The AlloBrain: an Interactive Stereographic, 3D Audio Immersive Environment

Graham Wakefield wakefield@mat.ucsb.edu
JoAnn Kuchera-Morin jkm@create.ucsb.edu
Marcos Novak marcos@centrifuge.org
Dan Overholt dano@create.ucsb.edu
Lance Putnam ljputnam@umail.ucsb.edu
John Thompson jthompson@georgiasouthern.edu
Wesley Smith whsmith@mat.ucsb.edu
Media Arts and Technology Program
University of California, Santa Barbara
Santa Barbara, California, USA

ABSTRACT

This document describes the AlloBrain, the debut content created for presentation in the AlloSphere at the University of California, Santa Barbara, and the Cosm toolkit for the prototyping of interactive immersive environments using higher-order Ambisonics and stereographic projections. The Cosm toolkit was developed in order to support the prototyping of immersive applications that involve both visual and sonic interaction design. Design considerations and implementation details of both the Cosm toolkit and the AlloBrain are described in detail, as well as the development of custom human-computer interfaces and new audiovisual interaction methodologies within a virtual environment.

Author Keywords

Immersive Environments, Sonification, Human-Computer Interaction, Wireless Controller, Ambisonic Spatialization.

ACM Classification Keywords

H.5.1 Multimedia Information Systems (Artificial, augmented, and virtual realities), H.5.2 User Interfaces (Input devices and strategies, Prototyping), H.5.5 Sound and Music Computing (Systems)

INTRODUCTION

We were commissioned to create content for the CNSI AlloSphere [1] to indicate its capabilities and provoke inspiration for future AlloSphere projects. The result of this commission was the AlloBrain project that explores brain imaging data as an immersive environment (Figure 1).

The CNSI AlloSphere

Housed in the California Nanosystems Institute (CNSI) at the University of California at Santa Barbara, the AlloSphere is a 10-meter diameter spherical perforated aluminum projection surface suspended within a three story near-anechoic cube. The environment accommodates approximately fifteen people on a bridge that runs through the middle of the space (Figure 2). The AlloSphere will implement high-resolution stereographic video projection across the entire sphere using multiple projectors, and include several hundred loudspeakers suspended behind the sphere surface[1].

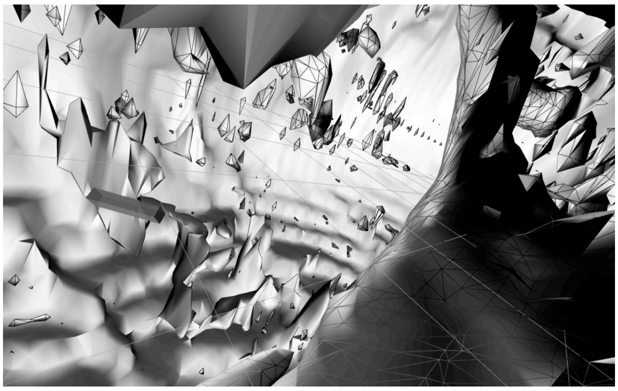

Figure 1. Exploring fMRI data inside the AlloBrain.

Figure 2. On the CNSI AlloSphere bridge.

The AlloSphere's primary function is the analysis, synthesis, simulation, and processing of complex multi-dimensional data within an interactive/immersive environment. It is an instrument for gaining insight and developing bodily intuition about environments into which the body cannot venture: higher-dimensional information spaces, the worlds of the very small or very

---

[1] At the time of writing the configuration is not finalized. The current system consists of an Ambisonic array of 22 speakers.

large, from nanotechnology to cosmology, from neurophysiology to new media.

A Toolkit for Immersive Sonification/Visualization

The nature of the AlloSphere presents a number of technical requirements in order to support the development and presentation of 3D audio, stereographic, immersive environments. We define immersive environments as virtual spaces that create worlds featuring spatial, temporal, and navigable congruence between audio and image.

Rather than building a software solution specific to the AlloBrain project, we designed a generalized system, the Cosm toolkit, to support the rapid development of many different kinds of projects within the AlloSphere and similar spaces.

The AlloBrain Project

Following our desire to meld data from the sciences with the artistic pursuits of new media art, we chose to create an immersive world through the aesthetic sonification and visualization of brain imaging data. Our goal is not to interpret the data in a scientific way, but rather to demonstrate what immersive three-dimensional media offers in terms of new insights and interaction with data sets from other disciplines.

The AlloBrain project has been exhibited many times in the AlloSphere, including during the ACM Multimedia conference in October 2006. The AlloBrain was also exhibited at the opening of the iWeb/Protospace Lab of the Department of Architecture at Delft University of Technology in the Netherlands in March 2007 [5], and simultaneously in the AlloSphere, with both installations telematically connected and sharing viewer data via avatars.

THE COSM TOOLKIT

In this section, we describe the current form of the Cosm toolkit along with key challenges and responses through its development. The technical hurdles we encountered and the solutions chosen may be relevant to immersive content development in the wider community.

We identified the following core requirements for the Cosm toolkit:

1. Rapid development of project-specific auralizations/visualizations
2. Spatialized audio over potentially hundreds of loudspeakers
3. Stereographic projection over numerous projectors
4. Six-degrees of freedom navigation
5. Adaptability and scalability to distinct or changing hardware configurations

Choice of Development Tools

Whilst there is active research in the AlloSphere to develop in-house software solutions to the above requirements, in the interim it was desirable to demonstrate content in the AlloSphere itself (Cosm requirement 1). Given that software development is a lengthy process, and that the hardware configuration of the space was not completely settled (requirement 5), we opted to make use of Max/MSP/Jitter [17] as a flexible rapid prototyping platform in which to develop the Cosm toolkit.

The benefits of using this development tool were numerous. In particular, the strong existing support for audio, 3D graphics, networking, and interface I/O within a unified environment enabled a rapid development cycle. We authored a number of extensions (externals) in C/C++ in order to satisfy certain requirements of the Cosm toolkit not already fulfilled by the Max/MSP/Jitter environment. While it allows rapid development cycles, Max/MSP/Jitter did present some drawbacks; primarily, the closed-source nature made it difficult to debug the toolkit at critical points in its development, and the visual patching paradigm was not always well suited to this large-scale project.

The Cosm Toolkit Architecture

A distributed architecture is necessary to support the large number of audio channels and video projectors (Cosm requirements 2 and 3). To maintain simplicity in development and flexibility of control, we opted for a simple star network topology of a single master control node, and any number of remote rendering nodes (Figure 3). All Ambisonic decoding and stereographic rendering is offloaded to the remote nodes.

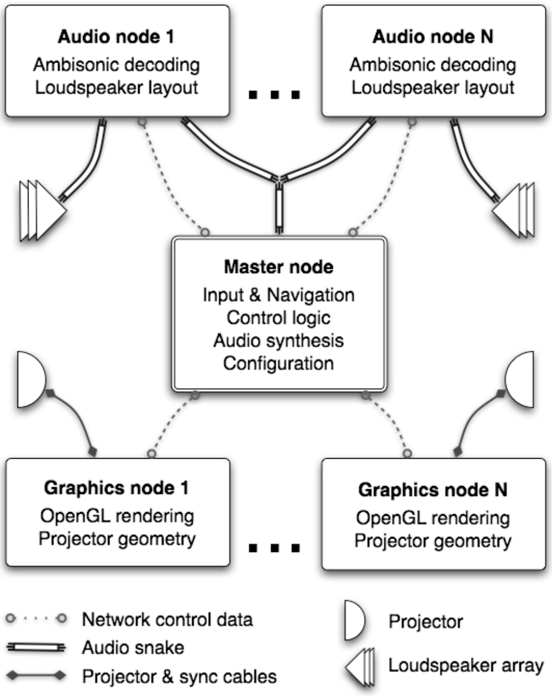

Figure 3. Cosm toolkit architecture.

The Cosm master node handles immersant navigation, audio/video rendering settings and overall system state, as well as broadcasting audio, data, and control messages. The Cosm graphics render node manages stereographic OpenGL rendering adjusted to match the attached projector's orientation. The Cosm audio render node decodes the Ambisonic domain signals according to the attached loudspeaker geometry.

Audio rendering settings include reverberation parameters, virtual air absorption coefficients, and the speed of sound for spatialization in realistic or non-realistic environments. Video rendering settings include stereographic mode, depth of field, and stereo separation.

A project utilizing the Cosm toolkit consists of two components implemented as Max patches (Figure 4). The project-specific master node incorporates all project control logic, interaction, and audio synthesis. The project-specific render node incorporates all project-specific graphical rendering objects.

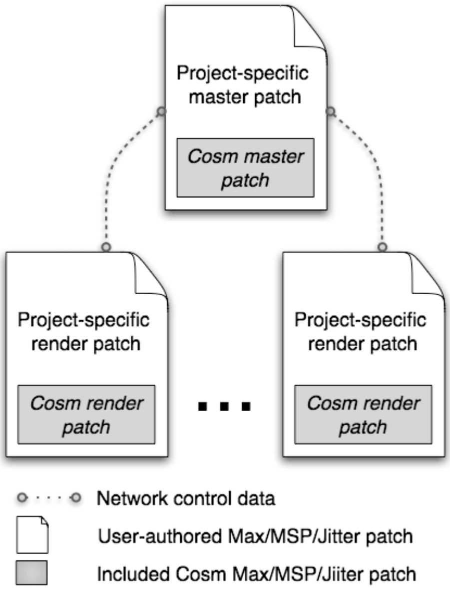

Figure 4. General architecture of a Cosm project.

Communication

The master node communicates to all render nodes via UDP network multicast on a group address. System state is maintained on this group multicast address, such that any audio or graphic render nodes can correctly represent the current project state to the immersants.

All nodes self-register their IP addresses with the master node upon initialization, allowing users to also direct control messages to specific nodes where necessary, removing the need for static IP addresses, and allowing remote configuration of node-specific parameters.

All Cosm messages (both across the network and internal to the patches) belong to a simple namespace system, in which names are prefixed with *_cosm* to avoid collision with project specific messages. In addition, a basic namespace is provided for project-specific distributed control over the same multicast network.

To support portability and rapid adaptability (requirement 5), all parameters specific to a render node, such as projector geometry, are stored in local configuration files on each rendering computer. A special patch was developed to quickly and visually define and store remote configurations from the master node computer. This is useful for installations in which render node machines are not easily accessible.

Spatial Audio

To make the best use of the AlloSphere's pluriphonic capabilities, the Cosm toolkit employs third-order 3D Ambisonics for audio spatialization, using Ambisonic extensions to Max/MSP/Jitter developed at MAT/CREATE [18]. Software for 3D Wavefield synthesis [7] is currently being developed for a future implementation.

Canonical Ambisonics models spatial orientation well [14], but does not inherently model distance. We have extended our implementation to incorporate multiple distance cues for point sources using standard techniques (amplitude attenuation, medium absorption/near-field filtering, Doppler shift and reverberation [8]). We implemented a rudimentary (sound-cone) radiation pattern simulation by mixing distinct filtered outputs for different directions according to the orientation of the object relative to the listener (Figure 5).

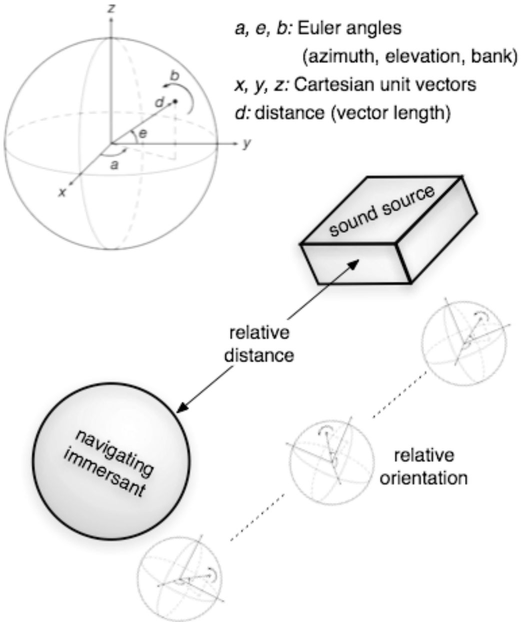

Figure 5. Overview of geometry used in spatial audio source localization

Point source synthesis, distance coding and Ambisonic encoding take place in the master node. The encoded Ambisonic domain channels (16 channels for 3^{rd} order 3D) are distributed to multiple rendering nodes for decoding to potentially hundreds of loudspeakers.

Navigation

The geometry of the AlloSphere suggested a freedom of virtual world navigation in any axis (Cosm requirement 4). Naïve approaches to navigation quickly show limitations; for example, Euler angles for orientation suffer from gimbal lock when two axes become congruent. This applies equally to the camera as to any mobile object within the space. Additionally, the trigonometry to render multiple views from a single point (to support 360° projection) or points (for stereography) quickly becomes nontrivial when the viewer orientation is flexible in six degrees of freedom.

Our solution was to employ quaternions to represent orientation in 3D space. Quaternions are used extensively for this purpose in flight dynamics, computer graphics and gaming development, and offer advantages of compactness, stability and smooth rotation interpolation over matrix representations. A set of extensions to Max was developed to support quaternion rotation and conversion to/from axis/angle, Euler angle, matrix and Cartesian unit vector representations.

As a result, we can derive unit vectors X, Y, Z relative to any orientation rather than the absolute axes; this makes it very easy to place content consistently within the viewer's perspective or to mobilize objects in a trajectory based upon their current orientation for natural, smooth movements. These features were used in the AlloBrain project for immersant navigation, overlays, and mobile agent behavior.

An Ambisonic sound field in its entirety can be simply rotated around three axes using equations based upon spherical harmonics [13]. Unfortunately, if the listener is a mobile point in a virtual world, the relative directions to each point source are constantly changing, and sound field rotation is no longer a great advantage. Thus the calculation of sound source direction and distance for each source must be recalculated whenever the viewer moves, changes orientation, or the source itself moves.

Sound source direction in Ambisonics is expressed using Euler angles of azimuth and elevation. The third angle, bank, is not directly relevant to Ambisonics, but could be important for raycast reverberation simulation. A custom external (*xyz2aed*) was written to simplify the process of deducing the appropriate azimuth, elevation and distance for Ambisonic encoding and distance simulation in audio from the relative object position and immersant navigation orientation. The unit vectors of the navigation orientation are multiplied by the viewer-to-object vector. The azimuth and elevation are derived using trigonometric arctangent and arcsine of the resultant up, side and forward scalars. Distance is calculated simply using the Pythagorean formula.

Graphical Rendering

The AlloBrain project and the Cosm toolkit implement active stereographic projection. Active stereo uses one graphics window that alternates left and right frames, requiring twice the frame rate and additional configuration of the graphics context. For an active stereo system a sync signal is generated by the graphics card to drive the shutter rate of the active stereo glasses and synchronize the appropriate left and right eye images. To enable this type of configuration in Max/MSP/Jitter, the OpenGL rendering and windowing objects (jit.window and jit.gl.render) were modified to enable quad-buffered stereo. Eventually these changes were folded back into the standard Max/MSP/Jitter distribution as of Jitter 1.6.3 beta 2.

Calculation of the left and right eye viewpoints is done using a quaternion camera. Given the camera's orientation, a local coordinate system is derived. From this, the left and right viewpoints are projected out into space based on the interocular distance parameter, providing each eye with a unique image that when viewed through the 3D glasses is perceived as existing in space.

THE ALLOBRAIN PROJECT

Digital artist and transvergent architect Marcos Novak undertook fMRI brain scanning while researching the neurophysiological basis of aesthetic appreciation. During this process a generative algorithm produced unanticipated visual stimuli that was presented to the viewer while inside the fMRI machine. Both conscious and neurophysiological data were captured during this presentation. The fMRI data were stored in 3D volumes in the Analyze file format. A custom object was built for reading the files and outputting them as 3D Max/MSP/Jitter matrices.

The structural component of the fMRI data provides an intricate spatial architecture, which is rendered visually as a navigable space that can be experienced as a world. It should be recalled that the project goal was to indicate rather than confirm the potential of the AlloSphere for scientific research, thus our design decisions were ultimately driven by questions of aesthetic experience.

Agents

The visualization or sonification of such a large, static data set is not sufficient to create a compelling interactive experience. A narrative for the installation was therefore introduced, based upon the concept of computer assisted data mining. A notion of dynamic search is embodied by a number of semi-autonomous explorer agents sharing the data space with the immersants, and supervised by the navigator using interactive physical interfaces (Figure 6).

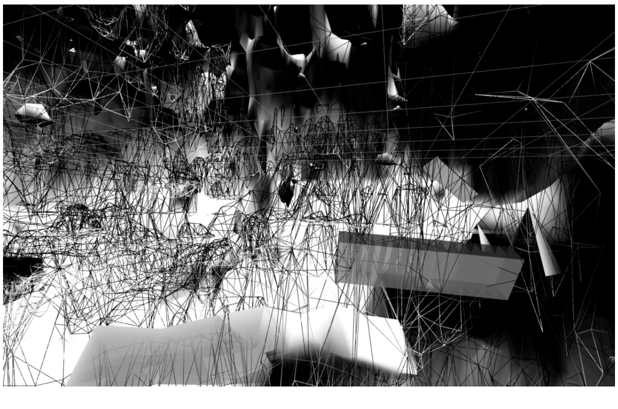

Figure 6. View inside the AlloBrain with cuboid agent.

In this narrative the explorer agents navigate the brain measuring blood flow densities, changing color according to brain region, and the navigator can call specific agents to report the status of their findings. Sonification of the data set becomes the responsibility of each of the agents, providing a complex and evolving sound-scape of multiple discrete spatialized paths. The agents allude to the possibility of a richly dynamic mode of human computer interaction, merging the best use of human and digital pattern-matching capabilities.

Sound Mappings

Several audio synthesis techniques were used to inform the immersant about the agents' current actions in the environment. It was important to create sounds that would allow the immersant to localize the mobile agents. Short noise bursts were used as spatial cues since wideband signals provide more precise elevation cues [12]. Here we capitalize on the advantage of using omnidirectional aural feedback versus unidirectional visual feedback.

In addition, we created a bed of ambient sound serving to draw the immersant into the environment. We found that in this sonic atmosphere immersants felt more inclined to spend longer stretches of time within the virtual world.

Interfaces

We developed two interfaces (Figure 7) for the AlloBrain project in order to provide audiovisual navigation within the MRI data and control the explorer agents via custom Inertial Measurement Units (IMUs) based on the CREATE USB Interface (CUI) [16]. The Ambisonic sound spatialization and visual navigation are controlled by two collaborating immersants. The CUI circuits were enhanced with Bluetooth wireless data transmission and the integration of multiple MEMs (Microelectromechanical systems) sensors. The sensors include 3-axis accelerometers, gyroscopes, and magnetometers. Both absolute and relative orientation can be derived from the real-time data through algorithms such as Kalman or washout filters.

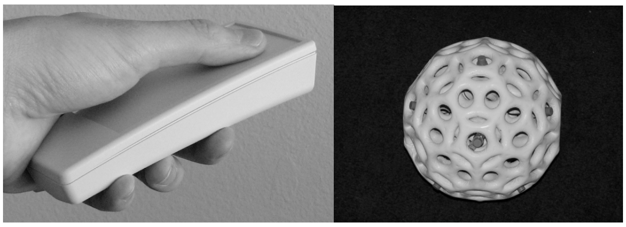

Figure 7. The Visual Navigator for the AlloBrain project, and The Sphere Spatializer interface.

The raw sensor data from each interface is combined and properly integrated to provide relative Euler angles (azimuth and elevation) from the IMU sensor arrays embedded in the controllers. In this way, the interfaces can control both movement and virtual orientation within the simulation space. The visual navigation device is a small hand-held unit with two buttons that activate each mode of control, one for movement and one for orientation changes via the user’s gestures.

The form factor of the second interaction interface, the Sphere Spatializer [2], is based on the hyperdodecahedron (a 4-dimensional geometric polytope) with the final form representing its shadow projected in to 3 dimensions (Figure 8). It was developed using procedural modeling techniques, algorithmically sculpting the mathematical shape to provide an organic and ergonomic look and feel that is aesthetically linked to the content in the AlloBrain project, and was constructed with a 3D printer capable of building solid objects [4].

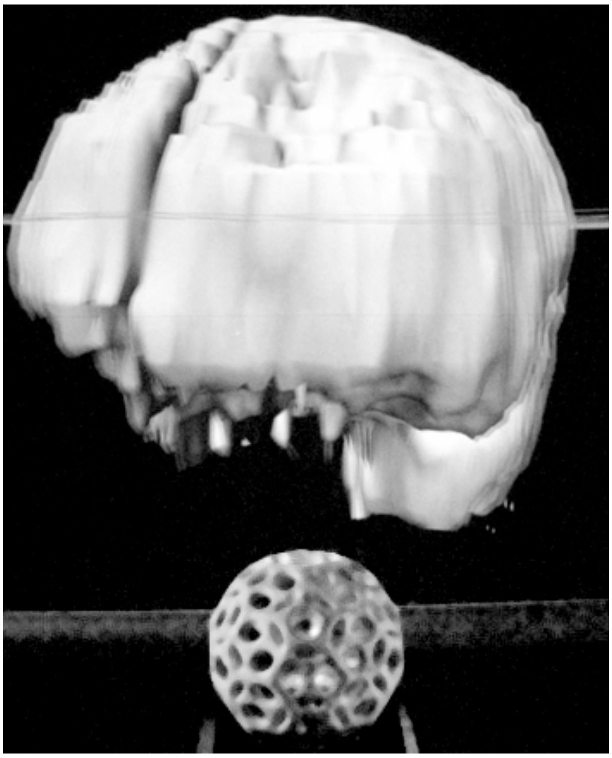

Figure 8. Sphere Spatializer on the bridge of the AlloSphere.

The Sphere Spatializer includes twelve buttons evenly spaced around its surface. These buttons provide control of the explorer agents within the simulation by calling them towards the current field of view, and shifting their spatial focus in order to specify a new region to explore.

FUTURE WORK

The Cosm toolkit continues to be developed and is currently used to develop multiple projects. It provides an initial code-base for future stereographic, 3D audio, immersive environment development relating to the AlloSphere. Certain key areas for improvement are outlined below.

Currently the Cosm distributed communication system does not include any kind of tight synchronization assurance. Installations to date have relied upon dedicated router hardware to minimize network latency and jitter, however this is not considered a scalable solution, particularly for telematic installations. Synchronization of distributed audio is also an open question. Latency and jitter are both dependent upon hardware implementations, however the distribution of Ambisonic domain signals in the current Cosm architecture minimizes some of these constraints.

The higher-order Ambisonics (HOA) employed in Cosm would benefit by replacing the encoding filters with the near-field coding format (NFC-HOA) proposed by Daniel et al [9], to better represent finite distance sources by synthesizing with spherical rather than plane waves. The gain is support for positional rather than directional encoding, including virtual sources inside the speaker radius. Likewise spatial width & characteristic radiation patterns may be supported using Menzies' W-panning & O-format techniques [15], though the relative merits and trade-offs for the AlloBrain project when compared to the existing sound-cone technique remain to be evaluated.

It may be fruitful to formally investigate user experience in the AlloSphere and within the AlloBrain project, however it would be considered premature given the fluctuating status of the technology at present [11].

The AlloBrain is also an evolving project. As our current fMRI data set is static, we recently initiated interface experiments proposing to explore how visualized and sonified EEG data can be integrated into a real-time interactive environment with biofeedback. Work to date has resulted in a Java-based extension to Max/MSP that captures real-time data from BIOPAC Inc.'s [3] line of bio-signal sensors (EEG, ECG, EOG, EMG, etc), allowing visualization and sonification of immersant's brain and body activity (Figure 9.)

Figure 9. Marcos Novak with 16-channel EEG test.

REFERENCES

1. The CNSI AlloSphere: http://www.mat.ucsb.edu/AlloSphere Retrieved April 2007.
2. D. Overholt, Sphere Spatializer: http://create.ucsb.edu/~dano/sphere Retrieved December 2007.
3. BioPac: http://www.biopac.com Retrieved December 2007.
4. ZCorp: http://www.zcorp.com Retrieved December 2007.
5. The AlloBrain at TU Delft iWeb and Protospace: http://www.tudelft.nl/live/pagina.jsp?id=429fa10f-a830-4944-90de-9a326e419b54 Retrieved December 2007.
6. D. Bergault, *3-D Sound for Virtual Reality and Multimedia*. Academic Press, Cambridge, MA, USA, 1994.
7. A. J. Berkhout, D. De Vries, P. Vogel, "Acoustic Control by Wave Field Synthesis", in *Journal of the Acoustic Society of America*, vol. 93, May 1993, pp. 2764-2778
8. J. Chowning, "The simulation of moving sound sources," in *Journal of the Audio Engineering Society*. Vol 19, No.1, 1971.
9. J. Daniel, " Spatial Sound Encoding Including Near-Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format", *Preprints of the 23rd Int. Conf. of the Audio Eng. Soc.;* Copenhagen, Denmark, 2003
10. R. Dannenburg, R. Fisher, "An Audience-Interactive Multimedia Production on the Brain" in *Proceedings of the Connecticut College Symposium on Art and Technology* 2001.
11. S. Greenberg, B. Buxton, "Usability Evaluation Considered Harmful," in *Proceedings of the 2008 SID-CHI Conference*, to appear.
12. F. Hollerweger, "Periphonic sound spatialization in multi-user virtual environments," Master's thesis, Austrian Institute of Electronic Music and Acoustics (IEM), 2006.
13. D. Malham, "Higher order Ambisonic systems for the spatialisation of sound," in *Proceedings of the 1999 ICMC,* 1999.
14. D. Malham, A. Myatt, "3-D sound spatialization using Ambisonic techniques," *Computer Music Journal (CMJ)*, vol. 19, no. 4, pp. 58–70, 1995.
15. D, Menzies, "Ambisonic Synthesis of Complex Sources", J. Audio Eng. Soc., October 2007
16. D. Overholt, "Musical Interaction Design with the CREATE USB Interface Teaching HCI with CUIs instead of GUIs," in *Proceedings of the 2006 ICMC*, 2006
17. M. Puckette, "Max at Seventeen,*" Computer Music Journal* 26, 4 (2002), 31-43.
18. G. Wakefield, "Third-order Ambisonic extensions for Max/MSP with musical applications," in *Proceedings of the 2006 ICMC*, 2006.