The Future of Spatial Computer Music

Eric Lyon
Virginia Tech
School of Performing Arts
ICAT
ericlyon@vt.edu

ABSTRACT

Composing computer music for large numbers of speakers is a daunting process, but it is becoming increasingly practicable. This paper argues for increased attention to the possibilities for this mode of computer music on the part of both creative artists and institutions that support advanced aesthetic research. We first consider the large role that timbre composition has played in computer music, and posit that this research direction may be showing signs of diminishing returns. We next propose spatial computer music for large numbers of speakers as a relatively unexplored area with significant potential, considering reasons for the relative preponderance of timbre composition over spatial composition. We present a case study of a computer music composition that focuses on the orchestration of spatial effects. Finally we propose some steps to be taken in order to promote exploration of the full potential of spatial computer music.

1. INTRODUCTION

The creation and manipulation of sound, and the performance of sound into space constitute two fundamental areas of interest that bridge the transition from instrumental music, through analog electronic music, to most recently, computer music. Early interest in the manipulation of musical spatial experience in Classical music is attested to by the antiphonal music of Giovanni Gabrieli and Adrian Willaert, the practice of cor spezzati, and Thomas Tallis's famous 40-part antiphonal motet "Spem in Alum." Interest in spatialization continues in instrumental practice with offstage instruments, in such works as Beethoven's "Leonore Overture No. 2," and Charles Ives's "The Unanswered Question." Charts indicating precise spatial placement of instruments in the symphony orchestra may be found in the music of Bela Bartok, Elliott Carter, Karlheinz Stockhausen, and Henry Brandt, all with the intention of creating striking spatial musical effects.

With the advent of electronic reproduction of sound, the spatial aspect was exploited in short order, with spatial distribution of speakers for the Telharmonium, hidden behind ferns [1], multichannel performances by Theremin in the 1920s [2], and famously with the projection of Xenakis's "Concret PH" and Varese's "Poème Électronique" through the hundreds of speakers in the Philips Pavilion at the 1958 World Fair [3]. Other notable 20th century examples include Osaka Expo 1970 [4], and the still-active Audium performance space [5]. Spatial deployment of sound is often an important aspect of sound art in works such as Janet Cardiff's "The Forty Voice Motet," a reworking of Tallis's "Spem in Alium" for forty loudspeakers. It is evident that sound in space remains a potent research area for contemporary composers and sound artists. Audiences often grasp the special nature of spatial music as compared to more standard stereo or proscenium presentations of music. Music for large numbers of speakers also commands a social presence that stereo music does not require, since spatial electroacoustic music must be performed in special venues with appropriate multichannel sound systems, whereas much stereo electroacoustic music could be fairly appreciated at home on a good stereo system.

2. TIMBRE ABOVE SPACE

Despite continued interest in creating and theorizing about striking spatial experiences as part of digital music as evidenced by such writing as Denis Smalley's "Spaceform and the Acousmatic Image" [6], interest and energy directed toward sound synthesis and processing (referred to here as *timbre composition*) has remained widely predominant over spatial concerns in contemporary computer music practice. There are many reasons for the emphasis of timbre composition over spatial composition. Both analog and digital timbre design techniques achieved striking triumphs early on, with the musical and theoretical works of the competing schools of Elektronische Musik and Musique Concrète. The push into digital music in the late 1950s quickly produced the Music 5 model allowing for broad timbre experimentation, Frequency Modulation, LPC, and the numerous sound transformations made possible through applications of the Fast Fourier Transform. Timbre composition retains considerable momentum from these early successes.

3. TIMBRE TOOLS ARE UBIQUITOUS

When fixed-media electronic music first emerged as an area for compositional research in the late 1940s, opportunities for composers to work in this field were severely limited by the paucity of electronic music studios such as WDR, ORTF, and the Milan Studio di Fonologia. Over the course of the next two decades, analog electronic music studios proliferated to a growing number of university music departments and other institutions throughout the world, such as Columbia, Princeton, University of Illinois at Urbana-Champaign, and Goldsmiths, University of London. Equally important was the development of independent studios, such as The Cooperative Studio for Electronic Music in Ann Arbor, and the San Francisco Tape Music Studio. From the 1960s forward, commercial manufacturers such as Moog and Buchla enabled individuals to purchase electronic music equipment for experimentation in private studios.

The development of computer music facilities followed a similar trajectory, with computer music systems originally developed at Bell Labs migrating to such institutions as Princeton, Stanford, UC San Diego, and IRCAM. By the late 1980s, it was possible, if still alarmingly expensive, for individuals to purchase personal computers with built-in DACs and ADCs, such as the NeXT Cube, and SGI Indigo. By the end of the 20th century, it was quite common for composers to own notebook computers vastly more powerful than the mainframes to which they would have been tethered at institutions only 15 years prior. Today in much of the world, it is rare to find a composer who does not own a computer.

4. TIMBRE INNOVATION MAY BE NEARING SATURATION

The ubiquity of hardware and software tools for sound synthesis and processing is reflected in the music heard at ICMC since its inception in 1974. The prevalence of a concern with timbre in the computer music community is such that as early as 1994, Agostino Di Scipio could propose a two-category model for computer music consisting of *algorithmic composition*, and *timbre composition*. [7] Indeed the concern with timbre composition extends well beyond the domain of computer music into the areas of popular music and sound design for cinema, among many others. However, I would suggest that we now have a timbre problem. Innovations in the tools and techniques for timbre composition have been slowing for some time now. No new computer music technique in the last 20 years has had nearly the impact of Frequency Modulation in the 1970s or FFT-based processing in the 1980s. And our recent computer music seems to reflect this fact. I have heard many fine computer music pieces in the last few years, all of which sounded to me as if they could have been composed using timbre tools available 20 years ago. If this seems like a bold assertion, consider the aggregate evolution in the sound/timbre world of computer music between 1965-1985 as compared with 1994-2014. A slow down of innovation in timbre composition seems evident. This is not necessarily a bad thing. The stabilization of the constitution of the modern orchestra by the early 20th century left room for decades of fruitful musical creation in that medium. The same may hold true for timbre composition. But practice-based computer music researchers may need to look elsewhere than timbre composition for the kinds of fundamental breakthroughs heard in Chowning's early FM pieces, Paul Lansky's LPC composition "Six Fantasies on a Poem of Thomas Campion," Barry Truax's granular synthesis composition "Riverrun," or Christopher Penrose's spectral processing composition "Fraud." Spatial composition is a domain of computer music where such fundamental advances may still be possible. At the same time, spatial research requires sounds to spatialize, and the remarkably rich timbre palette of digital sound developed over the last half century presents a wholly adequate basis for grounding the coming developments in spatial computer music.

5. WHAT IS HOLDING SPATIAL COMPUTER MUSIC BACK?

If spatial composition is so promising, why do we still hear relatively little of it? There is certainly no shortage of papers on spatialization in recent years. And it is significant that the call for music in this ICMC specified an opportunity to perform pieces with as many as 24 channels. In fact we do hear a fair number of pieces for eight channels, which has become something like a new standard at electronic music conferences. But spatial music for large numbers of speakers is still relatively rare.

There are significant obstacles that might explain the current situation. Composers are faced with quite limited access to performance spaces providing installed multichannel systems with large numbers of speakers. There are still relatively few such studios and performance spaces in the world that support composition of computer music for 24 or more speakers. If a composer is not fortunate enough to work at an institution that provides such a space, access can be very difficult indeed. Some institutions that support large multichannel systems do not have a dedicated space for them, such that their speakers must be rigged for every performance (of which there are few, because of the exhausting setup and teardown efforts required), and otherwise remain in storage. Compared to the huge variety of software tools available for timbre composition, the tools for spatial composition are few, and relatively undeveloped. Commercial DAWs like ProTools have very little to offer composers who wish to work in larger multichannel formats than 7.1. And more progressive acoustic compilers like Csound and SuperCollider can present hard-coded limits on channel numbers for reading and writing sound files. Such software obstacles can be overcome, but the solutions require coding skills that not every computer music composer possesses. And finally, composing for large numbers of speakers is considerably more difficult than composing in stereo.

6. HEDGING SPATIAL BETS

In the face of the above-mentioned obstacles, composers often add their own forms of resistance to full spatial practice. When faced with an opportunity to present a piece in a multichannel space with a large number of speakers, a commonly accepted solution is to compose for a smaller number of channels that are available at one's local institution or studio, and then perform a live diffusion into the multichannel space. While not an unreasonable solution, this does limit the ability to make full use of the multichannel space since, as Natasha Barrett, an expert electroacoustic composer and diffusion artist, has pointed out, there remain spatial effects that require more degrees of control than our hands can produce in real-time performance [8]. But even composers who are fortunate enough to enjoy a significant amount of direct access to multichannel performance spaces often wonder if it makes sense to compose a piece that can only be heard in one performance space in the entire world. A piece that settles on the lowest common denominator of what several multichannel spaces have to offer would seem to have more performance opportunities than a piece that takes full advantage of the capabilities of a single performance space. Indeed it must seem a singularly perilous endeavor to compose music for large numbers of speakers at the present time. But the riskiness of the undertaking can hardly be considered an objection in the context of experimental music. Experimentalism is embedded deeply in the DNA of computer music. And as argued above, timbre composition may be a domain showing diminishing returns for experimentation. By contrast, computer music composition for large numbers of speakers is an area ripe for experimentation and discovery. All the elements that make spatial composition such a difficult prospect might well make the pursuit all the more attractive to the most daring composers.

7. SPIRITS: A CASE STUDY

In 2011 I was awarded a Giga-Hertz prize from ZKM, which brought with it an opportunity to compose a piece in the ZKM Klangdom, a multichannel space with 47 installed speakers: 43 suspended speakers and four subwoofers. The Giga-Hertz residency afforded me a total of five weeks in two separate visits of nearly uninterrupted time in the Klangdom in April and August 2012, during which I composed a work for 43 speakers, 30 minutes in duration, entitled “Spirits,” which was premiered at the IMATRONIC extended Festival of Electronic Music at ZKM on November 22, 2012.

7.1 The Technical Setup

The first few days of the residency were devoted to selecting an appropriate software environment to compose with. I first decided not to use the four subs, and to work with only the 43 suspended speakers, which would provide the clearest point source localization. ProTools was immediately rejected as inadequate to the intricate mixing schemes I had in mind. I reviewed Zirconium, an inhouse ZKM tool that is quite useful for automating spatial trajectories. However I decided not to use Zirconium either, as I wanted to work with incoherent, as well as coherent spatial images. Csound was tempting, but I was planning to mix multiple multichannel files with 43 channels each, and the Csound disk file reader Soundin had a documented hard-coded limit of 24 channels of input, later increased to a still-inadequate 40 channels. SuperCollider was then chosen as the primary compositional platform. The algorithmic capabilities of SC, along with its flexible audio bus routing scheme were ideal as a compositional environment. My initial plan was to create smaller sections as 43-channel interleaved CAF files and mix them down in SC. This plan soon ran into a snag, since the DiskIn unit generator of SC, as recently as version 3.6, has an undocumented, hard-coded limit of 32 channels, and fails silently when attempting to read sound files with 33 or more channels. [9] At that point, I wrote a small utility in C with libsndfile in order to demultiplex sound files with an arbitrary number of channels to sets of mono sound files. [10] This added an unwelcome additional layer of complexity, but allowed me to mix freely in SC, using groups of 43 mono files for each sound source.

I mention all of this not to criticize SuperCollider, which is an excellent and unique acoustic compiler, but to make a larger point about the current state of affairs in audio software for spatial composition, where unnecessary channel limitations are still common. At a certain point in time when the DiskIn ugen was written, 32 channels must have seemed like a perfectly reasonable limit, but that’s the problem with hardcoded limits. They seem reasonable, until they don’t. The surprising limitations for spatial composition of even highly progressive computer music tools like SuperCollider indicate how much room there remains for spatial exploration. Lacunae in computer music software are de facto markers of areas that have not yet been extensively explored.

7.2 The Music

In order to focus on spatial composition over timbre, a set of sampled piano notes from the University of Iowa Musical Instrument Sample distribution was used as the exclusive sound source for the composition. [11] From the very first experiments, the aesthetic power of spatialization with large numbers of speakers was evident. The work starts with the lowest piano notes, convolved with 40-second duration trapezoidal noise impulse responses, to smooth off the attacks. Different notes are assigned to different speakers, which then pan across the room. The effect of multiple notes panning in parallel in a plane is strikingly different from a stereo pan in front of the listener. In stereo panning there is a sense of distance, as the listener observes from outside the area of virtual motion. In the multi-speaker panning, the listener feels as if inside a slowly moving wave of sound.

Elevation worked well too. The Klangdom speakers are hung at five different elevation levels, counting a single central speaker mounted on the ceiling. Panning melodies at different speeds and elevation levels in a contrapuntal texture produced envelopment and enhanced intelligibility for the counterpoint. An even more striking use of elevation was a dense granular texture of piano notes that were assigned to speakers according to pitch, with higher pitches assigned to higher speakers. With eyes closed this passage is capable of producing a sensation of levitation in the listener. Other uses of elevation included patterns panning in circles up and down opposite walls, and tremolo groups panning from the front, up, and over the audience to the back of the hall. All of these spatial patterns were clearly audible from numerous listening locations in the Klangdom.

7.3 Spatial Re-orchestrations

The first version of Spirits, for ZKM, was completed on August 19, 2012. The work was subsequently re-orchestrated for the Sonic Lab at SARC, and the iCAST system at Louisiana State University. Both systems provided different spatial affordances, while lacking some spatial features that I made use of in the ZKM version of the piece. The SARC version made use of 42 available speakers. Unlike the Klangdom, in the SARC Sonic Lab, the lowest speaker ring is situated below the audience, who are seated atop a metal grid. Placing the low opening tones underneath the audience was an obvious change to make for SARC. In SARC the rings contain at least eight speakers, whereas in ZKM, the number of speakers per ring varies, from 14 on the lowest level to six on the level just below the single ceiling mounted speaker. Point source panning configurations that relied on the number of speakers per ring at ZKM needed to be reconceived for SARC. In every case, assigned speaker numbers were different, so the entire piece needed to be recompiled for the SARC version, which was premiered on December 13, 2012 in the Sonic Lab. This SARC re-orchestration took a week to achieve. A binaural recording of the SARC spatial re-orchestration of Spirits may be heard online [12]. This recording was produced with a Neumann KU 100 Dummy Head microphone, situated on the ground floor in the middle of the Sonic Lab. While headphone listening is vastly inferior to being in a multichannel space with the ability to move around and hear the piece from multiple perspectives, nonetheless a significant amount of spatiality is captured on the binaural recording, including elevation cues.

The LSU iCAST system consists of 24 main speakers and 24 addition discrete sources on four centrally located hexaphonic speaker arrays on the floor. These speakers, although relatively low in output power, enabled the possibility of moving sounds through the hall on the floor. My spatial orchestration for LSU took advantage of that capability. Following the experience of spatial orchestration for SARC, preparing a new 48-channel version for LSU took one full day. The striking decrease in the time required to prepare a new spatial orchestration indicates that, although still labor-intensive, the process of spatially re-orchestration a multichannel work for large numbers of speakers will become more streamlined and efficient with practice.

8. FUTURE WORK

In order to consider the future of spatial computer music, it is worth recalling what the prospects of electronic art music looked like in the early 1950s: a very small handful of studios in the world, with cumbersome tools, but exciting theories about a new kind of music. At that stage, predicting that electronic music would become ubiquitous would have seemed wildly optimistic. If we today maintain some of that optimism, we might speculate that we are currently in the 1950s with respect to spatial computer music. There are relatively few spaces that support projection of sound over large numbers of speakers for experimental computer music. Most were built in the 21st century: SARC Sonic Lab (2004), iCAST (2005), ZKM Klangdom (2006). And more are being built. Cube at the Moss Arts Center of Virginia Tech will be active in late 2014, with upwards of 132 installed speakers.

At this stage, the role of institutions is crucial, and it is possible to glean some general tips from the most successful early electronic music studios, such as WDR and ORTF. These institutions had serious production values that were reflected in generous access policies for composers and state-of-the-art equipment. Musical research was supported by theoretical research that was disseminated alongside the music produced. Without carefully designed policies, there is always the danger that institutions housing facilities with unique spatial audio capabilities will nonetheless fail to turn potential into actual achievement. And of course artists will also need to take some risks by composing works that at present will have very limited opportunities for performance, as a price for aspiring to share remarkable spatial-sonic experiences with their audiences. For all the negatives and risks mentioned here, which in any case will only act as a stimulus for the most adventurous composers, general audiences often respond positively to electroacoustic music with a strong spatial aspect, whether in the Philips Pavilion, or in more recent spatially-focused sonic art such as Cardiff's "The Forty Voice Motet" as staged at The Cloisters of the Metropolitan Museum, or the multiple sold-out performances of Karlheinz Stockhausen's Oktophonie at the Park Avenue Armory in 2013. [13] A major advantage that music for large numbers of speakers will continue to hold over stereo music is that it is inherently social in nature, requiring an audience to come to a space, rather than be able to consume the music privately at home.

9. CONCLUSIONS

We have proposed that spatial computer music is at an important historical moment in which spatial aesthetic and research initiatives strongly suggest themselves to both institutions and creative individuals. It is hoped and expected that we will witness dramatic aesthetic advances and experiences in spatial computer music from these quarters in the coming decades.

Acknowledgments

I would like to thank Ludger Brümmer at ZKM, Steve Beck at Louisian State University, and Michael Alcorn, Chris Corrigan, Gary S. Kendall, and all my other Queen's University Belfast colleagues who were so supportive of my work in spatial computer music composition, during my time at SARC. I also thank Elizabeth Hoffman for an insightful reading of an early draft of this paper.

10. REFERENCES

[1] Dewan, B. 2002. *Thaddeus Cahill's "Music Plant."* http://cabinetmagazine.org/issues/9/dewan.php (14 April 2014)

[2] Zvonar, R. 2005. *An Extremely Brief History of Spatial Music in the 20th Century.* http://cec.sonus.ca/econtact/7_4/index.html (14 April 2014)

[3] Treib, M. *Space Calculated in Seconds*. Princeton University Press, 1996.

[4] Föllmer, G. Karlheinz Stockhausen: *Spherical Concert Hall.* 2004. http://www.medienkunstnetz.de/works/stockhausen-im-kugelauditorium (14 April 2014)

[5] Shaff, Stan. Audium. 2002. http://www.audium.org (14 April 2014)

[6] Smalley, D. "Space-form and the Acousmatic Image," *Organised Sound,* 12(1) pp. 35-58, 2007.

[7] Di Scipio, A., "Formal Processes of Timbre Composition Challenging the Dualistic Paradigm of Computer Music," in Proc. Int. Computer Music Conference, Denmark, 1994, pp. 207-208.

[8] Otondo, F. "Creating Sonic Spaces: An Interview with Natasha Barrett", *Computer Music Journal*. 31(2), pp. 10-19, 2007.

[9] Lyon, E., 2012. "Multichannel NRT advice sought," http://new-supercollider-mailing-lists-forums-use-these.2681727.n2.nabble.com/multichannel-NRT-mixing-advice-sought-td7456864.html (14 April 2014)

[10] Lyon, E. "Demux," 2012. http://disis.music.vt.edu/eric/LyonSoftware/demux/ (14 April 2014)

[11] Fritts, L. "Musical Instrument Samples" 2011, http://theremin.music.uiowa.edu/MIS.html (14 April 2014)

[12] Lyon, E. "Spirits," 2012. http://disis.music.vt.edu/eric/Data/Spirits/ (14 April 2014)

[13] Park Avenue Armory Programs, "Oktophonie," 2013. http://www.armoryonpark.org/programs_events/detail/oktophonie (14 April 2014)