EG UK Theory and Practice of Computer Graphics (2010)
John Collomosse, Ian Grimstead (Editors)

Audio-Visual Animation of Urban Space

P. Richmond[1] , Y. Smyrnova[2] , S. Maddock[1] & J. Kang[2]

[1]Department of Computer Science, University of Sheffield
[2]School of Architecture, University of Sheffield

Abstract

We present a technique for simulating accurate physically modelled acoustics within an outdoor urban environment and a tool that presents the acoustics alongside a visually rendered counterpart. Acoustic modelling is achieved by using a mixture of simulating ray-traced specular sound wave reflections and applying radiosity to simulate diffuse reflections. Sound rendering is applied to the energy response of the acoustic modelling stage and is used to produce a number of binaural samples for playback with headphones. The visual tool which has been created unites the acoustic renderings with an accurate 3D representation of the virtual environment. As part of this tool an interpolation technique has been implemented allowing a user controlled walkthrough of the simulated environment. This produces better sound localisation effects than listening from a set number of static locations.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Radiosity, Raytracing, Virtual Reality H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems

1. Introduction

In the past few decades acoustic simulation models that can accurately predict sound propagation in indoor and outdoor spaces have become very popular within the scope of room acoustics and, more recently, in urban acoustics design [Kro68, Nay93, MK04, Kan06, Men08]. Such models are, however, computationally expensive and are not suitable for integration with real-time interactive software such as interactive computer-aided architectural design (CAAD) walkthroughs, which are used extensively in architectural design and urban environment planning. Instead, pre-rendered video, silent interactive walkthroughs or walkthroughs which utilise simple, non-physically-accurate acoustics have been used.

For real-time, interactive graphical applications (e.g. games), the primary focus is on producing content extremely quickly, often sacrificing accurate physical simulation for faster approximation-based techniques. As a result, audio rendering within such entertainment-focused applications is not a major concern and instead a perceptually-based approach is widely used to provide sound simulation and effects [SHLV99, TGD04]. Since no modelling is required, the perceptually-based approach requires significantly less processing power, which is thus freed up for other aspects of a the game loop simulation, such as artificial intelligence, multiplayer networking and advanced rendering effects. Whilst the accuracy/performance trade-off is acceptable, and arguably necessary, for games, the design and planning of urban environments requires objective measures of sound fields to be calculated within a certain level of accuracy [Kan06]. Such accuracy cannot be obtained without using a physically-based approach for sound rendering, and therefore requires explicit modelling of the acoustic process.

Just as accurate acoustic modelling is important, it is also essential to consider that several studies have shown that listeners are unlikely to perceive a complex auditory environment such as a multi-source dynamic urban space in its entirety [TGD04, Tsi07]. Therefore, within the modelling process, some acceptable level of simplification may be considered in order to achieve the trade-off between the accuracy in acoustic modelling and rendering of the virtual auditory scenes (auralisation [KDS93]) and the speed of the simulation process. Some previous techniques have embraced this concept by modelling specular reflective sound propagation in real time though the use of ray tracing on GPUs [RKM07]. Likewise, the GPU has been used effectively to simulate a frequency domain approach to sound modelling using adaptive rectangular decomposition based on discrete cosine transformations [RNL09]. Although in some real-time auralisation-related software it is possible to

DOI: 10.2312/LocalChapterEvents/TPCG/TPCG10/183-190

provide a walkthrough sensation of a rendered sound, they only utilize direct sounds and no reflections are taken into account [Far05]. In such software, the sound effect is calculated with a temporal resolution similar to the resolution of the hearing mechanism, and thus no interpolation between points is required. In this paper, we favour an alternative offline acoustic modelling process which utilises online interpolation between a discrete set of pre-calculated acoustic renderings that take into account the complex phenomenon of reflections from the boundaries or an environment. These are combined with real-time visual rendering to achieve a realistic audiovisual experience within urban environments. The acoustic modelling is based on a combined ray-tracing and radiosity (CRR) model [MK04, Men08, SP94]. The advantage of this technique is that the quality of sound modelling need not be compromised as with previous GPU accelerated alternatives which do not consider diffuse sound propagation though the radiosity method. At the very least we can be guaranteed that at the discrete locations (which may represent key positions within the environment) the listener is always subject to accurate acoustic simulation with the approximation taking place between the locations. In the next section, we present the acoustic simulation process. Section 3 and 4 will present the visual modelling process and the environment used for the experiments respectively. Section 5 will then present the results of series of subjective experiments used to assess the auditory experience, with section 6 presenting conclusions.

2. Acoustic simulation

2.1. Scene modelling

The acoustic simulation of sound propagation with a combined ray-tracing and radiosity (CRR) model first requires a 3D geometric model of a simulated urban environment. This is a much-simplified version of the model used for visual simulation, which will be described in section 3. Since radiosity is employed in the acoustic simulation, the boundaries of the urban environment model are subdivided into patches [Kan06, Men08]. Autodesk 3DS Max 9.0 is used to do this, as it provides parameters to control the number of the patches to be used, based on time and computer processing capabilities. The geometric model, together with information about the size and location of the sound sources and receivers, and the absorption and diffusion coefficients of the boundaries, is stored in a text-based scene file [FvDFH90].

2.2. CRR model

The core of the acoustic simulation process is the CRR model, which combines ray tracing and radiosity models, allowing the modelling of different patterns of sound energy reflection from the boundaries. Ray-tracing is used to calculate specular reflections and radiosity deals with diffuse reflections. The consideration of diffuse reflections from boundaries is important in terms of acoustic objective indices and subjective perception [Kan00, SMK08b]. In the ray-tracing part of the calculation, a number of rays are emitted from the sound source in random directions. The density of rays radiated in a particular direction may be constant for the whole space or may in some way reflect the spatial characteristics of the sound source. Rays then travel through the scene and their energy decreases due to absorption in the air and at the boundaries during specular reflections. In the case study (see later), which uses a 35x32x15m space, the receiver is simulated as a transparent sphere of diameter 0.5m. The energy of sound is obtained in predefined time intervals by summing up the energy of all rays that have crossed the listener. For the purposes of this paper, a time interval of 10ms has been used, considering the time resolution of the hearing mechanism [Kut93].

With respect to simulating diffuse acoustic reflections, patches are used to emit and receive energy from each other, thus processing the radiative exchange [SP94]. The combination between ray-tracing and radiosity is shown in Figure 1. When a ray hits a surface patch, part of the incident energy is reflected specularly and carried further along the reflected ray. If the ray hits a receiver, its energy is recorded into the receiver. Another part (diffuse energy) is stored into the patch; then patches carry out the energy exchange, to calculate the diffuse reflection. Energy impulse responses are then obtained in four octave bands (125Hz, 500Hz, 2kHz and 4kHz). Typical objective acoustic indices such as sound pressure level (SPL), reverberation time (RT) and early decay time (EDT) can also be calculated [Kut79].

Digital signal processing to reconstruct the pressure impulse responses from the calculated energy responses is the last stage of the CRR software. It is realized by applying the method of "microscopic" structure of impulse response reconstruction, extensively used in room acoustics auralisation [Kut93]. The details can be found in [Men08]. The impulse responses are obtained in the so-called B-Format that takes into account the 3D directional information of the received sounds and is based on Farina and Ugolotti's approach [FU98]. The B-format impulse responses are presented as four "channels" of wide-band pressure (W) and particle velocity X, Y and Z components [Ger73] with a sampling rate of 44.1kHz and a bit resolution of 16 that satisfies CD quality.

The CRR model has been proved to provide accurate results through comparison with other simulation models and with a number of measurement results [MK07, MKS08, SMK08a]. A number of parametric studies have also been carried out for the investigation of parameters, and for the investigation of the relationship between specular and diffuse reflections [MK07].

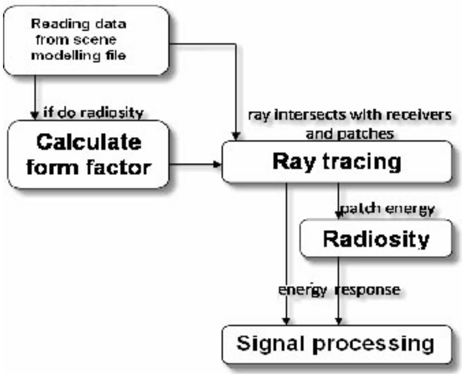

Figure 1: *Calculation procedure for the CRR model.*

2.3. Sound rendering and reproduction

For the purpose of sound rendering the impulse responses calculated through the CRR process are further convolved with dry signals [KB98] (recorded in the free field conditions without background noise and reflecting surfaces). The obtained signals can be further decoded for reproduction using a variety of surround technologies, such as binaural for playback via headphones, or 5.1 or Ambisonics [MM95] via various arrays of loudspeakers. We use the binaural technique for playback via headphones. In order to preserve the spacial sound information embedded in the obtained sound files, the four B-format signals are decoded into binaural stereo files using Ambisonic Player 1 software (www.muse.demon.co.uk/utils/ambiplay.html) and applying Head Related Transfer Function (HRTF) filters [Møl92].

3. Visual simulation

3.1. The visual simulation tool

The visual simulation tool has been created for the purpose of combining high fidelity urban scene rendering with the real time playback of acoustic renderings described in the previous section. The ultimate purpose behind creating a generic tool for doing this is to allow architects and urban planners to assess both acoustic and visual renderings of planned architecture simultaneously. In order to provide realistic rendering of the environment the visual tool requires a high fidelity model of the environment (in addition to the acoustic scene file). This can be produced by extending the acoustic model using any modelling software such as 3Ds Max or Google Sketchup. For the case study (see next section), we used Google Sketchup. Within the tool the user is able to interchange views between the detailed model and a visualisation of the underlying acoustic model which is used for the CRR simulations (Figure 2). The rendering of the environments is achieved with the Ogre 3D (ver. 1.6.5) graphics library which has been extended to include self casting texture based shadows which can be turned on or use using a configuration file. Within the visual simulation tool, playback of acoustic samples generated by the CRR approach is available using two differing methods. The first of these uses pre-computed location-direction playback. Here, the user stands in a particular location, looking in a particular direction, that corresponds to one of the receiver location-directions pre-calculated using the CRR approach. This provides a direct link between a single rendered acoustic sample and a visual representation of the scene from that location, giving the user an indication of the acoustics of a given location within the urban space. Using the visual simulation tool, the user is able to select from a number of receiver locations each of which may contain a number of acoustic samples from different directions. Selecting a sample will translate the user to the sample's location and orientate them accordingly. Whilst it is possible to move and look freely around the environment, playback is only available when located in one of the rendered positions.

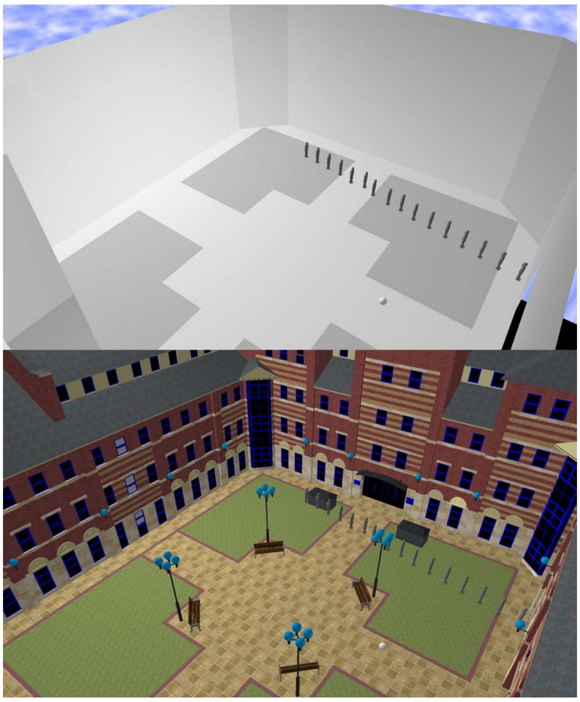

Figure 2: *An example rendering of an acoustic model (above) and visual model (below) within the visual simulation tool (The short vertical lines are representations of sound receiver locations that are used in the experiments in section 4).*

The second playback technique allows the user to move along a pre-defined animation path within the environment. This path links a number of acoustic samples. When a user is located at a sample point they will hear the acoustic sample in the same way as playback from a pre-computed location.

For movement between calculated samples (including both positional and directional aspects) playback is interpolated between the two nearest points using linear interpolation of the sample gains (as shown in Figure 3). This second playback technique is hardware accelerated through the use of the OpenAL API which is used for playback and mixing of the samples through control of the sample gains (volumes). When animated acoustic playback begins each of the acoustic samples has its own sound source and all samples begin playing simultaneously with all but the nearest two samples set to a minimal gain. Starting all sound sources simultaneously ensures an accurate synchronisation between samples and ensures there are no audio artefacts or delays which would have a detrimental effect. The user controls movement along the animation path by using a horizontal slider on the interface. It is also possible to jump to particular sample locations.

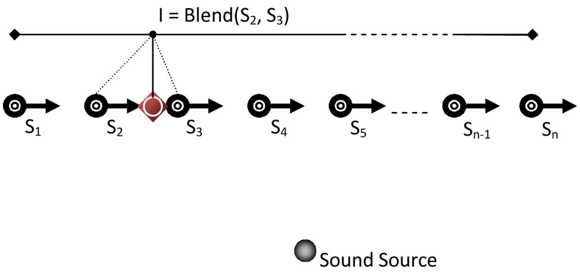

Figure 3: *The black double circles labelled S represent precalculated sound receivers. The position I represents an interploated position which may lie anywhere between S_1 and S_n. The interploation is based on the two nearest precalculated positions, in this case S_2 and S_3.*

In order to generalize the tool, configuration of the scene and sample sets is available by specifying an XML input file which is passed to the simulator as a command argument. The XML document (which can be validated by an accompanying XML Schema) briefly comprises of a model file specification, an audio source location, a set of sample locations and an animation path. The model file specification includes the filename of both the acoustic model file and the visualisation model file which must be located in the simulator's data folder. The audio source requires a position which indicates the location of the original sound source used in generating the acoustic samples. Within the simulation this is represented as a white spherical object. A sample location has a name, position and any number of acoustic samples. Each sample has an associated wav file location and a viewing direction from the sample position. An animation path is specified by listing sample names as key points to be used for the interpolated movement and acoustic playback.

4. A Case Study Site

In order to test the simulation tool, the Regent Court quad of the University of Sheffield has been selected to represent a typical open urban area. The courtyard itself is 35m long, 32m wide and 15m high and is surrounded by buildings. The boundary materials include brick walls mixed with windows, a glass panel located at each corner of the courtyard, and a ground plane of grass mixed with pavements of concrete tiles. Figure 4 presents a photographed view from inside the Regents Court quad as well as a screenshot from our detailed virtual representation. This high fidelity visual model was designed using Google Sketchup and consists of approximately 12,000 edges and 62 unique materials.

Figure 4: *A view from the Regents Court case study site (above) and our virtual representation (below).*

For the purposes of the CRR simulation, a relatively low fidelity model has been created consisting of box with 4 truncated edges. This boundary is divided into 8m^2 sized patches (giving a total of 463 patches) and each is labelled with particular absorption coefficients based on the material in that area of the model. Average boundary absorption coefficients were used and the air absorption was determined for each octave band under calculation. The sound source was placed at a static position and 16 receiver locations, each 1.5m apart from its neighbours, were arranged in a line, as shown in the detailed model in Figure 5. The sound source and the receivers were simulated to be 1.5 m above the ground.

The performance of the simulation tool with our case study site resulted in a steady 70fps on the development machine, an AMD Athlon 64 (2.5Ghz) with 2GB of RAM and a NVIDIA 9800GX2 GPU. It is worth noting that the performance rate of the tool is highly dependent on the complexity of the visual model with the interpolation between audio samples playing a relatively negligible part of the overall performance.

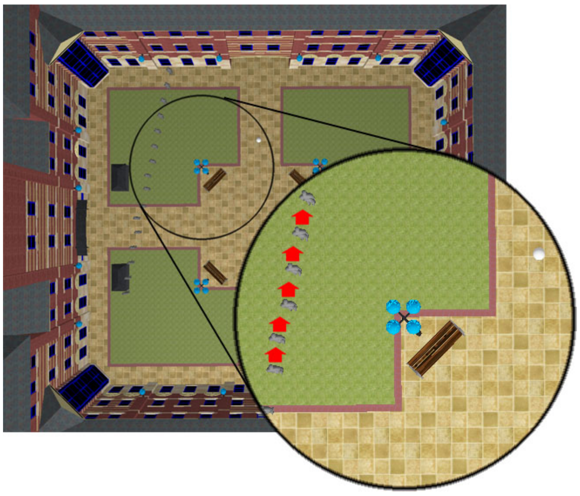

Figure 5: *Overhead view of regent court showing the position of the sample locations (grey), direction of view (red and wrt Experiment 1 with the correct orientation) and acoustic source position (white).*

5. Experiments

In order to assess the use of the visual simulation tool, we designed an experiment to test the link between visual rendering and sound localisation. This is described in section 5.1. The second experiment examines the use of interpolation to produce acoustic samples between pre-calculated locations. In each experiment the participant was given a maximum of 30 seconds decision time (after the end of playback) to complete their assessment for each question/task.

5.1. Experiment 1: Evaluating the use of visual animation

The aim of this experiment was to verify the importance of the visual animation for perception of the sound environment of an urban space. The experiment consisted of two parts. Experiment A tested whether participants could tell whether what they heard matched what they saw. Experiment B tested how accurately they could localise a sound when given either sound samples at discrete locations or as part of an animation path. Figure 5 shows an detailed version of the environment's sound setup for these experiments. In both these experiments, the simulation tool was run in a virtual reality laboratory which includes an 8 foot by 6 foot rear projected screen with stereoscopic rendering (achieved with active 3D glasses) of the environments. Participants were also provided with a set of high quality headphones to listen to acoustic outputs. In all experiments, participants were familiarised with the environment and the visual simulation of Regent Court before the experiments began.

In experiment A, 10 participants were presented with audiovisual walkthroughs along the animation path shown in Figure 5 and represented in Figure 3. The walkthrough maintained the pre-calculated view direction in the sound samples. Inbetween sound samples were interpolated from the nearest neighbours. In the first walkthrough, the correct path from start to finish in Figure 5 was presented. Thus the sound should appear to be coming from the correct position. (The participant controlled movement along the path using a slider bar on the interface.) In the second walkthrough, the incorrect path from finish to start was presented, i.e. with the viewer looking in the wrong direction as the pre-calculated acoustic information was played. All the participants identified both the correct and incorrect walkthroughs, thus demonstrating that they knew when the acoustics were correctly orientated to match the visual information they were receiving.

Five new participants took part in experiment B. The sound source was not visible and the aim was to investigate whether or not they could localise where a sound was coming from under two different sets of conditions, and whether or not one condition gave better results than the other. In the first condition, a participant was presented with pre-calculated acoustic samples from five locations (from the set of 16 shown in Figure 5). Here, the participant could switch between each of the five locations. In the second condition, a participant was presented with a real-time walkthrough along the path between the five pre-calculated acoustic sample locations, where acoustic information at inbetween locations was calculated using interpolation from the nearest neighbours. Here, the participant could continuously manipulate his position back and forth along the path using a slider bar on the interface.

To assess the sound localisation process for experiment B, each participant was given a 2D plan of Regent Court (shown in Figure 6) showing the position and direction of five sound receivers, but were not presented the position of the sound source. Each participant was asked to indicate on the plan (measuring 13cmx13cm) where they estimated the sound was originating from. Figure 7 shows the results of this. Whilst there is an issue regarding scale, since the participants are using a scaled map to indicate locations on, rather than indicate the position in the 3D world, the relative results between the two sets of experimental conditions show that, for four out of five cases, the participants were better able to estimate the position of the sound source when a continuous animation path was used. Whilst the data set is small, the indication is that a better experience is given if real-time audiovisual animation is used in a virtual urban space rather than only being able to listen to sound at a set number of locations.

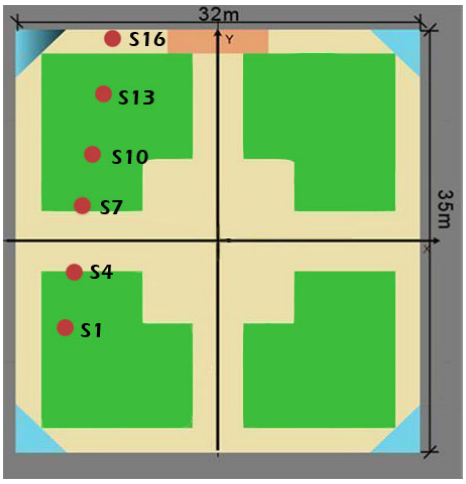

Figure 6: *Overhead plan used to indicate subjects perceived location of the sample source position in experiment 1B.*

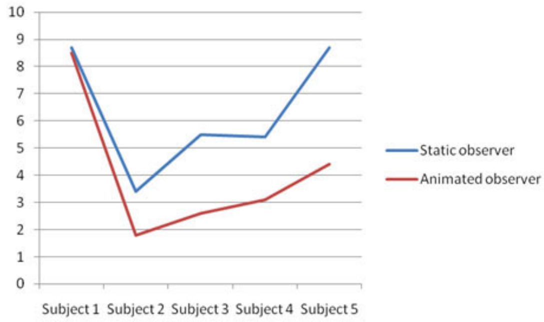

Figure 7: *Distance between the actual location and observed location of the sound source for static and animated observer (distance is measured in cm on the 2D printed representation).*

5.2. Experiment 2: Subjective evaluation of acoustic sample interpolation

The aim of the second experiment was to verify the interpolation procedure used to provide real time movement through the environment and to identify interpolation distances which produce an acceptable acoustic representation of positions between the pre-calculated points. This experiment did not involve the visual simulation tool and was restricted to just using the binaural sound setup. Four of the 16 pre-calculated locations shown in Figure 5 were chosen at unique position (S_i). We will call this A. Three further samples were then created using interpolation, with B = blend(S_{i-1}, S_{i+1}), C = blend(S_{i-2}, S_{i+2}), and D = blend(S_{i-3}, S_{i+3}). These samples were then put into 4 pairs: AA, AB, AC and AD.

The same 10 participants in experiment 1A, were used for this experiment. A participant was presented with each of the 4 pairs, in a random order to avoid any ordering bias. For each pair, the participant was asked to rate the similarity of the pair of samples on a scale of 0 to 10, where 0 meant the two samples were the same, and 10 indicated they were very different.

The results of the experiment were analysed (using SPSS software version 16) in terms of the differences between locations in subjective ratings for each of the interpolation method. As the 10-point scale is categorical, rather than continuous, and due to a relatively small sample size of subjects participated in the experiment, all obtained results were weighted by the number of observations (frequencies) that fell into each category prior to the further analysis. Figure 8 presents the graph of the distribution of the average (mean) scores of ratings of all interpolation distances over the locations.

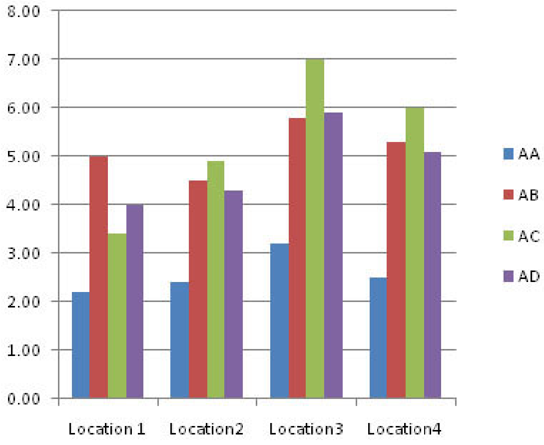

Figure 8: *Mean scores of ratings of all interpolation procedures over the locations.*

It can be seen from Figure 8 that the three pair samples involving interpolation give higher results than the uninterpolated pair of samples. Thus, it is not possible to conclude which of the tested interpolation procedure provides the sound sensation closest to the sound from the location for which accurate sound simulation was performed with CRR.

An Analysis of Variance (ANOVA) was further applied to test the null hypothesis that there is no significant difference between subjective ratings of the studied locations in terms of each of interpolation procedure. A 0.05 chance of making an error (confidence interval) was applied for rejecting the null hypothesis. The ANOVA revealed that a significant difference between the locations in terms of the interpolation procedures exists only for case AD ($p<0.029$). However, no significant difference was found in terms of other interpolation procedures. Considering the results obtained, two possible reasons of no significant difference between the interpolation procedures can be suggested. First, it might be due to an error in the mixing procedure (i.e. the mixing introduces artefacts which are obvious to the observer) or alternatively the spatial difference between locations introduces temporal artefacts (i.e. echoing) when the interpolation is applied.

In order to understand our results the two suggested sources of error have been explored by conducting further experimentation which again required subjects to compare differences between pairs of sounds using the same 10-point scale. Four pairs of sounds were again presented which consisted of the following;

1. A pair of pre-calculated samples from the same location (i.e. the baseline AA which was used previously),
2. A pair of sounds including an single interpolated sample, either AC or AD depending on which had the largest margin of error as indicated in Figure 8.
3. A pair of sounds AE where E = blend(S_i, S_i)
4. A pair of sounds AC' or AD' based on the same pair as in case 2, where an attempt to correct any temporal difference between the sample points S_{i-2} and S_{i+2} in the case of C' and S_{i-3} and S_{i+3} in the case of D', was achieved by delaying the playback of one of the samples according to geometric distance from the sound receiver position.

The experiment was repeated for three locations (location 1 to 3) and as previously the order of the four pairs of samples was randomised with the same participants taking part as in the second part of experiment 1.

The results of this extended experiment (shown in figure 9) were also analysed using SPSS software version 16. All obtained results were again weighted by the number of observations (frequencies) that fell into each category prior to further analysis. The error in the mixing of samples was evaluated by considering cases 1 and 3 above for all three locations. Applying an ANOVA test no significant difference ($p<0.518$) between these cases was found and thus it is possible to conclude that the mixing procedure provided no effect on the perceptual evaluation. The effect of temporal delays between interpolation points has been evaluated by considering the evaluative differences between the results of cases 2 and 4 above. The results indicated that subjects found fewer differences between interpolated samples when the calculated geometric delays were applied to playback (the mean value dropped from 0.49 to 0.45). This difference was however not statistically significant ($p<0.602$), indicating that using a single delay for sound propagation is not enough to improve the perception of interpolated acoustics. This leads us to the conclude that delays between reflections in various frequencies must instead be evaluated in future work if one wants to reproduce accurate audio animation using interpolation. The use of smaller interpolation distances may also reduce perceived difference and should be explored in future work.

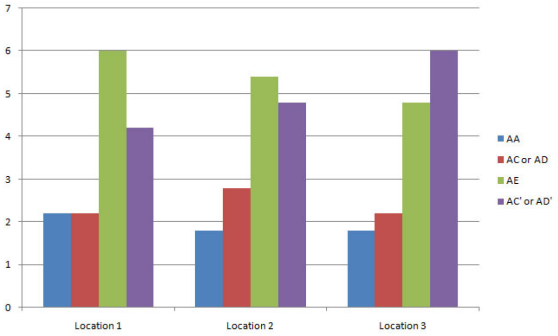

Figure 9: *Mean scores of ratings of all interpolation procedures for extended experiments used to assess the effect of apply temporal delays.*

6. Conclusions

The aim of this paper was to study the importance of the visual representation of an urban space for the subjective evaluation of a sound environment and to verify the applicability of an interpolation procedure based on a volume mixture between locations for which the received sound have been accurately calculated. It has been found that the visual representation benefits the subjective perception of an urban sound environment, especially with respect to sound source localisation. The benefit of real time animation and interpolation versus perception of acoustic samples from a limited set of locations in a virtual space has also been demonstrated.

In terms of the quality of the interpolation procedure it has been found that the simple mixture of the volume cannot be applied directly to provide results of equal quality to those that are modelled directly. A more complicated interpolation procedure which takes into account the frequency content of simulated sounds and delays between direct and reflected sounds is instead recommended as a likely avenue for future work. The density (or distance) of the pre-calculated locations used for interpolation also requires further investigation. The use of GPU acceleration to perform both the ray traced and radiosity methods in real time should also be explored.

References

[Far05] FARIA R.: Audience: Audio immersion experience by computer emulation project, 2005. 2

[FU98] FARINA A., UGOLOTTI E.: Software implementation of b-format encoding and decoding. In *Pre-prints of the 104 rd AES Convention* (1998), pp. 2–0. 2

[FvDFH90] FOLEY J. D., VAN DAM A., FEINER S. K., HUGHES J. F.: *Computer graphics: principles and practice (2nd ed.)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990. 2

[Ger73] GERZON M. A.: Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society 21(1)* (1973), 2–10. 2

[Kan00] KANG J.: Sound propagation in street canyons: comparison between diffusely and geometrically reflecting boundaries. *The Journal of the Acoustical Society of America 107* (2000), 1520–8524. 2

[Kan06] KANG J.: *Urban sound environment*. London ; New York : Taylor & Francis, 2006. 1, 2

[KB98] KAHRS M., BRANDENBURG K. (Eds.): *Applications of digital signal processing to audio and acoustics*. Kluwer Academic Publishers, Norwell, MA, USA, 1998. 3

[KDS93] KLEINER M., DALENBÄCK B.-I., SVENSSON P.: Auralization - an overview. *Journal of the Audio Engineering Society 41*, 11 (1993), 861–875. 1

[Kro68] KROKSTAD A.; STROM S. S. S.: Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibration 8(1)* (1968), 118–125. 1

[Kut79] KUTTRUFF H.: *Room acoustics*. Applied Science Publishers, London, 1979. 2

[Kut93] KUTTRUFF K. H.: Auralization of impulse responses modeled on the basis of ray-tracing results. *Journal of the Audio Engineering Society 41(11)* (1993), 876–880. 2

[Men08] MENG Y.: *Acoustic simulation and auralisation in urban open spaces (PhD Thesis)*. Master's thesis, University of Sheffield, 2008. 1, 2

[MK04] MENG Y., KANG J.: Fast simulation of sound fields for urban square animation. In *Proc. of the 33rd International Congress on Noise Control Engineering, Prague* (2004). 1, 2

[MK07] MENG Y., KANG J.: Combined ray tracing and radiosity simulation for urban open spaces. In *Proc. of the 19th International Conference on Acoustics (ICA), Madrid, Spain* (2007). 2

[MKS08] MENG Y., KANG J., SMYRNOVA Y.: Numerical modelling of sound fields with mixed specular and diffuse boundaries using combined ray tracing and radiosity method. In *Proc. of the Institute of Acoustics 30(2), UK* (2008). 2

[Møl92] MØLLER H.: Fundamentals of binaural technology. *Applied Acoustics 36*, 3-4 (1992), 171 – 218. 3

[MM95] MALHAM D. G., MYATT A.: 3-d sound spatialization using ambisonic techniques. *Computer Music Journal 19*, 4 (1995), 58–70. 3

[Nay93] NAYLOR G.: Odeon–another hybrid room acoustical model. *Applied Acoustics 38*, 2-4 (1993), 131 – 143. 1

[RKM07] ROBER N., KAMINSKI U., M. M.: Ray acoustics using computer graphics technology. In *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07), Bordeaux, France* (2007). 1

[RNL09] RAGHUVANSHI N., NARAIN R., LIN M. C.: Efficient and accurate sound propagation using adaptive rectangular decomposition. *IEEE Transactions on Visualization and Computer Graphics 15*, 5 (2009), 789–801. 1

[SHLV99] SAVIOJA L., HUOPANIEMI J., LOKKI T., VÄÄNÄNEN R.: Creating interactive virtual acoustic environments. *Jounral Audio Engineering Society 47(9)* (1999), 675–705. 1

[SMK08a] SMYRNOVA Y., MENG Y., KANG J.: Objective and subjective evaluation of urban acoustic modelling and auralisation. In *Proc. of the 124th AES Convention, Amsterdam, The Netherlands* (2008). 2

[SMK08b] SMYRNOVA Y., MENG Y., KANG J.: Subjective evaluation of urban soundscape auralisation based on combined ray-tracing and radiosity (crr) model. In *Proc. of the Institute of Acoustics (IOA) (UK), 30(2)* (2008). 2

[SP94] SILLION F. X., PUECH C.: *Radiosity and Global Illumination*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1994. 2

[TGD04] TSINGOS N., GALLO E., DRETTAKIS G.: Perceptual audio rendering of complex virtual environments. *ACM Trans. Graph. 23*, 3 (2004), 249–258. 1

[Tsi07] TSINGOS N.: Perceptually-based auralization. In *Proc. of the 19th International Conference on Acoustics(ICA), Madrid, Spain* (sep 2007). 1