Introduction

This paper provides an overview of approaches taken by the author, Amirbahador Rostami, as a studio assistant, co-supervised by Mark-David Hosale and Philip Beesley, to investigate the implementation of localized granular synthesis within large-scale immersive installations. This paper will discuss the implementation of localized audio in two large-scale sculptures, Amatria at the University of Indiana\(^1\) and Aegis and Noosphere at the Royal Ontario Museum\(^2\) in 2018. This discussion will also provide an overview of their microcontroller and audio design topology and discuss the pros and cons of centralized versus centralized audio distribution topologies.
Audio System Anatomy

Every component within an audio system carries a significant role in the production of clear, crisp Hi-Fi audio quality. In addition to cabling, power strategies, and amplification, one of the most important components is the loudspeaker. It was imperative for Amatria, Aegis and Noosphere to find a loudspeaker with the highest quality of sound production in relation to its footprint. Ideally the speaker would be small enough so that it could aesthetically blend in with the sculpture’s environment, while also providing a full spectrum frequency range with a linear response.

In speaker design, acoustic properties of the material, enclosure design and even how the environment reacts has a huge impact on the response of the loudspeaker. When it comes to the material used for speakers, a solid composite is preferable as it will mitigate undesirable effects, such as unwanted resonance, or phase reinforcement, and cancellations in the speaker itself. Placement is also a consideration, as external reflections in the room can cause phase distortions. Although most speaker boxes are rectilinear, it is better practice to avoid perfect cubes and move towards more water-drop designs for spherical enclosures. While the spherical model is more optimum, it is more challenging to precisely create the shape with solid and appropriate material.

3 DIY Video & Audio https://www.diyaudioandvideo.com/Guide/BuildSpeaker/ (last access October 16, 2018)

4 Ibid.
Image 2 illustrates a first generation of the enclosure design that was 3D printed for Amatria. This enclosure was based on sealed back design; however, it had a horn shaped cone at the front. This design resulted in higher directional dynamic range than usual, but also caused a lot of fluctuation in the overall frequency response. Another contributor to these fluctuations was the material used to print these components. 3D printed resin was not strong enough to contain the inner vibrations, causing a lot of out of phase signals.

Image 3 shows the final design for the speaker enclosure for Aegis/Noosphere. Designed by Mehreen Ali, with input from Amirbahador Rostami, the collaboration resulted in a ported spherical design inspired by the B&W 800 D3 model loudspeaker. The speaker is semi-spherical in shape, with a porthole at the back that is tuned specifically to capture low-end frequencies ranges.
Localized Vs. Centralized System Topologies

One of the main characteristics in Philip Beesley’s sculptures is the transformation of localized processes happening in smaller clusters leading into larger emergent phenomenon. This systematic approach is inspired by nature, where smaller, independent, and simpler components contribute to larger complex behaviours. This method of system design can afford modular design strategies where parts of the system can operate, update, and modulate independently. Unlike a centralized system, localized systems are more unpredictable, leading to more overhead in terms of managing complexity, especially in control and behavioural composition. In such systems effective communication is of high importance.

Centralized topologies create a uniform and, in some cases, more responsive experience, however, bottle necks in communications are still expected. There is a limited amount of data that can pass through a system as messages need to be sent from a central location.

Centralized systems and localized systems both have their advantages and disadvantages. To recreate a truly natural experience, there is a need for a certain degree of both strategies within a system. The integration of these kinds of systems needs to be handled with care. Communications and protocols in hybrid systems could lead to conflicts between localized behaviours and their centralized counterparts, resulting in message collisions and system malfunctions.

Granular Synthesis

Granular synthesis, also known as Granulation, is a synthesis technique that is a composition of many different grains of sound. Curtis Roads\(^6\) describes a grain as short audio sample as short as 20 milliseconds with an amplitude determined by a fixed envelope. Each grain could independently have a source, pitch and amplitude. The grains are mixed together to create complex soundscapes. There are three main classes of granulators based on the type of audio source: Tapped Delay Line Granular Synthesis, Stored Sample Granular Synthesis and Synthetic Grain Granular Synthesis.\(^7\)

---


7 Ross Bencina, Implementing Real Time Granular Synthesis Synthesis.
1. Tapped Delay Line Granular Synthesis: reads its grains from a delay line buffer. The delay and playback time create time stretching and time smearing effects on real-time input.
2. Stored Sample Granular Synthesis: uses pre-generated or sampled wave tables for grain source.
3. Synthetic Grain Granular Synthesis: makes use of other sound synthesis techniques such as oscillators or FM synthesis to create grains.

Each implementation method of Granulation comes with a set of pros and cons. It was determined that for these installations the most fitting option would be Stored Sample Granular Synthesis for two main reasons: one being it is the most efficient and flexible method which is optimum for microprocessors; and two, it gives the option of programming more intimate behaviors by allowing real-time recording of the surrounding environment.
Implementation Strategies for Granular Synthesis

The advantages of using a granular synthesis over other systems in Beesley’s large-scale installations are both perceptual and aesthetic. In granular synthesis, the audio is comprised of clouds of short sounds making it easier to locate the audio source, which is important in large complex installations with many loudspeakers. Clear audio location allows participants to easily see the cause and effect relationships between kinetic motion, light activity, and sounds being generated. Aesthetically granular synthesis has a large parameter that allows for the creation of sounds ranging from natural to purely synthetic, with states in-between being quasi-organic, potentially alien, and familiar at the same time.

Audio processing is computationally expensive; hence, when using microcontrollers, the absence of computational power is much more noticeable. A promising tool in microcontroller-based audio production is the highly powerful Teensy platform. In addition to its power, PJRC, the manufacturer of Teensy, has launched an open source audio library to facilitate their use for audio processing.

References:


9 "Teensy 3" PJRC. [https://www.pjrc.com/store/teensy3_audio.html](https://www.pjrc.com/store/teensy3_audio.html)
In *Amatria* a node was created to resemble a processing unit comprised of teensy 3.6, teensy audio shield1, an amplifying board2, and a microphone (*Image 4*). This set of hardware enabled the creation of a device capable of receiving one stereo input (or two mono inputs) and outputting stereo audio. The receiving inputs were set to two separate mono audio channels, one being a line-in input for external audio inputs, and the other as a microphone input.

The behavioral design in *Amatria* was selected to create a feedback system whereby the sculpture would record its surrounding environment and respond by repeating the recorded data in an altered form. The implementation of this concept was the first step in achieving real-time granular synthesis in the sculpture. Mapping every node to behave like a single grain, triggered by the localized sensors, would give out short bursts of recorded audio captured from the surrounding area as it rippled out from the moment of stimulation.

This approach seems more sympathetic with the overall system design logic; however, it does come with some drawbacks, as audio electronics are sensitive and complex. Every component within a large-scale installation requires extensive attention. The smallest miscalculation within the hardware design of the boards, including how every component is powered, how each audio processing node is connected to higher level processing units, and how external audio is routed to the board could lead to noise leaks, and/or ground hums. It would not take much noise in an audio system to render the entire system unusable.

While similar behavior was incorporated for *Aegis* and *Noosphere* project, the underlying system architecture was fundamentally different. The sculpture used a centralized system with each node being a dedicated simple audio output module sourced by a master laptop to generate all of the sound using a software environment called 4D Sound10, an ambisonics software developed to create immersive auditory environments. This approach resulted in a much more simplified audio modules, tailored towards the goal of delivering Hi-Fi audio throughout the sculpture. Each unit consisted of a Dante analog module - an audio through Ethernet protocol with its own dedicated hardware to translate digital audio from the Ethernet input to analog audio signals. This technology was much easier to work with, due to its incredible noise free performance, while providing flexibility with the number of active audio channels. This radical shift in systems architecture...
caused a change in the experience of the overall system. Stepping away from the audience’s interaction and acoustic properties of the environment being the key determinations of the audio record/playback, the technology allowed a professional musical composition to be at the center of work that could played out when the viewer stimulated the sculpture.

Conclusion

In conclusion, the paper explores the pros and cons of different approaches towards an appropriate implementation for an audio systems architecture. The answer is not binary. For the best outcome, a system should live with the prospect of modularity at its core while keeping a centralized hierarchy of command. For example, the node-based approach in Amatria illustrates generative audio processes happening at micro-controller level while a centralized command system keeps track of each node from the top. Amatria as a testbed has given the opportunity to implement underlying system to support high quality audio systems, paving the way for future updates to the granular engine at the heart of each node. Currently in research and development stages, the technology could lead to the creation of a stand-alone modular Hi-Fi audio system with more capabilities than a regular speaker. The outcome could provide artists, musicians and creative coders a new tool for musical expression or as parts of a larger system.

Amirbahador Rostami is an Undergraduate student currently studying Computational Arts at York University in Toronto, Canada. He has collaborated as a research and project assistant to Graham Wakefield on Conservation of Shadows (2017) and Worldmaking and Technolodge, a class collaboration with the University of Alicante, the Technical University of Delft, and York University (2017), mostly focusing on hardware design and production within installations; and Philip Beesley Architect Inc./Living Architecture Systems Group on Amatria, Aegis and Noosphere, focusing on audio systems development. Currently he is working within Dispersion-lab, an ambisonic research lab run by professor Doug Van Nort focusing on spatial sensing using Microsoft Kinects.
Mark-David Hosale is a computational artist and composer. He is an Associate Professor in Computational Arts at York University, Toronto, Ontario, Canada. He has lectured and taught internationally at institutions in Denmark, The Netherlands, Norway, Canada, and the United States. Mark-David’s work explores the boundaries between the virtual and the physical world. His practice varies from performance (music and theatre) to public and gallery-based art. His solo and collaborative work has been exhibited internationally at the SIGGRAPH Art Gallery (2005), International Symposium on Electronic Art (ISEA 2006), BlikOpener Festival, Delft, The Netherlands (2010), the Dutch Electronic Art Festival (DEAF 2012), Biennale of Sidney (2012), Toronto’s Nuit Blanche (2012), Art Souterrain, Montréal (2013), and a Collateral event at the Venice Biennale (2015), among others. He is co-editor of the anthology, Worldmaking as Techné: Participatory Art, Music, and Architecture (Riverside Architectural Press, 2018).

Poul Holleman is Co-Founder and Lead Software at 4DSOUND, and teaches Creative System Design at the Utrecht University of the Arts. His work revolves around interactive installations, instruments, and interface design. His particular field of expertise is spatial sound, the medium that he has been exploring with 4DSOUND since 2007. 4DSOUND develops custom systems for composition, performance, and interaction with spatial sound. An extensive software suite and specific hardware technology enable to create immersive sound experiences in a wide range of applications and contexts. 4DSOUND’s collaborations are interdisciplinary; e.g. in electronic music, interactive sculpture, sound as medicine, theater, scientific research, and more. In 2015 4DSOUND established the Spatial Sound Institute in Budapest, Hungary, an R&D facility that hosts Artist Residencies to develop and present projects in spatial sound. Recent examples of other projects are: Biblioteca de Ruidos y Sonidos, Simona (Wageningen University), and diverse interactive AV installations with Nick Verstand.