REAL-TIME AMBISONIC SIMULATION OF AUDITORIUM ACOUSTICS WITHIN A DIGITAL AUDIO WORKSTATION ENVIRONMENT

Barry Watson, Dr D F Clark

School of Computing,
University of the West of Scotland,
Paisley, PA1 2BE, UK
barry.watson@uws.ac.uk

ABSTRACT

This research investigates the feasibility of an ambisonic system to replicate the acoustic conditions experienced by solo musicians in a concert hall, allowing for the development of familiarity with a specific venue. Spatial impulse response measurements are taken at the hall and real-time convolution is employed to recreate the acoustics in a domestic setting. Reproduction is via four loudspeakers of a standard 5.1 configuration where digital audio workstation software and associated plug-ins perform routing and signal processing.

The work supersedes related research by offering a integrated consumer-biased approach, affording greater accuracy in regard to the balance of direct/reverberant sounds within the simulation and greater scope for switching between stage and auditorium aspects and simulating venue occupancy.

1. INTRODUCTION

It is well recognised that solo classical musicians interact closely with the specific acoustic conditions of the venues in which they perform [1, 2]. Such conditions are perceived to form an 'extension' of the instrument being played, with performers reacting to their environment via the careful modification of fundamental performance aspects including breathing, articulation, tempo, timbre and dynamic range.

However, financial constraints frequently hinder the amount of time available to the artist for practice and preparation within a concert venue, and in the case of less experienced performers this may adversely affect the quality of performance. To obviate this difficulty, this paper describes the use of an integrated system which is able to provide replication of acoustic conditions for the relevant auditorium via a digital audio workstation (DAW) environment.

Ambisonic systems for surround reproduction were developed in the 1970's by M. Gerzon et al, and afford a developed sense of immersion within a sound field [3]. Unfortunately, due to high hardware costs at that time commercial success was not realised. Advances in digital technologies have now greatly reduced obstacles of cost and complexity, and as a growing number of audio engineers and musicians have access to a digital audio workstation (DAW), multichannel audio interfaces and loudspeaker array, then so the practicality of 'domesticating' such exciting technology increases. This study, then, attempts to help alleviate the historical difficulties associated with the implementation of ambisonic reproduction [4] and to demonstrate that it can be put to creative use within a domestic listening environment.

The work presented compliments previous studies into the development of acoustic environment simulations [5, 6, 7] and presents new approaches for the implementation of such a system in respect of the ease of impulse response measurement, frequency response calibration of the reproduction array and integration into a DAW environment. Whilst the systems cited above tend to use highly specialised convolution processors and complex loudspeaker matrices, this investigation assesses the viability of a more straightforward horizontal ambisonic reproduction employing consumer level equipment. The ability to switch between the perspective of the player and that of the audience, and the careful level balancing of the direct and reverberant sounds within the simulation are fundamental features implemented here that have not been fully realised or explained in existing research.

2. SYSTEM SPECIFICATION

The system proposed involves two principal stages, namely the measurement of impulse responses in the concert venue and the subsequent real-time simulation of the acoustic conditions within a domestic listening environment.

2.1 Impulse response measurement

A Core Sound A-format microphone [8] is used to record impulse responses at the venue. For stage measurements this is positioned as to represent the players head, some 1.65m in height and 1.5 metres from the edge of the stage. For auditorium positions the microphone is positioned facing centre stage set at ear height of the seated listener (1.2m).

The room excitation signal is played over a Genelec 8020B active loudspeaker facing directly out into the auditorium at a height of 1.35m; the resultant impulse responses are designed to be used for instruments with similar directivity patterns, e.g. recorder, trumpet, voice etc. Other instruments would necessitate an alternative speaker angle/height due to differing directivity patterns.

Whilst it is anticipated that an application for the collection of impulse responses will be developed by the research

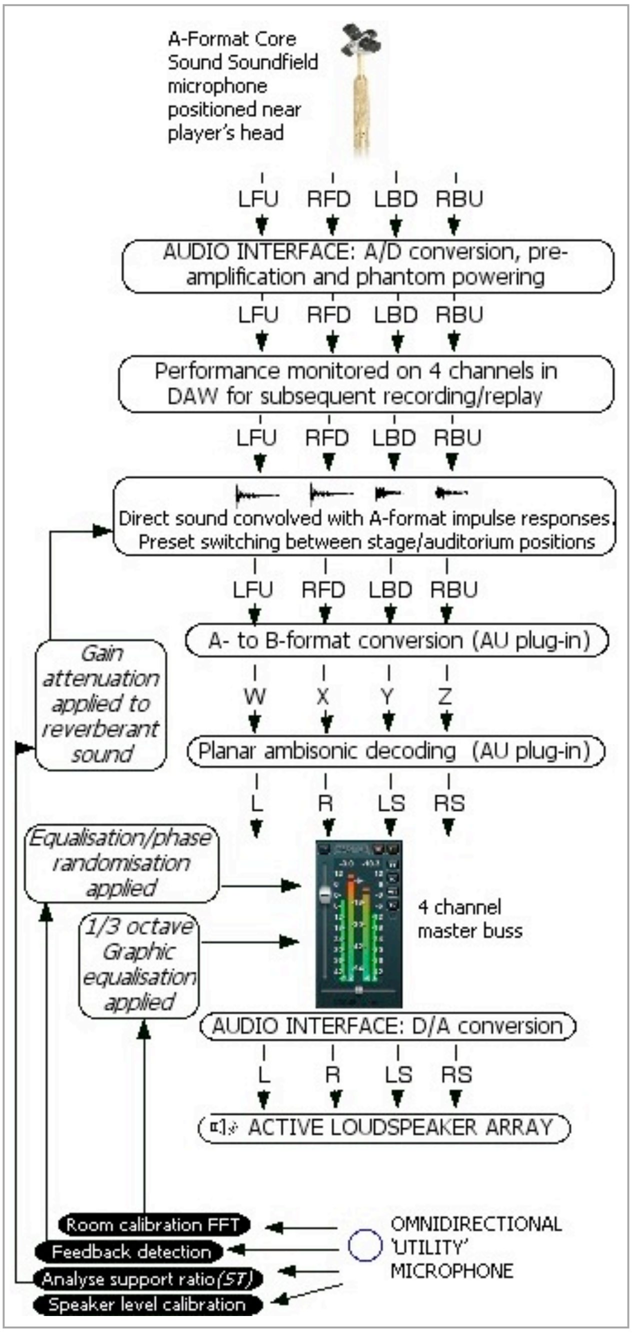

Figure 1: Flow chart outlining the real-time simulation system in its basic state.

group, measurements have thus far been carried out using the Apple Impulse Response Utility bundled with Logic Studio [9]. In the software, a four channel configuration is selected. A 10-second sine sweep is used as an excitation signal, and once de-convolved each impulse response channel can be unpackaged from the Impulse Response Utility project and freely loaded into the convolution within the acoustic simulation.

2.2 Real-time acoustic simulation

The real-time simulation utilises the A-format soundfield microphone as the receiver, placed near to the head of the performer. The four channels are passed through pre-amplifiers on the audio interface with matched gain and routed to separate tracks in the popular and inexpensive DAW application, Reaper [10]. A flow chart representing signal flow and processing for the system in its basic state is presented in Figure 1.

The four channels are then convolved with A-format impulse responses from the venue and routed onwards.

Conversion from A- to B-format and ambisonic decoding is handled by the B2X Ambisonic Studio plug-ins developed by Daniel Courville [11]. These will eventually be succeeded by cross platform plug-ins currently being developed by the research group. The output of the decoder is then sent to four speakers of an ITU 5.1 layout.

Additional audio unit and VST plug-ins developed by the author will be provided for calibration of loudspeaker levels, room equalisation, feedback prevention and perhaps most importantly, balancing of direct and reverberant sounds in the real-time simulation by means of support ratio analysis [1].

3. OBJECTIVE SCRUTINY

Early experiments have been conducted to test the reproduction accuracy of the real-time simulation system in its most simple state, without any room equalisation or feedback prevention. This initial study concentrates on accuracy of frequency response, replication of directional characteristics and reverberation time.

By way of preparation, A-format impulse responses were recorded at three concert venues in Paisley, Scotland. Employing the methodology outlined above, two positions for each the venue were measured, corresponding to the acoustic conditions of stage and auditorium. Details of each venue are presented in Figure 2.

The experiment was conducted in a listening room (measuring approximately 6.5m x 7.5m x 2.8m) featuring acoustic treatment that affords limited high frequency damping, thus approximating a domestic listening environment. Genelec 8020B loudspeakers were set at ear height in an ITU 5.1 layout at a distance of 2m from the listening position (with only the L, R, LS and RS loudspeakers being utilised in the simulation). Loudspeaker amplitude levels were calibrated using pink noise and an omnidirectional microphone placed at the listening position.

A 10-second sine sweep was fed in equal proportions to the four channels within the simulation and then each channel convolved with the corresponding impulse response measured at the venue using the proprietary 'ReaVerb' Reaper plug-in. Following conversion to B-format, the stream was output to the loudspeaker array.

The soundfield microphone was placed at ear height at the listening position, and the output of the simulation was recorded as an A-format stream. Finally, the sine sweep excitation signal was de-convolved from each of these recordings to produce impulse responses that could be analysed and compared to the actual venue responses.

Venue	Seated Capacity	Maximum Reverb time (RT60)	Position 1	Position 2
The Brough Hall	200	2.5s	Stage	3m from stage
Thomas Coats Recital Hall	300	2.8s	Stage	3m from stage
Thomas Coats Church	600	4.3s	Choir stalls	5m from choir stalls

Figure 2: Details of the three venues studied

4. SUMMARY OF FINDINGS

Preliminary results suggest a fair correlation of frequency response across the useable range between the impulse responses taken at the venue and those measured within the simulation. (Figure 3). It is anticipated that any significant observed differences in the response contour will be corrected by the room equalisation function, yet to be implemented.

Accuracy in the replication of directional characteristics by the system is assessed by comparing waveform peaks of A-format impulse response channels (see Figure 4). Once time-aligned, there is broad agreement in the spacing of amplitude peaks when comparing the actual and simulated response. In the actual acoustic response the reflections tend to me more defined with greater variance in amplitude, whilst the simulated responses tend to exhibit more modest spikes in level.

Analysis of reverberation time at 1/3 octave intervals reveals that the simulated acoustic exhibits in the most part, a shorter T20 value than the prevailing acoustic conditions within the venue and less variance. (Figure 5). Whilst there is broad proportional agreement in time values between the frequency bands from 630Hz upwards, there is noticeably less agreement in the lower portion of the spectrum. Fundamentally, this requires further scrutiny, and it remains to be determined – via subjective listening tests – whether or not such differences can be perceived by the listener and what bearing, if any, room equalisation may have upon reverberation time values.

5. PROGRESSION

During the second stage of experimentation, additional concert venues will be studied and a series of subjective listening tests performed with semi-professional classical musicians and audio engineers as subjects. Further investigation into the position and type of speaker used for impulse response measurement will be carried out (including a comparison of directional vs. omnidirectional sources).

Audio unit and VST plug-ins will be developed to facilitate the additional functionality described above in addition to all ambisonic encoding and decoding tasks. Furthermore, the system will be adapted to include modelling of the occupied state of a venue by means of the automated post-processing of measured impulse responses, and also the ability to switch between stage and auditorium responses when a rehearsal is recorded and subsequently played back.

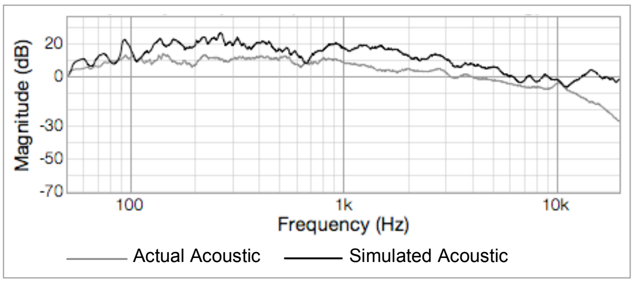

Figure 3: Comparing 1/12 octave frequency response of the original impulse response to that measured within the simulation (Thomas Coats Church, position 1, omnidirectional response).

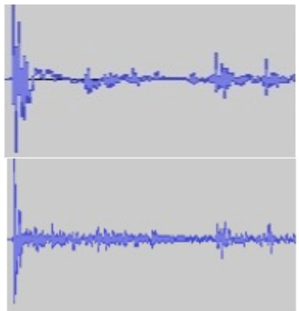

Figure 4: Waveforms representing actual (top) and simulated (bottom) acoustics for the left-front-up channel of an A-format impulse response (Thomas Coats Church, position 1).

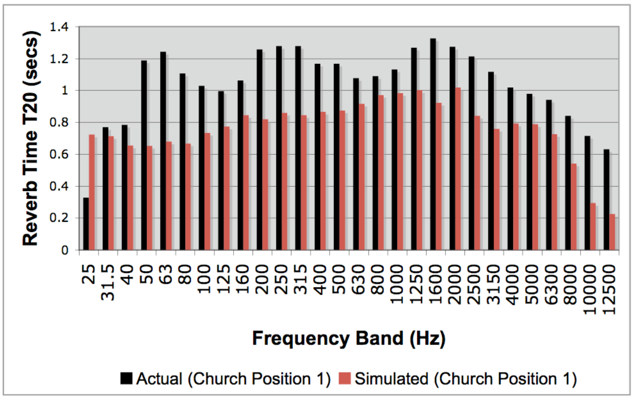

Figure 5: T20 reverb time measured in the venue plotted against that measured within the simulation (Thomas Coats Church, position 1, omnidirectional response).

The user experience will be refined via the development of a custom impulse response measurement utility and experimentation with custom templates, themes and 'skins' within the DAW software.

The fully developed system will be assessed via a range of objective measurements as well as exhaustive testing by professional solo musicians preparing for public performance.

6. REFERENCES

[1] A. C. Gade, "Investigations of musicians' room acoustic conditions in concert halls. Part I: methods and laboratory experiments," *Acustica,* vol. 69, pp. 193-203, 1989.

[2] K. Ueno, H. Hachibana, "Cognitive modelling of musicians' perception in concert halls," *J. Acoust. Sci. & Tech.,* vol. 26, no. 2, pp. 156-161, 2005.

[3] M. Gerzon, "Ambisonics in multichannel broadcasting and video," *J. Audio Eng. Soc.*, vol. 33, issue 11, pp. 859-871, 1985.

[4] R. Elen, "Getting Ambisonics Around," http://ambisonic.net/pdf/ambisonics_around.pdf

[5] K. Ueno, K. Yasuda, H. Tachibana, T. Ono, "Soundfield simulation for stage acoustics using 6-channel system," *J. Acoust. Sci. & Tech.,* vol. 22, no. 4, pp. 307-309, 2001.

[6] K. Ueno, H Tachibana, "Experimental study on the evaluation of stage acoustics by musicians using a 6-channel sound simulation system," *J. Acoust. Sci. & Tech.,* vol. 24, no. 3, pp. 130-138, 2003.

[7] W. Woszcyzk, W. Martens, "Evaluation of virtual acoustic stage support for musical performance," in *Proc. Of Acoustics '08 Conference,* Paris, France, 2008, pp. 1041-1046.

[8] http://www.core-sound.com/TetraMic/1.php (accessed 25th March 2010)

[9] http://documentation.apple.com/en/impulseresponseutility (accessed 25th March 2010)

[10] http://www.cockos.com/reaper (accessed 25th March 2010)

[11] http://www.radio.uqam.ca/ambisonic/b2x.html (accessed 25th March 2010)