ACHIEVING CONVOLUTION-BASED REVERBERATION THROUGH USE OF GEOMETRIC ACOUSTIC MODELING TECHNIQUES

Stephen Oxnard
Audio Lab, Department of Electronics
University of York
York, UK
s1152327@sms.ed.ac.uk

Damian Murphy
Audio Lab, Department of Electronics
University of York
York, UK
dtm3@ohm.york.ac.uk

ABSTRACT

The following paper describes a work which sought to revive and investigate the soundscape of a 14th century Abbey. Geometric room acoustic modelling methods were employed to realise this aim and particular attention was given to monitoring the influence of the level of geometrical detailing on acoustical characteristics. Three virtual acoustic models, of varying detail, were constructed and simulated in order to obtain room impulse responses which were then harnessed for analysis and the creation of auralizations. The results given in this document were obtained through direct examination of impulse responses gathered and subjective testing. Reverberation times were shown to be more or less consistent at high frequencies in all three models and increasingly dissimilar as frequency decreases. Subjective testing was employed to gain insight into the audibility of the contrast in acoustical qualities possessed by each model. The conclusions drawn from analysis are, in part, supported by the subjective testing results.

INTRODUCTION

Acoustic modelling programs such as ODEON and CATT-Acoustic are often utilised to construct and simulate virtual models of sound fields of architectural interest [1], [2]. The applications of undertakings of this kind include the revival of acoustical environments which are no longer accessible and the recording of acoustical characteristics of spaces as part of a “hybrid” architectural account [2].

Geometric acoustic modelling approaches approximate the propagation of sound through an enclosed space by allowing a large number of infinitely thin rays of acoustical energy to emanate from a source. The rays interact with boundary surfaces according to the laws of absorption and scattering. If a ray arrives within the vicinity of a receiver position, data relating to energy levels and directivity are recorded. The data collected then contributes to the formation of an impulse response, providing insight into the nature of the modelled space in terms of standard room acoustic parameters and facilitating auralization through audio convolution.

The subject site modelled in this study was St. Mary’s Abbey, York. This Abbey was destroyed during the dissolution under the rule of Henry VIII and currently lies in ruin (see Figure 1). As such, it is impossible to capture room impulse responses (RIRs) from this site and, therefore, comparison between results of practical acoustical measurement and those virtually generated is also not possible. However, results taken from virtual models of increasing geometrical complexity were compared during this study in an attempt to gain an understanding of the effect of geometric detailing on room acoustic characteristics.

Figure 1: *The ruins of St.Mary’s Abbey, York as viewed in its current state. Image sourced from* [3].

In addition, subjective testing was carried out by allowing listeners to experience auralizations of each virtual space to ascertain whether the variance in room topologies, and therefore the acoustical qualities, are perceptible. This process exposed the sonic revival of the soundscape once present within the Abbey and supports the notion of using virtual RIRs as a means of obtaining realistic reverberation effects for creative audio processing.

2. THE ACOUSTIC MODELS

ODEON 10.1 Auditorium was employed to create and simulate virtual models of St. Mary’s. Prior to the modelling process, it was necessary to research and obtain room dimensions of suitable accuracy. To this end, architectural evidence was sourced from relevant texts (i.e. [4], [5]) and also from the remains of the subject site itself. Dimensions of the floor plan of the Abbey’s interior were derived from scaled diagrams created during excavations and surveys of the site, made available for this work from the archives of York Museums Trust. Data relating to the height of roofing sections and archways were taken from these surveys and further evidence of this kind from a scaled drawing depicting the South elevation of the building, developed by Everard Ridsdale Tate [5]. Amalgamation of the sourced measurements ensured that resulting room dimensions possessed a sufficient

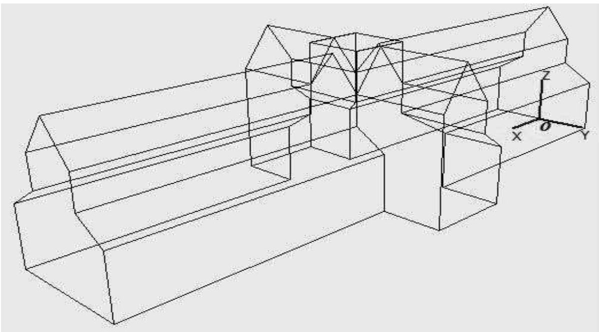

Figure 2: *'Wire Frame' view of phase 1 model as displayed in ODEON 10.1 Auditorium.*

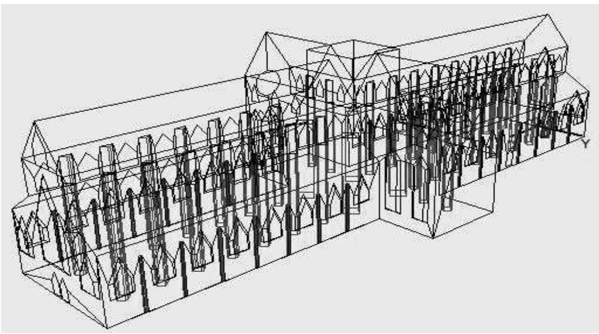

Figure 3: *'Wire Frame' view of phase 2 model showing inclusion of column and window structures.*

level of detailing for the purposes of the project. All measurement data was collated and combined, first in drawn plans, and then applied to the creation of geometry files within ODEON. The central spire of the Abbey was omitted from all models as its inclusion was believed to have negligible effect on the overall acoustics of the interior space (as in [6]).

In total, 3 virtual models were implemented (as shown in Figures 2-4). The first phase of modelling incorporated a basic room geometry consisting of large ceiling, floor and wall surfaces defined by 56 corner points and 40 surfaces. Interior columns, window bays and windows were included in the second phase model and the final phase saw the inclusion of interior archways and roofing structures. The second and third phase models consisted of 1144 and 2114 surfaces respectively.

In addition to room geometries, surface materials were also defined using the most suitable approximations available in the ODEON material library.

3. SIMULATION AND RIR GENERATION

Prior to the simulation of each acoustic model it was necessary to define certain parameters to ensure impulse responses were captured correctly. The values chosen for these parameters were the same for each of the three models to ensure a valid comparison of results obtained.

3.1 Simulation Parameters

Basic simulation parameters that can be altered in ODEON include the transition order, scattering type and impulse response length. Reflections of an order equal to or less than the transition order are modelled using a combination of the image-source and

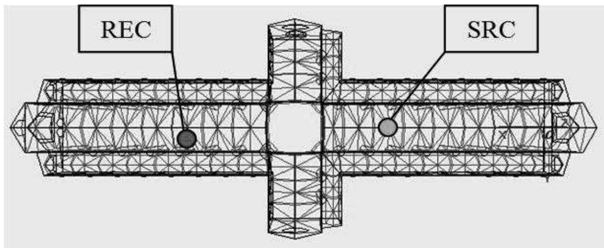

Figure 4: *Plan view of phase 3 model highlighting the addition of archways and ceiling structures. Source ('SRC' right) and receiver ('REC' left) locations, as defined in all three ODEON models, are shown.*

ray-tracing methods. Reflections of a higher order are calculated through use of ray-tracing alone. This quantity was set to a value of 2. The scattering type was set to 'Lambert Scattering'. Lambert scattering applies a probability governing the amount that ray reflections may deviate from specular reflection paths. This provides a more realistic approximation than that given through application of Snell's law. The impulse response length was set to 20 seconds (at least 2/3 of the highest estimated reverberation time) to ensure that the reverberant tail was captured in its entirety.

3.2 Impulse Response Parameters

In addition to the basic simulation parameters, options relating to the nature of the impulse responses are also available in ODEON. These include bit depth of the resulting .wav files, impulse response type and surround sound impulse response calibration. In order to ensure the audio files produced were of sufficient quality for further use, 24-bit PCM .wav files were chosen to store RIRs. It is possible to generate a number of different RIR types simultaneously during simulation, and so, both Ambisonic B-format and 2D surround sound impulse responses were selected for production. The playback system itself consisted of two Genelec 8030As in a standard stereo configuration, presented to listeners in an acoustically treated recording studio control room environment.

3.3 Sound Source and Receiver Locations

An omni-directional virtual sound source was positioned in the choir altar area of the Abbey and a receiver position was defined in the Nave (both at a height of 1.75m). This combination of source and receiver placement was intended to represent a likely scenario and configured to mimic a vocal performance as experienced in the audience area. A plan view of these locations is provided in Figure 4.

4. RESULTS

Results relating to the reverberation time and early decay time of each acoustic model were computed by applying ISO3382 T_{30} and EDT [7] calculations to the RIRs generated from ODEON. Mono RIRs were obtained by extracting the W-channel from each B-format RIR .wav file. The B-format W-channel corresponds to the output of a single omni-directional receiver. Direct analysis of the Mono RIRs was realised through use of [8]. The ISO parameters measured in each virtual model are presented in Figures 5 & 6. EDT and T_{30} times obtained from RIRs measured in York Minster [9], a large cathedral of similar size and shape to

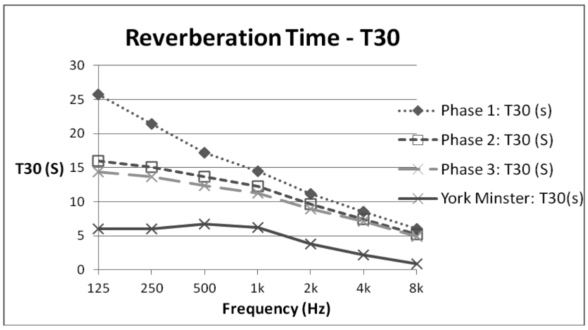

Figure 5: *T_{30} values derived from RIRs obtained from each virtual model and calculated using [8]. Values obtained from measurements taken in York Minster are included for comparison.*

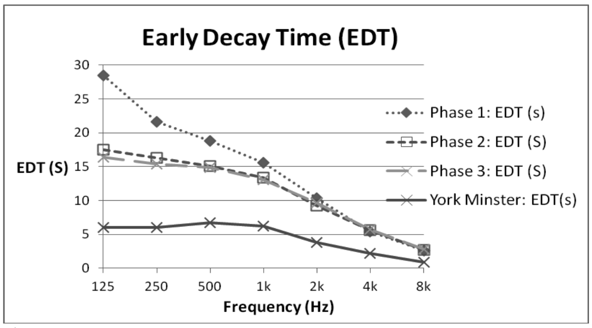

Figure 6: *EDT values derived from RIRs obtained from each virtual model and calculated using [8]. Values obtained from measurements taken in York Minster are included for comparison.*

that of St. Mary's, are included by way of comparison. The findings detailed in this section present a justification for the nature of the subjective test results given in section 5.

4.1 Reverberation Times, T_{30}

The T_{30} results (Figure 5) correlate with predicted results regarding a reduction in reverberation times as frequency increases. In practical measurement of real spaces this is generally the case, owing to the fact that low frequency sounds possess a higher level of energy than that of high frequencies. The phase 1 model possesses the longest reverberation times across all measured frequency bands. This is particularly apparent in the first 3 octaves. The large differences in T_{30} values between phases 1 and 2 are present due to the insertion of column structures within the model. These columns scatter sound creating a diffuse sound field much earlier than in the basic model, as more acoustical energy is lost due to an increase of surface reflections and subsequent absorption. A less extreme variation in reverberation times is shown between phases 2 and 3. As such, it may be assumed that the addition of archway and roofing structures acts to slightly lower reverberation times due to a decrease in model volume and the scattering effects of non-planar ceiling detail. When compared with results obtained from York Minster, noting obvious differences in measurement and construction (e.g. source/receiver locations, more wooden furnishings and interior features, a larger height and volume), T_{30} (and EDT) values are considerably higher, indicating that a completely bare interior, even with some considerable geometrical detail, is not sufficient in objective terms to arrive at a realistic result.

4.2 Early Decay Times (EDTs)

The results derived from EDTs (Figure 6) also show a correlation with predictions as reverberation times decrease with increase in frequency. This measurement is heavily influenced by direct sound and early reflections as it is calculated from the first 10dB decrease of the sound pressure level decay curve. It is for this reason that the EDT measured in the phase 1 model is much greater at lower frequency bands than that measured in phases 2 and 3. This further supports the conclusion that the sound fields in phases 2 and 3 become diffuse earlier than phase 1 due to the scattering effects of the interior geometrical detailing. A smaller difference in EDT is apparent between phases 2 and 3 at lower frequency bands. This shows that the inclusion interior archways and roofing structures act to further disperse sound, but only marginally in the first two frequency octaves measured. At higher frequencies, all three models show increasingly consistent EDTs in agreement with the T_{30} measurements.

5. CONVOLUTION PROCESS AND SUBJECTIVE TESTS

The stereo RIRs were generated using the parameters and source and receiver positions as detailed in section 3. These RIRs were convolved with mono anechoic excerpts sourced from recordings of choral performances (one male only, plainchant; two mixed choirs, excerpts from arrangements of the sung mass) obtained at the Audio Lab at the University of York. This resulted in a number of 2-channel stereo auralizations suitable for stereo playback.

5.1 Convolution using SIR Plug-in

Three two-channel RIRs were employed for auralization. The .wav files containing the impulse responses were imported into the Steinberg Nuendo audio processing suite. A project file was then created using SIR v1.011 audio tools VST plug-in [10] for the audio convolution. The anechoic material was edited such that three suitably short phrases (lasting in the range of 2-5 seconds) were obtained. Each excerpt was then convolved with each 2-channel RIR producing 9 auralizations in total.

5.2 Subjective Testing and Results

A simple subjective test was devised to assess the level of audibility in the variance of acoustical properties possessed by each model of St. Mary's Abbey. The test material, played back in stereo, consisted of audio tracks created using the auralizations obtained. Tracks were arranged such that an auralization of one model (using a particular anechoic excerpt) was played, followed by an auralization of a different model (using the same anechoic excerpt). Then one of the two auralizations was repeated, thus each test case contained three auralizations, one of which was dissimilar to the other two. A total of 9 cases were compiled in this manner, in such a way that all model combinations (phases 1 and 2, phases 2 and 3, phases 1 and 3) were presented using the three anechoic audio excerpts. Each track was played twice during testing, giving a total test length of 18 cases with the order of tracks randomized. During testing, subjects were asked to select the excerpt on each track that they believed to be different from the other two. Correct selection would suggest audible differences between the acoustic qualities of the models experienced. The test tracks were played once only with a pause between each,

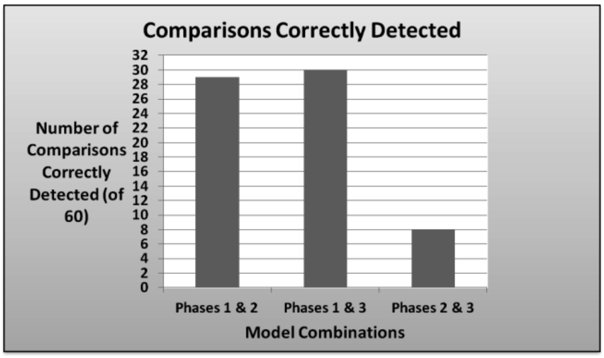

Figure 7: *Bar chart displaying the number of correct answers given for each combination of models over all 10 subjective tests.*

allowing the subject time to arrive at a considered decision. The track and auralization arrangement (previously described) allowed 6 comparisons of each model combination to be made by the subject per test. In total 10 subjects carried out testing, their cumulative results are presented in Figure 7.

The chart above shows that, on average, subjects correctly detected less than 50% of the presented model comparisons. However, these results possess a correlation with those gained analytically. The variation in the audible nature of the soundscapes is shown to be more apparent between the phase 1 model and the remaining two detailed models. Also, these results suggest that the acoustical differences between model phases 2 and 3 are difficult to perceive with less than 14% of all comparisons correctly detected. The numerical data relating to reverberation times (section 4) supports the findings of these tests. This indicates that the larger differences in the low to middle frequency reverberation levels between phase 1 and phases 2 and 3 are more perceptible than that of the smaller variance present between phases 2 and 3 (for this given combination of source/receiver positions and as determined by the source material used). These findings are further supported given that the just noticeable difference (JND) for EDT is 5% [11]. The difference in EDT between phases 2 and 3 varies less than 5% for all but the lower two frequency bands, a much smaller variance than that shown between phases 1 and 2. However, the majority of comparisons were incorrectly detected suggesting that the geometrical detailing added to the model does not profoundly affect the perception of these changes, accepting the limitations of the complexity of this model and the nature of the subjective test carried out as a first means of validation.

6. CONCLUSION

The studies described in this paper formed an investigation into the effect of altering levels of geometrical detailing within a virtual acoustic model on results obtained analytically and empirically. Findings show that the addition of interior structures (such as columns, archways, roof detailing etc.) cause overall reverberation times to decrease as predicted. However, the audible differences between basic and detailed models have been shown to be less apparent through subjective tests. The large reverberant tails present in the auralizations of the phase 1 model provide a listener with sufficient cues allowing an aural comparison between the basic and detailed models to be correctly detected with a success rate of only 50%. Subjects were mostly unable to perceive small changes in reverberation times created by the insertion of detailed roof vaulting. This indicates that certain geometrical features inserted in the detailed models may have a negligible effect on auralizations created in order to sonically revive the Abbey's acoustics.

It should, of course, be noted that this model is not an exact replication of the original Abbey, given that it now only exists in part, and that much of the architectural detail that would have been present has been approximated with much simpler planar surfaces. However, the overall geometry and construction has been built based on available evidence obtained either on-site, or from documentation from a variety of sources including archaeologists expert in the period. The subjective test was limited in its scope, and designed to provide some additional validation of the changes observed in the calculated and measured reverberation times. Given that only one type of sound source has been used, it is interesting to note that, in general, a listener could not identify the changes made with anything more than a 50/50 chance. Further work should conduct a more rigorous test based on different source/receiver location combinations, and use a more varied range of anechoic source material. This short study has demonstrated that convincing reverberant effects can be synthesized from even simple geometrical models of otherwise complex spaces. Certainly, this would be suitable for a first approximation of the acoustic properties of a historical space that no longer exists in its entirety. Apart from use of these RIRs in music production for generating reverb effects, other applications include on-site audio guides, web presentations and the potential for interactive virtual reality walkthroughs, all of which help in terms of facilitating enhanced access and engagement with historic sites.

7. REFERENCES

[1] J. H. Rindel, "The ERATO project and its contribution to our understanding of the Acoustics of Ancient Theatres", *Proc. of the Acoustics of Ancient Theatres Conference, Patras, Greece*, 2011.

[2] The CAHRISMA Project Website, Accessed April 12th, 2012, [Online], http://server.oersted.dtu.dk/www/oldat/cahrisma/.

[3] Doc Brown's Science Website, Accessed April 12th, 2012, [Online], http://www.docbrown.info/docspics/yorkscenes/

[4] N. Pevsner, "Yorkshire: York and the East Riding", Pevsner Architectural Guides; illustrated edition, 1975.

[5] B. Wilson and F. Mee, "St Mary's Abbey and the King's Manor York: The Pictorial Evidence", York Archaeological Trust, 2009.

[6] Foteinou, A., Murphy, D.T., and Masinton, A., "Investigation of Factors Influencing Acoustic Characteristics in Geometric Acoustics Based Auralization", *Proc. of the 13th Int. Conference on Digital Audio Effects (DAFx-10)*, pp. 178-181, Graz, Austria, Sept. 6-10, 2010.

[7] ISO3382, International Standard 3383, "Acoustic– measurement of reverberation time of rooms with reference to other acoustical parameters", ISO, Second Edition, 1997.

[8] AURORA Plug-ins, Accessed April 12th, 2012, [Online] http://pcfarina.eng.unipr.it/Aurora_XP/index.htm

[9] OpenAIR Website, Accessed July 17th, 2012, [online] http://www.openairlib.net/auralizationdb/content/york-minster

[10] SIR Audio Tools, Accessed April 12th, 2012, [Online] http://www.knufinke.de/sir/index.php

[11] Prodi, N. and Velecka, S., "The evaluation of binaural playback systems for virtual sound fields", *Applied Acoustics*, 64, 147-161, 2003.