A VIRTUAL ACOUSMONIUM FOR TRANSPARENT SPEAKER SYSTEMS

Elliot Kermit-Canfield
Center for Computer Research in Music and Acoustics, Stanford University
kermit@ccrma.stanford.edu

ABSTRACT

An acousmonium, or loudspeaker orchestra, is a system of spatially-separated loudspeakers designed for diffusing electroacoustic music. The speakers in such a system are chosen based on their sonic properties and placed in space with the intention of imparting spatial and timbral effects on the music played through them. Acousmonia are in fact musical instruments that composers and sound artists use in concerts to perform otherwise static tape pieces. Unfortunately, acousmonia are large systems that are challenging to maintain, upgrade, transport, and reconfigure. Additionally, their sole task is limited to the diffusion of acousmatic music. On the other hand, most computer music centers have incorporated multichannel sound systems into their studio and concert setups. In this paper, we propose a virtual acousmonium that decouples an arbitrary arrangement of virtual, colored speakers from a transparent speaker system that the acousmonium is projected through. Using ambisonics and an appropriate decoder, we can realize the virtual acousmonium on almost any speaker system. Our software automatically generates a GUI for metering and OSC/MIDI responders for control, making the system portable, configurable, and simple to use.

1. INTRODUCTION

An acousmonium is traditionally a system with multiple, characteristically sounding speaker groups positioned in space. A performer mixes a stereo audio piece through the system in real-time, imparting spatial and sonic coloration effects on the music as it is diffused. Acousmatic music is specifically written to be rendered through these speaker systems. Composers will deliberately mix their electroacoustic works to be two channels so that there is still the possibility of "interpreting" the music through an acousmonium—the acousmonium is in fact the instrument on which a composer or sound mixer breathes life into a live performance of a fixed media work.

Acousmonia have one purpose: to diffuse sounds through their highly colored speakers. Although they contain many loudspeakers, each speaker's frequency response might be different. Other systems, which we will refer to as transparent speaker systems, contain two ore more loudspeakers that are meant to amplify electroacoustic signals without imparting any timbral change on the signals played through them. That is to say signals passed through these speakers should contain no audible distortion or filtration effects, and should simple amplify the sounds. [1] While an acousmonium can not be repurposed to project amplified audio signals without coloration, a transparent speaker system can render audio for a multitude of applications. In fact, one can use digital signal processing to mimic the colored response of an acousmonium.

Acousmonia hold an important status in computer music history. Even though the concept of the acousmonium is quite old and there have been many technological advances in spatial audio and signal processing, acousmonia are still relevant today. Composers and sound artists still write music to be performed on acousmonia and hold strong opinions on the entire process. An acousmonium has immense creative possibilities for diffusing music concréte and we would be at a great loss without these systems.

Unfortunately, acousmonia are usually large and complex systems that are generally location specific and not particularly portable. Furthermore, one cannot depend on finding an acousmonium in a concert venue. These days, transparent speaker systems for spatial audio are almost guaranteed in concert venues, studios, and computer music centers. For these reasons, we need a way to properly diffuse acousmatic music on a more general, transparent speaker system.

In this paper, we describe a software solution for a portable and reconfigurable virtual acousmonium. Our proposed system allows a user to configure a collection of virtual, colored speakers and place them in space. It automatically generates MIDI and OSC hooks for controlling the system as well as a GUI for metering and audio file playback. By rending the output as an ambisonic encoded signal, this system decouples the virtual acousmonium from the speaker system it is rendered through.

2. ORIGINS

The first acousmonium was designed in 1974 by François Bayle and used by the Groupe de Recherches Musicales (GRM) to diffuse musique concréte through more than 80 loudspeakers [1]. Over time, other acousmonia have appeared, most notably the Gmebaphone and Cybernéphone at the Institut international de musique électroacoustique de Bourges and the Birmingham ElectroAcoustic Sound Theatre (BEAST) at the University of Birmingham [2–4]. All of these systems have been through many reconfigurations and

[1] This is of course impossible.

renovations and are still in use today. Like church organs, these systems are often large and complex. The sheer number of loudspeakers and their associated amplifiers, cables, and mixer channels mean that an acousmonium is typically complicated to setup. Additionally, the acoustics of the space in which an acousmonium is performed is inherently coupled to the timbral and spatial effects of the speakers. For these reasons, acousmonia are often housed in specific locations where the room acoustics and speaker systems compliment one other. [2] Moreover, since these systems are complex, maintenance requires intimate knowledge of the system and they are often built and upgraded over many years. Acousmonia have played an important role in shaping the development of the electroacoustic music tradition, especially in western Europe. For some reason, there seem to be very few such systems outside of Europe. Parallel to the tradition of acousmatic music, electroacoustic music diffusion has also tended towards 3D sound rendering systems. Various techniques have been employed for positioning sound in space (e.g., wave-field synthesis, ambisonics, vector based amplitude panning, etc.). The systems designed for projecting sound using these techniques are highly varied. Simple systems have speakers in a single pantophonic ring around the listeners while more complex systems involve speakers that have height displacement as well. Larger number of speakers in these systems roughly translates to more sophisticated spatial processing and a higher ability to localize sounds in space.

Acousmonium means speaker orchestra. Like a conventional orchestra, groups of speakers have characteristic tonal qualities, spatial locations, and radiation patterns. An artist diffuses a stereo, concréte electroacoustic work through the system much the way a simple melody could be adapted to make use of the full sonic and creative capacity of an orchestra. An acousmonium system contains carefully tailored groups of speakers that have specific purposes for coloring or effecting the sound passed through them. Speakers are characterized by their roles, and include ensemble speakers that produce band-filtered outputs in different frequency ranges, highly colored solo speaker instruments, and effects speakers (e.g., ones for spatial panning, or extreme vertical displacement) [5, 6]. The music is performed live, often from a notated score, and unites a stereo music concréte composition with the interpretation through the speaker system and room. This last step is crucial—composers feel that a composition with more than two channels already has a spatial aspect and that there is no more room for interpretation.

3. JUSTIFICATION FOR A VIRTUAL ACOUSMONIUM

Acousmatic music has a deep musical tradition, but it also has severe limitations. The fact is that these systems are so complex that reconfiguration and transportation become arduous tasks. Maintenance requires detailed knowledge of the system to diagnose problems and upgrade components. Moreover, acousmonia are usually analog instruments that make use of custom mixers, equipment, and electrical components acquired and built for specific systems. Not only does this complicate the upkeep of an acousmonium, but also means that that knowledge does not necessarily translate from one acousmonium to another.

A virtual acousmonium is not bound to a specific space or hardware setup. Our system is designed in a way that it is agnostic to the final diffusion system and can be reconfigured virtually on the fly. Last, this system transcends the traditional acousmonium paradigm. The first goal is to replicate the behavior of an acousmonium—in this case the ability to diffuse a stereo audio piece to a system of virtual, timbrally-colored and spatially separated speakers. We also expand the capabilities of the system so that one can diffuse multichannel works through the system and create hybrid systems of colored and transparent speaker systems.

Our system outputs ambisonic encoded signals allowing us to insert our virtual acousmonium into the signal path of any transparent speaker system. This means that not only can we diffuse acousmatic music in the same concert as music meant to be played through a transparent system, but we can also reconfigure the acousmonium—changing its speaker configuration—between musical works. More importantly, we can reproduce *specific* acousmonium setups, meaning that we can virtually reproduce existing and historic acousmonium systems. This is important for archival purposes and allows us to revive musical compositions that could not be correctly performed due the lack of an appropriate system. It also untethers music written for a specific acousmonium allowing the music to be performed anywhere.

Acousmonia are not simply speaker systems—they are musical instruments in their own right that require years of training to master. Because they are not portable and live in specific spaces, one can not necessarily access the acousmonium in order to practice or experiment with the diffusion of one's music. With our virtual acousmonium, one can approximate specific real-world acousmonia over headphones in order to practice the instrument. Naturally this does not replace working with the true system, although it can serve as a proxy much like a practice organ stands in for the one in the recital hall.

While one can easily up-mix an audio file to a large, transparent speaker system, the same is not true for an acousmonium despite the large number of speakers. The characteristic sound of the speakers renders the these systems useless for applications where the system is not supposed to impart coloration on the signal. At the same time, 3D sound systems have become prominent in computer music institutions. Octaphonic and larger systems, often with speakers displaced at various elevations, are almost guaranteed for concert presentations. Since the virtual acousmonium autogenerates its routing matrix and MIDI/OSC responders, the entire system can be treated as a plug-and-play module that is inserted into a larger system.

Because of the wide availability of transparent speaker arrays, we implemented a framework for performing acousmatic music, in the traditional sense, without a physical acousmonium.

[2] Although some institutions have the means to transport smaller systems for festivals.

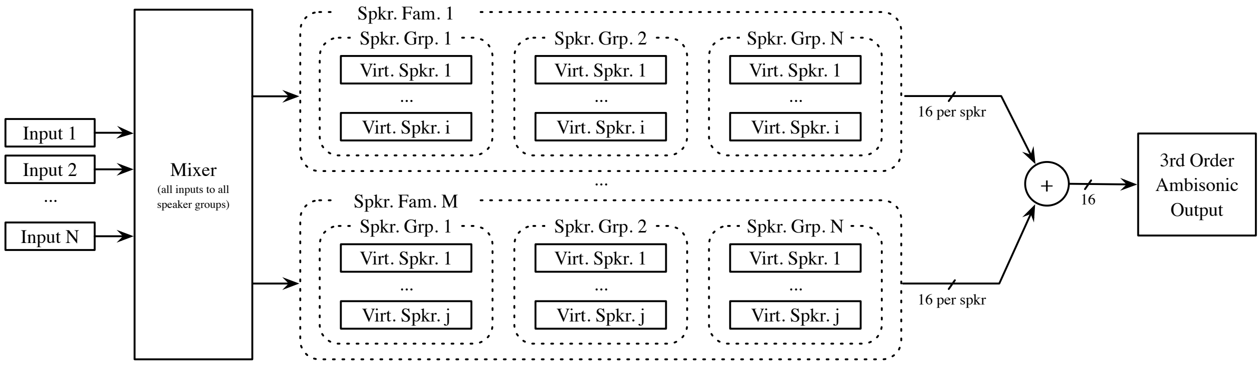

Figure 1. Acousmonium system diagram.

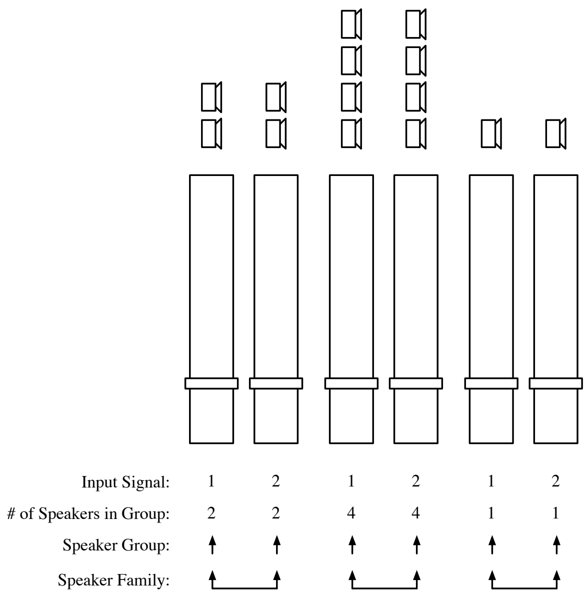

Figure 2. Auto-generated fader arrangement for two source channels and three speaker groups.

4. THE VIRTUAL ACOUSMONIUM

At its core, the virtual acousmonium is a configurable diffusion tool that sits between the audio input and the reproduction system. It is written in SuperCollider to be cross-platform compatible and work with an arbitrary speaker arrangement. We provide a simple configuration for interfacing the software with audio and MIDI hardware and default/example configurations for virtual speaker descriptions. Our virtual acousmonium provides convenient methods to load and playback audio files as well as a mechanism for reading input directly from the sound card. When you load a speaker configuration, the system auto-generates the appropriate virtual speakers, audio routing matrix, OSC/MIDI wrappers for controlling the system, and a GUI for monitoring the system.

4.1 Software

We wrote this virtual acousmonium in SuperCollider because it provides the right combination of efficient signal processing and configurability. SuperCollider is an open source audio programming language and environment with a strong developer community and user base. Additionally, SuperCollider handles multi-channel audio data in a convenient way for implementing complex audio routing.

In our benchmarks, running a complex system with 2 input channels and 80 virtual sources (each with its own ambisonic panner), the peak cpu load was less than $\frac{1}{4}$ of a single core of a 2011-era i7 processor, and the average cpu load close to 10%. This tool free and open source software distributed under a GNU GPL. [3]

4.2 Speaker Descriptions and Routing

Virtual speakers in our system are described by SuperCollider SynthDefs that accept a monophonic input and output 3^{rd} order ambisonic encoded signals. Each virtual speaker has an associated angle (θ), azimuth (ϕ), and relative gain (γ) that describe where in space the virtual speaker should appear. We use open source SuperCollider plugins for higher-order ambisonic encoding.

Since the speakers are defined in SuperCollider code, custom speaker responses can be achieved by extending the speaker definitions we have included. As long as a speaker definition respects our speaker definition format, arbitrary code can be evaluated to model the desired speaker responses. This processing can take the form of simple signal processing algorithms like filtering, distortion, or compression, but could also include a full physical model of a specific speaker, horn, or driver. The format for the speaker definition includes several reserved arguments (e.g., spkrAzmth, spkrGain, etc.), a method for sending OSC messages to update meters, and a unit generator graph that conforms to a "monophonic input $\rightarrow$ processing $\rightarrow$ ambisonic panner" paradigm. In addition to the reserved arguments, the system has appropriate mechanisms for interpreting extra parameters that are not generalized for all

[3] This virtual acousmonium software can be downloaded at `https://ccrma.stanford.edu/~kermit/website/acousmonium.html`.

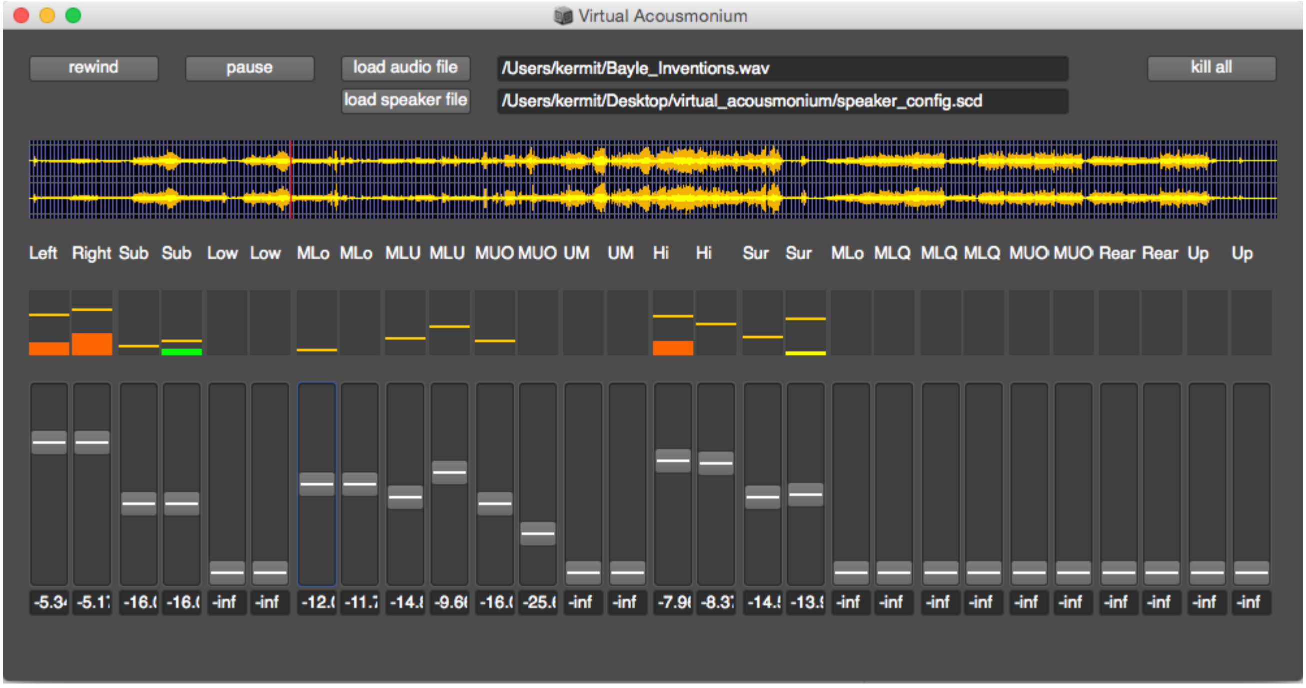

Figure 3. SuperCollider QT graphic interface.

speaker descriptions (e.g., a speaker with a filter might need a mechanism for setting the filter cutoff frequency and Q value).

The speaker routing in our system is described by a hierarchical configuration file. Virtual speakers can be grouped together to create the impression of more complex speaker arrangements, and multiple groups of speakers can be collected into "families" in order to facilitate control over multiple, similar-but-spatially-separated speaker sets. A full system block diagram can be seen in Fig. 1. The speaker configuration is stored as an array of speaker families, where a speaker family is an array of speakers and each speaker is a dictionary of associated parameter keys and values, see code snippets 1 and 2.

Code 1. Data structure for the speaker configuration

```
[               // spkr family containing
  [(spkr), (spkr), ...], // spkr grps
  [(spkr), (spkr), ...],
  ...
]
```

Code 2. Single virtual speaker description

```
(
  \spkrType  : \speaker_hp,
  \name      : \Mid_Range_L,
  \params    : (
    \spkrAzmth  : (-pi/4),
    \spkrElev   : 0.0,
    \spkrGain   : 1.0,
    \cutoff     : 2000,
    \rq         : 1.0
  )
)
```

The output of all the virtual speakers gets summed together and sent out of the system effectively decoupling the virtual speakers from the physical, real-world speaker system. Since the system description is just a piece of software, they speaker arrangement is flexible and can be reconfigured on-the-fly to accommodate compositions that are written to be diffused through specific speaker setups.

4.3 Interfaces

From the speaker configuration, SuperCollider will automatically generate a QT GUI and set of OSC and MIDI responders for the system. The system provides both an interface for processing external inputs to the SuperCollider program as well as a way to load an audio file into an internal playback system. If the later is used, basic transport controls (play/pause, seek, rewind) and a waveform plot with playhead are provided for navigating the sound file. [4]

Inherently, all inputs are mapped to all speaker groups. The number of speaker groups determines the number of fader groups and the number of inputs the actual number of faders, see Fig. 2. This is then exposed to the users as a GUI fader bank. Each fader can be controlled with OSC or MIDI allowing the use of external interfaces. The system automatically generates faders and level meters as well as OSC and MIDI responders for each speaker group, based on the hierarchical speaker description in the config file.

An example of the GUI generated for a stereo input signal and 24 virtual speaker groups is shown in Fig. 3.

4.4 Limitations

One major issue with using virtual sources placed in space with ambisonics instead of real speakers is the fact that

[4] Very soon the framework will also support computing and displaying a spectrogram view of the sound file as well.

it is challenging to accurately simulate speaker radiation patterns. Physical speaker enclosures have nonlinear radiation patterns that are impossible to reproduce in a virtual ambisonic environment. Even so, we can use reverb, and filters, as well as a cluster of point sources with anti-phase components to approximate speaker orientation, depth, and radiation patterns.

5. CONCLUSIONS AND FURTHER WORK

In this paper, we have introduced a tool for generating arbitrarily complex arrangements of spatially-located and colored virtual speakers and interfaces for diffusing audio through the system. By encoding the output of this virtual acousmonium using ambisonics, our system is modular and can be used with any transparent speaker system. Our system automatically generates a graphic user interface for interacting with and monitoring the system as well as providing OSC and MIDI responders so one can control the system with external devices.

In the future, we intend to add a graphical interface for creating the speaker configuration files. While convenient for programmers, the current textual method does not provide any visual feedback that speakers are placed in the correct locations. We envision a tool where a system designer can enter speaker locations and types in a context where there is no ambiguity of the virtual speaker's positions.

This tool does not seek to replace acousmonia, but rather complement these impressive systems. The portability and configurability of a virtual system increases the umbra of this musical tradition, and the signal processing possibilities of this tool will benefit composers and diffusion artists alike.

6. REFERENCES

[1] F. Bayle, *Musique acousmatique: propositions... ...positions*. Buchet/Chastel, 1993.

[2] C. Clozier, *Proceedings volume III of the 1997 works of the International Academy of Electroacoustic Music*. Editions Mnemosyne, 1998, ch. Composition/Diffusion in Electroacoustic Music, pp. 233–281.

[3] J. Harrison, "Diffusion: theories and practices, with particular reference to the beast system," *eContact!*, 1999.

[4] H. Tutschku, "On the interpretation of multi-channel electroacoustic works on loudspeaker orchestras: Some thoughts on the grm acousmonium and beast," http://www.tutschku.com/content/interpretation.en.php.

[5] J. Prager, "L'interpretation acoumatique: fondements artistiques et techniques de linterpretation des œuvres acousmatiques en concert," 2002.

[6] A. V. Gorne, "L'interprétation spatiale. essai de formalisation méthodologique," *Revue Déméter*, December 2002.