TRACKING MOVING SOUNDS: PERCEPTION OF SPATIAL FIGURES

Cédric Camier $^{\star(a,b)}$*, François-Xavier Féron*[c]*, Julien Boissinot* [b]*, Catherine Guastavino* [a,b]

[a]Multimodal Interaction Lab
School of Information Studies,
McGill University
3661 Peel St. Montreal, Canada H3A1X1.
[b] Center for Interdisciplinary Research in Music Media and Technology
School of Information Studies,
McGill University
527 Sherbrooke St. West Montreal, Canada H3A 1E3.
[b] CNRS, Laboratoire Bordelais de recherche en Informatique
LaBRI, UMR 5800, F-33400
33400 Talence, France *cedric.camier@mail.mcgill.ca

ABSTRACT

With the emergence of electroacoustic music, spatial figures have become part of the musical vocabulary of many composers. But the perception of auditory trajectories has received scant attention in the scientific literature. This study aims at determining under which conditions simple common spatial figures (such as circles, squares and triangle) can be perceived by a listener positioned in the center of a loudspeaker arrays. In a series of listening tests, we investigate the effect of rendering techniques (VPAB vs. WFS), reverberation (dry vs. modeled reflections) and sound velocity on spatial figure identification performance.

1. INTRODUCTION

While spatial considerations in music date back to the Renaissance, space only earned its place among critical musical attributes in the second half of the 20th century with the development of spatial sound reproduction and the emergence of electroacoustic music. However, the extent to which trajectories conceived in the mind of the composer, implemented by sound engineers, musical assistants and performers can be perceived by listeners remains an open question. The present study focuses on closed spatial sound figures inside of a circular array of loudspeakers based on a review of musical works using dynamic sounds localization and spatial rendering techniques. We investigate the conditions under which these figures can be perceived as a function of the rendering technique, reverberation of the room and the velocity of the moving sound. Previous trajectory studies have used other report methods, such as asking subjects to draw the trajectory [1]. Investigation of perceived trajectories in spatial audio is still in need of formalisation, so the three-alternative forced choice method intended for use here should add to that conversation.

2. SPATIAL TRAJECTORIES IN CONTEMPORARY MUSIC

With the emergence of electroacoustic music in the 1950s, composers became interested in positioning and moving sounds in space. Since the development of loudspeaker orchestras in the beginning of the 1970s (gmebaphone [2] and acousmoniums [3]), the interpretation of electroacoustic composition consists in a spatial interpretation which tends to explore spatial effects mostly based on manual amplitude panning [4]. According to Van de Gorne [4], the ideal room acoustics to perform spatial interpretation is a dry room or a open field, as reverberation is hypothesized to have a detrimental effect on spatial interpretation precision. In 2008, Peters [5] conducted a web-based survey with 52 composers and sonic artists to better understand how they use spatialization, what spatial aspects are essential and what functionalities spatial audio systems should strive to include or improve. Immersion was reported as one of the most desirable effect produced by a spatial reproduction system which was linked to source width and spatial reverberation simulation. Furthermore, respondents highlighted the need to consider room acoustics in spatial rendering software tools.

3. DYNAMIC SOUNDFIELD SYNTHESIS

Static soundfield synthesis has been extensively studied over the last decades and allows accurate and robust simulation results [6]. However, dynamic soundfield synthesis implies the reproduction of physical alterations of the sound waves produced by moving sources. The sound waves experience compression or expansion related to the direction of motion, which leads to a Doppler effect consisting in a pitch shift and an amplitude modification. Typical implementation of soundfield synthesis do not take the Doppler effect into account. Rather, dynamic virtual soundfields are discretized as a sequence of stationary snapshots. Depending on the duration of each snapshot, this discretization may lead to a Doppler-like frequency shift. Depending on the technique used, this results from a compression/depression of successive loud-

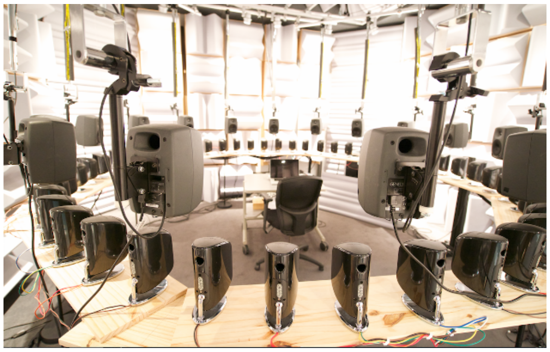

Figure 1: Experimental setup in the CIRMMT Spatial Audio Lab. Both loudspeaker arrays consist of equally spaced loudspeakers on a 3.5-m circle in the horizontal plane (16 Genelec for the top one, 48 B & W for the bottom one).

speaker contribution radiation (e.g. VBAP) or from time warping (e.g. WFS) rather than from the actual Doppler effect between the virtual source and the listener. As shown by Franck [7] and Ahrens and Spors [8, 9] for WFS, these artefacts occur in conventional implementations of WFS, but can be avoided by taking into account the physics of soundfield generated by moving sources.

4. APPARATUS

The experimental setup is shown in Fig. 1 with two circular loudspeaker arrays with a diameter of 3.5 m in the hemi-anechoic Spatial Audio Lab of the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) in Montreal (Canada). The dry room is 5.40 m (W) $\times$ 6.40 m (L) $\times$ 3.60 m (H) with a measured Reverberation Time (RT60) and Early Decay Time of 0.09 s and 0.28 s respectively. The first circular array consisted of 16 Genelec 8040A (Genelec, Iisalmi, Finland, frequency range 48 - 20,000 Hz) regularly spaced in the horizontal plane. The second loudspeaker array is located 10 cm below and consists of regularly spaced 48 M1 Bowers & Wilkins loudspeakers.

Stimuli are generated using a Max/MSP object using Vector Base Amplitude Panning (VBAP) [10] implemented in a previous study [11] in order to guarantee smooth angle variations at very high velocities. The MAX/MSP object was programmed in JAVA language (with the MXJ functionality) to sample the angle parameter at the audio sample rate and allow modeled primary and secondary reflections upon four virtual walls while omitting specular reflections.

5. PROCEDURE

On each trial, participants are presented with a spatial figure around them and asked to indicate which of three figures they perceived (triangle, square and circle) using a three-alternative forced choice. We manipulate 3 independent variables in a series of experiments, namely reverberation (dry, vs. modeled 1st and 2d reflections), the spatial rendering techniques (VPAB vs. WFS), as well as the velocity of the moving sound. We hypothesize that reverberation will have a detrimental effects on the perception of spatial figures as suggested by Van de Gorne [4]. In addition, the artifacts introduced by the spatial rendering systems could interfere with dynamic localization, especially at high velocities.

Binomial tests will reveal which figures can be correctly identified in each experimental condition. The findings will determine conditions under which closed spatial figures can be perceived by a listener in the center of a circular array. Incorrect answers will inform us on misidentifications and confusions between figures.

Depending on the results of these experiments with the 3 basic figures, we will extend our investigation to a wider range of figures at various distances from the listener. Another extension will involve manipulating the spectral content of the sounds based on the observation that low frequency sounds moving in a circle around the listener can be tracked at higher velocities than higher frequency sounds [12]. Furthermore, we will conduct acoustical measurements with a binaural mannequin to complement the analysis of the perceptual tests and determine which psychoacoustical cues are critical to track sound trajectories.

6. ACKNOWLEDGMENTS

This research was supported NSERC (RGPIN 327392-13) grant and a William Dawson award (McGill University). The authors would like to thank Olli Rummukainen for his assistance with the experimental setup.

7. REFERENCES

[1] M. Gorzel, G. Kearney, C. Masterson, H. Rice, and F. Boland, “On the perception of dynamic sound sources in ambisonic binaural renderings,” in *Audio Engineering Society 41st International Conference: Audio for games.*, London, 2011.

[2] C. Clozier, “Un instrument de diffusion : le gmebaphone,” *Musique et recherche*, vol. L’espace du son, pp. 56–57, 1988.

[3] F. Bayle, *Musique Acousmatique, propositions positions*, Buchet/Chastel, Ed. INA, 1993.

[4] A. Van de Gorne, “L’interprétation spatial. essai de formalisation méthodologique.” *Demeter*, 2002.

[5] N. Peter, “Sweet [re]production. developing sound spatialization tools for musical applications with emphasis on sweet spot and off-center perception,” Ph.D. dissertation, Department of Music Research Schulich School of Music McGill University, 2010.

[6] R. Rabenstein and S. Spors, *Multichannel Sound Field Reproduction*, springer ed., J. Benesty, M. Sondhi, and Y. Huang, Eds. Berlin: Springer Handbook on Speech Processing and Speech Communication, 2007.

[7] A. Franck, A. Gräfe, T. Korn, and M. Strauss, “Reproduction of moving sound sources by wave field synthesis - an analysis of artifacts,” in *The proceedings of the AES 32nd international conference*, Hillerod, Denmark, Sept. 2007, pp. 188–196.

[8] J. Ahrens and S. Spors, “Reproduction of moving virtual sound sources by wave field synthesis: An analysis of artifacts,” in *32nd Int. Conference of the AES*, 2007.

[9] ——, “Reproduction of moving virtual sound sources with a special attention to the doppler effect,” in *124th Conv. of AES*, 2008.

[10] V. Pulkki, "Virtual sound source positionning using vector base amplitude panning," *Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.

[11] C. Camier, J. Boissinot, and C. Guastavino, "Does reverberation affect upper limits for auditory motion perception," in *ICAD*, vol. Submitted, Graz, Austria, 2015.

[12] F.-X. Féron, I. Frissen, J. Boissinot, and C. Guastavino, "Upper limits of auditory rotational motion perception," *Journal of Acoustical Society of America*, vol. 128, no. 6, pp. 3703–3714, 2010.