AN ECOLOGICALLY VALID EXPERIMENT FOR THE COMPARISON OF ESTABLISHED SPATIAL TECHNIQUES

Hugh Lynch
Digital Media and Arts Research Centre (DMARC)
Department of Computer Science and Information Systems
University of Limerick
Limerick
Ireland
0614866@studentmail.ul.ie

Robert Sazdov
Digital Media and Arts Research Centre (DMARC)
Department of Computer Science and Information Systems
University of Limerick
Limerick
Ireland
Robert.sazdov@ul.ie

ABSTRACT

The paper presents an ecologically valid experiment design for a perceptual study of electroacoustic music. This involves the perceptual evaluation of 2D and 3D multi-channel audio produced from established spatial techniques. The spatial attributes of envelopment and engulfment are used to determine the perceptual ratings of presented 2D and 3D audio produced from identified spatial techniques. It is hoped that by undertaking perceptual testing, specific aesthetic observations made by composers regarding the effect that their spatial techniques produce can be quantified in established spatial terminology.

1. INTRODUCTION

The majority of articles, studies and research undertaken of electroacoustic music focus on theoretical and conceptual rather than empirical investigative research [28]. Sazdov et al. [20] argues that perceptual evaluation experiments of electroacoustic music for the purpose of formulating compositional approaches, is currently absent. This paper outlines identified spatialisation techniques and will attempt to understand how they are perceived by presenting a research method for the perceptual evaluation of electroacoutic music composed utilizing these techniques. The proposed study will consist of ecologically valid perpetual experiments, which will be used to elicit the perceptual rating of composed electroacoustic music referred to as spatial scenes. These will be presented on both 2D and 3D multi-channel configurations and participants will be asked to rate the spatial attributes of envelopment and engulfment. The study will also compare spatial scenes presented to determine if particular attributes are unique to specific loudspeaker configurations.

2. RESEARCH IN ELECTROACOUSTIC MUSIC

The majority of research in electroacoutic music is concerned with conceptual or practical methods and approaches to composition [14, 20, 26]. A number of studies have concentrated on quantitative research in the form of perceptual experiments to evaluate multi-channel sound [1, 20]. Sazdov et al. [20] and Adaire et al. [1] have undertaken studies involving the testing of multi-channel music in the form of ecologically valid experiments. The studies involve listener evaluation of multi-channel space using research established from other related field, namely concert hall acoustics and reproduced audio research [20].

Research undertaken in these fields has established specific spatial attributes, which can be physically measured, rated and predicted by listening subjects [3,4,15]. Envelopment, spatial clarity, apparent source width (ASW) and the proposed 3D attribute of engulfment are some of the spatial attributes that have been identified by researchers [1,3,4,15,20].

Envelopment refers to the perception of being surrounded by sound [4]; spatial clarity is the perceived clarity of a sound [25]. ASW refers to the perceived size of sound in the forward direction [3]. Finally, engulfment refers to the perception of being covered in sound and is believed to be a spatial attribute unique to the elevated dimension [20]. The perceptual studies undertaken by Sazdov et al. [20] and Adaire et al. [1] have used these spatial attributes as quantitative spatial reference terms for the perceptual evaluation of spatial attributes inherent in multi-channel electroacoutic music.

3. RESEARCH CONSIDERATIONS - ESTABLISHING COMMON SPATIAL TERMS

Some composers have stated that electroacoustic music does not have a specific universal terminology when it comes to describing the spatial aspects within a composition [10]. This has resulted in much ambiguity when trying to interpret the observations made by composers leading to confusion and misinterpretation [10, 20]. Therefore, can commonly understood terms be established to describe specific observations and evaluations made by composers regarding their spatial techniques? Can these observations be assessed as having 2D or 3D spatial characteristics?

In order to establish qualitative descriptions of observations made, an ecologically valid perceptual study is proposed. This study will have a similar framework to the perceptual studies undertaken by Sazdov et al. [20] and Adaire et al. [1]. In doing so, it is hoped that commonly understood terms could be established to describe the perceptual observations made by composers and listeners of electroacoutic music. However, before a study can be established, a review of the current spatialisation practices by composers of electroacoustic music needs to be undertaken. These are grouped into three main areas (i) Spectral Techniques, (ii) Granular Techniques, and (iii) Decorrelation and Panning Techniques.

4. SPECTRAL TECHNIQUES

The following techniques refer to methods, which use frequency domain processes for the purpose of spatialisation. This approach to spatialisation of sound is referred to as spectral spatialisation [12]. A number of spectral techniques are outlined, along with a documentation of aesthetic observation made by the creators of each technique. The spectral techniques outlined are: timbre spatialisation [16], spectral delays [12], frequency domain processing [22] and spectral splitting [8].

4.1. Timbre spatialisation

Normandeau [16] proposes a technique that consisted of dividing up the spectrum of a sound into individual frequency registers. The registers are then independently diffused throughout a multi-channel loudspeaker configuration. Even though the spectrum of the sound material is fragmented into discrete registers, the entire spectrum of the sound source is spatialised through out the configuration [16].

Normandeau uses this technique in his composition *StrinGDberg* (2001-03). How this technique is used in this piece is by firstly allying four dynamic band pass filters to twenty-four sound excerpts that are then spatialised through out a network of speakers. The size of each of the filters bandwidth is programmed to dynamically change through out the duration of the piece. Initially, each filter is set as a notch filter where by all the frequencies of each sound are filtered except for a specific central frequency, which is left unfiltered. Using the central frequency as a staring position a gradual un-filtering of each of the sounds frequency spectrum occurs. This un-filtering continues until the entire frequency of the sound is revealed [16].

Normandeau has used modified versions of this spatial techniques on a number other works, but the process outlined above forms a representation of the underlying approaches used by the composer each time. According to Normandeau, the ideal listener experience enables the audience to move within the listening space so that movement into the sound is possible and for the listener to become completely 'immersed' in the sound [16].

4.2. Spectral delays

Kim-Boyle [12] proposes a spatial technique called spectral delays. A spectral delay is a technique, which involves slowing the re-synthesis of discrete FFT bins of a short -time Fourier transform [12]. By applying this to individual channel in a multichannel system particular frequency ranges within a sounds spectrum can be moved. For instance, a panning of frequency below 860Hz will occur from one speaker to the other in stereo pair if the first twenty FFT bins are delayed, the frequencies above that point will remain static [12]. Spatialisation is achieved by re-synthesizes the entire frequency of the sound source [12]. In order for this technique to be perceived by the listener a psychoacoustic phenomenon is used, namely the *Precedence effect* [23].

4.3. Frequency domain processing

Frequency domain processing is a technique proposed by Yorchia and Lippe [22]. This technique uses real-time frequency processing to spatialise sound material [22]. Max/MSP software is used to apply real-time spectral processing. FFT analysis/re-synthesis processes are constructed in the form of Max patches. These patches are used to visualize and spatialise the sound material over a multi- channel loudspeaker array [22]. A number of spatial techniques are presented using this process. They are (i) circular spatial distribution, (ii) spatialisation based on signal analysis and (iii) stereo spatial cross-synthesis [22].

Circular spatial distribution spatialises sound material around the centre of the listening space in a circular pattern, this involves controlling 64 FFT bins in 64 spatial locations whereby the shape and energy of the pattern can be manipulated [22]. Spatialisation based on signal analysis of audio signals involves obtaining and using the spectral properties of an audio signal to generate a spatialisation pattern. The signal that is analyzed can be a different signal or the same signal as the signal being spatialised. This technique can be used to produce difference energy levels within the spatial field [22]. Stereo spatial cross-synthesis involves obtaining a stereo signal energy dispersion properties and administering those properties to a selected mono signal. The stereo signal acts as a modulator to the monophonic carrier signal [22].

Yorchia and Lippe [22] conclude that when applying these techniques, sonic material that is rich in high frequency content spatialises more effectively then lower frequency material. They also state that lose of coherency is perceivable when an auditory event is spectrally dissected and distributed past a certain point [22].

4.4. Spectral splitting

Wilson and Harrison [8] propose a technique called spectral splitting. This technique consists of presenting sound material through a non-homogenous loudspeaker configuration (i.e. BEAST) [8]. This type of configuration causes varying frequency responses and onset times resulting in spatial decorrelation [8]. Different frequency ranges of the sounds spectra are perceived as emitting from different locations throughout the speaker array [8]. For example, when the tweeter tree loudspeakers are used within a configuration (i.e. BEAST), the splitting of the sounds spectrum occurs because of the high pass filters used on those specific speakers. Wilson and Harrison observe that sound is perceived as separating out spatially to different locations in the performance space [8].

5. GRANULAR TECHNIQUES

Granular Synthesis is a sound synthesis method used by electroacoutic composers to spatialise sound material [24]. Spectral and granular spatialisation with Boids [11] swarm lab [6] and swarm granulation [24] are some such techniques that employ this method.

Reynolds [18] formulated a number of parameters describing the behavior and interaction of flocks of birds in the wild. The parameters are (i) separation, (ii) alignment and (iii) cohesion. Using these parameters he formulated an algorithm called the boids algorithm. Reynolds created a computer application that generated an accurate simulation of the behavior of a flock of birds [18]. All of the granular techniques outlined in this paper integrate Reynolds's Boids algorithm within their spatialisation systems to some degree [6, 11, 24].

The boids algorithm has been used and further developed within a numbers of spatialisation systems [11]. Singer's boid object is an implementation of the algorithm in the form of a Max/MSP object [21]. This object allows for boids to be moved in virtual space by controlling various parameters. The object allows for multiple parameters to control the movement of boids. These boids are represented on an x and y Jitter graphical window. The jitter window provides visual representation and feedback of the boids spatial location and movement. A mouse or track pad controls the movement and direction of the boids. The boids sensitivity response to these controllers can be influence by changing additional parameters such as speed, attraction and inertia [11,21].

Kim-Boyle [11] uses a granular sampling patch from MaxMSP 4.5 patch library in conjunction with Singer's Boids object to create a spatialisation system. The sampling patch controls the x and y co-ordinates which are mapped to a quadraphonic loudspeaker configuration. To create an impression of depth additional z co-ordinates are used to control the amplitude of each grain particle. This mapping allows for the control of spatial location of each grain. Spatialisation is achieved by mapping the grain particles to co-ordinate within the configuration. A spatialisation effect such as the Doppler effect can be realized using this system [6,11]. Kim-Boyle states that sonic material is perceived as being more spatially scattered as the sound material increases in harmonics.

5.1. Swarm Lab

Swarm lab is a spatialisation system developed by Davis at the SARC's Sonic Laboratory [6]. It is presented through an eight-channel loudspeaker configuration using Max/MSP. Two Max objects are used within this system, the Vbap~ object and a Singer's boids Object [21]. The Vbap~ object is a vector based amplitude panning object used for spatialisation of boids [17]. As well as spatialisation, various effects such as Doppler, reverb and volume can be controlled. The boids object is used to control the relationships between the boid agents. A graphical representation of the boids, produced using MaxMSP Jitter, corresponds to the boids location in virtual space [6]. Additional spatial effect can be perceived which are not visible in the graphical representation. These effects are caused by timing and amplitude differences between boids. According to Davis and Rebelo [6] when listening to a flocking motion the listener becomes an inhabiting agent as the sound envelops them within this listening environment [6]. Similar spatial attributes are perceived when a swarm effect is presented, but not as strongly as the flocking motion [6].

5.2. Spatial Swarm Granulation

Wilson [24] has developed a spatial swarm granulation technique involving the use of swarming algorithm's to dynamically control the spatial location of sound. This technique is implemented using SuperCollider DSP and synthesis software. Three SuperCollider DSP tools have been developed, BMSwarmGranulator, BMGrainBoid, and BMGrainBoidSpace. BMSwarmGranulator is a sample sound generator which controls varies parameters such as grain pitch, time stretch and grain duration. BMGrainBoidSpace defines the parameters of the multi-dimensional space within which the boids can move. BMGrainBoid is based on the boids algorithm and controls the position and movement of individual boids within the defined space [24].

The individual boids represent streams of grains within the virtual space. By sampling the boids, the proximate location of the grains can be determined in relation to the loudspeaker closest to it [24]. According to Wilson [24] an impression of diffuseness and/or increased physical volume can be perceived. He also states that the spatial feature that is most perceived is diffused but localized effect [24]. Granular spatialisation with boids and Swarm lab are presented in a two-dimensional speaker format [6, 11]. The spatial Swarm granulation technique proposed by Wilson has the ability to spatialise sound material in both 2D and 3D speaker configurations [24].

6. DECORRELATION AND PANNING TECHNIQUES

Decorrelation refers to sound source that is processed into a number of waveforms, which are different from one another but are perceived as sounding the same [9] Many types of sound synthesis such as granulation, reverberation and chorusing is produce as a result of decorrelation. Decorrelation can be used for timbral coloration, diffusion of a sound field, the externalization of sound (headphone listening), image shift and precedence effect elimination [9].

The process of producing a decorrelation signal involves convoluting an input signal with by two correlated signals. The level of correlation of the two correlated signals is specific and is determined by filtering the input signal with a finite-impulse-response (FIR) filter. This creates signals with different correlation amounts. The correlation amounts determined are the amounts applied to the two signals used to convolute the original input [9]. This technique is used to achieve a diffused, reverberant and broad enveloping sound image [2, 9].

6.1. Amplitude point source panning

The process of point source panning consists of spatialising monophonic sound signal across a speaker configuration. The sound originates from individual speakers with no panning occurring between them. Perceived spatialisation of sound between speakers is achieved with changes in individual amplitude levels of each speaker [27]. Wyatt [27] presents a set up for the presentation of this technique using a configuration called *Discrete Eight System* or the *D-8 System.* The *D-8 system* consists of an eight channel system position in a circle around the listening position at ±45° angles [27].

7. ECOLOGICALLY VALID EXPERIMENTAL DESIGN

This form of testing involves the presentation of spatial scenes to participants for perceptual evaluation. The spatial scenes will consist of one-minute electroacoutic music excerpts and are considered to be compositional studies of the above-mentioned spatial techniques. This will provide ten different experimental conditions for perceptual evaluation.

The techniques to be used are timbre spatialisation [16], spectral delays [12], spectral splitting [8], Spectral and granular spatialisation with Boids [11], Swarm lab [6] swarm granulation [24], decorrelation [9] and amplitude point source panning [27]. The spatial studies will be presented using multi-channel loudspeaker configurations in an environment similar to that of a concert hall. Participants will rate these spatial scenes for levels of envelopment and engulfment using a five level Likert scale. The composed music will be consisting of complex stimuli to reflect the varying sonic complexity inherent in electroacoutic music [20].

The loudspeaker configuration used to present the studies will include both horizontal and elevated loudspeaker placement. This will allow for the presentation of the spatial scenes in both two and three dimensional speaker configurations where comparison can be made. The horizontal loudspeaker configuration used to present will consisted of an eight-channel configuration. This is a popular configuration used to present electroacoutic music. Each speaker is placed at a ±45° angles in relation to the listening position. The second configuration will be made up of an additional four speakers added to the existing eight channel set up, this will consisted of twelve speakers in total. The extra four speakers will be placed elevated at ±45°and ±135° azimuth as recommended by Sazdov et al. [20].

The listening test will consist of randomly presented spatial scenes and will be presented at consistent intervals. The participants involved will consist of experienced and non-experienced listeners of electroacoutic music. Before the listening test is carried out, the spatial attributes will be explained and an audible example of the each attribute will be presented to the participants. This will provide the participants with a reference to what particular spatial attributes sounds like. It is hoped that this will result in a more accurate rating of the spatial scenes.

8. CONCLUSION

It is believed that the findings from the study will demonstrate which of the mentioned spatial techniques will be rated higher for envelopment and engulfment. This will therefore provide valuable insight into which technique might be better suited for 2D spatialisation or 3D spatialisation and provide a quantitative terminology of the descriptive observations that are used to describe the effects of different spatial techniques.

9. REFERENCES

[1] Adair, S. Alcorn, M. Corrigan, C, (2008), A study in to the perception of envelopment in electroacoustic music, Queen's University, Belfast.

[2] Baalman, M. (2010). Spatial Composition Techniques and Sound Specialization Technologies. Organized Sound, 15, pp 209-218

[3] Barron, M., Marshall, A., (1981). Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure. J.Sound and Vibration, 77(2): 211– 232.

[4] J.S. Bradley G.A. Soulodre, 1995). "Objective measures of listener envelopment," J. Acoustic. Soc. Am., vol. 98, no. 5, pp. 2590- 2597.

[5] Chowning, J. (1971). "The simulation of moving sound sources" Journal of the Audio Engineering Society, 19(1).

[6] Davis, T., Rebelo, P (2005) 'Hearing emergence: Towards sound based self- organization' in Pro-Ceding of the 2005 International Computer Music Conference, Barcelona, pp. 463-466

[7] Harrison, J. (1998) "Sound, space, sculpture: some thoughts on the "what", "how", and "why" of Sound diffusion." Organized Sound vol. 3, no. 2, pp 117-127.

[8] Harrison, J., Wilson, S. (2010) Organized Sound, Volume 15, Issue 03, Dec, pp 183-184

[9] Kendall, G. (1995). The Decorrelation of Audio Signals and its Impact on Spatial Imagery. Computer Music Journal 19(4): 71– 87.

[10] Kim-Boyle, D. (2005) "Sound Specialization with Particle Systems" Proceedings of the 8th International Conference on Digital Audio Effects (DAFX-05), Madrid 2005, pp. 65-68 and in the Proceedings of the Journeys d'informatique musicale Conference, Paris 2005, pp. 143-146

[11] Kim-Boyle, D. (2006) "Spectral and Granular Specialization with Boids" Proceedings of the 2006 International Computer Music Conference, New Orleans, pp. 139-142.

[12] Kim-Boyle, D. (2008) 'Spectral Spatialization - An Overview', Proceeding of the 2008 international Computer Music Conference, Belfast.

[13] Lossius, T., Balthazar, P. and De La Hogue, Th. (2009). DBAP: Distance Based Amplitude Panning. Proceedings of the International Computer Music Conference (ICMC), Montreal, and QC: 489–92. Princeton University Press.

[14]]Malham, D., (2001). "Toward Reality Equivalence in Spatial Sound Diffusion", Computer Music Journal 25(4): 31-8.

[15] Nakayama, T., Muira, O., Okamoto, M., and Shiga, T., (1971). "Subjective Assessment of Multichannel Reproduction", J. Audio Eng. Soc vol 19, pp. 744-751.

[16] Normandeau, R. (2009) Timbre Spatialization: The medium is the space*Organized Sound, 14: 277-285 Cambridge University Press.

[17] Pulkki, V. (2001). Spatial sound generation and perception by amplitude panning techniques. Doctoral dissertation, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Report 62.

[18] Reynolds, C (1987), "Flocks, herds and schools: a distributed behavioral model", Computer Graphics, 21(4), pp. 23-34.

[19] Rumsey, F. (2002). "Spatial Evaluation for Reproduced Terminology, Meaning, and a scene-based paradigm", J. Audio Eng. Soc 50, 651-666.

[20] Sazdov, R., Paine, G., Stevens, K. (2007) "Perceptual Investigation into envelopment, spatial clarity, and engulfment in reproduced multi-channel audio", AES 31st International Conference.

[21] Singer, E. "Boids for OSX" Available Online: http://www.ericsinger.com cyclopsmax.html. [03, 05].

[22] Torchia, R & Lippe, C (2003), "Techniques for multi- channel real-time spatial distribution using frequency-domain processing", in Proceedings of the 2003 International Computer Music Con-, Singapore, pp. 41-44.

[23] Wallach, H., E. Newman and M. Rosenzweig. (1949) "The Precedence Effect in Sound Localization", The American Journal of Psychology, Vol. 62, No. 3, pp. 315- 336.

[24] Wilson, S. (2008) Spatial Swarm Granulation. Proceedings of the 2008 International Computer Music Conference, Belfast, UK: 1– 4. Quality Sound.

[25] Zacharov, N., and Koivuniemi, F., (2001). "Unraveling the Perception of Spatial Sound Reproduction", in Proc. AES 19th Int. Conf, pp. 272-286

[26] Zvonar, R. (2004)."A History of Spatial Music", eContact, vol.7, no. 4. Retrieved, June 7, 2005 from http://cec.concordia.ca/econtact/index.htm

[27] Wyatt, S. "Investigation studies on Sound Diffusion/Projection" eContact, 4 Retrieved, January 20, 2011 from http://cec.concordia.ca/econtact/fiffusion/investigative.htm

[28] Weale, R. 2005. The Intention/Reception Project: Investigating Composer Intention and Listener Response in Electroacoustic Compositions. Conferences ET Seminaries Series, No. 19. Paris: Observatories Musical François. Pp. 71-81.