using video, audio, and networked computers to support interaction between distributed groups of people. When placed in public or semi-public spaces, they are often designed for casual encounters among people within that community. Thus far, communicating via these systems has not met expectations [6][8]. Some drawbacks to such systems have been lags in interaction time, gaze ambiguity, lack of privacy, spatial incongruity, and fear of appearing too social in a work environment [8][15].

We believe that current systems are also affected by a lack of mobility in the interaction space and a *window effect*, whereby the edges of the transmitted rectangular video window frame further emphasize the distance in the interaction.

Our approach to creating sociable media spaces is to blend the communication interface into the space of the environment through the use of a physical interface - in this case, a

**Figure 1**. The first *Chit Chat Club* Installation.

physical telepresence sculpture. This physical object acts as interface, medium, and focus for interaction. In doing so, it becomes a social catalyst for interaction [9][10]. We illustrate this approach with the iterative creation of the installed system, the *Chit Chat Club*.

RELATED WORK

There have been a number of "media space" projects that connect geographically distinct locales with some combination of audio and video [1][2][4] as well as studies of the relative affordances of audio, video, and other media [8].

One of the main goals behind the creation of the first media space project at Xerox Parc was to find means to support cross-site work and to maintain the necessary social connection between remote research labs [1]. There has also been relevant work focusing in the audio-only or audio-dominant media spaces. One such media space, the *Somewire* project, is one of the most relevant, since it was designed to foster casual interactions among colleagues [13]. Here, Singer et al experimented with a number of visual interfaces in conjunction with an audio-only media space. They found that control over such features as localization or other attributes was not needed, but that information that supplemented users' knowledge of the social aspects of the space, such as awareness of the presence of others, was quite useful.

Few audio-video media spaces have ventured into the physical realm. One early example is *Hydra*, a project that does facilitate spatial social cues including position relative to others and gaze direction. Small distinct physical modules -
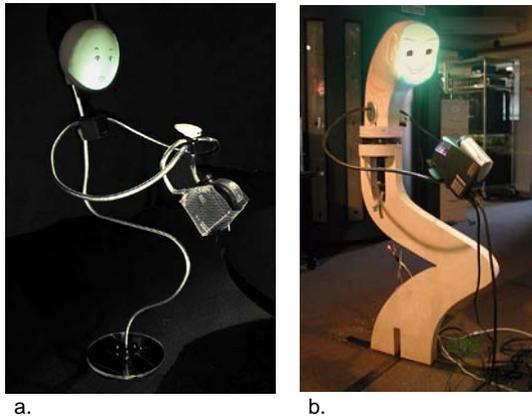
Figure 2. The two *Chit Chat Club* telepresence sculptures. (a) first version. (b) second version.

each containing a camera, microphone, speakers, and a video screen - were used to allow a four-remote-person meeting with the participants positioned around a table [2]. The modules are quite small relative to the people they embody, and have the appearance associated with rectilinear computer displays rather than sociable physical gestural presences. These perhaps seemingly unimportant details do carry social, emotional, and cognitive cues to the participant experiencing her or his colleagues as little talking heads on her or his desk.

Chit Chat Club moves further into the scale of the physical realm. The social interface in this project is designed to blend the interface and space of the interaction.

THE CHIT CHAT CLUB

Chit Chat Club is an experiment in bringing people together in a mixed physical and virtual environment. One of the major inspirations for creating this exploratory installation is the blending of the online and the physical. From the online world, we wanted to maintain the ease of interaction with strangers, the ease of entry into public spaces, and the low risk associated with interaction. From the physical world, we wanted to maintain the affordances (navigation, sense of touch, visual field, etc.) of the physical space surrounding us, the presence of physical, tangible forms, and the sensory perception of space.

Our goal is to create, through careful design of the physical environment and computer interface, a place that gracefully combines these two cultures; the analysis of how well this space actually functions will further our understanding of social interaction, both online and in person.

The First Chit Chat Club Installation

Physical Avatar

The telepresence sculpture was designed to be of human scale. The idea is to communicate with an interface that is at the same head level with similar proxemic codes as sitting face to face as opposed to a monitor. We found that if the seat is bigger and looks down on the person, it is intimidating; if it is much smaller, it is often ignored. This way, the remote participant occupies a similar space as that of the physical participants.

We designed the telepresence sculpture to be anthropomorphic to a degree, but not so anthropomorphic that one would expect human movement and human expression (see Figure 2a). The body frame is meant to resemble a relaxed figure of human proportion. The head resting on the frame has some curvature and is painted white to make a good projection surface. A projector hangs in a wire basket beneath the crossed "hands" of the telepresence sculpture. This projector is aligned to project moving faces onto the sculpture's head. Above the crossed "hands", rests a camera facing away from the telepresence sculpture. This camera captures video of the companions in the physical *Chit Chat Club* so the remote participant logging into the telepresence sculpture can "see" from the point of view of their telepresent self. The faces are designed to complement the physical form. They are abstracted human faces rather than photorealistic faces. As an experiment we tried photorealistic faces, but they appeared disconcerting both because of the disjuncture with the very abstracted physical form and because participants focused much more on lag discrepancies in the interaction.

The Local Space

The local space of the *Chit Chat Club* is a cafe setting. The telepresence sculpture, "sits" at a table and people entering the cafe may choose to sit near him or her. Figure 1 shows a picture of one *Chit Chat Club* table.

The Remote Interface

The online site is a portal for the remote visitors to enter the *Chit Chat Club*. Using the *Chit Chat Club* website, the remote cafe attendees can create an appearance for their visit, they can choose where to sit, and they can converse with the cafe's local participants using audio. While they are conversing with the visitors of the physical cafe, they see an abstracted video representation of their table companions from the point of view of the telepresence sculpture they occupy.

Getting Started

When an online participant first encounters the online *Chit Chat Club*, they are presented with an interface for creating a face for their respective telepresence sculpture (see Figure 3a). Using this interface, the visitor can customize the appearance of their projected face. They may choose face shape, eyes, lips, as well as the color of each feature from a
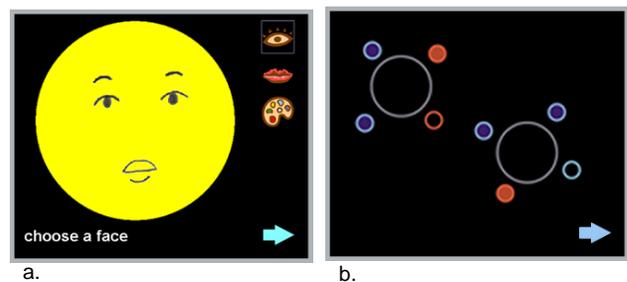
Figure 3. (a) The face selection interface.The user can cycle through different eyes, lips, and face color by clicking on the representative icons on the right. (b) Birds-eye view of the cafe. Blue represents walk-in visitors and red represents remote visitors respectively. A vacant seat is depicted by an outlined circle; an occupied seat is depicted by an opaque circle.

Figure 4. The first remote *Chit Chat Club* interface.

selection of hand-drawn, claymation, and cartoon facial components.

After the visitor creates their appearance, they are ready to proceed to the *Chit Chat Club* entrance. Here they see a graphical birds-eye view of the layout of the physical space (see Figure 3b). This representation shows the location of the tables, chairs, and telepresence sculpture. It also shows which seats - regular and telepresent - are occupied. The participant uses this image to select which telepresence sculpture to occupy.

The Communication Interface

Once a seat is selected, a two-way audio and video connection is established. In the cafe, the remote participant's created face appears in the chosen sculpture. At the remote location, the participant sees a live, processed image of the cafe as seen from that chair (see Figure 4) and can hear, see, and participate in the conversation at that table.

While connected, remote visitors can communicate by talking and can momentarily change their sculpture's expression to appear happy, bored, disgusted, sad, or angry. The faces animate to the chosen expression and then back to a neutral position.

This interface was ultimately flawed. The remote users would spend a large portion of their time selecting facial animations; this deterred from the conversation. They also continued to click on the facial expressions so as not to appear inattentive. Remote users also wanted to direct the gaze of their avatar themselves as apposed to negotiating with the visitors to the physical cafe for rotation of gaze. We concluded that for these reasons, we had to make facial expression selection easier and effort-free, and allow the remote user to see how they appear at the physical cafe end.

The Second Chit Chat Club Installation

Physical Avatar
We began to design a new telepresence sculpture to complement the needs of our new proposed remote interface. Allowing the remote user to control their gaze into the space was a priority. The new seat, therefore, would require a redesign with a motor.

The new chair would also possess an anthropomorphic form, yet would not be a human sculpture. The final design

and implementation of the second telepresence sculpture is shown in Figure 2b.

The Remote Interface
Entering the virtual component of the cafe is the same as in the previous version. What has changed in the second version, is the addition of a set of facial components, the communication interface, and the design of the telepresence sculpture.

The Communication Interface
In the new remote communication interface, the remote participant could view the expression that was being projected onto their sculpture's head and control the motorized gaze of their sculpture (see Figure 5). Novel to this interface was a new, comprehensive expression palette in the form of a wheel.
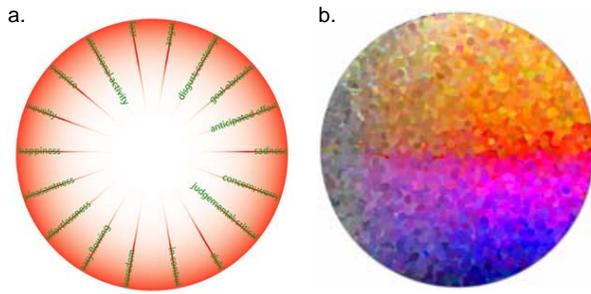
Expression Wheel
The expression wheel is designed to be a simple, intuitive interface for performing facial expressions. Facial expressions composed of component elements [14] such as the lowering of eyebrows and the raising of lip corners are mapped in smooth transitions around the circle clockwise (see Figure 6a). Higher intensities of these expressions are at the perimeter of the circle and blend to more neutral expressions towards the center of the circle. To make it easier on the user, there was no clicking involved. One simply had to mouse over the area of the expression.

Even with this new continuous wheel, users continued to focus on the written expressions and not on the conversation. To shift the attention to the interaction while maintaining expression cues during the conversation, we decided to abstract the representation of the expression wheel and add an autonomous component to the expression selection (see Figure 6b). We do this by tracking the pitch of the remote user in real-time and using simple heuristics to alter facial expression [12][14]. One example of this is the correlation of rise in pitch to rise in eyebrows. Although, the expression selection is semi-autonomous, the user always has the ability to override the system and select their own expression independently.

The abstracted expression wheel allowed the remote user to focus more of their attention on the conversation and less

Figure 5. The Second *Chit Chat Club* communication interface. This implementation incorporates the abstracted expression wheel while maintaining gaze control and remote expression visibility.

Figure 6. (a) the continuous labeled expression wheel. ex. expressions: happiness, novelty, surprise, attention, fear, anger, disgust, etc. (b) the continuous abstracted expression wheel.

on moving the mouse. There was, however, a trade-off in the expressiveness of the faces.

With the automated expression selection, we were cautious not to deduce a false expression. Hence, the arc of the expressions did not always reach the extremes. In retrospect, it was the extreme, cartoon-like animations that provided more of a catalyst for interaction than the subdued ones.

DISCUSSION

Chit Chat Club was presented to hundreds of users and video was collected of the interactions. This was conducted in a laboratory setting. The next iteration will take place in a more typical cafe setting.

We found that users preferred the extremity of the expressions with the first installation. As mentioned earlier, this is most probably due to our being conservative in our autonomous classification of expression when using the second emotion wheel. We also found that users felt more comfortable with the sculpture when the faces fidgeted and the eyes blinked. A completely still face was very disconcerting.

Users preferred the ability to control the motor in the second sculpture, and hence, their gaze. However, the majority of users found the first telepresence sculpture "friendlier". This suggests we incorporate the motor into the first design and re calibrate the expression extremities.

Our following explorations will focus on studying the nature of the interaction with varied abstraction levels and cue representations in the face projection.

Symmetry

A final note on *Chit Chat Club*: Although the installation emphasizes togetherness, there is an asymmetry in the public and the private. The local physical *Chit Chat Club* occupants perceive more of the social catalysts and the physicality, although they only see an abstracted representation of the remote cafe-goer. In contrast, the remote user sees a fuller view of the participants albeit at a smaller scale and in a physically remote setting.

Summary

The physicality of *Chit Chat Club* makes for interaction far different from what happens while staring at a computer screen. In contrast to the *Hydra* system mentioned earlier, the human-scale in *Chit Chat Club* provides for gestural behaviors at eye-level and not to several chess-piece-like screens. The similarity in scale blends the physical and vir-

tual worlds together to emphasize togetherness versus remoteness.

In creating this hybrid world, we no longer interact as we do in an online-only space or in a physical-only space. We have created a new style of space that brings people together in unfamiliar and serendipitous ways.

REFERENCES

1. Bly, S. and Irwin, S. Media Spaces: Bringing people together in a video, audio and computing environment. *Comm. ACM* 36,1, 28-47, 1993.

2. Buxton, W. Telepresence: integrating shared task and person spaces. *Proceedings of Graphics Interface 1992*.

3. Dobson, Kelly. *AgoraPhone*. Masters Thesis. Massachusetts Institute of Technology. 2002.

4. Finn, K., Sellen, A., and Wilbur, S. *Video-mediated communication*. Erlbaum. 1997.

5. Goffman, E. *Behavior in Public Spaces: Notes on the social organization of gathering*. New York: The Free Press. 1963.

6. Grudin, J. Why CSCW applications fail: Problems in the design and evaluation of organizational interfaces. *Proceedings of CSCW 1988*.

7. Hollan, J.and Stornetta, S. Beyond Being There. *Proceedings of CHI 1992*.

8. Isaacs E. and Tang J. What Video Can and Can't do for Collaboration: A Case Study. *Multimedia '93*.

9. Karahalios, Karrie. *Social Catalysts: enhancing communication in mediated environments*. PhD Thesis. Massachusetts Institute of Technology. 2004.

10. Karahalios, Karrie and Donath, Judith. Telemurals: Linking Remote Spaces with Social Catalysts. *CHI 2004*.

11. Pederson, E.R. and Sokoler, T. AROMA: abstract representation of presence supporting mutual awareness. *Proceedings of CHI 1997*.

12. Scherer, K.R. Vocal Affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143-165. 1986.

13. Singer, A., Hindus, D., Stifelman, L., and White, S. Tangible Progress: Less is more in Somewire audio space. *Proceedings of CHI'99*.

14. Smith, Craig A. and Scott, Heather S. A Componential Approach to the meaning of facial expressions. *The Psychology of Facial Expression*, eds. Russel, James A. and Fernando-Does, Jose Mogul. New York: Cambridge University Press, 1997.

15. Whittaker, Steve. Rethinking Video as a technology for interpersonal communication: Theory and design implications. *International Journal of Human-Computer Studies*, 42, 501-529. 1995.

16. http://www.tonyoursler.com