

Visualizing Audio in Group Table Conversation

Karrie Karahalios and Tony Bergstrom

University of Illinois, Urbana-Champaign
{kkarahal, abergst2}@cs.uiuc.edu

ABSTRACT

In this work, we create social visualizations of aural group conversation. These visualizations are tailored to a table setting. Examples of such settings are a family sitting about a table eating dinner or colleagues sitting about a table in a meeting. The form of the visualization is highly coupled to its function. People sit about the table structure and see their conversation visualized on the table surface as they speak. Using this physical structure, we present two graphical table interfaces. The first interface depicts the rhythm and conversational patterns in table top interaction; the second visualization extends this theme by incorporating a voting mechanism to highlight agreement and disagreement in spoken interaction. The visualizations evolve over time to create an evocative, graphical, interactive snapshot of the entire conversation within that table space.

INTRODUCTION

In this work we are visualizing aural conversations of people sitting about a table onto the table surface itself. In the area of social visualization, there has been an emphasis on visualizing textual archives of information such as chat, usenet, and email between users that are remotely located [3]. Similarly, there has been work in visualizing audio between spaces in the form of “media spaces” and artistic installations [2][5][6][7].

Visualizing audio conversation in a single shared physical space has thus far not been a focus for research. One reason may be that people in the same space can see each other and decipher expressions that evoke agreement, disagreement, anger, boredom, etc. more easily than people connected remotely via a mediated channel. The physical proximity and visual cues allow for accountability among the participants [5].

MOTIVATION

Why then visualize conversation of participants in the same space when they can see and hear each other? Our motivation and interest is in highlighting the intricate negotiation and conversational patterns that are not obvious in face-to-face interaction.

With a previous piece, *Visiphone*, we discovered the interface allows you to “see things that you know, but do not realize that you know.” [2] A common pattern was conversational dominance. During one interaction, one person was dominating the conversation and not allowing the other person to speak. By seeing the act of “domination” on the *Visiphone*, both users acknowledge the same phenomenon in the same form. In some cases behavior was altered. Other salient conversation patterns were turn-taking, interruption, and matching of volume level of participants. Although the *Visiphone* was meant to connect two remote spaces, we received unexpected levels of feedback from marriage

counselors suggesting it could be a useful negotiation and reflection tool¹.

Abstraction for Visualization

There are many methods for visualizing audio. Traditional audio visualizations represent audio amplitude along a time axis. Our approach for visualizing audio is to abstract the audio data into simple graphical elements. The design of these visualizations focuses on the evocative rather than the technical. The abstract representation divorces the technical connotations from the audio channel and provides a new frame of reference for the conversations.

The horizontal surface is ideal for this, for it blends the physical form and function of the table with social interactions associated with it such as conversation. The table top is the interface, the display, the gathering point, and the focus of attention. We illustrate this concept with two audio visualizations: *Conversation Clock* and *Conversation Votes*.

DESIGN

Conversation Clock and *Conversation Votes* are implemented upon the same physical table interface. They are described in the following sections.

Conversation Clock

Conversation Clock visualizes conversation of up to four participants around the same table. Each person is represented as a distinct color. As users speak over an interval of time, their audio participation is represented as a rectangle where the length corresponds to the average amplitude of their volume. This stream of rectangles grows clockwise over time; a complete circle is formed in one minute. As time progresses, the outer circles move towards the center, and the current conversation circle is at the periphery of the table (see Figure 2). The effect is reminiscent of a clock, only in this case, it is a *Conversation Clock*.

Conversation Votes

Conversation Votes follows a similar rectangular theme. In this case, different people are represented as different colored rectangles, however, the rectangles in this visualiza-

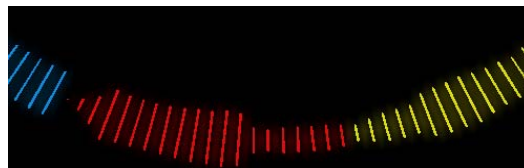


Figure 1. Close-up of three people speaking in sequence using *Conversation Clock*.

1. *Visiphone* was on display for several months in a laboratory setting and was on exhibit at Siggraph Emerging Technologies '99.

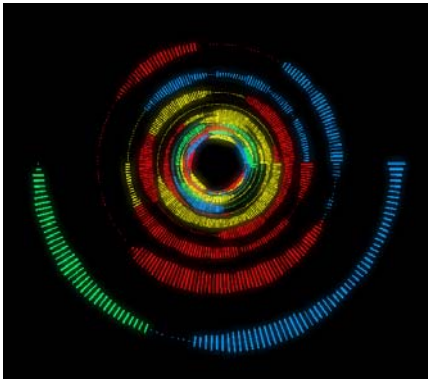


Figure 2. Projection image of the Conversation Clock visualization after several cycles.

tion are not filled in. During conversation, each participant has two buttons they can select on a small hand held device. Pressing one button implies agreement with what is being said; pressing the other implies disagreement. The handheld device can be held discretely during conversation. When a user agrees with a comment, that series of rectangles become filled in and the rectangles grow in length. When one disagrees, the rectangles again fill in to imply voting but they shrink in length. When people speak at the same time, they are represented alongside each other. Users cannot vote on their own audio.

The rectangles progress linearly across the diameter of the table. When each line traverses the table completely, it is laid out in miniature, perpendicular to the ongoing conversation line (see Figure 3).

The design for this visualization was motivated by studies showing that employees in an organization are afraid to disagree with employers if they are in a face-to-face meeting, but will voice their opinion in mediated conversation such as email [8]. With this interface, we aim to blend the affordances of physical and virtual group interaction.

The Physical Interface

The physical interface for these visualizations is a table onto which the visualization is projected. The table has four directional microphones embedded into the surface of the table top, each separated at right angles. The current implementation supports four users about a single table. The display surface has been coated to provide a clear projection with good contrast. The graphics are projected onto the table from a projector mounted above the space. The table with projection becomes an interface in the round: one can view it from any side. This is essential for an object meant to convey audio information from all users, for all users. The design of the table itself is a key element in this multi-modal interface, and its size, location, and appearance influence its use and its ability to portray the sense of activity and awareness.

DISCUSSION

We are familiar with various forms of volume representation whether they be bar chart visualizations such as in Quicktime™ or waveforms and drawings in many sound recorders and players. Although these representations are

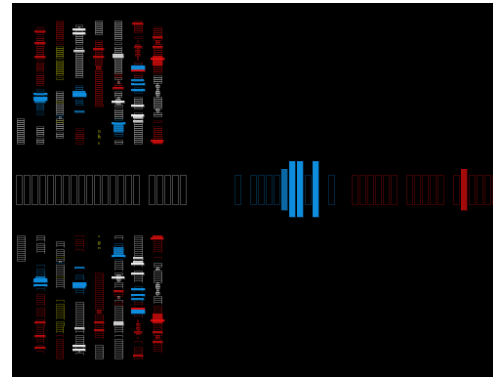


Figure 3. Image projected onto Conversation Votes table. Current conversation is the center line. Past conversation is arranged orthogonal to it.

readable, they have a technical connotation and their representations are fleeting. The abstract audio renderings in this paper separate themselves from traditional technical representations in design and form; the social visualizations convey conversation patterns among people - not a single audio stream.

The result is an abstract communication sculpture that blends the real and virtual into the task environment. Future work involves using pitch and pattern recognition techniques to extract features that we can use to create a graphical visualization toolkit for spoken text.

Memory, Patterns, and Surveillance

Unlike email which is persistent, it is difficult to uncover explicit patterns of audio and convey them to participants because the spoken work is ephemeral in nature. We speak and our words drift away. The same conversation is remembered differently by different people. We have memory, but no accountability. There is a beauty in the ephemerality of spoken language and our intention is not to alter that. Conversations in these interfaces are not recorded unless all involved parties consent. The graphical representations as a picture of the conversation patterns and rhythms exist on their own — we find that participants frequently request graphical renderings of meaningful conversations as evocative souvenirs.

REFERENCES

1. Arnheim, Rudolf. *Visual Thinking*. Berkeley: University of California Press. 1969.
2. Donath, J., Karahalios, K., and Viegas, F. Visiphone. *Proceedings of ICAD2000*.
3. Donath, J., Karahalios, K., and Viégas, F. Visualizing Conversation. *Journal of Computer Mediated Communication* 2 (4), 1997.
4. Goffman, E. *Behavior in Public Spaces: Notes on the social organization of gathering*. New York: The Free Press. 1963.
5. Karahalios, Karrie. *Social Catalysts: enhancing communication in mediated environments*. PhD Thesis. Massachusetts Institute of Technology. 2004.
6. Levin, G. and Lieberman, Z. 2003. <http://tmema.org/messa/>
7. Singer, A., Hindus, D., Stifelman, L., and White, S. Tangible Progress: Less is more in Somewire audio space. *Proceedings of CHI'99*.
8. Sproull, L. and Kiesler, S. *Connections: new ways of working in the networked organization*. Cambridge, MA: MIT Press. 1991.