Sonic Scenography - Equalized Structure-borne Sound for Aurally Active Set Design

Otso Lähdeoja
University of the Arts,
Sibelius Academy,
Helsinki, Finland
otso.lahdeoja@free.fr

Aki Haapaniemi
Aalto University
Department of Media Technology,
Espoo, Finland
aki.haapaniemi@aalto.fi

Vesa Välimäki
Aalto University, Dept. Signal
Processing and Acoustics,
Espoo, Finland
vesa.valimaki@aalto.fi

ABSTRACT

This paper suggests the use of a plywood panel, which is also a scenographic element in a dance performance, as a flat speaker. The sound emanating from the panel is subjectively different from a traditional loudspeaker, since the sound appears to originate from behind the panel. However, its frequency response is severely coloured by the panel modes, and the panel has a displeasing low pass-filtered sound. We propose a digital equalizing filter to improve the sound quality of the panel. The panel response is measured at various angles using the sine-sweep method, and a smoothed average response is formed. A minimum-phase FIR equalizing filter is then designed using an FFT-based technique. Applying this filter to the input signal of the panel alleviates the spectral imbalance. As the measurement and filter design can be conducted online on the scene, the proposed equalized structure-borne sound now becomes an attractive possibility for modern performances.

1. INTRODUCTION

This paper discusses a research effort on the use of structure-borne sound in an artistic context. The project is an interdisciplinary “art-science” cooperation, bringing together researchers from the Signal Processing and Acoustics Department of Aalto University (Espoo, Finland) as well as a composer from the Sibelius Academy (Helsinki, Finland). The results of the scientific study are directly applied into art praxis. The article is divided in two larger parts. The first one (Sections 3 & 4) discusses the use of structure-borne sound as a full audio-range tool for sound diffusion. The basic principles of the technology as well as the major challenges encountered are presented. A case study of a plywood panel speaker is exposed in order to demonstrate a filtering method for the enhancement of structure-borne sound quality. The second (Section 5) part presents an artistic application of structure-borne sound, namely an acoustically active scenographic design for a contemporary dance performance. Compositional strategies and sonic percepts arising from the structure-borne sound technology are discussed.

The term “structure-borne sound” is used to signify sound waves induced into solid elements via acoustic transducers. The resulting vibrating solids act as loudspeakers, giving rise to air-borne sound diffusion via the structures of the performance space (e.g. walls, seats, windows, scenographic elements), as well as to audiotactile perception when these elements are brought in direct physical contact with the spectators. A large variety of surfaces can be transformed into loudspeakers with structure-borne sound drivers. In our opinion, sound emitting solids present an interesting alternative to the traditional cone speaker, especially well suited for creating new spatial effects and impressions. A panel speaker radiates sound from its whole surface, whereas the loudspeaker is a point-source. The spatial acoustic image created by the two types of speakers is very different. Cone speakers are well suited for producing well localizable sound sources, while panel speakers are interesting for creating less localizable, diffused sonic percepts. We are aiming to create spaces where the sound may move in the location's architecture as well as in specific scenographic elements. A theatre stage can become the venue for the movement and presence of sound. Equally, everyday architectural spaces can become sonically active. Loudspeakers can be used in conjunction with sound emitting surfaces, offering a wide scope of possible sound sources. Together, the vibrating surfaces and speakers constitute a heterogeneous set of spatially distributed sound sources, a complex 3-dimensional acoustic space termed here as an aurally active space.

2. BACKGROUND

Structure-borne sound is a widely studied phenomenon in industrial design and the acoustics of vibrating solids are well understood. However, the existing corpus of engineering-related research is primarily concerned with the reduction of structural vibration, for example in order to attenuate machine-related noise levels. The actual use of vibrating solids has received some attention in the domain of music and sound design. Recent psychoacoustic experiments have shown that the subjective quality of the listening experience is enhanced with the pairing of structural vibration with air-borne sound diffusion [1]. Also, the perceptual characteristics of the audiotactile sense have been studied, showing that humans possess the capacity to haptically discern frequency differences independently from hearing. [2]. Cinema has been the main area of development for the use of structure-borne sound for the general public. Enhanced realism and spectator immersion have been achieved with audio bass frequencies conducted into the public seating. Also, various flat panel loudspeaker systems have been studied with various materials. A full audio range glass speaker has been developed, but not commercialized [3]. A wooden speaker has been studied in order to enhance room acoustics by adding reverberation with an aurally active panel [4]. Public address systems are based on audio driven into solid surfaces are currently finding a viable market, for ex. the "whispering window" trademark [5]. Panel-type loudspeaker acoustics and equalization have been studied [6], and methods for enhancing the sound output via signal processing in plane-wave producing dipole speakers have been proposed [7]. In the last decade, structural vibration has stirred interest in the artistic domain, manifested by a growing number of installation art pieces. For example, sound conducted into the bone structure of the body has been used to explore new sonic percepts [8].

3. TRANSFORMING SURFACES INTO SPEAKERS

The project aims to create scenographic and/or architectural elements that are capable of delivering high quality full-spectrum audio output. A wide selection of commercially available structure-borne audio drivers allows for the transmission of audio-rate vibrations into solid surfaces. The available transducer units come in a variety of sizes and power ratings, from one watt up to a kilowatt, making it possible to transform virtually all but the most massive and rigid surfaces into sound sources. Structure-borne audio driver technology is variation of the traditional speaker design, using either electro-dynamic or magnetorestrictive methods to transduce electric current into kinetic and then into acoustic energy. Many available sound drivers are capable of full audio-range performance. However, the exact acoustic properties are not readily communicated by the manufacturers. In order to elucidate their suitability for a use in a high-fidelity musical context, we have carried out perceptive and acoustic tests of a selection of transducers in a previous research effort [9]. In our tests, we found the HiWave and Clark Synthesis transducers to be the most suitable ones for our purposes.

While the transformation of solid surfaces into speakers is easily performed, the issue of sound quality stands out as the main challenge. The audio signal driven into the surface excites the object's resonant modes. The solid-surface speaker's air-borne audio output is dependent of the object's acoustic properties. From a practical point of view, the surface can be seen as a hardware filter: the audio signal is filtered by the object's frequency response. The resulting sound is thus defined by the object's physical characteristics: its material, dimensions, and thickness, as well as by the way it is attached and driver placement. Moreover, a panel-type surface radiates sound over 360 degrees, according to a dipole pattern. As the whole surface vibrates, the sound emission varies according to the listener's angle. The frequency output differs considerably from the frontal position towards the sides. As a result, the perceived sound quality of audio-rate excited solids is often limited and spatially inconsistent. The audio signal comes out as heavily filtered with the object's resonant modes are over-emphasized, bass response might be missing or blurry and the dynamic range is reduced.

This initial limitation in audio quality is responsible for the little interest stirred by structure-borne sound technologies in the context of music and sound art. Especially in the field of standardized high-fidelity sound reproduction, advanced cone speakers display much better performance than any solid surface and structure-borne audio driver combination. However, our argument is that with careful design, structure-borne audio is capable of delivering high-quality audio perfectly adapted to many broadcasting needs. Moreover, the fact that architectural, scenographic, sculptural or natural elements can be transformed into sound sources offers a fecund resource for the sonic arts as a whole, opening a wide range of possibilities for creative work with sound. Also, one has to bear in mind that the situation is not "either or"; structure-borne sound can be used in conjunction with cone speakers and acoustic instruments, multiplying the possibilities of sound sources available for a composer or sound designer.

4. OPTIMIZING SOUND QUALITY IN A PLYWOOD PANEL SPEAKER

A key issue of our project is to optimize the sound quality of solids transformed into speakers. Our working hypothesis has been that it is possible to improve sound quality by balancing the object's frequency response via active filtering. This is performed by measuring the object's impulse response (IR), designing an appropriate equalizing filter, and applying the equalizing filter upstream in the signal chain. Theoretically, the equalization should flatten the frequency response curve and significantly improve the sound quality. The same approach has been used to correct the frequency response of loudspeakers [10] and headphones [11].

This basic assumption is challenged by the change of perceived frequency response as a function of the angle. A measured and corrected IR is only valid for the spot of measurement; as the listener's angle changes, the filter correction loses its relevance. In order to take the spatial inconsistency of the solid-surface speaker's output, we designed an initial experiment to investigate a method of inverse filtering with an IR obtained by averaging a series of measures from different angles.

A series of experiments was conducted with a panel of plywood (81 × 122 × 0.65 cm) and a HiWave HI-AX32C30-4/B transducer (freq. range 100 - 20000 Hz) in an anechoic chamber, as shown in Figure 1. The transducer was positioned off-center (2 cm in diagonal from the centre) and the panel was attached on its lower sides with two clamps. The frequency response of the panel was measured according to the angle. Ten measurements were performed from frontal (0˚) to lateral position (90˚), one every 10 degrees. An exponential sine sweep function from the HISS Tools Max/MSP toolbox was used

[12]. Figure 2 shows the smoothed magnitude responses of the measured impulses from various angles.

The initial frequency response measurements portray a wildly uneven profile with a general tendency of high frequency emphasis and low/mid frequency loss with the growth of the angle. The region between 3 kHz and 8 kHz becomes predominant from 50° onwards. The measurements from 0° to 50° show a high degree of consistency up to about 4 kHz. However, there is more variation in the high frequency responses, and the low frequency responses are severely attenuated for the extreme angles.

Figure 1. Photo of the measurement situation

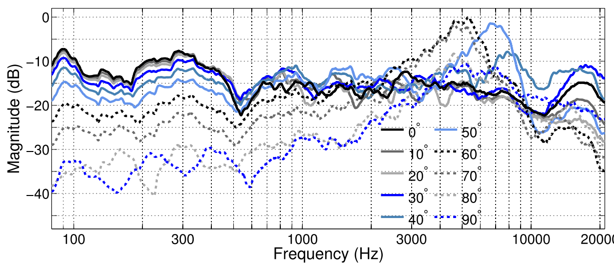

Figure 2. Magnitude responses (with 1/3 octave smoothing) for various receiver angles.

A straightforward inverse filter design scheme was devised in order to equalize the average frequency response of the object. The process takes the measured impulse responses from various angles as input and produces a set of FIR filter coefficients as output. The design procedure is the following:

1. FFT's are computed for the impulse responses from each measured angle and an average is taken of the resulting magnitude spectra.

2. The magnitude average spectrum is smoothed with a 1/3 octave wide Hamming window.

3. The low frequency response below approximately 90 Hz is flattened in order to avoid excessive bass boost created by inverting the magnitude response (the nominal low frequency cutoff of the transducer is at 100 Hz).

4. The magnitude response is inverted and transformed to time-domain with the inverse FFT, and the real part of the time-domain response is folded to create a causal IR. The imaginary part of the impulse response contains numerical noise and is neglected.

5. A minimum phase reconstruction of the IR is made with cepstral processing and tapering is applied to avoid creating a discontinuity at the end of the IR. The resulting sample values are used as the coefficients of the FIR equalizing filter.

Different combinations of measurements were tried as a basis for the filter design. This was done in order to explore whether a practical compromise between accuracy and simplicity would be feasible. Inverse filters with 1024 taps were computed using the 0° measurement (FIR 1), average of 0° and 40° (FIR 2), and average of all measurements from 0° to 60° (FIR 3). The extreme angles were left out of the averaging process because the responses are increasingly irregular. Figure 3 shows the magnitude responses of the correction filters and Figure 4 shows examples of how the filtering works in theory and practice for three measurement angles.

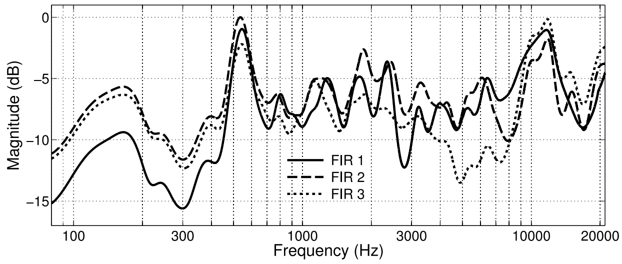

Figure 3. The magnitude responses of the three filters FIR 1, FIR 2 and FIR 3.

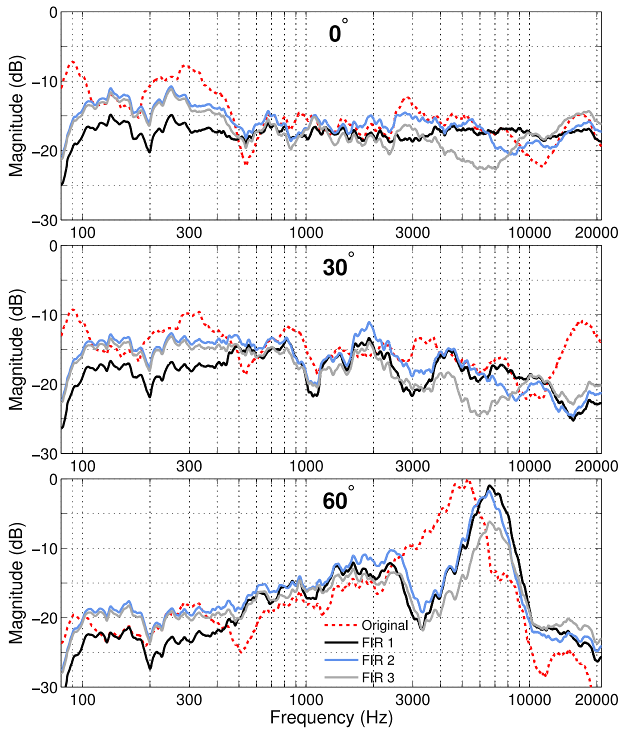

Figure 4. The original and filtered magnitude responses for three measurement angles: 0° (top), 30° (middle) and 60° (bottom), and three set of filter coefficients: FIR 1 (based on 0° measurement), FIR 2 (average of 0° and 40°), and FIR 3 (average of all the measurements from 0° to 60°). The magnitude responses are 1/3 octave smoothed.

As expected, FIR 1 performs best in flattening the panel's frequency response in the 0° direction. However, it is also evident that there remains up to 5 dB fluctuation in the low frequency response, which suggests that there may be effects at play that are unequalizable. These may be e.g. edge diffraction from the panel edges, and the low frequency limitation of the small anechoic room used for the measurements.

The similarity of the shapes of the equalized responses with different averaging suggests that FIR 1 may well suffice for a rudimentary equalization. However, the low frequency cut seems too much for directions other than 0°. FIR 3 is seen to compensate for the high frequency emphasis evident in the original measurements for angles 40° to 60°, and therefore results in an attenuated level between about 2 and 8 kHz compared to FIR 2. Out of the presented filters, FIR 2 seems to present an attractive practical compromise.

Subjective aural evaluation shows that the equalizer significantly enhances the sonic impression. The obstructing low-mid modes lose their predominance and the sound becomes more detailed and clear. A slight effect of thinness or high-end over emphasis is also noted, suggesting the necessity of an overall calibration of the signal with a low-self/high-self equalizer. A recorded example of the unfiltered vs. filtered panel is available at http://otsola.org/?page_id=743.

To enable flexible use of the filtering, a real-time implementation was realized in the form of a VST plugin that performs time-domain convolution. The user interface of the plugin is presented in Figure 5. The plugin enables trying out different FIR filters (e.g. with different averages or smoothing) and allows various degrees of filtering by enabling crossfading between the unfiltered and filtered signals. For further fine-tuning an additional first-order low/high shelf filter is included.

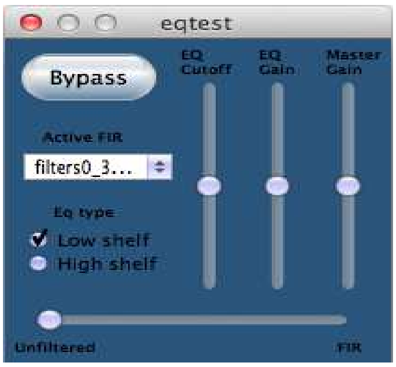

Figure 5. Screenshot of the equalizer VST plugin. The plugin displays a cross-fader between unfiltered and filtered signal, a menu for filter choosing coefficients, and an additional low/high shelf EQ with cutoff and gain functions.

5. A CASE STUDY OF AURALLY ACTIVE SCENOGRAPHY

As a wide range of solid surfaces can be transformed into sound sources by using structure-borne sound drivers, sound may be fused with scenographic or architectural elements. The performance space itself can become a macro-scale sonic instrument. The two main sound source paradigms in music and related art forms are the acoustic instrument or voice, and the cone loudspeaker. These two types of sound production are omnipresent and their sonic characteristics are immediately recognizable. Transforming ordinary solid surfaces into sound sources widens the scope of possibilities to create sonic experiences. The listener is easily puzzled by a sound coming seemingly from nowhere; for example from a wall, a table or a window. Aurally active objects are not part of the general repertoire of "things that make sound". For this reason, they hold a strong expressive potential for music and sound art: they can convey ideas of sonic space, timbre and narration in a novel and intriguing manner. Since 2012, Lähdeoja has been investigating the artistic potential of structure-borne sound in the field of music-related art forms. In the following, we discuss the sound design for a contemporary dance piece.

5.1 Sonic scenography for a contemporary dance performance

"Riisuttuna" ("Bare") is a contemporary dance piece by the Finnish choreographer Satu Tuomisto, premiered at the Oulu City Theatre, Finland, on March 14, 2014. The composition and sound design for the piece involve structure-borne sound driven into the scenographic elements on stage. Figure 6 shows a schematic representation of the sound sources in the theatre space.

At the center of the stage, a tall (3 × 3 m) wooden structure dominates the scene. In one corner stands a table, manipulated by the dancers during the show. Both the structure and the table are custom-made from 4 mm plywood. They are hollow and have concealed soundholes, mimicking the design of a wooden string instrument's resonant chamber. Five channels of structure-borne audio are used on stage in the piece. Three HiWave HIAX32C30-4/B transducers are attached to different parts of the central construction, and a fourth one under the table. A freestanding sound driver is operated by the dancers during the performance in order to induce sound into the dancing bodies themselves. The initial audio work for "Bare" was conducted in parallel to the filtering application presented in section 4, and thus did not incorporate the IR measurement and correction phases. Traditional equalizing was used instead. However, "Bare" is scheduled to tour in late 2014, and the frequency response method presented in this paper will be fully applied.

In addition to the five audio channels broadcasted through the scenographic elements, a four-channel loudspeaker square surrounds the audience, as well as two movable speaker cones. The loudspeakers are used either as monophonic point-sources, or as a four-channel ambisonics ring using the ICST Ambipanning algorithm [13]. A total of 11 sound sources are thus used in a mixture of four-channel PA and on-stage sonic scenographic elements. Figure 7 is a photo of the stage set.

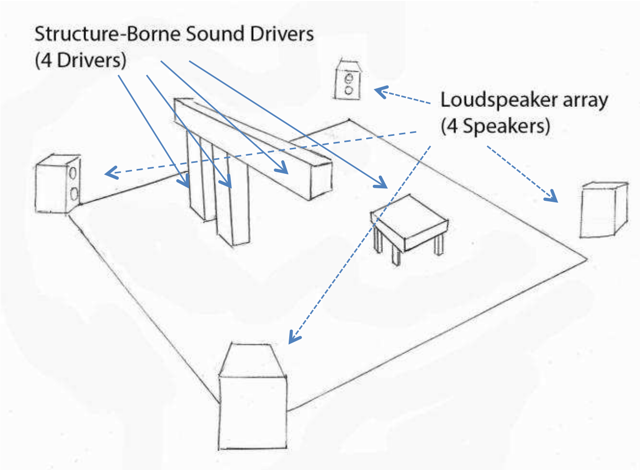

Figure 6. Sound sources distributed in the theatre space

Figure 7. Photo of the "Bare" performance's stage set

Together, the loudspeakers and scenographic elements create a 3D soundfield where the sound can be spatialized. Sonic elements can move in the loudspeaker soundfield and in the structures on stage, as well as between them. The composition and sound design of the piece take full advantage of the setup. Sound localization and movement are used as a key element of the compositional process, in relation with the moving bodies.

For example, in one sequence, a dancer literally pushes the sound between the scenographic elements, before throwing it behind the audience into the loudspeaker soundfield. In another sequence, a sonic dialog is established between the on-stage structures and the loudspeaker ring. Using percussive sounds (easily localized by the ear), the listener is engaged into a narration of multiple "sound points" localized in the performance space. A sense of spatial depth is achieved, radically different from a loudspeaker-created soundfield.

As a third example, the piece employs the notion of "embodied sound" in a literal sense: a structure-borne audio driver is used to induce sound into the bodies of the dancers. Using the bone structures and lung cavities as resonant chamber, the body itself is transformed into a sound source. An intimate relationship between movement and sound is thus established, the sound being inside the dancer's body (Fig. 8).

Figure 8. Two dancers in "Bare", using a movable transducer in order to induce sound into the body, transforming it into a sound source.

5.2 Sound design strategies

The sound design of "Bare" employs loudspeakers and aurally active surfaces in parallel. The main challenge of the sound work is the integration of the two systems into one perceptually convincing sonic tool.

We ran experiments with the ICST ambisonics-equivalent panning [13] controlling the loudspeaker soundfield in combination with the Distance-Based Amplitude Panning algorithm [14] for the scenographic elements. This double technique was found to be very cumbersome to use. The project was finalized as a composition for "fixed sounds" with the loudspeaker array treated as an ambisonic soundfield with monophonic additions for the active surfaces. Nevertheless, the research points to the direction of an integrated spatialization module capable of combining a loudspeaker array and individual sound sources into one tool where spatial localizations and trajectories could be composed and controlled.

The loudspeaker - active scenography combination is suited for producing intriguing aural percepts. We explored the use of reverberation disconnected from the sound source in order to modulate the impression of depth and space. For example, a monophonic "dry" signal was sent to one of the scenographic elements, completed by a 100% "wet" reverb signal sent to the loudspeaker array. A sense of spatial expansion of the object was achieved. We also experimented with the use of the Doppler effect on the movement of sounds in the 3D soundfield. Adding the Doppler effect's characteristic frequency shift to sounds with trajectories proved to radically enhance the sense of movement within the soundfield. Our experiment suggests the necessity to implement a Doppler effect simulation into the module that controls the sound trajectories.

6. CONCLUSIONS AND FUTURE WORK

In this article we have presented an alternative speaker design with a plywood panel and structure-borne sound driver. A method for an equalization of the panel's frequency response has been explored, with encouraging results. An application of the technology into a contemporary dance performance's scenography and sound design has been presented, and the compositional strategies enabled by this approach have been discussed.

Future work points towards the refinement of the equalization, namely the process of measuring and calculating the FIR coefficients, as well as tests with other materials, such as plastic and higher grade wood soundboards. Also, a long-term project incorporates the implementation of a software tool for sound spatialization in a heterogeneous, multi sound-source environment. The projected tool should include options for creating trajectories, spatial percepts (via reverberation and delay), as well as Doppler-like effects for simulating moving sound sources.

7. ACKNOWLEDGEMENTS

Otso Lähdeoja would like to thank the Emil Aaltonen Foundation (Finland) for its generous support of the research work presented in this article.

8. REFERENCES

[1] E. Altinsoy, Auditory-Tactile interaction in Virtual Environments, Shaker Verlag, Germany, 2006.

[2] S. Merchel, M. E. Altinsoy, M. Stamm, Touch the Sound: Audio-Driven Tactile Feedback for Audio Mixing Applications, Journal of the Audio Engineering Society, 60(1/2), pp. 47-53, 2012.

[3] O. Mal, M. Novotn, B. Verbeeren, Neil Harris, A Novel Glass Laminated Structure for Flat Panel Loudspeakers, Audio Engineering Society Convention Paper 7343, 124th Convention, Amsterdam, The Netherlands, 2008

[4] G. Berndtsson, Acoustical Properties of Wooden Loudspeakers Used in an Artificial Reverberation System, Applied Acoustics 44, pp. 7-23, 1995

[5] Whispering Window: http://www.feonic.com/

[6] L. Hörchens, D. de Vries, Comparison of Measurement Methods for the Equalization of Loudspeaker Panels Based on Bending Wave Radiation, Audio Engineering Society Convention Paper 8325,130th Convention, London, 2011

[7] A. Kelloniemi, K. Kynnös, Plane Wave Loudspeaker with Signal Processing Enhancements, Audio Engineering Society 30th International Conference, Finland, 2007

[8] Pook, Lynn, Stimuline, audio-tactile installation. http://www.lynnpook.de/english/stimuline/index.htm

[9] O. Lähdeoja, L. Reboursière, Augmented Window: Structure-Borne Sound Drivers for Sound-Emitting Solid Objects and Surfaces, QPSR of the numediart research program, Vol. 4, No. 4, December 2011

[10] M. Karjalainen, E. Piirilä, A. Järvinen, and J. Huopaniemi, Comparison of Loudspeaker Equalization Methods Based on DSP Techniques, J. Audio Eng. Soc., vol. 47, no. 1-2, pp. 15-31, Jan./Feb. 1999.

[11] J. Rämö and V. Välimäki, Signal Processing Framework for Virtual Headphone Listening Tests in a Noisy Environment, Proc. AES 132th Convention, Budapest, Hungary, 2012.

[12] A. Harker, P-A. Tremblay, The HISSTools Impulse Response Toolbox: Convolution for the Masses. Proceedings of the International Computer Music Conference. 2012

[13] M. Neukom, Ambisonic Panning, Audio Engineering Society 123st Convention, New York, NY, USA, 2007

[14] T. Lossius, P. Baltazar, T. Hogue, DBAP - Distance-Based Amplitude Panning, Proc. International Computer Music Conference, Montreal, 2009