Eric Lyon,[*] Terence Caulkins,[†] Denis Blount,[**] Ivica Ico Bukvic,[*] Charles Nichols,[*] Michael Roan,[††] and Tanner Upthegrove[*]

[*]School of Performing Arts
Institute for Creativity, Arts, and Technology
Virginia Polytechnic Institute and State University
195 Alumni Mall
Blacksburg, Virginia 24061, USA
{ericlyon, ico, csnii, tannerupthegrove}@vt.edu
[†]Arup
77 Waters St.
New York, New York 10005, USA
terence.caulkins@arup.com
[**]Arup
719 Second Avenue, Suite 400
Seattle, Washington 98104, USA
denis.blount@arup.com
[††]Department of Mechanical Engineering
Virginia Polytechnic Institute and State University
111 Randolph Hall
Blacksburg, Virginia 24061, USA
mroan@vt.edu

Genesis of the Cube: The Design and Deployment of an HDLA-Based Performance and Research Facility

Abstract: The Cube is a recently built facility that features a high-density loudspeaker array. The Cube is designed to support spatial computer music research and performance, art installations, immersive environments, scientific research, and all manner of experimental formats and projects. We recount here the design process, implementation, and initial projects undertaken in the Cube during the years 2013–2015.

The Cube is a performance and research space featuring a high-density loudspeaker array (HDLA) comprising approximately 150 loudspeakers. Its dimensions are 50 × 40 × 42 ft (length × width × height, about 15 × 12 × 13 m) inside its surrounding catwalks. The grid level, which suspends the highest loudspeakers in the space, is at a height of 32 ft (9.8 m). The space comfortably supports a seated audience of 80, with a maximum occupancy of 198. The Cube is housed in the Moss Arts Center (MAC) at Virginia Polytechnic Institute and State University (Virginia Tech), a major renovation completed in 2013 on the site of Shultz Hall, a dining hall completed in 1962. The MAC was designed by Snøhetta, Arup, and Theatre Projects Consultants, all subcontracted to STV Architects. After completion of the MAC construction, the Cube's 3-D audio system was designed by a working group consisting of faculty and staff members at Virginia Tech and acoustic consultants from Arup. Managed by the Institute for Creativity, Arts, and Technology (ICAT) and the MAC, the Cube supports computer music, arts installations, immersive environments, scientific research, and open-format experimentation. By articulating the context, motivations, design, implementation, evaluation, and uses of the Cube, we hope to encourage the building of more spaces with HDLA capabilities, which can promote the creation of new forms of spatiomusical experience through

Computer Music Journal, 40:4, pp. 62–78, Winter 2016
doi:10.1162/COMJ_a_00394

the expansion of contemporary computer music practices.

Conception

The Cube was initially conceived of as a digital and experimental performance environment that could demonstrate new approaches to audiences through cross-disciplinary collaboration between faculty, students, guest artists, and visiting researchers. It would require a noise-free environment with a neutral, variable room acoustic and a robust, flexible technical infrastructure. The laboratory component would require the ability to change from performance mode quickly and easily, and provide an "equipment-agnostic" infrastructure to support evolving new technologies. High-bandwidth connectivity would allow both digital and analog content to be aggregated from multiple spaces on campus, as well as from the outside world.

Program requirements for the Cube were developed to anticipate the needs of a diverse artistic and research community, including the following use cases: live performance including theater, dance, and music; 3-D sound installations; immersive audio for multiuser virtual- and augmented-reality experiences; 3-D audio research into new modes of sound spatialization; and multidisciplinary research incorporating large-scale visualization, auralization, and sonification of big data sets. Designing a system to these requirements would involve finding a balance between spatial audio resolution, maximum required sound-pressure levels, ease of rigging, and cost, all while maintaining high sound quality.

Arup produced a number of test fits to demonstrate the range of activities implicit in the program requirements including cinema mode, immersive performance and 3-D audio, theater with thrust stage, immersive video and audio installation, amplified music, orchestra rehearsal, and recording. These test fits became the basis of a design that, in collaboration with Snøhetta and Theatre Projects Consultants, resulted in the final version of the Cube shell that is in use today. The Cube was officially opened on 21 November 2013, with a commissioned work by video artist Joan Grossman, entitled *This Edge I Have to Jump*, with multichannel music provided by Eric Lyon and Charles Nichols, and video programming by Carol Burch-Brown, a Virginia Tech professor of studio art and creative technology (see Figure 1).

Designing the Technological Infrastructure for the Cube

The design for the Cube's technological infrastructure was built on a vision articulated by Ivica Ico Bukvic in 2008 for a multidisciplinary, multiuser, collaborative research environment that would combine visualization, motion tracking, and full-scale immersive 3-D audio. The envisioned infrastructure would support multiple audio spatialization approaches, including wave-field synthesis (WFS) and Ambisonics. It would also utilize motorized ultrasonic loudspeakers, Holosonics Audio Spotlights, with the goal of allowing the space to serve as a blank canvas for artists and researchers in support of 3-D audio research and artistic expression. In spring 2013, the Cube's audio infrastructure team, consisting of ICAT director Ben Knapp, Virginia Tech faculty members Bukvic, Lyon, and Nichols, ICAT's media engineer Tanner Upthegrove, and Arup acoustic consultants Denis Blount and Terence Caulkins, began to design the audio system in detail. By April 2014, the team finalized the setup and moved towards implementation.

Spatial Audio System: Layout and Spatial Resolution

The high-density spatial audio system was designed to provide full-bandwidth audio for listeners located inside the Cube. This audio comes from a 3-D HDLA, currently comprising 124 satellite JBL SCS-8 loudspeakers, 10 floor-standing JBL LSR6328P loudspeakers, 4 Meyer UMS-1P subwoofers, and a JBL 5628 subwoofer, for a total of 139 independently addressable channels. The system was designed to be fully reconfigurable to suit the needs of future research by using expandable, IP-based audio networking.

Figure 1. The Cube opening with four-way projection and multichannel audiovisual installation by Joan Grossman. (Photo courtesy of Virginia Tech.)

The initial layout of loudspeakers, comprising 124 SCS-8s and 4 UMS-1P subwoofers, is shown in Figures 2 and 3. The high-resolution 64-channel loudspeaker array laid out on Level 1 (L1) is designed to support WFS playback in the main ground-floor listening area. The remaining 64 channels are evenly distributed as a cuboid periphonic array around L2, L3, and Grid level, and are designed to supplement the horizontal array for rendering high-order Ambisonics (HOA), 3-D WFS (Rohr et al. 2013), and vector base amplitude panning (VBAP) rendering. All channels of the system can be used in conjunction with each other to create hybrid spatialization approaches depending on the needs of the application.

The higher spatial resolution of the array near the horizontal plane serves to limit spatial aliasing artifacts arising from typical WFS rendering (Wittek, Rumsey, and Theile 2007; Corteel, Kuhn-Rahloff, and Pellegrini 2008). It is also intended to provide a higher degree of accuracy for virtual source rendering near the horizontal plane, and addresses the greater ability of humans to localize in the

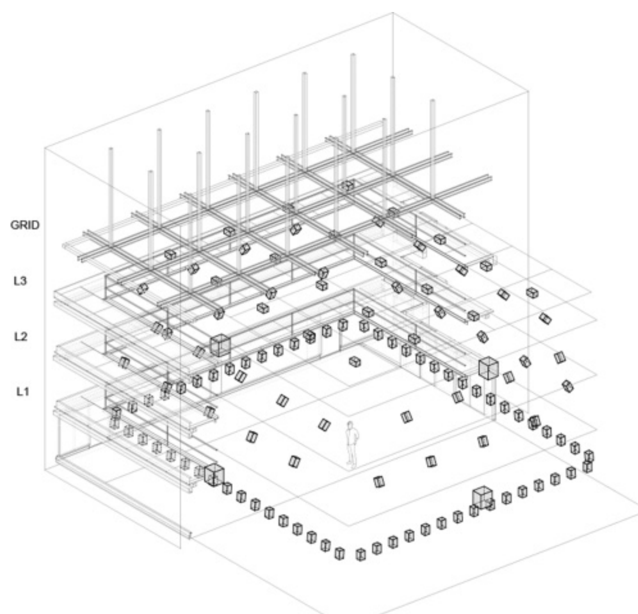
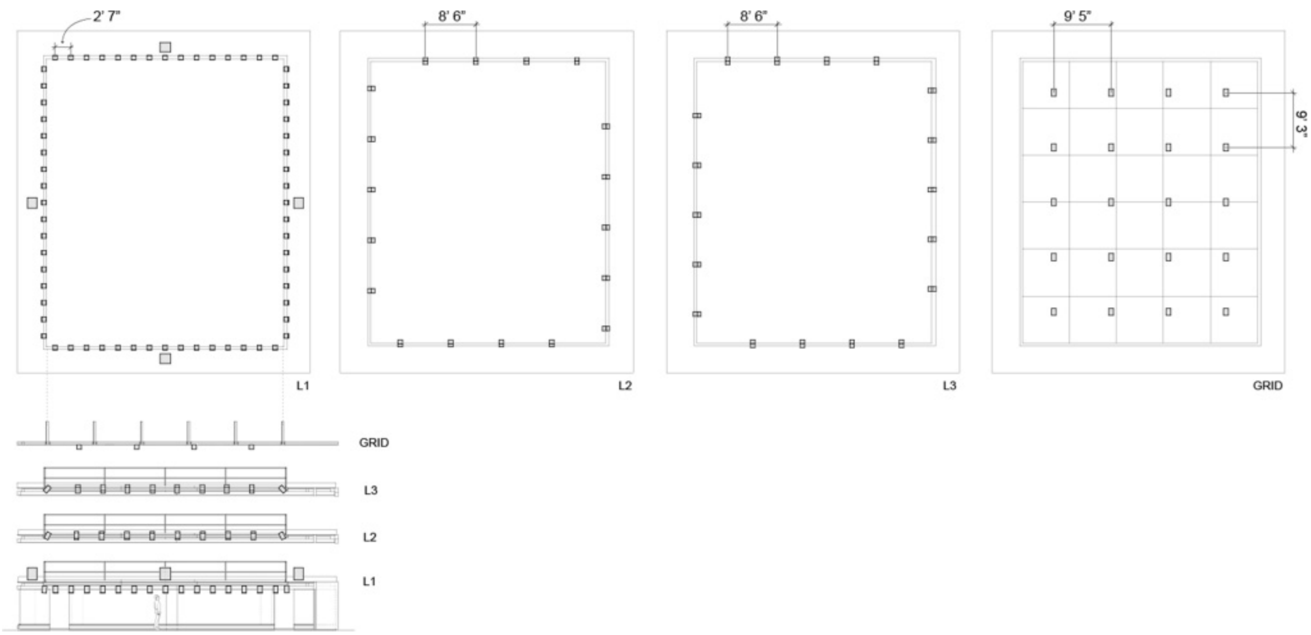


Figure 2

Figure 2. Isometric view of the main audio system.

Figure 3. Plan and section view of the main audio system.



horizontal plane compared with the median and frontal planes (Blauert 1996, pp. 40–47). The Cube is designed to accommodate both traditionally seated audiences as well as standing audiences that are free to move around the space. For standing audiences, the L1 array is nearly in the horizontal plane, as safety regulations prevented locating the loudspeakers directly on the horizontal plane. For seated audiences, the floor-standing array (not shown in Figures 2 and 3) renders sound in the horizontal plane.

Spatial Audio System: Maximum Sound Pressure Levels

The maximum full-range sound-pressure level (SPL) produced by the main audio system at ear height in the center of the room depends on the number of loudspeakers addressed, which varies depending on the type of spatialization technique being used. Some estimates of maximum SPLs produced by the system for different types of spatial sources are given in Table 1. The full set of 139 loudspeakers was used for the maximum SPL estimate.

Table 1. Cube Audio System SPL Targets for Various Playback Configurations

Virtual Sound Source Location	*Estimated Maximum Sound Pressure Levels*
3-D WFS plane wave at ear height emanating from shortest dimension (6.5 m distance)	96 dB LAeq
3-D WFS plane wave angled down from upper corner (16 m distance)	90 dB LAeq
All loudspeakers active at 0 dBFS (uncorrelated)	110 dB LAeq
Subwoofer system—all loudspeakers active (uncorrelated)	115 dB LAeq

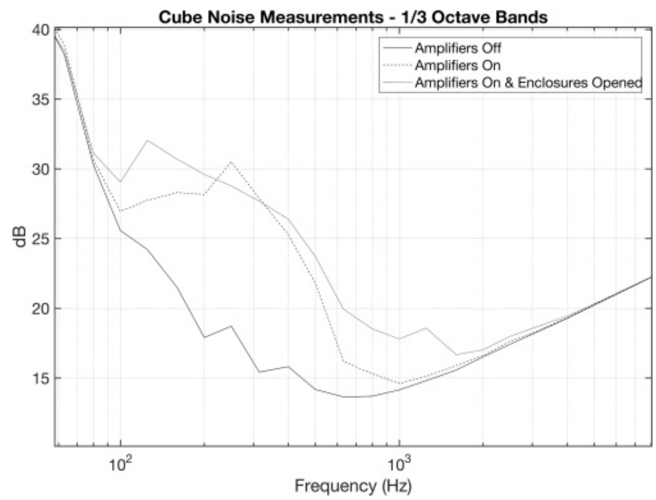
Cube Acoustical Measurements

Acoustical measurements were conducted in the Cube using a class 1, four-channel analyzer. The system's main components were two GRAS 40AQ microphones and a National Instruments 9234 DAC with a sample rate of 50 kHz at 24-bit resolution. The National Instruments Sound and Vibration Measurements Suite was used to calculate

Figure 4. Background noise in the Cube.

Table 2. Sound Level Measurements in the Cube

Measurement	*SPL*	*SPL: Two Microphones Averaged*
Background noise level: all equipment off	31 dB LAeq	31 dB LAeq
Background noise level: power amplifiers on, curtains closed	33 dB LAeq	33 dB LAeq
Background noise level: power amplifiers on, curtains opened	34 dB LAeq	34 dB LAeq
Pink noise: maximum SPL with DSP limiting on	89 dB LAeq	89 dB LAeq
Pink noise: maximum SPL with DSP limiting off	109 dB LAeq	108 dB LAeq



time-averaged, A-weighted sound pressure levels (overall as well as in third-octave bands). One microphone measured sound levels at the center of the room at a height of 69 in (1.8 m). The other microphone was randomly placed in the room at the coordinates (144 in, −109 in, 53 in), corresponding to *x*, *y*, and height relative to the center microphone (about 3.7, −2.8, and 1.3 m). Table 2 provides SPLs for the center point as well as the average of the two microphones. Microphones were calibrated with a GRAS 42AA pistonphone and were within 0.25 dB of the calibration tone level (114 dB at 250 Hz). Serge wool curtains were included, using 250 yards of 8-ft material (about 230 m of 2.4-m fabric). The main noise source is from 16 Yamaha XMV8280-D power amplifiers inside custom enclosures; measurements were made with the enclosure doors closed. The custom enclosures were built by Sound Construction and Supply. Ten amplifiers are located on the catwalk of L2, three amplifiers are located on the catwalk of L3, and three amplifiers are situated on the Grid.

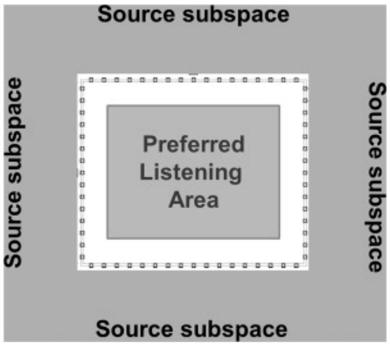
Lastly, the A-weighted third-octave bands were calculated at the center of the room to provide insight into the frequency content of the background noise with the power amplifiers both on and off, and with the amplifier enclosure doors open. The noise source is from variable-speed fans in the XMV8280-D amplifiers. The test was conducted with the amplifiers idling at room temperature for several hours until the standby temperature was attained. The result is shown in Figure 4. The dotted curve represents the normal operating condition for the Cube with the amps on. Most of the increase in the overall level due to the amplifiers is in the range from 100 to 700 Hz.

The last measurement made was the reverberation time of the room. For this measurement, pink noise at a level of 89 dB LAeq was used to drive the room. The cutoff method was used to determine the RT30 value for the Cube with the curtains open and curtains closed. These were extrapolated to determine the RT60 values as follows: RT60 with curtains open: 0.84 seconds, RT60 with curtains closed: 0.66 seconds.

Spatial Audio Rendering Capabilities

The Cube's spatial audio system is designed to facilitate implementation of a variety of existing spatialization techniques while affording the flexibility to develop new techniques and approaches. Audio computation takes place primarily on a quad-core Mac Pro, referred to as the Cube Spatial Audio Renderer (CSAR). Wave-field synthesis is provided by a Sonic Emotion Wave I 3-D sound processor. We

Figure 5. Preferred listening area and subspace of possible source positions for the main loudspeaker array.



discuss here the main techniques implemented in the first instantiation of the Cube's audio system.

3-D Wave-Field Synthesis

Wave-field synthesis is a holophonic technique for spatial audio that provides accurate and stable localization cues over an extended area, as opposed to most other spatial audio techniques that are dependent on a "sweet spot" (Corteel and Caulkins 2004). The Cube's audio system supports 64-channel WFS rendering in the horizontal plane located around head height on the main floor of the room. By using WFS plane waves as virtual playback channels, a high-quality stereophonic experience can be offered throughout the listening area, not just in a single sweet spot or sweet line as would be the case with conventional point-source stereo.

Preferred Listening Area

For optimal WFS audition, listeners should be located in the preferred listening area at the center of the room, at least 5 ft away from the loudspeaker array (about 1.5 m), as shown in Figure 5. In our informal experience, this is required so that each listener is in the field of at least three loudspeakers, such that their acoustic fields blend and that they become unnoticeable as individual loudspeakers and offer a true sound-field reproduction. WFS virtual sources, both point sources and plane waves, can be positioned anywhere within the source subspace, which spans the portion of the horizontal plane located behind the loudspeaker array from the perspective of listeners located in the preferred listening area.

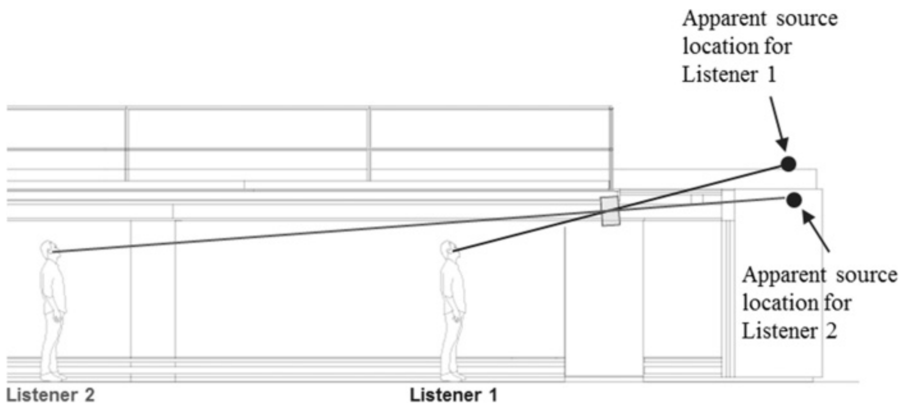
Vertical Localization Shift

Given the vertical rigging location of the WFS array at roughly one foot (30 cm) above ear height, a perceived shift in vertical localization of virtual sources occurs. This source shift is inherent to the cylindrical symmetry of the wave front produced by linear WFS arrays (Start 1997). It increases for seated listeners and as the listener approaches the edge of the preferred listening area as shown in Figure 6. Previous studies have investigated audio reproduction used in congruence with visual displays and found that the presence of a matching video image will significantly shift the perceived virtual source towards the desired location as compared with an audio-only situation, because of the so-called "ventriloquism effect" (De Bruijn and Boone 2003). In this case, the vertical source shift is least apparent for standing listeners equipped with head-mounted displays, and it is most apparent for a seated audience in the center of the room. With programs for seated audiences, a moveable, floor-mounted system consisting of ten JBL LSR6238P loudspeakers is used to synthesize sound sources located at ear height. (See the section "Supplemental Audio Systems," below.)

Spatial Aliasing

Audible spatial aliasing artifacts occur within the listening area due to the unconventionally large spacing between loudspeakers, about 3 ft (close to 1 meter), rather than a more traditional 6-in (15-cm) spacing between loudspeakers. This large spacing is what drives the 5-ft (1.5-m) minimum distance designed to maintain a "rule-of-thumb" distance between horizontal loudspeakers such that the distance to a listener located within the preferred listening area is always approximately at least twice the distance between two consecutive loudspeakers at each level of the system. This rule of thumb applies, provided appropriate techniques are implemented to reduce audible colorations, localization errors, and excessive virtual-source

Figure 6. Vertical source delocalization due to elevation of the main WFS array on L1 above ear height.



width as is the case in the Sonic Emotion WFS Rendering engine used in this project (Wittek, Rumsey, and Theile 2007; Corteel, Kuhn-Rahloff, and Pellegrini 2008).

High Order Ambisonics

The Ambisonics format is widely used within the spatial-audio research community, in particular in the context of capturing and reproducing high-resolution spatial audio. Beyond the original first-order Soundfield microphone, high-order microphones have been developed to capture sound fields for HOA playback, and this makes this format particularly interesting compared with other 3-D audio formats, from a recording standpoint (Meyer and Agnello 2003; Rafaely 2005). The main audio system has been designed to support 2-D and 3-D HOA playback, for up to 31st-order 2-D HOA and up to 9th-order 3-D HOA. In its conventional formulation, HOA is restricted to the synthesis of virtual sound sources outside the listener area, though recent investigations have indicated it could also eventually be used to synthesize sound sources within the listening area (Ahrens and Spors 2008).

Vector Base Amplitude Panning

Vector base amplitude panning is an amplitude panning method that allows the user to position virtual sources within triangles formed by adjacent triplets of loudspeakers in either two or three dimensions (Pulkki 1997). Compared with HOA, VBAP can offer improved localization accuracy. With this increased localization accuracy comes an increased awareness of the timbral characteristics of each individual loudspeaker superimposed onto the virtual source. Higher-order Ambisonics, which typically involves many more loudspeakers than VBAP when decoded to an HDLA, produces a greater sense of immersion in a sound field and less of an impression of listening to sounds through individual loudspeakers. In our experience, however, this sense of immersion comes at the cost of increased source width and lack of precision in localization, at least at lower orders of HOA.

Virtual sources rendered via VBAP cannot be positioned outside the region of space defined by the active triplet, which holds true independently of listener position, such that maximum error in virtual source localization is proportional to the dimensions of the active triangle. The main Cube audio system is designed to support 2-D and 3-D VBAP over the full hemisphere. The initial layout shown in Figure 3 offers balanced triangle sizes above the horizontal plane by staggering consecutive loudspeaker rings at L2, L3, L4, and Grid level. This balance in triangle size is geared towards achieving smooth transitions for VBAP sources moving in the vertical dimension. Given the cuboid distribution of loudspeakers, not all VBAP triangles will be the same size, both as 2-D triplets approach the corners of the room, and in various 3-D triplets. We have observed that techniques intended to produce precise localization of sound, including VBAP, are highly effective in the Cube. This excellent localization capability may be largely thanks to the high density of the 64-loudspeaker array on L1, though psychoacoustic experimentation would be required to validate this informal observation. Over the course of numerous computer music performances, none of the presenting artists have found the cuboid structure of the Cube's HDLA problematic for achieving their artistic goals.

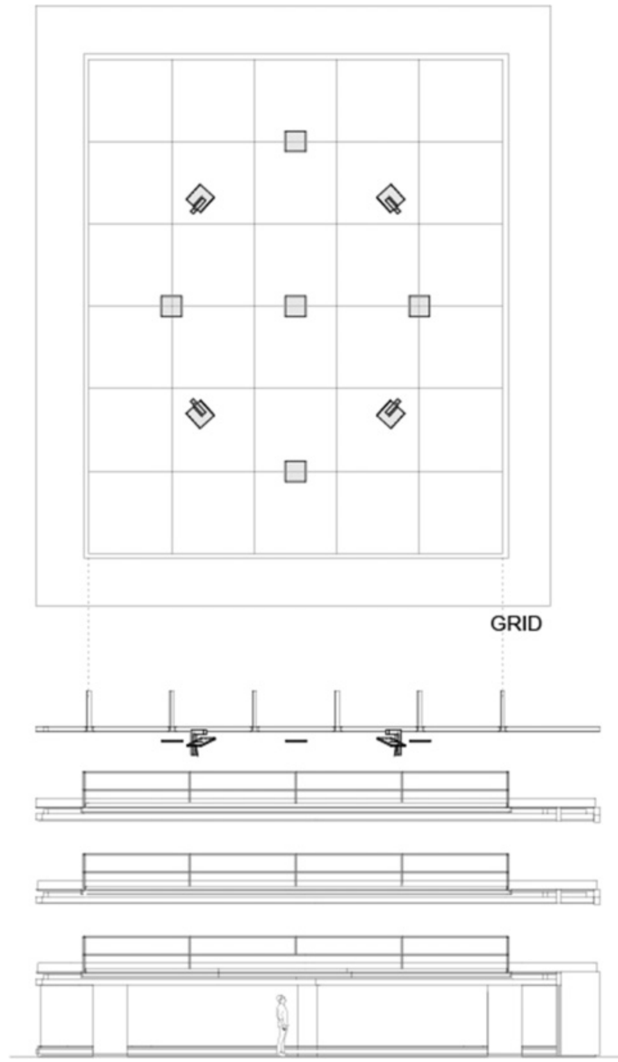


Figure 7. Ultrasonic loudspeakers mounted on pan-tilt heads.

Audio System Signal Distribution

The spatial audio system utilizes tie-line infrastructure to distribute multichannel audio. A low-latency system for distribution of digital audio, with less than 10-msec system delay, was chosen to minimize points of failure in cabling, to allow for easy reconfiguration of signal routing and to facilitate future expansion. The audio system signal is distributed by a Dante digital audio network system, making use of a CAT6A tie-line infrastructure. The main playback system comprises 124 passive loudspeakers with Dante-enabled amplifiers and four active subwoofers, all driven by BSS BLU-806 DSP processors and individually addressable. The BLU-806 units provide low-latency EQ and delay adjustments to each individual loudspeaker of the spatial audio system. All supplemental audio systems (described next) utilize the Dante audio network and Dante-to-analog breakout boxes to easily route and configure the various sources with the different reproduction systems. Individual user machines can play back multichannel audio content via virtual Dante soundcards, preferably using Dante-optimized network interface cards. A gigabit network switch connects playback machines to all of the Dante-enabled endpoints, such as amps, breakout boxes, and DSPs. Latency between the CSAR host computer and the BLU806 was measured to average 88 μ sec over five days. Latency summed across host computer, BLU806, and Dante-enabled Yamaha XMV8280-D amplifiers has been measured at constantly less than 10 msec and typically less than 3 msec. Analog input latency has not yet been formally tested, but in live electronic performances it has not added problematic delay.

Supplemental Audio Systems

An array of nine Holosonic Labs ultrasonic loudspeakers supplements the main audio system and provides discrete sound beams to target specific individual users or areas within the space. Four AS-24 loudspeakers are mounted on Apollo Right Arm pan-tilt heads to permit physical displacements of the target beam within the space, e.g., for use in conjunction with the tracking system. Five AS-24i loudspeakers are loose on the grid, allowing for manual repositioning (see Figure 7). The pan-tilt heads are controlled by DMX with custom code to control an ENTTEC DMX interface. Custom DMX controller applications were developed in-house by Denis Gracanin, Tanner Upthegrove, and Stephen Whitehead using C++ and Max/MSP. The AS-24i loudspeakers receive analog audio from a BSS BLU-806 on the DANTE network into the audio tie-line infrastructure.

An array of ten floor-mounted JBL LSR6328P loudspeakers is used for events where a circular array of loudspeakers is required at ear height for seated listeners. These loudspeakers can be tied into the main audio system via the audio network to be used concurrently with the main 128-channel array. The current configuration places three loudspeakers in front of the audience, three in back, and two on either side of the audience. The floor-mounted system is on wheels and can be rapidly reconfigured.

System Evaluation and Installation

Before installing the system, multiple loudspeakers were compared and contrasted, taking into account cost per channel, frequency range, sound power, subjective sound quality, and rigging options. A listening test was organized in the Cube, with the following loudspeakers: Community CS8, Genelec 8030, JBL6328, JBL SCS8, Tannoy Di8DC, Tannoy DVS8, Yamaha HS8, and Yamaha VSx8. The JBL SCS8 was selected for the main spatial audio array, based on a number of contributing factors, including timbre coloration, ease of mounting, maximum SPL, and passive versus powered amplification. All loudspeakers were tested on a variety of program materials, with the results discussed among Bukvic, Lyon, Nichols, and Upthegrove. Time constraints did not allow for a double-blind listening test, but given that all participants converged on the SCS8, it is unlikely that a double-blind test would have resulted in a different decision in this case.

Virginia Tech staff, led by Knapp and Upthegrove, procured, installed, and configured all the equipment for the main spatial audio system and supplemental systems described in this article, with support from MAC production staff. The multichannel Dante-enabled Yamaha XMV8280-D audio amplifier was selected, in part for ease of integration with the existing networked audio infrastructure available on each level of the Cube. One such amplifier powers eight of the 124 channels. Each of the 16 needed amplifiers is located on or below catwalks within the space. In order to maintain the low background noise in the listening area, custom vented enclosures were designed and built for the project. Caulkins helped conduct the initial system calibration along with Upthegrove and support staff (see Figure 8). Delay compensation was not included in loudspeaker calibration, but is introduced through software solutions as needed. Multiple BLU-806 signal processors were configured to apply parametric equalization to each SCS-8 loudspeaker for a flat response, as measured by an Earthworks M23 microphone on axis to the tweeter at a distance of 39 in (1 m). The system as a whole was then corrected to a flat response, as measured in the center of the room, at a height of 69 in (1.8 m). In the final setup, CSAR running Max/MSP renders spatial audio and routes signals from the Sonic Emotion WFS renderer such that any combination of WFS, HOA, or VBAP sources can be rendered in real time.

Computer Music Research and Creative Work in the Cube

Computer musicians Bukvic, Lyon, Nichols, and Upthegrove have all pursued both computer music experimentation and composition in the Cube since the installation of the audio system described in this article. We now discuss several computer music projects undertaken in the Cube during the period 2013–2015. During the same period, several scientific research projects were also successfully conducted in the Cube, though that work is beyond the scope of this article.

Spatial Orchestration

Spatial orchestration (Lyon 2008) comprises a set of strategies to adapt spatially complex computer music from one HDLA space to another. This approach allows the composer to fully exploit the

Figure 8. Initial system calibration of the main spatial audio system. (Photo by Terrence Caulkins.)

unique properties of a particular HDLA facility, while accepting that some spatial features might not be reproduced at a different HDLA facility. On the other hand, new spatial features could be introduced in a future spatial reorchestration. Lyon composed *The Cascades* in 2014 for the Cube, using tetrahedral A-format Ambisonic field recordings of a waterfall and stream in the Jefferson National Forest, produced by Michael Roan. Early in the piece, excerpts from the original four-channel recordings are reproduced, first statically, with each channel routed to one of the corner loudspeakers on L1, and then dynamically, by rotating these four channels around the 64-loudspeaker ring on L1 with equal-power panning. Both of these 2-D deployments produce a satisfying sense of "surround," with excellent coverage throughout the listening area. When *The Cascades* is presented to a standing audience, listeners often choose to walk around the Cube in order to spatially explore these naturalistic passages. In the succeeding sections, music with a 3-D spatial conception is derived from the source recordings primarily using granular algorithms to address individual loudspeakers. Spatial effects—such as waterfall sounds gradually descending from ceiling to floor, spatial canons, spatially articulated harmonic structures, and spatially articulated rhythmic patterns—are all produced in this manner. Because most of this music focuses on texture, rather than placing coherent images within a virtual sound field, Ambisonics was not used. Reorchestrating *The Cascades* involves reassigning materials to available loudspeakers in different HDLA configurations. One striking demonstration of the value of the dense array on L1 is heard in a passage that canonically articulates a sample around the 64 loudspeakers in point-source fashion, with no use of virtual sources. No other HDLA tested provided the loudspeaker density of the Cube's L1 array, and therefore equal-power panning was required to simulate the successive placement of sound into 64 distinct locations. In all cases of reorchestration to other HDLAs, the use of virtual sources resulted in a striking loss of perceived localization accuracy, and a reduction in timbral clarity as well.

When spatially orchestrating *The Cascades* for the Birmingham Electroacoustic Sound Theatre (BEAST; see also Wilson and Harrison 2010), a new feature was introduced by projecting flowing water sounds across loudspeakers situated on the floor, among the audience, resulting in a different spatial experience than in the original Cube version. The simulation of a waterfall gradually coming down on the audience was realized on both systems, but the effect was more pronounced in the Cube, because of the greater height of the Cube's grid, compared with the rigging of the elevated BEAST loudspeakers. When adapting *The Cascades* to the Sonic Lab at University of Belfast's Sonic Arts Research Centre (SARC), flowing water sounds were directed to loudspeakers located one floor below the audience, which is seated on a grid, creating the striking effect of water flowing under the audience. The development of the compositional practice of spatial orchestration requires sustained access to an HDLA facility such as the Cube, as well as access to alternative HDLAs. To this end, Lyon has worked at the AlloSphere (Höllerer, Kuchera-Morin, and Amatriain 2007), BEAST, the Stanford Giant Radial Array for Immersive Listening (GRAIL), Klangdom (Ramakrishnan, Goßmann, and Brümmer 2006), SARC, and the Immersive Computer-controlled Audio Sound Theater (ICAST) at Louisiana State University (Beck et al. 2006).

The D^4 Spatialization Tools and *Tornado*

D^4 is a Max library created by Ivica Ico Bukvic for 4-D audio spatialization, with the fourth dimension being time. D^4 consists of a set of tools that take advantage of a new layer-based amplitude-panning algorithm (Bukvic 2016). These tools are complemented with abstractions designed to streamline the process of time-based audio spatialization in a way that is accessible and easily comprehended and that offers integration with such external tools as DAWs and video production suites.

Tornado, a project headed by Virginia Tech Geography professor Bill Carstensen and premiered in the Cube in 2014 at the fall Tech or Treat event, calls for immersive visualization and sonification of a tornado. Bukvic was tasked with providing an immersive sonification of what a tornado might sound like if it were to pass through the Cube. The project also called for a demonstration capability that would allow for the components of the ensuing composition to be altered in real time within an intuitive user interface. D^4 provided the rapid prototyping tools for the complex spatialization required by the *Tornado* project.

Tornado is a two-minute immersive simulation of one of the most destructive natural phenomena on our planet. The project builds upon a recording of the tornado that struck Guin, Alabama, in 1974. It was taped by Richard Alan Lindley, who made a cassette recording as the powerful F5 tornado approached. The ensuing recording, preserved and provided by Jim Metzner (1994), was used as an inspiration and a foundation for the immersive experience of *Tornado*.

The D^4 library allows users to easily spawn point sources, source clusters, and moving sources with varying volumes and trails produced with the library's motion blur tool. Akin to a visual editor, D^4 introduces spatial masks and facilitates the design and editing of these masks using a built-in 2-D (see Figure 9) and 3-D (see Figure 10) editor and viewer. It is designed to seamlessly adapt to any speaker arrangement that offers information on speaker azimuth and elevation, and does so without any changes to the compositional spatialization parameters. Its functionality is further complemented with helper abstractions that enable, among others, circular motion at an angle, bouncing motion, and motion blur.

Finally, D^4 provides critical low-latency, real-time rendering of spatialized audio sources even in scenarios having large numbers of speakers and audio streams. This makes it particularly useful in interactive environments. The Tornado demo features 1,018 internal 24-bit, 48-kHz audio streams or channels that are mixed down and played through 128 loudspeakers in real time with a computer audio-processing latency of less than 11 msec.

Figure 9. The 2-D editor for D^4.

Figure 10. The 3-D editor for D^4.

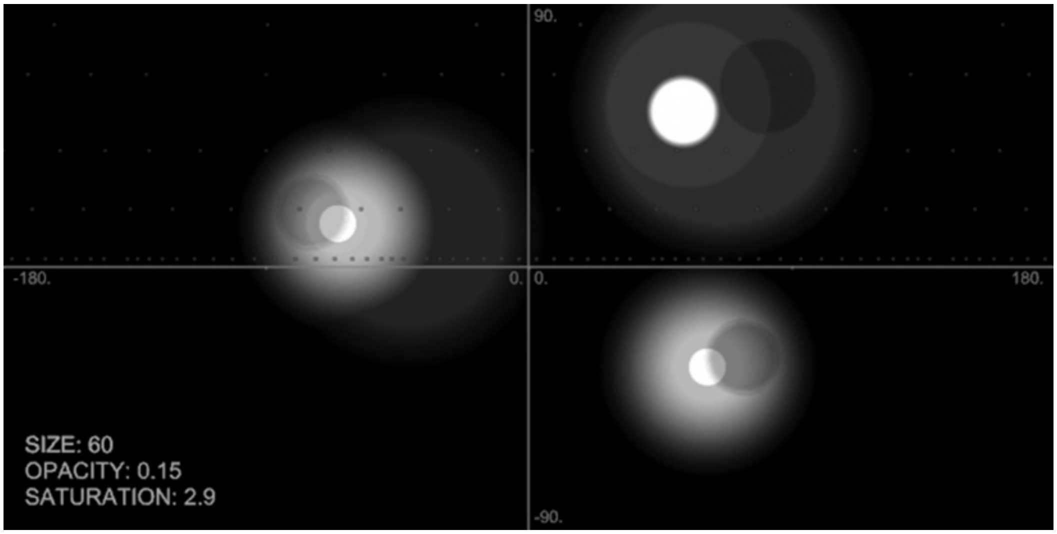


Figure 9

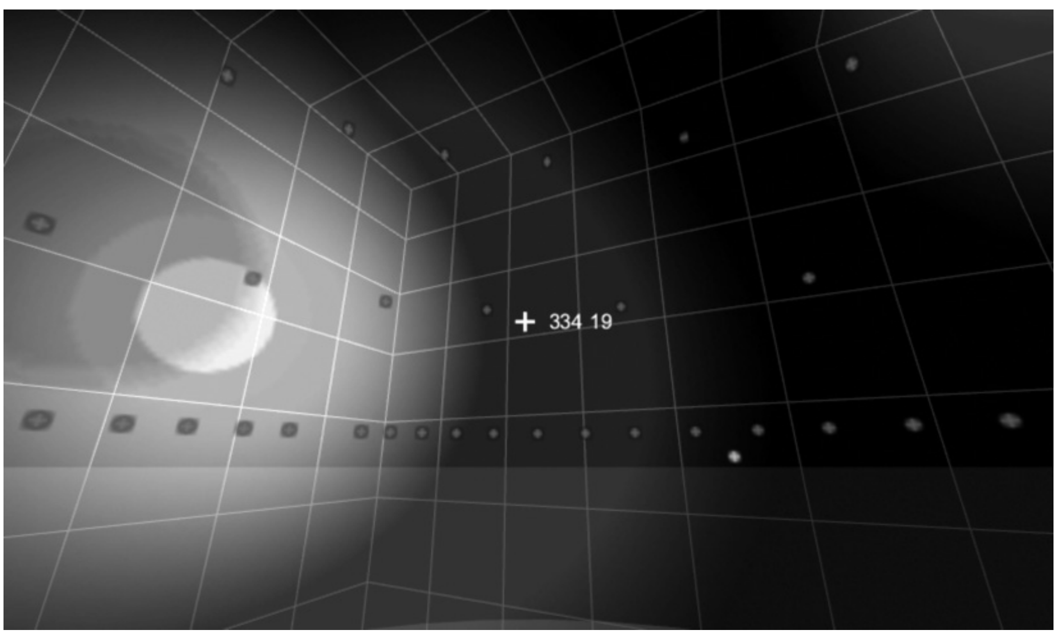
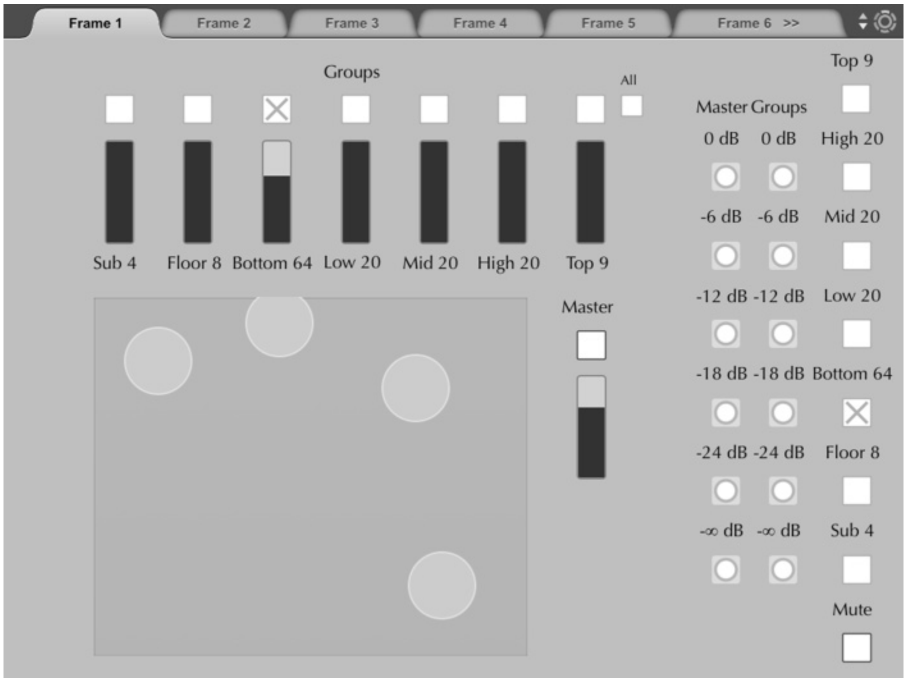


Figure 10

Live Spatialization with the Leap Motion

Tanner Upthegrove integrated the Leap Motion with CSAR as a control device for composition and sound design. Building upon Masayuki Akamatsu's Max external `aka.leapmotion`, real-time input from hand gestures transposes to spatial position. Three spatialization tools were controlled with the Leap

Figure 11. The master panel for Charles Nichols's Mira-based spatialization interface.



Motion: the Sonic Emotion Wave I, the spatialization external spat.spat~ for Max/MSP (developed by the Institut de Recherche et Coordination Acoustique/Musique, IRCAM), and the ICST Ambisonics externals for Max (Schacher and Kocher 2006).

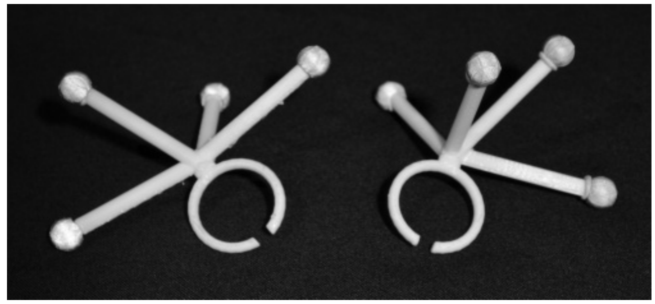
The Sonic Emotion Wave I achieves WFS with proprietary hardware, addressable by Open Sound Control (OSC). Upthegrove scaled finger movements to polar coordinates in Max, then packaged and sent OSC messages over UDP to the Wave I. The finger movements controlled audio point-source locations for two-dimensional WFS. The Wave I system includes a visual interface to verify point source locations, which was used during performances.

Upthegrove used the Leap Motion to control the Spat library for real-time VBAP and 3-D HOA spatialization in Max. VBAP panning moves around and overhead on the cubic perimeter, whereas 3-D HOA movements travel through 3-D space. Akamatsu's aka.leapmotion was embedded in the patch with spat~. Hand gestures controlled audio source positions by converting tracking data to spherical coordinates, which were packaged as messages and sent to spat~.

Mira-Based Spatial Control

Nichols has developed a series of Max patches to control the Cube's loudspeakers through CSAR, both individually and in groups, and to spatialize audio sources using the matrix~ and matrixctrl objects. Using Mira, a touch controller app for the iPad developed by Cycling '74, these patches incorporate Max gain~ audio fader objects in multiple mira.frame interface-mirroring objects that display as tabs of panels with touch-controlled faders on an iPad. The tabs are divided into a master panel and group panels for each layer of loudspeakers in the Cube. On the master panel tab, a main fader controls the overall level of the system and subgroup faders control the total level for layers of loudspeakers (see Figure 11). The group tabs are divided into panels for 8 loudspeakers on the floor (before they were expanded to 10) and 4 subwoofers, 64 loudspeakers on the first catwalk, 20 loudspeakers each on the second and third catwalks and grid, and the 9 ultrasonic loudspeakers on the ceiling, with an individual fader for each loudspeaker. On each panel, toggles in a column enable and disable audio

Figure 12. Custom motion-capture rings, produced with 3-D printing. (Photo by Tanner Upthegrove.)



for every layer, a toggle mutes the entire system, columns of buttons select default decibel levels for all faders in a group of loudspeakers, and toggles enable and disable audio for each individual fader on the panel. The Mira-controlled Max patches allow users to audition work from the floor of the Cube while manipulating the CSAR computer located in the control room.

These Max patches spatialize audio sources, mixed with the panels of faders, in a variety of ways. At the heart of each patch is a six-row, 64-column control matrix that interpolates to individual single-row matrices for the rings of loudspeakers in the different layers of the Cube. Controlling a common matrix whose width is the maximum number of loudspeakers in a layer allows the user to easily map controller data to any loudspeaker position. Horizontal position on the matrix wraps to circumference position on the rings of loudspeakers, and vertical position maps to loudspeaker layer.

Hardware-Based Spatial Control

Nichols has experimented with commercial controllers and with custom controllers that he designed to control 3-D spatialization with physical gesture. For proof of concept, a Logitech Extreme 3D Pro joystick was used to move the azimuth position of a single source, and its paddle was used to move the height of the source. For more expressive gestural control through a wider range of physical motion, a Bourns PEC11-4215F-S24 rotary encoder and a Sharp GP2Y0A21YK infrared proximity sensor, wired to an Arduino microprocessor sending serial data to Max over USB, were used to control the azimuth position and height of a single source. With a tracking range of 10 to 80 cm, the infrared proximity sensor follows a wide sweep of hand motion to control the elevation of a single sound source. To spatialize multiple sources, the `mira.multitouch` gesture-capture object was used to receive the position data of up to five fingers from the Mira app on an iPad, spatializing up to five sound sources on the Cube's L1, using the Wave 1 WFS system. Custom-designed rings (see Figure 12), produced using 3-D printing, were used to spatialize multiple sources with a dancer's expressive gestures. Each ring holds four 9.5-mm spherical reflective markers, which are tracked by the Cube's 24-camera Qualisys motion-capture system. The reflective markers are affixed to 4.5-cm posts positioned at different angles from the center of the form, so that each ring is captured as a separate solid body. The position data of the rings, sent to CSAR via OSC, spatialize one audio source per ring in the control matrix of Nichols's Max patch.

Il Prete Rosso

Nichols recently composed an interactive piece, *Il Prete Rosso*, for amplified violin, motion sensor, and computer, which has been performed in a number of spatial audio systems. For the Cube version of this piece, Nichols spatializes live violin around the 20 loudspeakers of Level 2, and four live-sampled violin tracks around the 64 loudspeakers of L1 of the Cube (see Figure 3), using a custom-programmed Max patch. In the first half of the piece, as violinist Sarah Plum performs, the patch records her part to be used later as accompaniment, and the amplified live violin circles around L2, reversing direction at the spatial location where each of the four recorded tracks starts playing on L1. This creates the impression that the live violin sets the recorded violins into their spatial positions. In the second half of the piece, the amplified violin remains statically positioned in the center front of the hall, while four interlocking rhythmic patterns of sampled violin, spatially equidistant from one another, gradually join a hocketing texture that swirls around the audience with increasing speed. The piece ends with a brief coda returning to the material of the first half, with the live violin again slowly orbiting the audience, while the four accompaniment parts progressively fade out.

Outreach

Outreach is a key element in our strategy to share the benefits of the Cube with both artists and researchers outside of Virginia Tech, and with the general public. Faculty researchers are able to develop the possibilities of the Cube in a sustained manner, but it is equally important to disseminate ideas through the sharing of skills, knowledge, and the irreplaceable experience of hearing spatial sound in the Cube. In 2015, Virginia Tech hosted the conference of the Society for Electro-Acoustic Music in the United States (SEAMUS), inviting over 200 delegates to directly experience the spatial possibilities of the Cube. The conference concerts in the Cube focused on spatial projection of music, with a special concert dedicated exclusively to works composed for HDLA systems.

The MAC regularly hosts artist residencies, including residencies for sound artists in the Cube. Visiting artists are paired with a resident researcher to facilitate the creation of new work that utilizes the unique spatial opportunities of the Cube. In 2015, Stephen Vitiello created the spatial audio installation *A Scuttering across the Leaves*, based on closely recorded insect sounds that were spatialized in the Cube in collaboration with Upthegrove. In 2016, Pamela Z undertook a residency, and created new spatial music for the Cube in collaboration with Lyon.

The Spatial Music Workshop was inaugurated in 2015 to disseminate ideas and practical experience in composing for HDLA systems. A small group of sound artists is invited to spend five intensive days working on a project in both the Cube and the Perform Studio, which contains an installed 24.4 HDLA system with two wall-mounted rings of twelve Genelec 8030A loudspeakers each, and four floor-mounted Genelec 7060B subwoofers. Each participant spends multiple hours in each space. A small number of instructional sessions are given on various spatial techniques. In addition to their individual work, participants meet with the workshop leader as a group to discuss their experiences and share problems and ideas. A public presentation of the outcomes of the workshop is given on the final evening.

Future Work

The Cube is an organically evolving facility, driven by aesthetic and engineering research needs that are addressed both by in-house software development and by ongoing audiovisual infrastructure development and improvement. A few near-term projects and plans are discussed in this section.

Binaural Head Tracking

A head-tracked binaural audio system is under development to allow individual researchers to experience localized "in-space" audio sources, given that the Cube's spatial-audio loudspeaker system is not designed for focused-source synthesis. The proposed system will utilize wireless open-ear cup headphones to allow researchers to move freely around the space and hear not only individualized sound in headphones, but also room sound projected from the Cube's HDLA loudspeakers. Multiple pairs of wireless headphones will be provided, each with an individual binaural feed. The binaural audio feeds will be transmitted to the wireless emitters via the Dante audio network. Headphone position and orientation will be tracked by an installed 24-camera Qualisys motion-capture system.

A New Control Surface

Nichols is designing a physical control surface to manipulate all Cube loudspeakers and group levels and to spatialize multiple sources. The control surface will use slide potentiometers and pushbuttons to interface with the faders and toggles of his Max patch (as shown in Figure 11), rotary encoders and infrared proximity sensors to manipulate spatialization in the control matrix, and RGB LED lights and an organic light-emitting diode graphic display for visual feedback of settings. The control surface sends and receives data through multiplexers to Arduino microprocessors. The sensors, lights, and display will mount to laser-cut acrylic panels, assembled into a mixing console.

Audio System Improvements

We plan to double the density of the L1 array to 128 loudspeakers. We would also like to introduce a more traditional, high-density WFS system (Spors, Rabenstein, and Ahrens 2008), which would extend the Cube's capabilities for projecting sound inside the space (Caulkins, Corteel, and Warusfel 2003).

Although the noise floor is quite low, there remains room for improvement. The main noise source is fan noise from the Yamaha XMV8280-D amplifiers and their enclosures (Figure 4). Moving the amplifiers to another space and running tie lines to the Cube would robustly address the fan noise issue. The XMV8280-D amps have a built-in protection circuit that turns off any channels that do not receive audio signal for about 15 minutes. Addressing a channel after it has gone to sleep results in a gradual fade-in, which is unsuitable for high-fidelity playback. To overcome this problem, a short burst of low-amplitude noise is played through each channel just before concert performance. In the event that the protection circuit issue does not have a solution, eventually the XMV8280-D amplifiers will need to be replaced. As more Dante-enabled amplifiers become available, it is likely that we will be able to source amplifiers that are both quieter and lack the protection-circuit problem.

Psychoacoustic Testing

Psychoacoustic experimentation is being undertaken to establish the limits of localization and other aspects of spatial perception in the Cube. The Cube is considered a work in progress, and the results of perceptual experimentation may lead to further expansion and refinement of its HDLA system.

Conclusions

Within a two-year period, a complex HDLA system was designed and installed to the Cube, and then utilized for a wide variety of purposes, including scientific research, computer music composition, intermedia installations, public performances, education, and outreach. Given the straightforward process by which the Cube's HDLA was designed, installed, and put into production, we anticipate the inauguration of many new HDLAs in the coming years, and a concomitant exploration of new aesthetic possibilities for spatially immersive computer music.

Acknowledgments

Funding for the Cube and its audio infrastructure was provided by ICAT and the State Council of Higher Education for Virginia (SCHEV).

References

Ahrens, J., and S. Spors. 2008. "Focusing of Virtual Sound Sources in Higher Order Ambisonics." In *Proceedings of the 124th Convention of the Audio Engineering Society*. Available online at www.aes.org/e-lib/browse.cfm?elib=14508 (subscription required). Accessed October 2016.

Beck, S. D., et al. 2006. "The Immersive Computer-Controlled Audio Sound Theater: Experiments in Multi-Mode Sound Diffusion Systems for Electroacoustic Music Performance." In *Proceedings of International Computer Music Conference*, pp. 649–655.

Blauert, J. 1996. *Spatial Hearing: The Psychophysics of Human Sound Source Localization*. Cambridge, Massachusetts: MIT Press.

Bukvic, I. 2016. "3D Time-Based Aural Data Representation Using D^4 Library's Layer Based Amplitude Panning Algorithm." In *Proceedings of the International Conference on Auditory Display.* Available online at www.icad.org/icad2016/proceedings2/papers/ICAD2016_paper_10.pdf. Accessed October 2016.

Caulkins, T., E. Corteel, and O. Warusfel. 2003. "Wave Field Synthesis Interaction with the Listening Environment, Improvements in the Reproduction of Virtual Sources Situated Inside the Listening Room." In

Proceedings of the International Conference on Digital Audio Effects. Available online at www.eecs.qmul.ac.uk/legacy/dafx03/proceedings/pdfs/dafx34.pdf. Accessed October 2016.

Corteel, E., and T. Caulkins. 2004. "Sound Scene Creation and Manipulation Using Wave Field Synthesis." Technical Report. Paris: Institut de Recherche et Coordination Acoustique/Musique. Available online at recherche.ircam.fr/equipes/salles/WFS_WEBSITE/Documents/WFS_overview.pdf. Accessed October 2016.

Corteel, E., C. Kuhn-Rahloff, and R. Pellegrini. 2008. "Wave Field Synthesis Rendering with Increased Aliasing Frequency." In *Proceedings of the 124th Convention of the Audio Engineering Society*. Available online at www.aes.org/e-lib/browse.cfm?elib=14492 (subscription required). Accessed October 2016.

De Bruijn, W. P .J., and M. M. Boone. 2003. "Application of Wave Field Synthesis in Life-Size Videoconferencing." In *Proceedings of the 114th Convention of the Audio Engineering Society*. Available online at www.aes.org/e-lib/browse.cfm?elib=12606 (subscription required). Accessed October 2016.

Höllerer, T., J. Kuchera-Morin, and X. Amatriain. 2007. The Allosphere: A Large-Scale Immersive Surround-View Instrument. In *Proceedings of the Workshop on Emerging Displays Technologies: Images and Beyond: The Future of Displays and Interacton*. Available online at doi.acm.org/10.1145/1278240.1278243 (subscription required). Accessed October 2016.

Lyon, E. 2008. "Spatial Orchestration." In *Proceedings of the Sound and Music Computing Conference*, pp. 84–85.

Metzner, J. 1994. *Pulse of the Planet: Extraordinary Sounds from the Natural World*. Berkeley, California: Nature Company.

Meyer, J., and T. Agnello. 2003. "Spherical Microphone Array for Spatial Sound Recording." In *Proceedings of the 115th Convention of the Audio Engineering Society*. Available online at www.aes.org/e-lib/browse.cfm?elib=12353 (subscription required). Accessed October 2016.

Pulkki, V. 1997. "Virtual Sound Source Positioning Using Vector Base Amplitude Panning." *Journal of the Audio Engineering Society* 45(6):456–466.

Rafaely, B. 2005. "Analysis and Design of Spherical Microphone Arrays." *IEEE Transactions on Speech and Audio Processing* 13(1):135–143.

Ramakrishnan, C., J. Goßmann, and L. Brümmer. 2006. "The ZKM Klangdom." In *Proceedings of the Conference on New Interfaces for Musical Expression*, pp. 140–143.

Rohr, L., et al. 2013. "Vertical Localization Performance in a Practical 3-D WFS Formulation." *Journal of the Audio Engineering Society* 61(12):1001–1014.

Schacher, J. C., and P. Kocher. 2006. "Ambisonics Spatialization Tools for Max/MSP." In *Proceedings of the International Computer Music Conference*, pp. 274–277.

Spors, S., R. Rabenstein, and J. Ahrens. 2008. "The Theory of Wave Field Synthesis Revisited." In *124th AES Convention*. Available online at www.deutsche-telekom-laboratories.de/_sporssas/publications/2008/AES124_Spors_WFS_Theory.pdf. Accessed November 2016.

Start, E. W. 1997. "Direct Sound Enhancement by Wave Field Synthesis." PhD dissertation. Delft University of Technology, The Netherlands.

Wilson, S., and J. Harrison. 2010. "Rethinking the BEAST: Recent Developments in Multichannel Composition at Birmingham ElectroAcoustic Sound Theatre." *Organised Sound* 15(3):239–250.

Wittek, H., F. Rumsey, and G. Theile. 2007. "Perceptual Enhancement of Wavefield Synthesis by Stereophonic Means." *Journal of the Audio Engineering Society* 55(9):723–751.