THE ART AND SCIENCE OF SOUNDSCAPE AURALIZATION FOR HIGH DENSITY LOUDSPEAKER ARRAYS

Tanner Upthegrove *and* Charles Nichols

Virginia Tech Institute for Creativity, Arts, and Technology; Virginia Tech School of Performing Arts, Blacksburg, Virginia, United States of America
email: upty@vt.edu

Michael Roan

Virginia Tech Department of Mechanical Engineering; Virginia Tech Institute for Creativity, Arts, and Technology, Blacksburg, Virginia, United States of America

Virginia Tech hosts four spatial audio labs utilized by artists, engineers, designers, and scientists. The largest spatial audio lab is the Cube, with an individually addressable 140 loudspeaker spatial audio system. Representatives from Virginia Tech propose a lecture and demonstration session comprised of strategies for soundscape auralization, specifically for high density loudspeaker arrays, which are systems of 24 or more loudspeakers. Discussion topics include system infrastructure design, real-time sound environment rendering, spatial audio strategies for composers and performers, control systems for live spatial audio, and virtual acoustics. Technical demonstrations include real-time convolution reverberation for arbitrary loudspeaker arrays, virtualizing loudspeaker systems, and tools to build soundscapes. Compositions and performances presented would include *Shakespeare's Garden*, an immersive theatrical experience with immersive spatial audio soundscapes, recitations played through spotlight speakers, multiple projected graphic designs, and theatrical lighting, *Badstar*, a concert of immersive audio and video, performed on electric violin, banjo, guitar, and drums, with multiple interactive videos projected onto custom architecture, and What Bends, three movements for electric violin and recited poetry performed in immersive spatial audio, 360º interactive video, and motion capture dance controlling audio spatialization and video parameters.
Keywords: 3-D audio, spatial audio, ambisonics

1. Introduction

In 2014, Virginia Tech's Institute for Creativity, Arts, and Technology (ICAT) produced the first sounds through a newly designed High-Density Loudspeaker Array (HDLA) in the Cube, a 15.85m long by 14.63m wide by 15.85m tall black box theatre housed in the Moss Arts Center in Blacksburg, Virginia [1]. The system was specifically designed for the Cube with consultation from Terence Caulkins and Denis Blount of Arup. The original design consisted of 128 individually addressable loudspeakers. From the design and usage of the Cube, three other spatial audio studios emerged with comparable technologies and workflows to render soundscapes.

2. Facilities

The Cube opened in October 2013. Since the installation of the permanent HDLA in 2014, ten more loudspeakers have been added on stage as well as two subwoofers, each with two 45cm drivers for low frequency effects.

In parallel to the opening of the Cube, the Perform Studio in the Moss Arts Center came online with a 24.4 channel HDLA. The Perform Studio is approximately 9.14m square and acoustically treated. Perform Studio is a shared studio for student and faculty work. It shares similar infrastructure to the Cube so that projects can develop in the Perform Studio and transition seamlessly to the Cube.

The Digital Interactive Sound & Intermedia Studio (DISIS) in Newman Library hosts a 16.2 channel system, with plans to upgrade to a 24.2 system during Summer 2019. DISIS is primarily used as a teaching space and features computing workstations, controllers, and audio equipment for student to develop spatialized audio projects.

The Acoustics Signal Processing and Immersive Reality (ASPIRe) Lab is the fourth permanent spatial audio lab. ASPIRe features a 48.2 system. Currently, the lab is adding wave field synthesis capabilities designed by Roan. ASPIRe is also acoustically treated and has audio and virtual reality workstations for student projects.

2.1 Signal Distribution

All four spatial audio labs have macOS or Windows computers connected to Audinate Dante audio networks. A host computer has software rendering real-time audio to the network. Other computers or networked devices can send control data or audio to the host computer over the network.

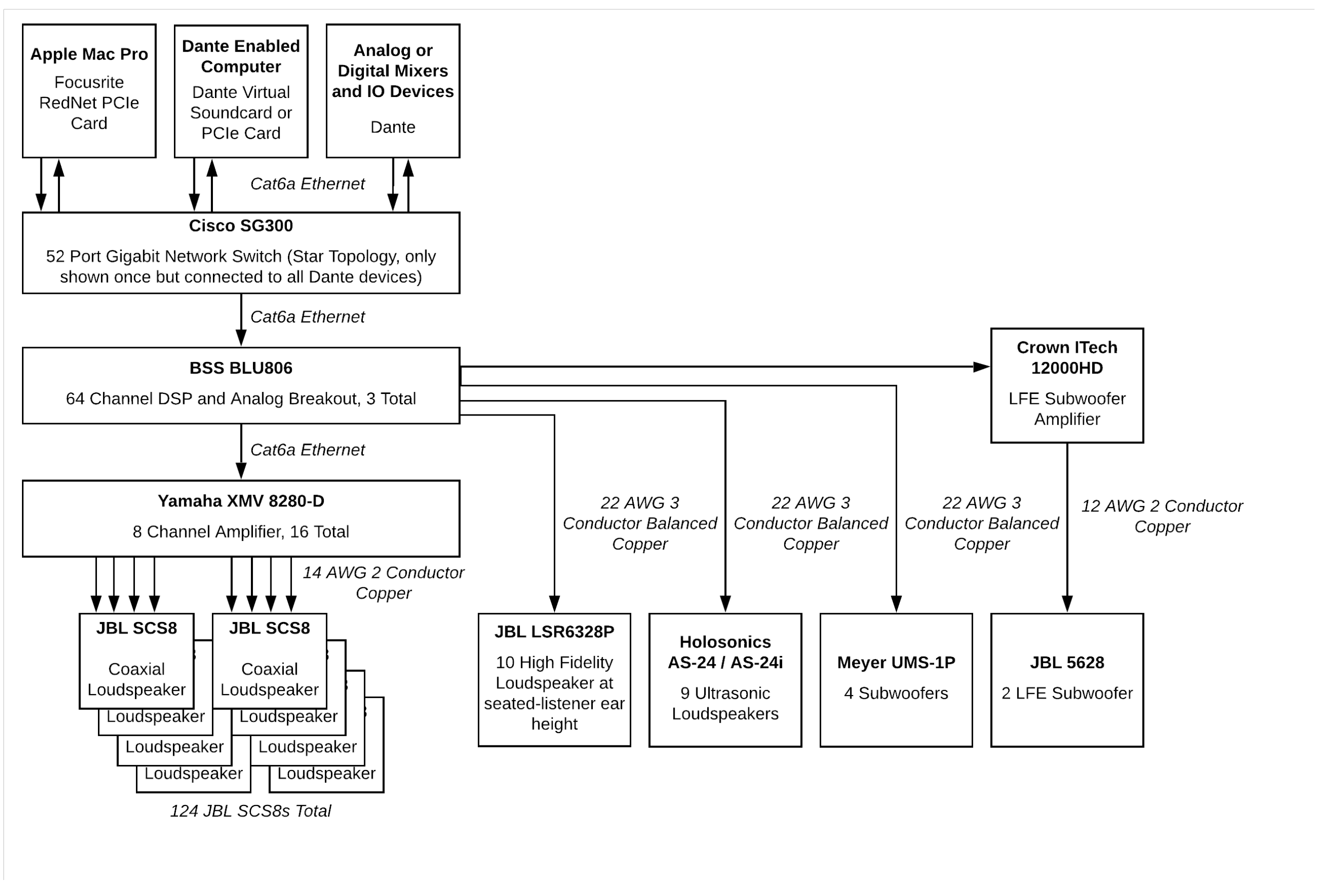

Figure 1: Cube Signal Flow Diagram.

2.2 Local Rendering

Host computers can generate and render digital audio with latencies under 20ms. Common applications include spatializing point-source audio with plug-ins within an audio workstation and programming languages like Max.

2.3 Distributed Computing for Signal Generation

Remote computers can send 128 channels of uncompressed audio to the host computer. The host computer can act as a passthrough, or leverage encoding and decoding formats, such as ambisonics, to address arbitrary loudspeaker arrays.

ASPIRe and DISIS also have round-robin functionality. The computers can be networked together, in serial or parallel circumstances, for distributed signal processing.

Another common application for distributed soundscape auralization is streaming live encoded audio from one studio for decoding in another venue. An example would be capturing a live performance in the Cube and using the Perform Studio as an overflow seating venue. Sound is captured with a sound field microphone, then sent as encoded audio over the network to the Perform Studio where it is decoded to the loudspeaker array.

3. Soundscape Auralization

Virginia Tech audio spatialization labs are used by faculty and students from disciplines in science, engineering, arts, and design. The host computers have several methods for real-time soundscape auralization.

3.1 Representational Auralization

Distributed and local content generation both utilize representational spatialization techniques. Representational auralization relies on software knowing the locations of the loudspeakers in a system. Ambisonics panning and Vector Base Amplitude panning are examples of representational techniques. Ambisonics has workflows for reproducing audio recorded with sound field microphones as well as virtual encoders. Ambisonics workflows are becoming ubiquitous and desirable for the low computational cost and reliability across disparate loudspeaker arrays [2].

3.2 Real-time Signal Processing for HDLAs

Encoded formats like ambisonics are opportune for efficient signal processing. For example, applying equalization of an encoded ambisonics B-Format signal globally effects the soundscape, instead of processing signal for individual loudspeakers. Convolution of B-Format signals means spatial information is built-in to the encoding and decoding schema. This method delivers a computationally inexpensive, highly effective global reverberation for applications in architectural acoustics [3].

4. Applications

Scientists, engineers, artists, and designers utilize auralization to achieve a variety of goals. Virginia Tech's ICAT recognizes all domains and the importance of transdisciplinary research. The following are examples of how soundscape auralization has been applied to projects at Virginia Tech.

4.1 Motion Capture as Live Control

To control the HDLA in the Cube, with the 24-camera Qualisys motion-capture system, Nichols designed and 3D printed rings, with four posts that emanate from the center loop, at different angles. Attached to the end of the four posts are spherical reflective markers, that the motion-capture software identifies as a rigid body, because of their unique combination of angles. For three projects, Nichols collaborated with dancers, who wore the rings in the Qualisys system, to send motion capture location data in Cartesian format to a Max patch, that translated ring location to speaker position, to spatialize audio with performance gesture. For *Satisfaction Guaranteed*, his structured improvisation for amplified tap dancer, motion capture system, and laptop ensemble, dancer Ann Kilkelly wore the rings to spatialize audio from the Linux Laptop Orchestra (L2Ork). When Nichols performed at Cube Fest, with electronic musician Jay Bruns, guitarist Clark Grant, and keyboardist Ben Weiss, in their band Modality, dancers Claire Constantikes and Cambria McMillan-Zapf spatialized the audio from his processed electric violin and Grant's electric guitar, by moving the rings in the motion-capture system. For his composition *What Bends*, for electric violin and interactive computer music, accompanying recited poetry, motion capture dance, animation, and processed video, a collaboration with poet and reciter Erika Meitner, choreographer and dancer Rachel Rugh, and video artist Zach Duer, Rugh used the rings to spatialize computer music and erase a digital shroud that occluded the 360º video.

Figure 2: *What Bends,* photo courtesy Zach Duer.

4.2 Shakespeare's Garden

For the *Shakespeare's Garden* immersive art installation, comprised of recorded student performances of sonnets, soliloquies, and scenes played through directional speakers, graphic design of text projected onto custom scrims, and theatrical lighting, a collaboration with directors Amanda Nelson and Natasha Staley, graphic designer Meaghan Dee, lighting designer John Ambrosone, and Upthegrove as media engineer, Nichols composed an immersive soundscape of processed environmental sounds and recited poetry. For each quadrant of the Cube, he spatialized granulated bird samples in the upper level of speakers, particularized resynthesis of habitat sounds in the middle level, and harmonically filtered landscape sounds in the lower level of the HDLA, each corner depicting a different season. While the components of the soundscape oscillated across the quadrants in each layer, recordings of students performing a scene from *A Midsummer Night's Dream* rotated in parallel Lissajous curves around the HDLA, evoking Puck and Fairy chasing each other around the garden. The result was an immersive theatrical experience in a multimedia Elizabethan garden.

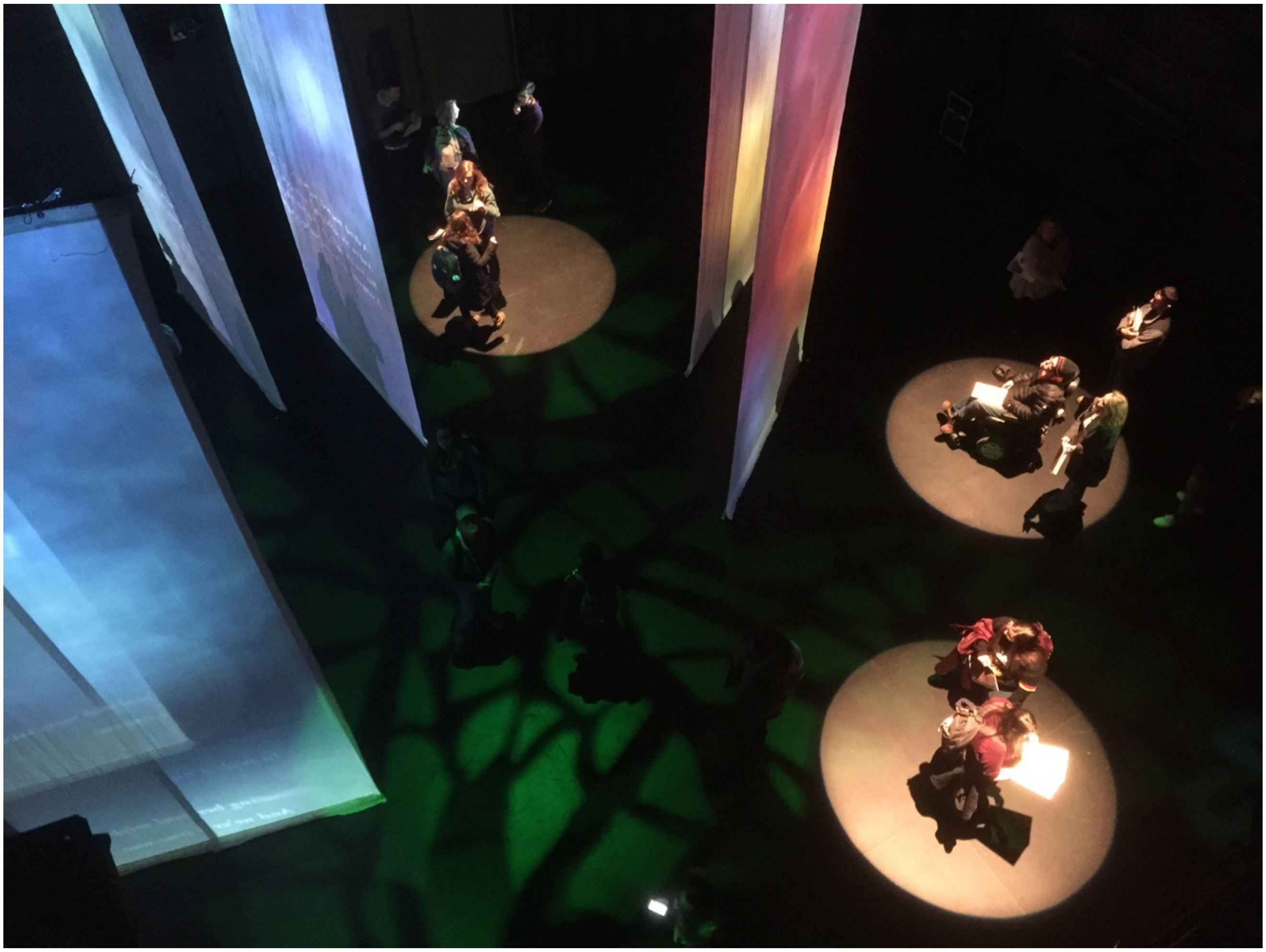

Figure 3: *Shakespeare's Garden*.

4.3 Badstar

Badstar was an hour-long show in three movements, premiered in the Cube HDLA, with music inspired by fusion, old-time, noise, and drone metal, collaboratively composed by Nichols, with Holland Hopson, and André Foisy, who performed on processed electric violin, electric banjo, electric guitar, and computers, with Denver Nuckolls on electronic drums, in immersive spatial audio. Along with the immersive musical performance, video artist Zach Duer performed interactive multichannel video, also responsive to the amplitudes and frequencies of the three electric instruments, projected on custom architecture surrounding the audience. The architecture, designed and constructed by architect Jon Rugh, was inspired by cave stalactites, cathedral flying buttresses, and funnel spider webs, immersing the audience, standing on the floor, with video projection, and drawing attention to the performers, playing on the first catwalk. Using the ICST Ambisonics objects in a master Max patch, the audio from the four instruments and three computers was spatialized in different automated trajectories for the three movements, moving in oscillating patterns across the walls, ascending and descending at different speeds, and spinning in accelerating and decelerating spirals, from the physical positions of the instrumentalists. Nichols programmed the immersive spatial audio to complement the framing of the architecture around the performers and the dynamism of the projected interactive video. The project was supported by an ICAT SEAD (Science, Engineering, Art, and Design) Major Initiative Program grant.

Figure 4: *Badstar,* photo courtesy David Franusich.

4.4 Lane Stadium Acoustics and Fan Simulation

Virginia Tech's Lane Stadium can host over 60,000 fans for each football game. To get a better idea of what the players experience, Roan and Upthegrove recorded games on the field with a Core Sound TetraMic. The recordings were converted to B-Format for reproduction in the Cube to simulate the experience. The B-Format signal was rendered to individual mono files for playback in an audio workstation The Cube can almost recreate the 113dBA sound pressure level of the actual event.

Roan and Upthegrove also recorded basketball games in Virginia Tech's Cassell Coliseum with Alexander Kern. Todd Ogle captured 360º equirectangular video. The audio and video were synchronized and played back in both the Cube and ASPIRe. In both cases, the B-Format was decoded live.

Roan and Upthegrove hope to use this information to inform future stadium designs and for situational training.

4.5 Soundscape Auralization Relationships to Virtual and Augmented Reality

The ICAT funded *Belle II VR* project takes subatomic particle physics simulation data and visualizes it in virtual reality for multiple users in the Unity game engine [4]. Spatial information about events in the simulation generates spatialized audio events, which tie together the visual and aural experience. Moreover, the Cube affords multi-user virtual reality as a social experience and the HDLA provides a common audio scene for all users.

REFERENCES

1 Lyon, E., Caulkins, T., Blount, D., Bukvic, I. I., Nichols, C., Roan, M. and Upthegrove, T. Genesis of the Cube: The Design and Deployment of an HDLA-Based Performance and Research Facility, *Computer Music Journal*, **40** (4), 62– 78, (2016)

2 Upthegrove, T., Nichols, C., Roan, M. Spatialization Pipeline for Digital Audio Workstations with Three Spatialization Techniques, *Proceedings of the Audio Engineering Society International Conference on Spatial Reproduction – Aesthetics and Science,* Tokyo, Japan, 7-9 August, (2018)

3 Ermann, M., Hulva, A., Upthegrove, T., Rehfuss, R. J., Haim, W., Kanapesky, A., Mcmillon, T., Park, C., Reardon, A., Rynes, J., Ye, S., Nichols. C. Subjective listening tests: Perception and preference of simulated sound fields, *The Journal of the Acoustical Society of America*, **141** (5), (2017)

4 Piilonen, L., *Belle II in Virtual Reality.* [Online.] avallabile: https://indico.cern.ch/event/658060/contributions/2925825/attachments/1624286/2586479/Belle2VR-HSF.pdf