Proceedings of the International Symposium on Room Acoustics, ISRA 2010

29–31 August 2010, Melbourne, Australia

Room acoustics simulation for multichannel microphone arrays

Andrew Wabnitz, Nicolas Epain, Craig Jin and André van Schaik

Computing and Audio Research Laboratory (CARLab)
School of Electrical and Information Engineering
The University of Sydney, NSW, Australia

PACS: 43.55.Ka, 43.58.Ta & 43.60.Fg

ABSTRACT

A new and efficient room acoustics simulation software package for MATLAB is presented which can simulate recordings of arbitrary microphone arrays within an echoic room. This simulator supports research related to developing and experimenting with multichannel microphone arrays and higher order ambisonic playback. Based upon the work by Schimmel *et al.* [1], this new simulation software package models both specular and diffuse reflections in a shoebox type environment. It is an improvement over previous work as it can simulate microphone arrays with arbitrary directional impulse responses and large numbers of sensors. The spherical harmonic expansion up to a specified order at any point within an echoic room can also be simulated by specifying a microphone array with custom directional gains that match the spherical harmonic functions. Furthermore, this simulator provides realistic phase information for the sound recorded by the microphone array, including accurate inter-sensor time delays for the sources and early specular reflections. The room acoustics simulator is implemented as a C program that interfaces with MATLAB and is freely available from the authors.

INTRODUCTION

Most, if not all, three-dimensional acoustic applications such as sound scene capture, sound spatialisation and high-resolution room acoustic analysis tools involve techniques and algorithms that rely on sound captured from a multichannel microphone array. The performance of these techniques and algorithms are usually evaluated using synthetic anechoic sound scenes or by conducting test measurements. In addition, the performance of various microphone arrays or microphone techniques and their suitability to their intended application are also analysed by conducting test measurements, as in [2]. However, as the aim of a three-dimensional acoustic application is mostly focussed on reliable operation in an echoic environment, such as a concert hall or meeting room, then evaluation of the relevant techniques and algorithms in an anechoic environment can be unsuitable or unrealistic. As well, conducting measurements with a microphone array is not only a time consuming process, but also requires significant effort in order to create reproducible test conditions. Therefore an efficient and accurate room acoustics simulator which has the advantage of creating reproducible results and can handle complex multichannel microphone arrays should serve as a valuable tool to the audio community.

Currently there are a variety of tools available to simulate the room acoustics of echoic environments such as the commercial architectural simulators Odeon [3] or EASE [4]. However the focus of these software packages is on the auralisation and architectural analysis of geometrically complex rooms such as churches, concert halls, etc. This level of sophistication is not necessarily required for the analysis of three-dimensional audio techniques/algorithms or evaluation of microphone arrays in echoic environments. A room acoustics simulation tool which caters more for evaluating signal processing algorithms is the freely available shoebox room acoustic simulator by Schimmel *et al.* for MATLAB [1], which simulates specular and diffuse reverberant sound. This tool is efficient but does not provide the ability to simulate arbitrary microphone arrays of multiple sensors. Also any phase inherent in the directional gain of a microphone is ignored. For example, a dipole microphone would appear to have two positive lobes.

In this paper a new room acoustics simulator is presented, the Multichannel Room Acoustics Simulator (MCROOMSIM) which can simulate recordings of arbitrary microphone arrays within an echoic shoebox room. Furthermore, this simulator provides realistic phase information thus resulting in accurate inter-sensor time delays, which is required for accurate simulation of microphone arrays. The possibility to simulate sources with arbitrary directional characteristics is also provided by this simulator. This software is written in the C programming language and is designed to interface with MATLAB. Versions of MCROOMSIM have been compiled for Microsoft Windows and Mac OSX platforms. The software is freely available from the authors.

DESCRIPTION OF THE SOFTWARE

Features

MCROOMSIM simulates both specular and diffuse reflections in an empty shoebox environment. Effects of air absorption and distance attenuation on sound are also modelled by the simulator.

The empty shoebox room geometry simplifies the necessary calculations used during simulation, which in addition to the simulator being implemented in C, enables a significant reduction in the computation time of room impulse responses, resulting in a fast simulation package. The simplification of the room geometry to a shoebox model is suitable when the importance is in creating realistic room impulse responses for evaluating three-dimensional audio applications in echoic environments, rather than modelling the architectural complexities of the room. The empty shoebox room model is described as a three-dimensional space bounded by six surfaces (walls). Each room surface is modelled using a configurable set of frequency dependent absorption and scattering (diffusion) coefficients. The scattering coefficients set the ratio as a function of frequency between the amount of specular and diffuse reflection. Sound scattering due to furniture and other objects in the room can be approximated by higher levels of overall room diffuseness.

MCROOMSIM supports the simulation of any number of source and receiver combinations, with arbitrary locations and orientations within the room. Each source and receiver under simulation can be configured with its own directivity. A set of idealised directivity patterns such as omnidirectional, dipole, cardioid, etc. are provided to the user, thus enabling the quick configuration of specific directivities for sources and receivers. One of the key features of MCROOMSIM, however, is its ability to simulate sources and receiver arrays with arbitrarily defined directivity patterns. The user can configure a custom directivity pattern for a source or receiver by supplying MCROOMSIM with a list of directional impulse responses or, alternatively, a list of real-valued gains constant across frequency for a number of directions in space. A list of directional impulse responses is useful when specifying arbitrary microphone arrays, such as spherical microphone arrays, or the head related impulse responses associated with the ears of a listener. A list of real-valued gains is useful for simulating simple directivity patterns constant across frequency as well as the real-valued spherical harmonic functions as used in Higher Order Ambisonics.

This simulator provides realistic phase information for the sound recorded by the receiver or receiver array. For a single channel receiver, this means that the phase of an idealised or custom pattern is considered, such as the positive and negative lobes of a dipole. For a multichannel receiver array, the simulator constructs accurate inter-sensor time delays for sound arriving from the source(s) and early specular reflections. The property of the diffuse sound field being uncorrelated between different sensors of a multichannel receiver array is also accurately represented.

MCROOMSIM is designed to interface with MATLAB. The simulator receives all setup data including directional gains or impulse responses for custom sources/receivers from MATLAB. Once the simulation has completed, MCROOMSIM returns the simulated room impulse responses for all configured source and receiver combinations back to the MATLAB workspace. A set of high level MATLAB functions are provided to configure the simulator, making it easier for the user to setup the room, sources and receivers with specific properties, as well as configuring specific properties of the simulation routines.

Room Simulation Method

The MCROOMSIM software was originally inspired by the software developed by Schimmel et al [1] but has been completely re-written to enable complex operations in the frequency domain. This allows the phase of a source or receiver's directional response to be included during simulations, thus providing a much more accurate representation of the source or receiver in echoic environments than was previously possible. Support for arbitrary multichannel receivers with user defined directional gains or directional impulse responses was also added.

Impulse responses are simulated between each source and receiver combination configured in the echoic room. MCROOMSIM uses two geometrical acoustic simulation algorithms to generate the impulse response between a source and receiver, with each algorithm simulating a different part of the reverberant sound field. The main or specular reflections are simulated using the image source algorithm [5] and the diffuse sound field is simulated using the diffuse rain algorithm [6, 7]. The output from both algorithms are combined to make the final output response.

Image Source Algorithm

The image source algorithm provides accurate direction and timing of the main reflections. The algorithm starts with the zeroth order sound which corresponds to the direct sound. For first order reflections, the sound source is mirrored in the room's surfaces to create a series of virtual sources, which are then traced back to the receiver along with the inclusion of distance attenuation and air absorption. These virtual sources are then expanded to create more virtual sources of higher order. By knowing the distance of the virtual source from the receiver and absorption of the surfaces that it reflects off, the contribution of the corresponding sound propagation path to the room impulse response can be calculated exactly. This contribution and the receiver's directional response are then convolved, with the output being the room response to that propagation path as 'recorded' by the receiver. The image source algorithm continues to expand the order of virtual sources up to an maximum order set by the user or ceases when the energy of all virtual sources of the same order drop below a predefined threshold.

Diffuse Rain Algorithm

The diffuse rain algorithm is a fast stochastic ray tracing method. It aims to provide a good approximation of the diffuse sound field in the room. This simulation technique models the propagation of sound from the source to the receiver by a series of discrete rays that are traced around the room. Each ray trajectory is traced and is reflected in a random direction every time it hits a wall. A ray's energy decreases in time due to wall absorption, air absorption and distance attenuation. The process of tracing a ray is continued until the ray's energy falls below a predefined threshold.

Whenever a ray hits a wall, its remaining energy after wall absorption is traced back to each receiver in the room. At each receiver location, the ray's energy (distributed over frequency), its time of arrival and angle of incidence are recorded in a time frequency space histogram. The final result of the ray tracing process is a directional time-frequency-energy map of the room's diffuse sound field at each receiver location.

A diffuse response for each receiver is constructed by generating a random set of impulses, corresponding to diffuse reflections for each directional group of the time-frequency energy map. As in [7] the rate of reflections is fixed to 10000 reflections per second for the overall diffuse output signal. For each impulse within a directional group, the time-frequency-energy map is used to interpolate an energy-frequency response at that time instance. This energy-frequency response is then used to generate a minimum-phase diffuse impulse response that is convolved with the directional responses of the receiver which are nearest to the current directional group. The result is then added to the overall diffuse response.

MATLAB Interface

MCROOMSIM is configured using a group of high level MATLAB functions that ease the setting of parameters for simulation. The parameters comprise four main parts:

1. Room setup: physical characteristics of the room such as

the size, frequency dependent absorption and scattering coefficients of the walls/ceilings, room temperature and humidity.

2. General simulation options: here the user is provided with an opportunity to configure various features of the simulator such as the maximum order that the image source algorithm will iterate up to, the minimum energy threshold for virtual image sources or rays, etc.
3. Receiver setup: the number of receivers to simulate, their locations, orientation, number of channels and directivity. For receivers with custom directivity, a list of directional gains or impulse responses along with the matching direction list is also included here.
4. Source setup: same as the receiver setup with the limitation that sources can only be single channel.

Once the MCROOMSIM function is invoked, MATLAB combines all of the data, including all user-defined directional responses (if provided), into a single structure which it then passes to MCROOMSIM. Once the simulation has completed, all of the room impulse response data are provided as output to the MATLAB workspace. The output of MCROOMSIM is a time domain room impulse response for each source and receiver combination. In the case of a multichannel receiver array, a separate response is provided for each channel.

Flow Of Software

Once the input configuration data is received by MCROOMSIM, the simulator first performs a validity check on the parameters to ensure that the configuration is not erroneous. Once this has completed successfully, the simulator uses these parameters to configure itself. All operations are performed in the frequency domain, which improves efficiency of convolutions. Therefore if a source or receiver has custom directivity defined with directional impulse responses, MCROOMSIM transforms these responses into the frequency domain. Next the simulator estimates the maximum possible length for the room impulse response by calculating the longest time for the virtual sources to decay below the energy threshold, as set in the simulation general options. This is performed so that simulation memory can be pre-allocated.

MCROOMSIM can be configured to run in one of three ways: specular only simulation, diffuse only simulation or both. The simulation algorithms are then executed to generate the room impulse responses between the defined sources and receivers. Each algorithm outputs its results in the time domain, so there is no need for transformation. If both algorithms have been executed then the outputs of each will be combined to create the overall room impulse responses. Upon completion, the resulting simulated room impulse responses are provided as output to the MATLAB workspace. Fig. 1 highlights the flow of the software from configuration to the end of simulation.

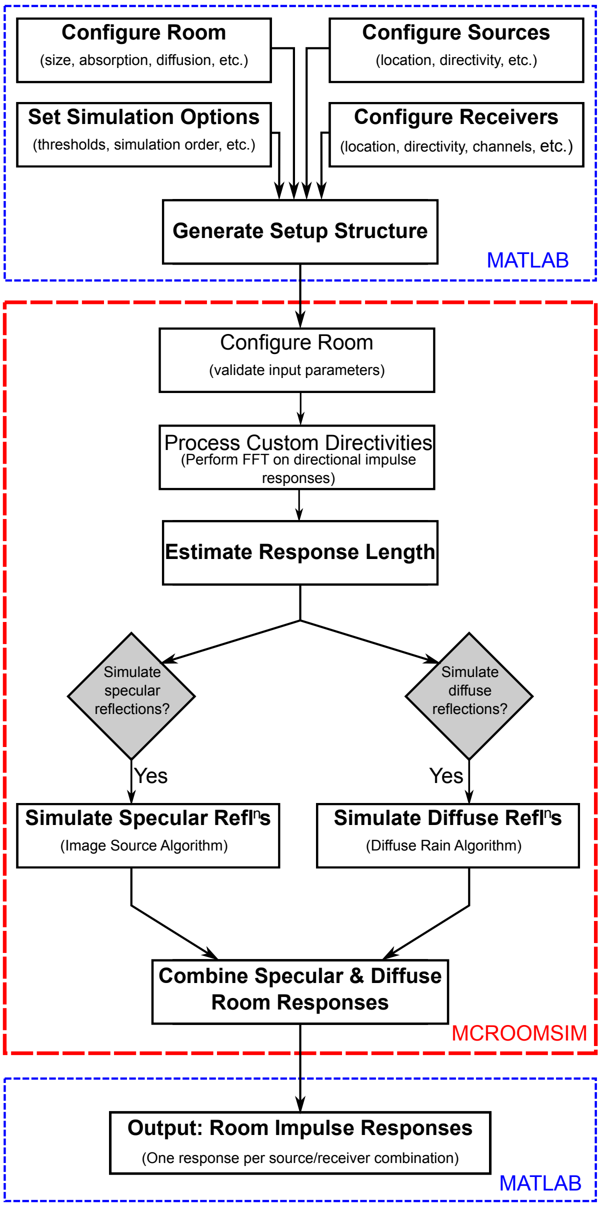


Figure 1: The flow of the MCROOMSIM software is shown from the configuration stage to the end of simulation, where the simulated room impulse responses are returned to the MATLAB workspace.

EXAMPLE APPLICATIONS OF THE SOFTWARE

Realistic in-room microphone array simulation

Microphone arrays are used in many three-dimensional audio applications. Some examples include spatial sound reproduction [2], measurement and analysis of directional properties of reverberant sound fields using beamforming techniques [8] or plane wave decomposition [9]. In all of these situations MCROOMSIM can be used to generate realistic three-dimensional room impulse responses for the microphone array under test. All that is needed are the anechoic directional impulses responses of the microphone array and the properties of the room under simulation, i.e. size, location of source(s), location of microphone array, etc. Moreover, sources with specific directivity patterns can also be simulated in the system, enabling the user to compare the effects of various sources on the reverberant sound field.

3D Binaural rendering of echoic sound scenes

MCROOMSIM can be used to directly simulate realistic Binaural Room Impulse Responses (BRIRs). This is done by configuring the simulation software with a pair of sensors whose directional impulse responses are anechoic head related impulse responses (HRIRs) measured for a number of directions in the free-field for the particular user. This is useful for quickly generating 3D BRIRs for many environments, without the need of performing time consuming recordings within each environment. From this the user could then for instance listen to the 3D reverberant sound fields created by sources with different directivities.

Spherical harmonic expansion of echoic sound fields

The spherical harmonic formalism is a powerful tool for describing the spatial properties of sound fields. An example of technique based on this formalism is the Higher Order Ambisonics method for spatial sound scene recording and reproduction. MCRoomSim can be used to calculate a far-field approximation of the spherical harmonic expansion for the sound field emitted by a source in a simple echoic room. This is done by configuring MCRoomSim with an array of co-located receivers whose directional gains are the values of the spherical harmonic functions up to a given order. The obtained spherical harmonic expansion can then easily be used for auralising rooms over an array of loudspeakers, as done in the LoRA auralisation system [10].

SOFTWARE VALIDATION

In this section we demonstrate how MCRoomSim can be used to produce a realistic room impulse response for a cylindrical microphone array. We compare the results of a real room impulse response measured with the cylindrical array to the room impulse response obtained from simulating the same microphone array using MCRoomSim.

Measurements

A small empty shoebox room with dimensions 4.6 x 2.9 x 2.6 m was used for measurement. Fig. 2 indicates the locations of the sound source and recording device in the room and Table 1 shows the materials that each of the surfaces of the room were made of.

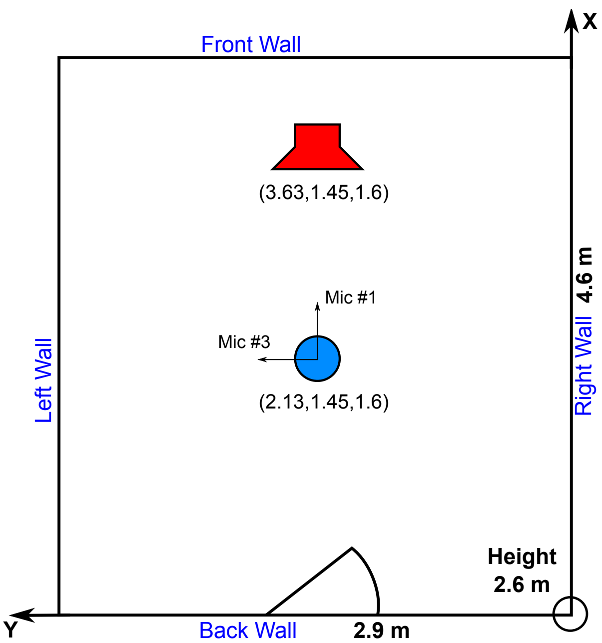


Figure 2: The dimensions of the room along with the location and orientation of the speaker and microphone are shown. The view is from the top.

Table 1: The material(s) for each surface of the room.

Surface	Material
Front wall	Brick, Window & Tontine sound batt
Back wall	Plasterboard & Wooden door
Left wall	Plasterboard
Right wall	Plasterboard
Floor	Carpet tiles on concrete
Ceiling	Plasterboard

A cylindrical microphone array with 12 sensors was used as the recording device. This device is shown in Fig. 3. The array consisted of 12 DPA 4060-BM omni-directional microphones mounted in an equally spaced arrangement around a rigid cylinder. The cylinder had a radius of 1.8 cm and a height of 11.6 cm. The ring of microphones were located 6.2 cm from the top of the cylinder. The array was oriented such that the microphones were located 1.6 m from the floor, with microphone channel 1 facing towards the speaker and microphone channels were increasing anticlockwise around the cylinder, if viewed from the top.

Figure 3: This figure shows a photo of the 12 sensor cylindrical microphone array used for the room measurements.

The sound source was a Tannoy V6 speaker. The centre of the speaker cone was located 1.6 m from the floor. The stimulus played was a log sine sweep, 12 seconds in length.

Signals from the 12 microphones were amplified by two 8-channel Digidesign PRE preamplifiers. The output from the preamplifiers is then digitised by a 16-channel Apogee AD-16X analogue-to-digital converter, with a sample rate of 48 kHz and 24-bit resolution. The output of the converter is in ADAT format and is sent to an RME Digiface sound card, which is connected to a standard personal computer (PC) running Steinberg Cubase software. At the same time as recording the signals from the microphones the PC software also outputs the stimulus signal to be played in the room. This stimulus is fed out of the RME Digiface sound card in ADAT format to a 16-channel Apogee DA-16X digital-to-analogue converter. The analog output from the convertor is then amplified by a 6-channel Ashley Powerlex 6250 power amplifier which powers the Tannoy V6 speaker.

Simulations

The room setup in MCRoomSim was configured to have properties matching that of the measurement room. Absorption coefficients matching each the room's surfaces were obtained from an absorption coefficient chart for standard building materials [11]. The absorption coefficients for the Tontine sound batts were obtained from [12]. Where surfaces are comprised of multiple materials, we averaged together the absorption coefficients for each of the materials. Table 2 shows the absorption coefficients used to model the measurement room.

Table 2: The absorption coefficients over frequency for each surface of the room.

	Frequency (Hz)					
	125	250	500	1000	2000	4000
Front wall	0.22	0.09	0.06	0.05	0.04	0.04
Back wall	0.17	0.30	0.11	0.14	0.24	0.27
Left wall	0.29	0.10	0.06	0.05	0.04	0.04
Right wall	0.29	0.10	0.06	0.05	0.04	0.04
Floor	0.01	0.02	0.06	0.15	0.20	0.25
Ceiling	0.29	0.10	0.06	0.05	0.04	0.04

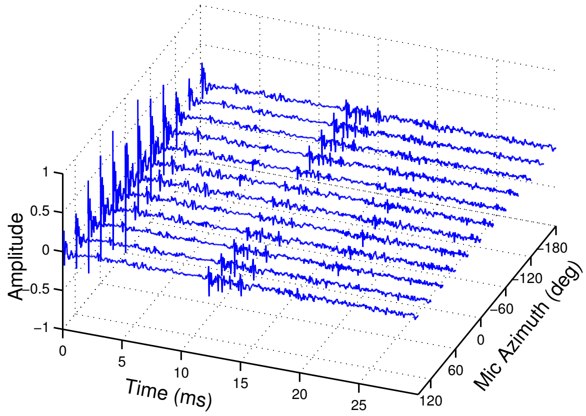
Unlike the absorption coefficients, we had no data for the diffuse scattering coefficients of the room's materials. However, as the room was empty, except for the measurement equipment (speaker, microphone array, stands, etc.), and the walls were made of smooth materials, it was assumed that there were very little diffuse reflections present in the room. Therefore we set the scattering coefficients globally at all low to mid frequencies such that 10% of all surface reflections were diffuse. We also made the assumption that the interaction of the sound waves with the measurement equipment would cause more diffuse reflections to occur at higher frequencies, hence we set the coefficients at the frequencies 1 kHz, 2 kHz and 4 kHz such that 15% of all reflections were diffuse. Table 3 shows the scattering coefficients used to model the measurement room.

Table 3: The scattering coefficients over frequency for each surface of the room.

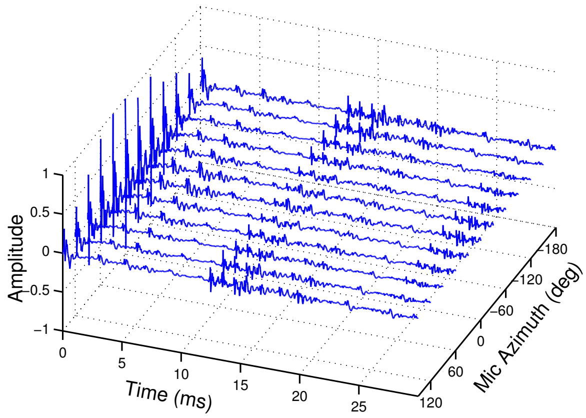
	Frequency (Hz)					
	125	250	500	1000	2000	4000
Front wall	0.10	0.10	0.10	0.15	0.15	0.15
Back wall	0.10	0.10	0.10	0.15	0.15	0.15
Left wall	0.10	0.10	0.10	0.15	0.15	0.15
Right wall	0.10	0.10	0.10	0.15	0.15	0.15
Floor	0.10	0.10	0.10	0.15	0.15	0.15
Ceiling	0.10	0.10	0.10	0.15	0.15	0.15

The microphone array was configured in MCRoomSim as a 12 channel receiver array with custom directional impulse responses. Anechoic impulse responses for the array were measured in an anechoic chamber from 393 directions, forming a sphere around the array. We could have instead used anechoic impulse responses of the microphone array obtained via a theoretical model of the array. This is useful in situations such as prototyping an array and comparing its performance in different reverberant environments.

In MCRoomSim we were also able to model the directional characteristics of the Tannoy V6 speaker used for playback. Using the directivity plots for the speaker that were obtained from Tannoy [13], we constructed a model of the speaker with directional impulse responses for 642 angular directions evenly distributed in space. The speaker was then configured in MCRoomSim as a source with custom directional impulse responses.



(a) Measured room impulse responses



(b) Simulated room impulse responses

Figure 4: This figure shows the first 30 ms of the measured (a) and simulated (b) room impulse responses for each sensor of the microphone array, highlighting the early reflections.

The time taken to perform the simulation on a laptop equipped with a 2.2 Ghz Intel Core 2 Processor and 4 GB of RAM was 110 seconds. Most of the simulation time was taken by the image source algorithm which iterated up to order 35 image sources. In most cases a lower image source order may be used without sacrificing the accuracy of the simulated results. Therefore the simulation time could have been significantly reduced by lowering the maximum order that the image source algorithm iterates up to. This is one of many options that can be configured by the user in MCRoomSim.

Results

Comparison Of The Impulse Responses

From examining the early part of the measured and simulated room impulse responses shown in Figs. 4(a) and 4(b), it can be seen that MCRoomSim accurately recreates the main reflections among all microphones of the array, both in time of arrival and in amplitude. For instance, the measured room impulse response shows that at 14 ms the reflection is primarily due to the sound reflecting off the back wall of the room. This is evident by the fact that microphone 1, oriented toward the front wall, has the smallest amplitude. While microphones 6, 7 and 8 have the largest amplitude as they are nearest the back wall. This feature of the room impulse response has also been recreated in the simulation.

In the measured room impulse response, there are a series of small reflections that arrive at approximately 22 ms for the microphones orientated towards the front of the room (microphones 1, 2, 3, 11 and 12). These reflections are due to sound travelling from the back of the room and then reflecting off the speaker, speaker stand and other measurement equipment located near the speaker. As MCROOMSIM cannot simulate reflections bouncing back from objects placed in the room, these reflections are not recreated in the simulated results.

Comparison Of The Energy Decay Curves

Schroeder's integrated energy decay curve [14] is traditionally used for evaluation of the reverberant properties in a room, such as the reverberation time. The Schroeder curve can be estimated by the formula:

$$E(n) = \sum_{i=n}^{N} h(i)^2 \quad , \qquad (1)$$

where n is the sample index, $h(i)$ is the room impulse response at sample index i and N is the total number of samples in h. Schroeder curves were evaluated to examine the rate of decay in energy for both the measured and simulated room impulse responses. Fig. 5 shows the comparison between the Schroeder curves for the measured and simulated room impulse responses of microphone array sensor 1.

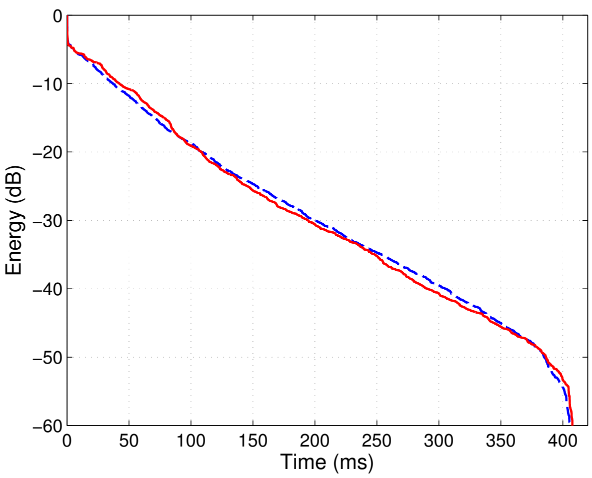


Figure 5: This figure shows the decay in energy over time for measured and simulated impulse responses. The dashed line represents the measured data and the solid line represents the simulated data. The results are from sensor 1 of the microphone array.

From observation of the energy decay curves in Fig. 5, it is clear that MCROOMSIM simulates with reasonable accuracy the decay in energy of the room reverberation.

CONCLUSION

We have presented a multichannel room simulation software for MATLAB that can simulate recordings of arbitrary microphone arrays within an echoic shoebox room. We have provided some example three-dimensional audio applications for the simulator.

Using a 12 sensor cylindrical microphone array, we recorded the impulse response of an empty shoebox room. We then used MCROOMSIM to create a simulated impulse response of this room with the 12 sensor cylindrical microphone array. A comparison of the measured and simulated results shows that MCROOMSIM accurately recreates the main features of the three-dimensional room impulse response.

REFERENCES

1 S. M. Schimmel, M. F. Müller, and N. Dillier. A fast and accurate "shoebox" room acoustics simulator. In *Proceedings of ICASSP*, Taipei, Taiwan, April 2009.

2 A. Parthy, C. Jin, and A. van Schaik. Evaluation of a concentric rigid and open spherical microphone array for sound reproduction. In *Proceedings of Ambisonics Symposium 2009*, Graz, Austria, 25-27 June 2009.

3 Odeon Room Acoustics Modelling Software. URL `http://www.odeon.dk`.

4 Enhanced Acoustics Simulator for Engineers (EASE). URL `http://www.auralisation.de`.

5 J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, 65(4):943–950, 1979.

6 R. Heinz. Binaural room simulation based on the image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail. *Applied Acoustics*, 38:145–159, 1993.

7 D. Schröder, P. Dross, and M. Vorländer. A fast reverberation estimator for virtual environments. In *Proceedings of the AES 30th International Conference*, Saariselkä, Finland, March 2007.

8 B. Gover, J. G. Ryan, and M. R. Stinson. Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array. *The Journal of the Acoustical Society of America*, 116(4):2138–2148, 2004.

9 M. Park and B. Rafaely. Sound-field analysis by plane-wave decomposition using spherical microphone array. *The Journal of the Acoustical Society of America*, 118(5):3094–3103, 2005.

10 S. Favrot and J.M. Buchholz. LoRA: A loudspeaker-based room auralization system. *Acta Acustica united with Acustica*, 96:364–375, 2010.

11 Absorption coefficients of common building materials and finishes. URL `http://www.sae.edu/reference_material/pages/Coefficient%20Chart.htm`.

12 Tontine sound batts Acoustic Design Manual. URL `http://www.spec-net.com.au/tontine/acoustic/designm5.htm#coeff`.

13 Tannoy V6 Specifications & Measurements. URL `http://www.tannoy.com/Downloads.aspx`.

14 M. R. Schroeder. New method of measuring reverberation time. *The Journal of the Acoustical Society of America*, 37 (3):409–412, 1965.