

IMMERSIVE AUDIO AND MUSIC IN THE ALLOSPHERE

Xavier Amatriain

UC Santa Barbara
xavier@create.ucsb.edu

Tobias Höllerer

UC Santa Barbara
holl@cs.ucsb.edu

JoAnn Kuchera-Morin

UC Santa Barbara
jkm@create.ucsb.edu

Stephen Pope

UC Santa Barbara
stp@create.ucsb.edu

ABSTRACT

The UCSB Allosphere is a 3-story-high spherical instrument in which virtual environments and performances can be experienced in a fully immersive fashion. It is made of a perforated aluminum sphere, ten meters in diameter, suspended inside a quasi-anechoic cube. The space is now being equipped with high-resolution active stereo projectors, a 3D sound system with hundreds of speakers, and with tracking and interaction mechanisms.

The Allosphere allows for the exploration of large-scale data sets in an environment that is at the same time multimodal, multimedia, multi-user, immersive, and interactive. This novel and unique instrument will be used for research into scientific visualization/auralization and data exploration, and as a research environment for behavioral and cognitive scientists. It will also serve as a research and performance space for artists exploring new forms of art. In particular, the Allosphere has been carefully designed to allow for immersive music applications.

In this paper, we give an overview of the instrument, focusing on the audio subsystem. We present first results and our experiences in developing and using the Allosphere in several prototype projects.

1. INTRODUCTION

The Allosphere is a novel environment that will allow for synthesis, manipulation, exploration and analysis of large-scale data sets providing multi-user immersive interactive interfaces for research into immersive audio, scientific visualization, numerical simulations, data mining, visual and aural abstract data representations, knowledge discovery, systems integration, human perception, and last but not least, artistic expression.

The space enables research in which art and science contribute equally. It serves as an advanced research instrument in two overlapping senses. Scientifically, it is an instrument for gaining insight and developing bodily intuition about environments into which the body cannot venture: abstract, higher-dimensional information spaces, the worlds of the very small or very large, the very fast or very slow, from nanotechnology to theoretical physics, from proteomics to cosmology, from new materials to new media. Artistically, the Allosphere is an instrument for the creation and performance of new avant-garde works and the development of entirely new modes and genres

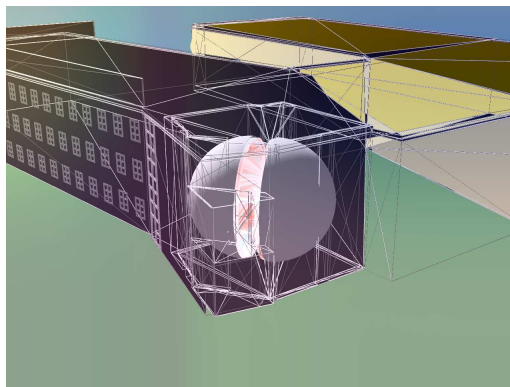


Figure 1: A virtual rendering of the Allosphere



Figure 2: Looking into the Allosphere from just outside the entrance.

of expression and forms of immersion-based entertainment, fusing future art, architecture, science, music, media, games, and cinema.

The Allosphere is situated at one corner of the California Nanosystems Institute building at the University of California Santa Barbara (see virtual model in Figure 1), surrounded by a number of associated labs for visual/audio computing, robotics and distributed systems, interactive visualization, world modeling, and media post-production. The main presentation space consists of a three-story near-to-anechoic room containing a custom-built close-to-spherical screen, ten meters in diameter (see Figure 3). The sphere environment integrates visual, sonic, sensory, and interactive components. Once fully equipped, the Allosphere will be one of the largest immersive instruments in the world. It provides a truly 3D 4π stereoradians surround-

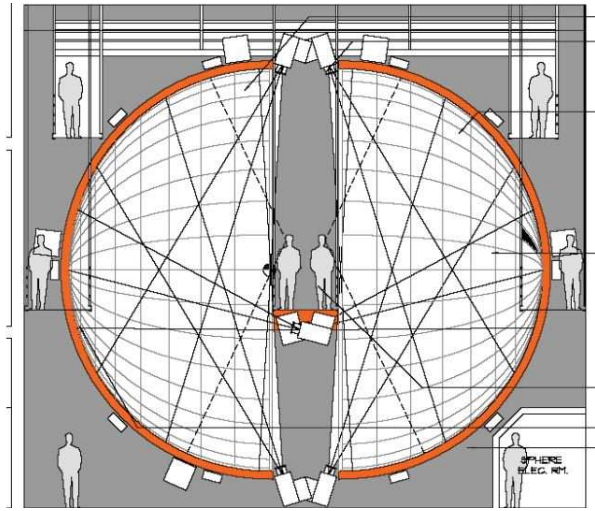


Figure 3: The Allosphere

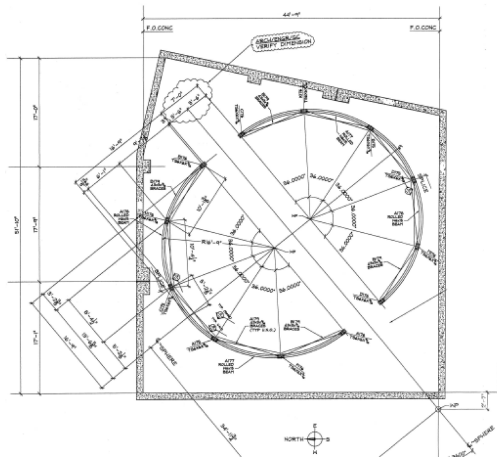


Figure 4: Horizontal section of the Allosphere

projection space for visual and aural data and accommodates up to 30 people on a bridge suspended in the middle of the instrument.

The space surrounding the spherical screen is close to cubical, with an extra control/machine room in the outside corner, pointed to by the bridge structure. The whole outer space is treated with sound absorption material (4-foot wedges on almost all inner surfaces), forming a quasi-anechoic chamber of large proportions. Mounted inside this chamber are two 5-meter-radius hemispheres, constructed of perforated aluminum that are designed to be optically opaque (with low optical scatter) and acoustically transparent. Figure 4 is a detailed drawing showing a horizontal slice through the Allosphere at bridge height. The two hemispheres are connected above the bridge, forming a completely surround-view screen.

We are equipping the instrument with 14 high-resolution video projectors mounted around the seam between the two hemispheres, projecting onto the entire inner surface. The loudspeaker array is placed behind the aluminum screen,

suspended from the steel infrastructure in rings of varying density.

The Allosphere represents in many senses a step beyond already existing virtual environments such as the CAVE [1], even in their more recent “fully immersive” reincarnations [2], especially regarding its size, shape, the number of people it can accommodate, and its potential for multimedia immersion. In this paper, we focus on a particular aspect of the multimedia infrastructure, the audio subsystem.

Although the space is not fully equipped at this point, we have been experimenting and prototyping with a range of equipment, system configurations, and applications that pose varying requirements. We envision the instrument as an open framework that is in constant evolution, with major releases signaling major increments in functionality.

2. A TRULY MULTIMEDIA/MULTIMODAL SYSTEM

An important aspect of the Allosphere is its focus on multimedia processing, as it combines state-of-the-art techniques both on virtual audio and visual data spatialization. There is extensive evidence of how combined audio-visual information can influence and support information understanding [3]. Nevertheless, most existing immersive environments focus on presenting visual data. The Allosphere is a completely interactive multimodal data mining environment [4].

The diagram depicted in Figure 5 illustrates the main subsystems and components in the Allosphere, as well as their interactions. The diagram is a simplified view of the integrated multi-modal/media system design. The exact interactions among the various media data (visual, aural, and interactive) is dependent on the particular individual applications to be hosted.

The remainder of this section will briefly introduce each of those components and subsystems, as well as the way they interact. In Section 3 we will then discuss the audio subsystem (*Allo.A*) in detail.

The main requirements for the Allosphere visual subsystem (*Allo.V*) are fixed both by the building and screen characteristics and the final image quality targeted. The sphere screen area is 320.0 m² and its reflective gain, FOV averaged, is 0.12.

The Allosphere projection system (*Allo.V.D.P*) requires image warping and blending to create the illusion of a seamless image from multiple projectors. We have designed a projection system consisting of 14 3-chip DLP active stereo projectors with 3000 lumens output and SXGA+ resolution (1400x1050) each. The projectors are being installed with an effective projector overlap/blending loss coefficient of 1.7.

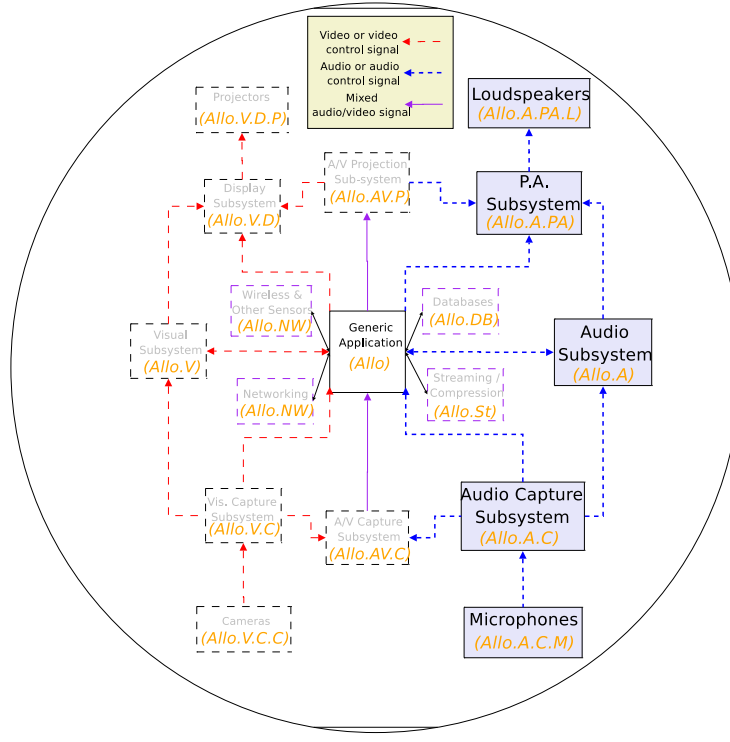


Figure 5: The Allosphere Components with a highlighted Audio Subsystem

A typical multi-modal application in the Allosphere (called Allo for short in Figure 5) will integrate several distributed components, sharing a LAN:

- back-end processing (data/content accessing)
- output media mapping (visualization and/or sonification)
- A/V rendering and projection management.
- input sensing, including real-time vision and camera tracking (related to *Allo.V.V*), real-time audio capture and tracking (related to *Allo.A.C*), a sensor network including different kind of regular wireless sensors as well as other presence and activity detectors (related to *Allo.SN*).
- gesture recognition/control mapping
- interface to a remote (scientific, numerical, simulation, data mining) application

It follows from our specification requirements – and our experiments have confirmed this view – that off-the-shelf computing and interface solutions are insufficient to power the sphere. Allosphere applications not only require a server cluster dedicated to video and audio rendering and processing, but also a low-latency interconnection fabric so that data can be processed on multiple computers (in a variety of topologies) in real time, an integration middleware, and an application server that can control the system in a flexible and efficient way.

The computation infrastructure will consist of a network of distributed computational nodes. Communication be-

tween processes will be accomplished using standards such as MPI. The Allosphere Network (*Allo.NW*) will have to host not only this kind of standard/low-bandwidth message passing but also multichannel multimedia streaming. The suitability of Gigabit Ethernet or Myrinet regarding bandwidth and latency is still under discussion. In our first prototypes, Gigabit has proven sufficient, but our projections show that it will become a bottleneck for the complete system, especially when using a distributed rendering solution to stream highly dynamic visual applications. We are considering custom hardware technologies as a possible necessity in the future.

3. THE AUDIO SUBSYSTEM

Our goal for the Allosphere is to provide “sense-limited” resolution in both the audio and visual domains. This means that the spatial resolution for the audio output must allow us to place virtual sound sources at arbitrary points in space with convincing synthesis of the spatial audio cues used in psychoacoustical localization. Complementary to this, the system must allow us to simulate the acoustics of measured or simulated spaces with a high degree of accuracy.

In a later stage we also plan to complement the audio subsystem with a microphone array in order to arrive at fully immersive audio [5]. However, this components is still at the very early stages of design and will therefore not be discussed in this section.

3.1. Acoustical Requirements

In order to provide “ear-limited” dynamic, frequency, and spatial extent and resolution, we require the system to be able to reproduce in excess of 100 dB sound pressure level near the center of the sphere, to have acceptable low- and high-frequency extension (-3 dB points below 80 Hz and above 15 kHz). We designed the spatial resolution to be on the order of 3 degrees in the horizontal plane (i.e., 120 channels), and 10 degrees in elevation. To provide high-fidelity playback, we require audiophile-grade audio distribution formats and amplification, so that the effective signal-to-noise ratio exceeds 80 dB, with a useful dynamic range of more than 90 dB.

To be useful for data sonification [6] and as a music performance space, the decay time (the “T60 time”) of the Allosphere was specified to be less than 0.75 seconds from 100 Hz to 10 kHz [7]. This is primarily an architectural feature related to the properties of the sound absorbing treatment in the quasi-anechoic chamber, which was designed to minimize the effect of the aluminum projection screen. The perforations on the screen have also been designed to minimize its effect across most of the audible spectrum. Initial experiments confirm that the absorption requirements have indeed been met.

3.2. Spatial Sound Processing

Since the AlloSphere is to foster the development of integrated software for scientific data sonification and auditory display, as well as artistic applications, it is essential that the software and hardware used for audio synthesis, processing, control, and spatial projection be as flexible and scalable as possible. We require that the audio software libraries support all popular synthesis and processing techniques, that they be easily combined with off-the-shelf audio software written using third-party platforms such as Csound, Max/MSP, and SuperCollider, and that they support flexible control via (at least) the MIDI and Open Sound Control (OSC) protocols. Due to the sophistication of the audio synthesis and processing techniques used in AlloSphere applications, and the expected very large number of final output channels, we require that the core audio libraries support easy inter-host streaming of large numbers of channels of high-resolution (24-bit, 96 kHz) audio, probably using both the CSL/RFS and SDIF networked audio protocols.

There are three main techniques for spatial sound reproduction used in current state-of-the-art systems: (1) vector-based amplitude panning [8], (2) ambisonic representations and processing [9], and (3) wave field synthesis (see [10] and [11]). Each of these techniques provides a different set of advantages and presents unique challenges when scaling up to a large number of speakers and of virtual sources. We have developed a flexible software framework based on the CREATE Signal Library (CSL, [12]), in

which different techniques, sets of psychoacoustical cues, and speaker layouts can be combined and swapped at run time (see Castellanos thesis [13]).

3.2.1. Vector-based Amplitude Panning

Practical VBAP systems allow interactive performance with multiple moving sound sources, which are mapped and played back over medium-scale projection systems. VBAP has been mainly promulgated by groups in Finland and France and is used effectively in 8-32-channel CAVE virtual environments.

The drawbacks of VBAP are that it does not directly answer the question of how to handle distance cues (relatively easy to solve for distant sources and low Doppler shift), and that it provides no spatialization model for simulating sound sources inside the sphere of loudspeakers. This is a grave problem for our applications, but also a worthy topic for our research. The question boils down to how to spread a source over more than 3 speakers without limiting the source position to the edges of the surface described by the chosen set of speakers.

The VBAP algorithm involves a search among the geometrical representations of the speakers defining the playback configuration, and then some simple matrix math to calculate the relative gains of each of the three chosen speakers. There are several open-source implementations of VBAP that support multiple sources (with some interactive control over their positions), and flexible speaker configurations involving up to 32 channels.

Members of our research group implemented a system in which the user can move and direct a number of independent sound sources using a data glove input device, and play back sound files or streaming sound sources through VBAP, using a variable number of loudspeakers specified in a dynamic configuration file (see McCoy thesis [14]). VBAP can be integrated with a spatial reverberator software, allowing early reflections from a reverberator to be individually panned, though this gets computationally very expensive with many sources, complex room simulations, or rapid source (or listener) motion.

Because VBAP is so simple, most implementations are monolithic 1-piece packages. This is obviously unacceptable for our purposes, so we needed to consider both (1) how the VBAP system scales to large numbers of sources, rapid source motion, and many output channels, and (2) how such a scaled-up application can best be distributed to a peer-to-peer server topology streaming data over a high-speed LAN. The scalability of VBAP encoding software is excellent, since the block-by-block processing is very simple, and the computation of new output weights for new or moving sources can be accelerated using well-understood geometrical search techniques. For the case of many sources or rapid source or listener motion, VBAP scales linearly, because each source is encoded into 3 channels, meaning that many mappers each write 3 channels

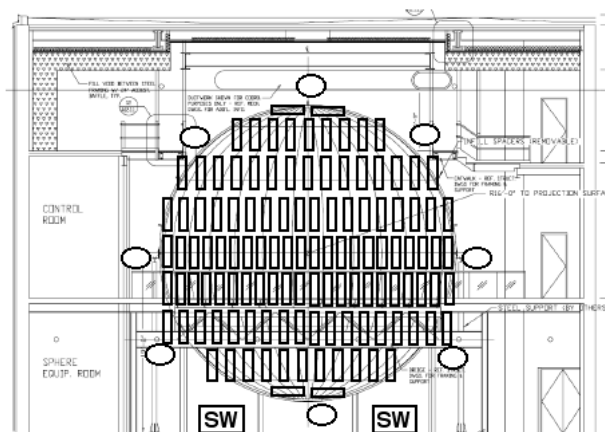


Figure 6: *Allosphere speaker placement design, initial iterations*

into a many-channel output buffer. Alternatively, if the servers are distributed, each mapper sends 3 channels over the LAN to its output server. If the output servers are themselves distributed (each taking over a subset of the sphere's surface), then most encoding servers will stream to a single output server.

Computational distribution of a VBAP-based spatial reverberator is more difficult, since by definition the individual reflections are not localized to a small number of channels; indeed, if you calculate a reasonable number of reflections (e.g., 64 or more) for a complex room model, you can assume that the reflections will approximate an even distribution among all channels, leading us back to a monolithic output server topology. We look forward to attacking this scalability and partitioning issue in the full system. For the time being, we run the reverberator on a single server.

The assumptions of the speaker elements and system configuration for playing VBAP are that the elements be identical full-range speakers, and that they be placed in triangles of more-or-less equal size in all directions. The speaker density can be made a function of height, however, leading to somewhat poorer spatialization accuracy above (and possibly below) the listener. All that being said, since VBAP makes so few assumptions about the constructed wave, it supports non-uniform speaker distributions quite well. Directional weighting functions to compensate for an uneven distribution of speakers can be built into the VBAP amplitude matrix calculations, and the fidelity of spatial impression is a directional function of both the speaker density and regularity of spacing. In our earliest designs for the sphere, we ran a set of programs to tessellate spherical surfaces, leading to the 80-channel configuration shown in Figure 6. Note the two regular rings above and below the equator; one can rotate the upper hemisphere by 1/2 the side length to form a zigzag pattern here (which handles VBAP better). Continuing this process, we can design and evaluate further regular subdivisions of a sphere.

3.2.2. Ambisonics

To implement the Ambisonic representation, one needs to write an encoder that takes a monophonic signal and a virtual source position and generates a multi-channel signal whose format and size depend on the order of encoding chosen. The decoder takes the encoded n -channel signal and a list of output speaker positions. For each speaker, it uses complementary trigonometric equations to determine the contribution of each harmonic signal to the output at the given speaker's location.

As with VBAP, graduate researchers from our group (see [15]) have implemented higher- (up to 11th-) order ambisonic processing and decoding in C++ using the CSL framework. The encoder and decoder are separate classes, and utility classes exist for processing (e.g., rotating the axes of) Ambisonic-encoded sound. We also implemented the algorithm for Max/MSP [16], and there are also open-source implementations in both SuperCollider and PD. Ambisonic encoders and decoders are all relatively simple, and can be decoupled from one another. For a simple scaled-up system, multiple 3rd-order encoders would run on machines in our server farm, each of them streaming a 16-channel signal to the output driver(s). These signal busses can be summed and then distributed to one or more output decoders. The scalability to higher orders is well understood, and scales with the number of channels required by the representation.

One of the main benefits of the Ambisonic representation is that it scales very well for large numbers of moving sources. The encoding is based on the order of the representation used. The decoding scales well to large numbers of speakers because decoders are independent of one another, each receiving the same set of inputs; there are no obvious scalability limits, either in terms of CPU processing or LAN bandwidth requirements.

Ambisonic decoders work best with a regular and symmetrical loudspeaker configuration; software and hardware decoders for 2, 4, 8, etc. channels are readily available. There is no way in the processing algorithms to compensate for irregular speaker placement. What is interesting is the fact that very large speaker arrays can especially benefit from higher-order ambisonic processing, using ever-higher orders of spherical harmonics to encode the sound field, and then decoding it using these factors to play out over a (regular and symmetrical) many-channel speaker array.

3.2.3. Wavefield Synthesis

Synthesis or processing of wave field signals involves solving an equation called the Kirchhoff-Helmholtz integral, which is done using computationally costly differential equation approximation techniques. Recent advances make this possible up to 128 channels on a current-model PC dual-processor server.

There are existing open-source implementations of WFS (e.g., [17]), and we are in the process of porting these to work within the CSL framework. WFS also requires compensation for speaker characteristics and room effects, but this process is well understood and computationally tractable (adding a level of FIR filters to the large convolutions involved in the implementation). Basic WFS processing for single sources and medium-range numbers of output channels can be implemented on a single processor. The scalability for the case of multiple sources is thought to be poor, requiring multiple processing servers, all of whose output channels are mixed. We are not aware of existing work attempting to partition a WFS server onto multiple hosts.

The processing of WFS math for large 2-D and 3-D systems is still a complex problem requiring efficient solutions to a set of matrix equations using techniques referred to as fast multipole methods (FMM), which is an active area of research in computational mathematics [18]. Implementing a large scale (> 200 speakers) 2-D system is straightforward, though we have yet to investigate the scalability and distribution issues. The development of distributed FMM-based 3-D solutions is planned as a graduate thesis project in our group. First experiments appear to indicate that a good solution is to combine wavefield synthesis in the horizontal plane at ear-level and ambisonics for the vertical axis, where the spatial resolution is not so critical.

The requirements on the loudspeaker configuration for WFS are simple; one needs the densest possible packing of speakers around the circumference of a circle, preferably less than half of the wavelength of the lowest frequency in the signal. The limits on speaker spacing for WFS are simple and physical. It is generally assumed that we do not perceive the spatialization for sounds below about 200 Hz, since the wavelength is so much larger than the distance between our ears. In order to get any subtlety in the spatial impression generated using WFS, we need the speakers to be spaced at intervals on the same order as the longest wavelength we desire to reconstruct with spatial accuracy. These two facts give us upper and lower bounds on the speaker spacing interval; if we want to have at least a few octaves where we get accurate spatial wavefront reconstruction, speaker spacing on the order of 1 foot or less is a requirement.

In any case, the Allosphere speaker count and configuration supports the use of any of these for sound spatialization. This implies high speaker density (on the order of one source per square yard of surface, or about 425 channels), and a semi-regular and relatively symmetrical speaker layout.

Although our final system will be designed mainly for work with wavefield synthesis, at the current time we still do not have a sufficient number of loudspeakers in the Allosphere to recreate a wavefield. Therefore, at this exploratory stage and for prototyping purposes in the space,

we are currently using ambisonics.

3.3. Speaker System

It has been a major project to derive the optimal speaker placements and speaker density function for use with mixed-technology many-channel spatialization software (see discussion and calculations in [15]). Our driver placement design comprises between 425 and 500 speakers arranged in several rings around the upper and lower hemispheres, with accommodations at the “seams” between the desired equal and symmetrical spacing and the requirements of the support structure. The loudspeakers will be mounted behind the screen.

We have projected densely packed circular rings of speaker drivers running just above and below the equator (on the order of 100-150 channels side-by-side), and 2-3 smaller and lower-density rings concentrically above and below the equator. The main loudspeakers have limited low-frequency extension, in the range of (down to) 200-300 Hz. To project frequencies below this, four large subwoofer(s) are mounted on the underside of the bridge.

At this moment, because of timing and construction constraints, we have installed a prototype system with only 16 full range speakers installed along the three different rings mentioned above and two subwoofers under the bridge. Those speakers are connected to Firewire interfaces that support 32 channels.

For the imminent growth of the prototype into the full system, we plan to switch to passive speaker elements wired to a set of 8-16 networked digital-to-analog converter (DAC) amplifier boxes, each of which supports in the range of 32-128 channels and has a Firewire interface. As an alternative, we are also considering building custom interface boxes consisting of a Gigabit Ethernet interface, digital/analog convertor, power amplifier, and step-up transformer (this would be based on a design developed at CNMAT for their 120-channel loudspeaker array [19]).

4. TESTBED APPLICATIONS

The Allosphere’s main function is the analysis, synthesis, simulation, and processing of complex multidimensional data in an interactive and immersive environment. Content and demand will drive the technological development just as it has driven its design and conception. For those reasons, specific application areas are essential in the development of the instrument as they define the functional framework in which the instrument will be used.

In the first iteration over the prototype we have set up an environment consisting of the following elements:

- * 4 active stereo projectors (Christie Digital Mirage S+2K), 3000 ANSI lumens, DLP

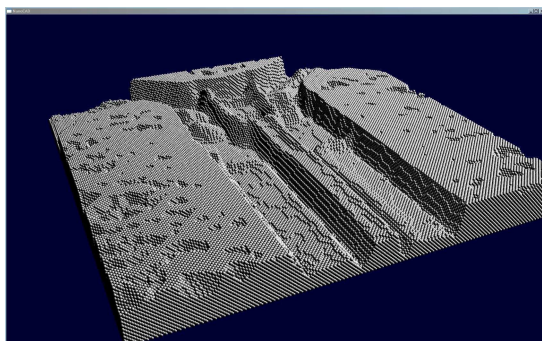


Figure 7: *Rendering of a 1M atom silicon nanostructure in real-time on a single CPU/GPU (Allosphere rendering occurs in stereo projection)*

- * 2 rendering workstations (HP 9400), AMD Opteron 64@2.8Ghz, NVidia Quadro FX-5500
- * 1 application manager + Audio Renderer (Mac Pro), Intel Xeon Quad Core @3Ghz
- * 2 10-channel firewire audio cards.
- * 16 full-range speakers + 2 subwoofers
- * Several custom-developed wireless interfaces.

The research projects described below make use of this prototype system to test the functionality and prove the validity of the instrument design.

In the first project, we are developing an immersive and interactive software simulation of nano-scaled devices and structures, with atom-level visualization of those structures implemented on the projection dome of the Allosphere (see figure 7). When completed, this will allow the user to stand in the middle of a simulation of a nano-scaled device and interact with the atoms and physical variables of that device.

Our science partners are implementing algorithms for nano-material simulations involving molecular dynamics and density functional theory using GPUs, transforming a single PC workstation into a 4 Teraflop supercomputer. This allows us to run nanoscale simulations that are 2-3 orders of magnitude faster than current implementations. We will also be able to use this extra computational power to solve for the physical properties of much larger structures and devices than were previously possible, allowing nano-system engineers to design and simulate devices composed of millions of atoms. Sound design will play an important role in such simulations and visualizations. The sound system will be used to bring important temporal phenomena to user's attention and to pinpoint it precisely with 3D sound. For instance, to alleviate the difficulty of finding specific molecules in the vast visual space of the Allosphere, subtle auditory clues can alert the user to the emergence or presence of a specific molecular event in a particular direction (which is especially relevant when the object is behind the user's back!).

In a second research project, called *AlloBrain* (see Figure

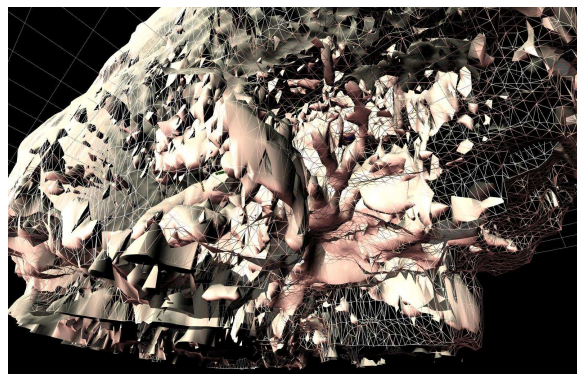


Figure 8: *Screen capture of the AlloBrain interactive recreation of the human brain from fMRI data*

2 and 8), we experiment with macroscopic, organic data sets, reconstructing an interactive 3D model of a human brain from fMRI data[20]. The current model contains two layers of tissue blood flow, and we have created an interactive environment where twelve “agents” navigate the space and gather information to deliver back to the researchers. The simulation contains several generative audio-visual systems. These systems are stereo-optically displayed and controlled by two wireless (Bluetooth) input devices that feature custom electronics, integrating several MEMs sensor technologies. Apart from navigating through the brain and controlling the agents, the controllers also allow you to move the ambient sounds spatially around the sphere.

This virtual interactive prototype illustrates some of the key agenda points of the Allosphere, such as multimedia/multimodal computing, interactive immersive spaces and scientific data understanding through art.

Finally, we are also starting a project focused on molecular dynamics. We are extending the Vmd [21] package through the use of Chromium in order to have seamless visualization of complex protein molecules and their interactions, immersively supported with direct manipulation and spatial sonification by the Allosphere.

5. CONCLUSIONS

Once fully equipped and operational, the Allosphere will be one of the largest immersive instruments in existence. But aside from its size, it also offers a number of features that make it unique in many respects. In particular, it features immersive spherical projection, multimodal processing including stereoscopic vision, 3D audio, and interaction control, and multi-user support for up to 30 people. In this paper, we have focused on the audio infrastructure in the Allosphere, discussing the requirements, approaches, and initial results.

We envision the Allosphere as a vital instrument in the future advancement of fields such as nanotechnology or

bio-imaging and it will stress the importance of multimedia in the support of science, engineering, and the arts. We have demonstrated first results in the form of projects of highly diverse requirements. These initial results feed back into the prototyping process but also clearly support the validity of our approach.

Although the Allosphere is clearly still in its infancy, we believe that the presented results are already meaningful and important, and will inform other integrative endeavors in the computer music research communities. The development of our prototype test-bed applications is geared towards an open generic software infrastructure capable of handling multi-disciplinary multi-modal applications.

6. REFERENCES

- [1] C. Cruz-Neira, D. J. Sandin, T. DeFanti, R. Kenyon, and J. Hart, "The cave: Audio visual experience automatic virtual environment," *Communications of the ACM*, no. 35, pp. 64–72, 1992.
- [2] J. Ihren and K. Frisch, "The fully immersive cave," in *Proc. 3rd International Immersive Projection Technology Workshop*, 1999, pp. 59–63.
- [3] H. McGurk and T. McDonald, "Hearing lips and seeing voices," *Nature*, no. 264, pp. 746–748, 1976.
- [4] E. Wegman and J. Symanzik, "Immersive projection technology for visual data mining," *Journal of Computational and Graphical Statistics*, March 2002.
- [5] H. Teutsch, S. Spors, W. Herbordt, W. Kellermann, and R. Rabenstein, "An integrated real-time system for immersive audio applications," in *Proc. 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, 2003.
- [6] J. Ballas, "Delivery of information through sound," in *Auditory Display: Sonification, Audification and Auditory Interfaces*, G. Kramer, Ed. Reading, MA: Addison Wesley, 1994, vol. XVIII, pp. 79–94.
- [7] J. Blauert, *Spatial Hearing*. MIT Press, 2001.
- [8] V. Pulkki and T. Hirvonen, "Localization of virtual sources in multi-channel audio reproduction," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 105–119, 2005.
- [9] D. Malham and A. Myatt, "3-d sound spatialization using ambisonic techniques," *Computer Music Journal (CMJ)*, vol. 19, no. 4, pp. 58–70, 1995.
- [10] S. Spors, H. Teutsch, and R. Rabenstein, "High-quality acoustic rendering with wave field synthesis," in *Proc. Vision, Modeling, and Visualization Workshop*, 2002, pp. 101–108.
- [11] R. Rabenstein, S. Spors, and P. Steffen, *Selected methods of Acoustic Echo and Noise Control*. Springer Verlag, 2005, ch. Wave Field Synthesis Techniques for Spatial Sound Reproduction.
- [12] S. T. Pope and C. Ramakrishnan, "The Create Signal Library ("Sizzle"): Design, Issues and Applications," in *Proceedings of the 2003 International Computer Music Conference (ICMC '03)*, 2003.
- [13] J. Castellanos, "Design of a Framework for Spatial-Audio Rendering," Master's thesis, University of California Santa Barbara, 2006.
- [14] D. McCoy, "Ventriloquist: A performance interface for real-time gesture-controlled music spatialization," Master's thesis, University of California Santa Barbara, 2005.
- [15] F. Hollerweger, "Periphonic sound spatialization in multi-user virtual environments," Master's thesis, Austrian Institute of Electronic Music and Acoustics (IEM), 2006.
- [16] G. Wakefield, "Third-order ambisonic extensions for max/msp with musical applications," in *Proceedings of the 2006 ICMC*, 2006.
- [17] M. A. J. Baalman, "Application of wave field synthesis in electronic music and sound art," in *Proc. 2003 International Computer Music Conference (ICMC)*, 2003.
- [18] N. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier Publisher, 2006.
- [19] A. Freed, "Design of a 120-channel loudspeaker array," CNMAT, University of California Berkeley, Tech. Rep., 2005.
- [20] G. Wakefield, W. Smith, J. Thompson, L. Putnam, D. Overholt, J. Kuchera-Morin, and M. Novak, "The allobrain: An interactive plurphonic, pluriscopic, immersive environment," in *Proceedings of the 2007 ICMC*, 2007, pending acceptance.
- [21] W. Humphrey, A. Dalke, and K. Schulten, "Vmd - visual molecular dynamics," *Journal of Molecular Graphics*, no. 14, pp. 33–38, 1996.