Demonstrating How to Construct a Sonic Ecology for Media Spaces through Cambience

Rob Diaz-Marino and Saul Greenberg
Department of Computer Science
University of Calgary, Calgary, Alberta CANADA T2N 1N4
+1 403 220- 6087
{robertod,saul}@cpsc.ucalgary.ca

ABSTRACT
Cambience is a sonic media space. Using a visual programming environment, people can map properties of a video stream into audio sequences, thus delivering sonic awareness cues of presence and activities to distance collaborators without the visual distraction. This demonstration illustrates how a Cambience sonic ecology is created.

Keywords
Cambience, Sonic Ecologies, Abstract Media Spaces, Video Awareness, Ambient Sound.

1. Sounds and Media Spaces
Media spaces are a common way of providing distance-separated collaborators with awareness of one another. The idea is that people can track the presence and perceived availability of others by viewing webcam video, and move into conversations at opportune times.

Of course, video has considerable problems. First, it may reveal too much, and some people – even those who want to stay in contact with others – perceive it as a privacy threat. Second, video requires foreground monitoring: a person can only be aware of the (usually silent) video contents when he or she is actually watching it. This leads to the third problem: unnecessary distraction. Because people are often focused on things of greater importance (ie. work), they have to break this primary focus every now and then to check if anything has changed in the video feeds that they are monitoring. The fourth problem is that of missed opportunities. Because people cannot watch the video continuously, they can miss events of importance.

As with several other researchers [2,3], we became interested in generating sound cues in response to changing video. Sound presents many promising possibilities for awareness maintenance. Because sound is perceived on a separate sensory channel, it can convey information in parallel to our visual sense. Many of us exploit this ability every day: listening to music while editing a document, for instance. While our visual awareness can only be concentrated on a cone of space in front of us, our hearing enables us to keep an awareness of our entire surrounding environment.

Copyright is held by the author/owner(s). Permission to copy, mirror, distribute, show and display some or all of the work (including associated video) for non-commercial and academic purposes is granted without fee provided that the work is attributed to the authors, and that copies bear this notice and a full citation on the first page.

CSCW'06, November 4-8, 2006, Banff, Alberta, Canada.

Away from the computer it is often the case that we redirect vision based on what we hear, to obtain more detail. For example, the sound of footsteps approaching an office door often compels a person to turn and look at the source. The regular sound of a clock ticking, on the other hand, would simply be ignored. In essence, our physical environment is an ecology of ambient sound that keeps us informed of events in our immediate surroundings. The nature and quality of a sound determines if we reassign our attention.

Within the context of a media space, we wanted to know if it is it possible to translate some visual properties of video streams from multiple users into sound variation in a sonic ecology. We pictured this resulting in an abstract sonification that would function as an attractor of visual attention only when necessary.

Sonification is "the use of non-speech audio to convey information or perceptualize data." We use the term to describe the act of creating audio cues from patterns that can be programmatically detected and measured from a video stream. By sonifying a video stream, we get the computer to do the watching for us, reporting to us on a continuous basis using audio - a channel that doesn't demand our visual attention. By using a medium that is not in such high demand, we hope to allow people to maintain awareness without causing them unnecessary distraction and cognitive overhead.

We quickly realized that it is quite challenging to create a meaningful mapping. There is no single solution because different media space cameras may be focused on quite different scenes, and different people may place greater importance on different cues from that scene. For instance, one person may prefer an awareness soundscape to comprise a mellow background of ambient nature sounds, where bird whistles and cricket chirps signal some type of activity in the video scene. Another person may favor a music mix, where the interplay of instruments is controlled by the video contents. Yet another may prefer a literal office mapping: visual motion over a keyboard is mapped to keyboard clicks, and activity around a doorway is heard as footsteps.

2. Cambience
Our idea was to give people a visual programming environment called *Cambience* that: (a) lets them define regions of interest on a video scene, (b) allows them to create a sonic ecology by mixing discrete sound samples, (c) lets them map visual properties of these regions onto properties of the sound samples, and (d) shares the regions of interest with other networked users, so that they may in turn create their own mappings.

Diaz-Marino, R. and Greenberg, S. Demonstrating How to Construct a Sonic Ecology for Media Spaces through Cambience. Adjunct Proc CSCW 2006, Demonstration and two page summary.

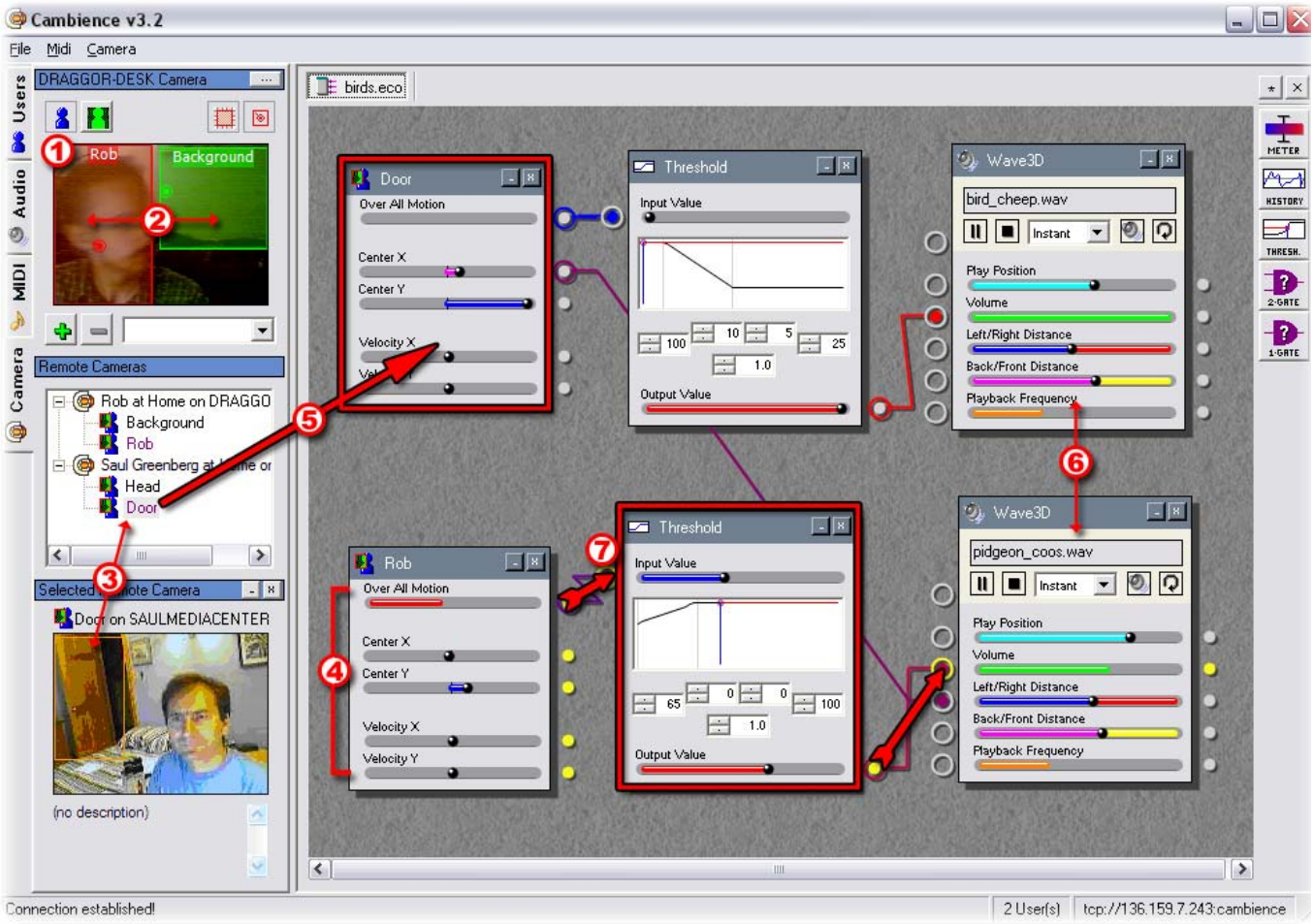

Figure 1: Creating a Cambience Sonic Ecology

The demonstration shows how Cambience works, while Figure 1 provides a snapshot of the Cambience visual programming environment. An illustrative video is also in submission [1].

1. Cambience captures the local live video stream from a webcam or other visual capture device.
2. Each user creates, names, and describes regions of interest over this video stream. For example, Figure 1 shows a region that bounds the head and torso of the seated office occupant (red), as well as a second region over their shoulder (green).
3. Several instances of Cambience can be connected over a network, and the regions from each user are made available for all to use in their ecologies.
4. Feature extraction is performed on these regions to provide basic but potentially useful measurements of visual change in the scene. These include: overall motion, the center x/y of the motion, and its x/y velocity.
5. A person can drag and drop one of the regions to the sonic ecology editor, where these properties and their current values are revealed.
6. Using the Audio tab (top left), a person can drag and drop different sound sources into the ecology. That sound can be played and its properties revealed, e.g., play position, volume, 3d left/right and front/back sound placement, and frequency. For example, Figure 1 contains a bird cheep and pigeon coo.
7. People can map video properties onto audio properties simply by joining them with connecting lines. Special thresholding tools can be dropped into the editor to further fine tune this behaviour.

Through these steps, Cambience allows people to customize the way in which the video is mapped onto a sonic ecology.

3. ACKNOWLEDGMENTS

Funding is partially provided by the NSERC NECTAR Research Networks grant and its industrial sponsors Smart Technologies and Microsoft Inc.

4. REFERENCES

[1] Diaz-Marino, R. and Greenberg, S. Cambience: A Video-Driven Sonic Ecology for Media Spaces. Video in submission to CSCW 2006.

[2] Mynatt, E., Back, M., Want, R., Baer, M., Ellis, J. B. (1998) Designing Audio Aura. Proceedings of CHI 98, 566-573.

[3] Pedersen, E. R., Sokoler, T. (1997) AROMA: Abstract Representation of presence supporting Mutual Awareness. Proceedings of CHI'97, 51-58.