Spaces @ CCRMA: Design and evolution of our 3D studio and concert diffusion systems

Fernando Lopez-Lezcano

To cite this version:

Fernando Lopez-Lezcano. Spaces @ CCRMA: Design and evolution of our 3D studio and concert diffusion systems. EAA Spatial Audio Signal Processing Symposium, Sep 2019, Paris, France. pp.109-114, 10.25836/sasp.2019.29. hal-02275190

HAL Id: hal-02275190
https://hal.archives-ouvertes.fr/hal-02275190
Submitted on 30 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
This paper describes the design and evolution of three multi-speaker systems that provide 3D surround sound diffusion capabilities and are used for research, composition and concert diffusion at CCRMA, Stanford University. The first of the two permanently installed systems is housed in the Listening Room, a full 3D studio with 22 speakers and 8 subwoofers, including 7 speakers mounted below the acoustically transparent floor. The second permanent system is in our small in-house concert hall, the Stage, which after a recent full audio remodel is equipped with 56 speakers and 8 subwoofers. Finally, our concert diffusion system, the GRAIL (the Giant Radial Array for Immersive Listening), is a “portable” system that can be set up and calibrated in a couple of days and can currently support up to 32 speakers and 8 subwoofers.

2. THE LISTENING ROOM

The Listening Room was built when our building, The Knoll, was remodeled in 2005. It was designed to be an acoustically dry, low noise floor, high quality full 3D studio. Part of the floor was opened and covered with an acoustically transparent grid, so that speakers could be installed below the listening position. 16 speakers were initially installed in 2006, 4 below the grid floor, 8 around the listening position and 4 more hanging from the ceiling, all of them driven by a 16 bus Tascam DM3200 digital mixer.

Figure 1. The Listening Room

After the system had been in use for a while it became increasingly clear that a digital mixer was not the best option for controlling the system. It provided functionality that was not included in the level of sound processing present in a digital mixer, such as equalization and compression, as those functions are best performed in the work computers.

The addition of Ambisonics modes with calibrated decoders for each space make it particularly easy to move pieces between them and fully use the capabilities that each space provides. Pieces can be composed and auditioned in Ambisonics in the Listening Room or the Stage, and then performed in concert with the GRAIL system with minimal or no changes.

In this paper we will describe the evolution of the three systems, and the current state of hardware and software that supports composition, experimentation, research and sound diffusion tasks at CCRMA.
that was not needed, it was unnecessarily complicated to operate and was not expandable. In 2008 we started to design and implement a fully custom sound diffusion system based on one of our fanless Linux workstations. The goal was to have a system that was easy to use and provided a very simple interface to the task of diffusing sound in the studio.

2.1 The OpenMixer hardware and software

The initial design [1] used a computer as the central hub to the diffusion tasks, with a simple user interface provided by two USB fader boxes. We selected SuperCollider as the computer language to use for its implementation, and also tried to leverage as much free software as possible to simplify the design task (Jconvolver, Ambdec and others). Everything was connected together through Jack, a low latency sound server, and the system was designed to boot unassisted into the control software.

![Figure 2. Listening Room: OpenMixer user interface](image)

Two PCI RME audio interfaces with MADI and ADAT I/O, and SSL converters, provided the audio I/O. Audio could be connected to the system through 16 analog inputs, two ADAT digital interfaces with word clock or an ethernet connection using Netjack Jack clients - the easiest way at the time to connect to the system 24 channel audio coming from external laptops and computers. A dedicated 24 channel ADAT connection was provided for a desktop workstation permanently installed in the studio, so that users would be able to use the system with minimal training.

The system was designed with two modes of operation. The Direct mode allows any input to be routed directly to any speaker, with arbitrary mappings. The Ambisonics mode routes the inputs to an Ambisonics decoder tuned to the speaker array. This makes it easy to work in Ambisonics without having to deal with designing and tuning decoders for the speaker array.

The system proved to be stable, and it was expanded in 2011 [2] to 22 speakers and 4 subwoofers. The spatial arrangement of the speakers was chosen to be $1 + 6 + 8 + 6 + 1$, selected so it could provide accurate decoding of 3rd order periphanic Ambisonics audio streams.

In 2014 digital speaker equalization was added to the system. In addition to matching the delays and levels for all speakers, digital equalization filters were designed for each speaker using the DRC (Digital Room Correction) software, to provide the best frequency and phase response possible. The sound quality improved significantly, specially for the speakers installed below the floor which had always been problematic due to the relatively small size of the pit in which they were housed and the coloration it created.

3. CONCERT SYSTEMS

![Figure 3. D.C Power Lab, preparing for a concert, circa 1980](image)

Through CCRMA’s history immersive music production and subsequent concerts have been part of daily life. Concerts were staged in the D.C Power Lab and later in the Frost open air auditorium using a four channel system, and also in Dinkelspiel and Campbell Recital Hall in the Music Department. The Ballroom, a small space that was used for teaching and concerts in our building, and later our backyard at The Knoll, hosted many events that used arrays of up to 8 speakers driven from analog and later digital mixers.

The planned construction of a new concert hall at Stanford, the Bing Concert Hall, spurred us to start planning events for the inaugural season which was to happen in 2013. Starting in 2009 we bought 8 new speakers for our concert system, then expanded to 16 and finally to 24 (with 8 matching subwoofers), which would enable us to diffuse sound in full 3D surround in the Bing spaces. Our Transitions outdoor concerts in 2011 and 2012 were testing grounds for new diffusion technologies, including higher order Ambisonics rendering and digital real-time simulation of real acoustic spaces. We started writing software inspired in our Listening Room control system to use one of our high performance custom fanless workstations [3] as the diffusion engine for concerts, including software that allowed us to use a 32 channel digital snake system as a low cost high channel count D/A and A/D [4]. In 2011 we successfully used this new system to diffuse a whole concert, including completely remoting the computer as all peripherals could interface with it through ethernet cables.

In 2013 we staged several events in the Bing Main Hall and its black box rehearsal space, the Studio. In both cases
Figure 4. Transitions 2011 concert diffusion control system

with dome speaker configurations of 12 + 8 + 4 + 1 which enabled us to render pieces with the same type of flexibility we enjoyed in our Listening Room, but for a large audience. Probably the most ambitious was the digital simulation of the acoustic space of Hagia Sophia [5] (part of the Icons of Sound project), in which the Cappella Romana group performed in real-time, singing byzantine chanting that was composed ages ago for that space. Their voices were input into our system which was mixing and processing multiple impulse responses, finally routing them to a 24 speaker dome, in effect creating a virtual space for the singers to perform in.

The diffusion system, later called the GRAIL (the Giant Radial Array for Immersive Listening) evolved over the years and currently can be deployed with up to 32 speakers and 8 subwoofers [6]. It is calibrated digitally after it is installed, with the goal of bringing high quality studio sound to the concert hall. In addition to the regular concerts at Stanford, we have staged events abroad, first in the context of the 2016 Sound Symposium in Newfoundland, and more recently in Taipei, Taiwan, as part of the 2018 season Innovation Series events in the National Theater and Concert Hall (NTCH).

4. THE STAGE

The Stage, our small in-house concert hall, also built during the 2005 renovation, is a cathedral ceiling room with very good acoustics and a low noise floor that can seat between 40 and 80 people depending on the type of event being staged. It is a multi-purpose space that can be used for concerts, lectures and classes. Its initial sound diffusion system was the same 8 channel ring of high quality Adam mid-field studio monitors that used to be housed in our Ballroom before the remodel.

The system was upgraded in 2009 to 16 speakers in a two ring configuration (8 + 8), with the upper speakers hanging from the ceiling pipes and with 8 subwoofers added to the low ring. The upper speakers were later repositioned for better rendering of 3d spatial sound, eventually hanging in an 8 + 6 + 2 configuration. A Yamaha DM1000 digital mixer is used to diffuse sound and is directly connected to the speakers.

4.1 The GRAIL comes to the Stage

While we had been using the Listening Room for many years, and the new GRAIL was being deployed for concerts several times every year, our diffusion systems had some shortcomings. In terms of spatial resolution both systems could accurately decode Ambisonics up to third order (and fourth order if some errors were ignored), but
composers and sound artists had started using higher orders routinely. We could not play their pieces in their original resolution, or compose HOA content with our existing facilities. We were also limited in physical size. Our only 3D capable permanent setup was the Listening Room, and the bigger concert spaces were only available for a few days before the concerts. Hardly what is needed to compose and audition new music.

At the end of 2015 we proposed to upgrade the diffusion capabilities of the Stage to convert it into a state-of-the-art 3D space which could be used for concerts, research and composition. Adding 32 small speakers to the existing system would transform it into a 48.8 system that would allow it to render 5th or 6th order Ambisonics pieces correctly, or the equivalent in spatial resolution using other systems [7].

But the Stage is not only a concert hall, it is also regularly used for classes, lectures, demos and other events that do not need or want a high spatial resolution speaker array. In fact, the majority of users require access to just stereo playback. As the CCRMA concert events combine live performers, touring musicians and researchers, many concerts do not deal with 3D surround sound, and use mostly stereo projection.

It quickly became apparent that the existing system could not be replaced. In fact a key requirement for the project to go forward was to leave the existing system in place, with no user visible changes. And while it was obvious that the full system would require a control computer and software similar to the one running in the Listening Room, the legacy system had to work even if the control computer was turned off.

All those requirements together made for quite a challenging project.

One of the key tasks of the design process was finding an audio transport technology that could meet these requirements at a reasonable cost. After much research we selected AVB (Audio Video Bridging) [8] [9] as a suitable technology that could interface with our open Linux-based systems, specially since there was a family of products manufactured by Motu that seemed to meet our requirements.

We created our audio system by interconnecting several Motu audio interfaces through ethernet, and using their USB interfaces as entry points for computers into the system. The internal routing matrices in the audio interfaces were used to reconfigure connections and switch operating modes. And since the Motu cards can be remotely controlled, a small Raspberry Pi powered computer with a touch screen could be used to switch operating modes.

Figure 9 shows the hardware of the minimal system in “Digital Mixer Mode”, in which the old legacy 16.8 system is controlled through the DM1000 mixer.

Figure 9. Stage audio system: signal routing in Digital Mixer mode

Figure 10 shows the expanded system in OpenMixer mode. The audio interfaces used for the legacy mode are now a subset of the full system and their audio routing is completely different. All signals are now processed by the control computer before being sent to the speakers.

The upgrade was done in several stages, it took almost two years, and there is still software work to do. During this time we also upgraded nearly all of the original speakers for a much better basic sound quality.

The current Stage audio system consists of 8 Motu audio interfaces interconnected through AVB and ethernet cabling, 56 main speakers of different sizes, and 8 subwoofers. Control and user computers are connected to the system through the USB interfaces of the Motu boxes, which can support up to 64 channels under Linux with the proper firmware. All speakers are digitally equalized and we can easily run 5th or 6th order Ambisonics decoders tuned to the system.

The inaugural concerts happened at the end of 2017, after the first stage of the upgrade, and the system has been operating successfully since then.

5. EXPANDING THE GRAIL

The experience we gained with AVB systems during the Stage upgrade has percolated to our other diffusion systems as well.

The GRAIL has moved away from our custom software and digital snake solution, which proved to be very reliable for years but limited us to 32 channels of I/O, to a
small network of between 4 and 6 Motu audio interfaces interconnected through ethernet cables. This has proved to be very flexible as it enables us to position the audio I/O where it is needed, makes all the analog connections shorter and enables us to connect multiple computers to the core system and route them appropriately as needed using AVB streams and the internal connection matrices of the interfaces. It also allows us to expand beyond the previous limit of 32 channels of I/O.

The control computer has changed as well. We had been using one of our high powered Linux fanless workstations as the control computer for both tasks: GRAIL core functionality and the diffusion of pieces in concert (this is a 4 or 6 core computer with 32 or 64G of RAM and fast SSD disks). This reduced the reliability of the core system as the computer had to deal with very dissimilar tasks that can interfere with each other.

We are now using two computers instead. The much smaller, more portable machines are based on Intel quad core NUC systems, and are housed in special purpose passively cooled fanless cases (the original NUCs are very small but use a noisy fan to cool the processor). The main control computer runs the core software (written in SuperCollider) which handles audio routing, master gain control, delay and level matching, speaker digital equalization, digital crossovers, and Ambisonics decoding. A second NUC system is used to diffuse pieces, either by playing back fixed media content through software such as Ardour, or acting as a digital mixer for live pieces, or combinations of the above. Both systems run Linux using real-time optimized kernels, and with proper tuning of all system priorities so that we can run at very low latencies.

Both computers connect to separate Motu audio interfaces, and their audio streams are routed through AVB and ethernet cabling. In fact there is nothing special about the diffusion computer, and with AVB we can have several computers, including performers’ laptops, connected to the system, and seamlessly switch between them.

6. UPGRADING THE LISTENING ROOM

The experience of using the new Stage system led us to upgrade the system in the Listening Room as well. One feature shared by the new Stage system and our updated GRAIL was the ability to connect multiple external computers to the system through USB audio interfaces. That is a much easier path to connecting to the system than the old NetJack ethernet interface in the Listening Room.

We have used this system in our last round of concerts and it has proven to be very configurable and reliable. We have also recently added a small single fader control surface to the control computer, which we configured as a master level controller. We can also select the operating mode, type and order of the Ambisonics decoder, and we are going to add presets for diffusion for common speaker configurations, such as a ring of 8 speakers or 5.1/7.1. This makes it much easier to control the system during a concert.
We upgraded the whole system, adding a new, faster rack-mounted control computer and replacing the audio hardware with Motu interfaces. The OpenMixer control software was upgraded to use the new hardware and moved to the new computer. The user interface did not change, and the only compromise was to only allow mixing of different sources up to a maximum of 64 channels (the old system could mix all sources at the same time).

Figure 13. Listening Room hardware and audio routing block diagram

We also replaced all the speakers by newer, much better sound quality Adam A7X/A5X units, and replaced the original four subwoofers with eight new new Dynaudio 9S units.

7. CONCLUSION

The three systems provide different full 3D surround calibrated environments for composers, sound artists, and researchers to work in. The Listening Room is a very dry full 3D experimental space, the Stage provides a larger space with higher spatial resolution that can also host concerts, and the GRAIL can travel to even bigger spaces and render pieces composed in the other spaces without much change, allowing us to host concerts for much larger audiences. All three systems are in constant use, and in particular our new Stage array has been used frequently in concert over the past 1.5 years.

8. THANKS

Many people contributed over the years to make all of this happen. Jason Sadural kick-started the initial OpenMixer code in 2008, and Elliot Kermit-Canfield helped transform the Listening Room in the 2011 upgrade. Many thanks to Carr Wilkerson for his help in caring for the GRAIL for many years, to Eoin Callery for taking over after Carr left, and to Constantin Basica for doing the same after Eoin started teaching. Without Christopher Jette the upgrade to the Stage to a state-of-the-art system would not have happened at all, and we can touch control it thanks to the work of Carlos Sanchez Garcia-Saavedra. Matt Wright helped immensely with the 2018 round of upgrades to the Listening Room. Thanks also to the many students that have coiled and uncoiled miles of cables over the years, shouldered speakers, and soldered many small connectors during our upgrades. And most of all thanks to the countless composers that have contributed pieces over the years, making it worthwhile to spend so much time on this. At the end of the concerts we are all smiles.

9. REFERENCES