Roomsim, a MATLAB Simulation of "Shoebox" Room Acoustics for use in Teaching and Research

Douglas R. Campbell[1], Kalle J. Palomäki[2] and Guy J. Brown[3]

[1]School of Information and Communication Technologies, University of Paisley,
High Street, Paisley PA1 2BE, Scotland, U.K. email: d.r.campbell@paisley.ac.uk
[2]Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing,
P.O. Box 3000, FIN-02015 HUT, Finland. email: kalle.palomaki@hut.fi
[3]Department of Computer Science, University of Sheffield,
211 Portobello Street, Sheffield S1 4DP, U.K. email: g.brown@dcs.shef.ac.uk

ABSTRACT

A simulation of the acoustics of a simple rectangular prism room has been constructed using the MATLAB m-code programming language. The aim in creating the program Roomsim was to provide a signal generation tool for the speech and hearing research community, and a teaching tool for illustrating the image method of simulating room acoustics and some acoustical effects. The program is menu driven, has a user guide and example data files, and will be made freely available over the Internet under a GNU General Public Licence.

1. INTRODUCTION

The image method of simulating room acoustics is often used to provide a means of generating signals incorporating "sufficiently realistic" reverberation and directional cues for the testing of audio/speech processing algorithms, or the demonstration of room acoustic properties [1]. Commercially available programs for simulating architectural acoustics can be used, but the cost and degree of geometric and acoustic realism is usually an overkill for such purposes. Many signal processing researchers now use the MATLAB technical computing language to develop their algorithms because of its ease of use, powerful library functions and convenient visualisation tools.

The Roomsim program is a simulation of the acoustics of a simple rectangular prism "shoebox" room, constructed using the MATLAB m-code programming language. The foundation on which Roomsim is built is the publication of a Fortran routine by Allen and Berkley [2] in 1979. Others have translated this core into various computer languages for use in general simulations of "shoebox" room acoustics. Roomsim uses a MATLAB translation of the core acoustic image calculation that has been comprehensively extended and features a menu driven system with various graphical displays and utilities. Our belief is that this freely available MATLAB implementation will be of advantage to digital signal processing (DSP) researchers and educators working in areas of acoustic signal processing, speech and hearing, because of their familiarity with what has arguably become the prime DSP algorithm development platform.

The authors had separately searched for a MATLAB implementation of room acoustics that would satisfy their experimental requirements for quickly and easily generating binaural speech signals sufficiently realistic for the demonstration and assessment of signal processing algorithms aimed at improving the performance of speech recognisers, hands-free telephones and hearing aids. A few candidate programs were found, but none incorporated the majority of features required. Palomäki et al. created a MATLAB program to satisfy the immediate needs of a research project [3]. Following the presentation of their results in 2002 at a Binaural Hearing workshop, Campbell was supplied with code by Palomäki et al. and used that as the core of the Roomsim program reported here.

The program will be released as free software, under the GNU General Public Licence, to encourage its dissemination and development. The m-code source files, for execution within MATLAB on a PC or Unix platform, and a compiled stand-alone executable for non-MATLAB Windows PC installations, will be posted on an appropriate web site e.g. the user contributed programs library at MATLAB Central. The existence and location of the code will be more widely disseminated via notices to various special interest groups in signal processing, binaural hearing, speech processing, audio processing and acoustics.

This paper consists of two main parts: a description of features and facilities of the Roomsim program and examples of educational use.

2. THE ROOMSIM PROGRAM

The program simulates the geometrical acoustics of a perfect rectangular parallelepiped enclosure using the image-source model to produce an impulse response from each omni-directional primary source to a directional receiver system that may be a single sensor, a sensor pair or a simulated head. The generation of an impulse response from a primary source to a receiver sensor by the method of image sources is very adequately described in reference [2] and will not be repeated here.

One of the extensions made to the Allen and Berkley algorithm is the incorporation of frequency dependent sound absorption coefficients. Using this, the Roomsim program provides the ability to select materials for each of the six major surfaces of the simulated room from a list of standard building materials that have had their frequency dependent absorption coefficients tabulated. Surface reflectances of 100%, 50% and 0% (anechoic) are also provided for test purposes.

Another desired feature was the ability to incorporate a range of receiver systems such as single sensor (e.g. mono microphone), sensor pair (e.g. two element microphone array) and simulated human head. The single and dual sensor receivers can be configured as directionally sensitive and the interpolation process recommended by Peterson [4] is incorporated. The simulation of a head utilises the Head Related Transfer Function (HRTF) data, actually Head Related Impulse Response (HRIR) data, provided from measurements made either on a Kemar mannequin at the MIT Media Lab. [5], or on real human subjects and a Kemar at the Center for Image Processing and Integrated Computing (CIPIC), University of California, Davis [6]. Both MIT Media Lab and CIPIC have agreed to the distribution of their data within the Roomsim package. All sensor systems may be oriented in the 3D space by the user.

A core aim was to provide user interaction having a suitable balance between ease of use for undergraduate students and flexibility for researchers. The m-code has been written to: (1) allow operation on PC and Unix platforms; (2) avoid the need for specialist MATLAB toolboxes; (3) avoid functions that cannot successfully be compiled to produce stand-alone executable code. Thus users need have only the basic MATLAB installation to run and develop the m-code, and users without a MATLAB installation may run the executable version on a Windows PC.

A comprehensive user guide, test data files, and example text and Excel set-up data files are provided. A typical Roomsim installation for running under MATLAB will require ~300MB of disk space of which about half is for the full CIPIC HRTF data (this may be reduced). The executable version requires ~100MB for Roomsim, its data files, three HRTFs and the MATLAB run-time libraries.

The program was developed on a medium specification PC (1.5 GHz Intel Pentium4, 512 MB RDRAM, 30GB HDD UDMA 100 IDE, AGP Graphics) running Windows 2000 with MATLAB v 6.5 rev. 13 installed. It has been successfully run under Windows 2000, Windows 98, Red Hat Linux and Mandrake Linux, and on desktop and notebook PC's down to 200 MHz clock frequency with 128 MB RAM. Reflection orders >10, impulse responses > 10,000 samples and multiple sources will all contribute to a large computational load resulting in slow response especially of the core impulse response calculation and the 2D and 3D graphics. The faster the CPU, memory, disk access and graphics, the better will be the response time.

2.1. Roomsim Operation

In operation the user specifies the dimensions of the room, its surface materials the type, location and orientation of the receiver system and the location of the primary source(s). This can be done interactively through the menu prompt system, or by submitting either a Microsoft Excel spreadsheet form, a text file, or by selecting a MATLAB *.mat file which saved a configuration from a previous run. The program performs various "sanity" checks and an estimate of the reverberation time (RT60) of the simulated room can be used to size the impulse response length to be computed.

The image-source to receiver responses are then computed using the method of images, modified to take account of the variation of surface absorption coefficient with frequency and for the attenuation due to acoustic path length. The frequency dependent attenuation due to air is included if desired. If a simulated head has been selected the response from each quantised image-source direction is convolved with the relevant HRIR data. The individual image-source responses are then accumulated to form the complete pressure impulse response from each primary source to the receiver and the results plotted and saved to file.

These two-channel impulse response file(s) can then be convolved within the program with the users' own monophonic audio files (*.wav, *.au or *.mat format). The resulting monaural, stereo, or "binaural response" can be saved as an audio or MATLAB file. If saved in *.wav or *.au format it can then be played using a compatible media player or sound editor. The user may elect to create a "cocktail party" by combining the reverberated response files related to each simulated primary source to produce the combined acoustic signal at each sensor/ear from speech and noise signals at different locations.

2.2. User Configurable Parameters and Facilities

User configurable parameters can be submitted as text files, Excel worksheets, or by prompted manual input. Previous scenarios may be saved as *.mat files for efficient repeat operations. Example files are provided.

The user configurable parameters are: Sampling frequency (at present 8 kHz to 44.1 kHz); Humidity of air (modifies air absorption coefficient); Temperature of air (modifies the speed of sound); Order of reflections calculated; Length of room impulse response; Air absorption model (ON or OFF); Effect of distance (ON or OFF); High-pass filter cut-off (for reduction of DC bias and low frequency ripple); Smoothing filter (Used for interpolation in sensor pair case); 2D and 3D plotting; Enclosure dimensions (Lx, Ly, Lz); Receiver coordinates (x, y, z); Receiver type (single sensor, sensor pair, HRTF); Separation of sensor pair; Sensor directionality (azimuth and elevation); HRTF from MIT Kemar or CIPIC subject (Kemar plus 42 human subjects); Receiver orientation (Yaw, Pitch and Roll); Source location(s) specified as polar with respect to Receiver origin; Surface absorption for each of the room's six surfaces (at present 24 frequency dependent surface types).

The display options are: 3D display of room, receiver and sources geometry; Plot of Surface Absorption vs. Frequency; Plot of Mean Reverberation Time (RT60) vs. frequency; Plot of Impulse Response(s) vs. Time or Sample number (Colour coded for left and right sensor); Magnitude Spectrum corresponding to above impulse response(s) (FFT length selectable and HRTF superimposed when MIT or

CIPIC data selected); 2D view of constant height slice through room, showing receiver, source(s), surrounding image rooms, and image sources (source intensity is indicated by colour code or stem plot height; 3D view of room, receiver, source(s), and image sources, with source intensity indicated by colour code. All displays can be zoomed and 3D displays rotated.

The processing utilities options are: An operation for building “cocktail party” scenarios; A Convolution operation that accepts Roomsim audio files, and displays the input file data and convolution result as time histories and as spectrograms; An Accumulation function for adding audio data files (Signal displays are provided); A Converter between the supported file formats (and one or two channels); A Display operation that plots an audio data file, plots its spectrogram and plays it; A Filtering operation using user supplied coefficients.

3. EDUCATIONAL USE

The Roomsim program can be of value to educators delivering courses or modules in acoustic signal processing, music technology, auditory perception, psychoacoustics, environmental acoustics, and digital signal processing. It can form the basis for standard demonstrations, exploratory classroom laboratories or undergraduate projects. For example, since the user has control of surface absorption, the image source method can be effectively illustrated by adding room surfaces one at a time and controlling the order of reflections displayed. The effect of surface and air absorption can be demonstrated and their dependence on room area, volume and air humidity. The simulation may be used as a reverberation chamber for generating audio effects or demonstrating the acoustic complexity of reverberant enclosures. By placing a sensor close to a reflective surface the acoustic comb filter effect can be demonstrated and the expected notch frequencies confirmed in the spectral magnitude plot. The difference in sound quality and localisation between a stereo microphone pair and a binaural “head” can be demonstrated. The 3D displays of room geometry, receiver and primary source locations, image rooms and image sources (location and strength) are powerful visualisations of the complexity of the reverberant environment and can be rotated and zoomed to reveal details.

An example follows of Roomsim's usage for teaching several fundamental issues in room acoustics, such as the linear decay of the room impulse response in dB, and estimation of RT60 reverberation time. The receiver is the MIT Kemar located in a “shoebox” room in which two sound sources are simultaneously active. The two sources are situated at azimuths of -10 degrees and -60 degrees relative to the listener, and at zero elevation.

Figure 1 displays the absorption characteristics of the surface materials as a function of frequency. Such a plot can be used to demonstrate the way in which the absorption properties of wall materials vary with frequency. The point can be made that this is still a simplification, as direction dependency of the absorption and the diffuse component of the reflection are neglected in the current version of Roomsim. The likely result of including these factors in the simulation would be a suitable tutorial question.

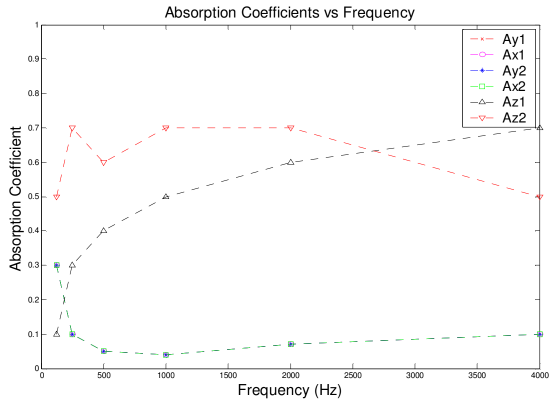

Figure 1: Absorption coefficients of the room surfaces

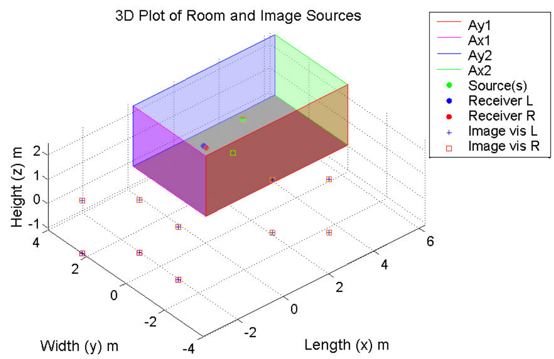

Figure 2: Room simulation with the MIT Kemar, two active simultaneous sound sources, and low order mirror images.

Figure 2 shows the display of geometry and the image sources corresponding to a subset of the first order wall reflections between both sources and the ears of the listener. Exemplifying room reverberation and reflections as mirrored sources is possibly one of the most powerful visual demonstrations of room acoustics, and could lead to further pedagogical demonstrations (e.g., demonstrating that speakers placed in close proximity to a wall generate reflections of the source that interfere with stereo imaging).

This simulation produced four room impulse responses (RIRs), plotted in dB in Figure 3, corresponding to the paths and reflections between the left/right ear and source 1 (top panel), the left/right ear and source 2 (bottom panel). Figure 4 shows the corresponding magnitude spectra. The impulse response plot also makes a useful pedagogical point, since it demonstrates several characteristics of a room impulse response such as the linear decay of the response in dB-domain. The RT60 reverberation time can also be estimated from this figure. An alternative estimate of reverberation time as a function of frequency is provided as shown in Figure 5. A suitable tutorial question could be to ask the student to apply a standard formula for reverberation time, given the dimensions of the room, and compare their answer against an estimate of RT60 from Schroeder's energy decay plot also provided in Roomsim.

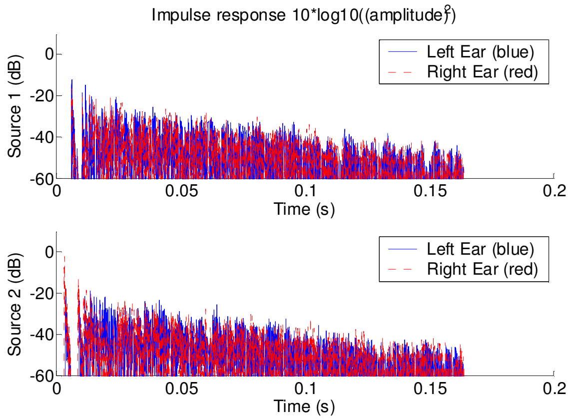

Figure 3: Impulse responses of the simulated room.

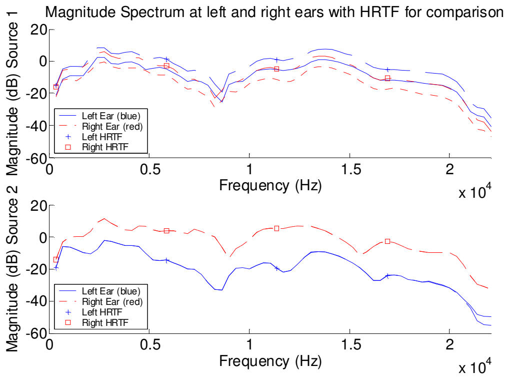

Figure 4: Magnitude spectra at each ear

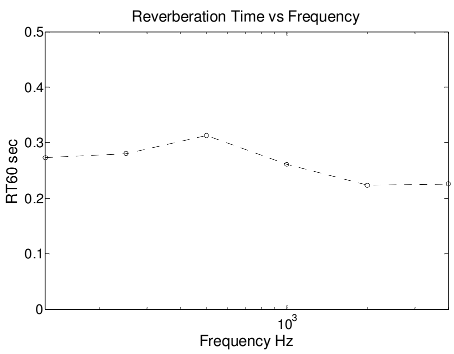

Figure 5:RT60 as a function of frequency (Sabine estimate).

When the simulation is finished, the room impulse responses are directly available for convolution with sound samples within Roomsim. The effect of different simulation parameters could then be assessed in listening tests. Students might record their subjective impressions and explain how these correspond to the simulation parameters e.g. a 'live' sounding room in which the surfaces have low absorption at high frequency. Time domain plots and spectrograms of speech and audio samples provide visualisations to complement the audio examples.

An example display from a more complex scenario with three omni-directional audio sources is given in Figure 6. This shows the images audible to a pair of simulated microphones one of which is uni-directional and the other is bi-directional.

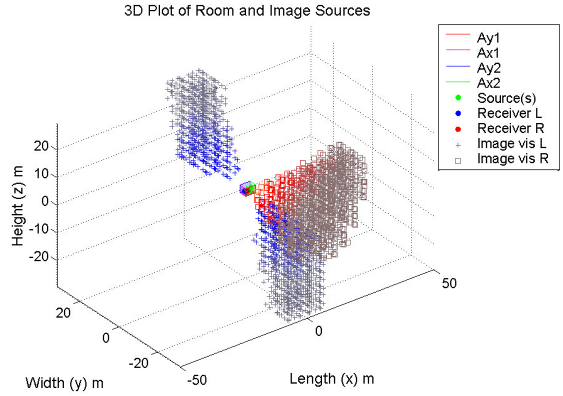

Figure 6: Room geometry and those images audible to a pair of microphones one of which is uni-directional and the other is bi-directional.

4. CONCLUSION

The Roomsim program provides an educational aid of interest to those teaching in areas of acoustic signal processing, music technology, auditory perception, psychoacoustics, environmental acoustics, and digital signal processing. It also is an effective tool for researchers working in areas of acoustic signal processing, speech and hearing, who require to generate binaural audio signals sufficiently realistic for the assessment of signal processing algorithms aimed at improving the performance of e.g. speech recognisers, hands-free telephones and hearing aids.

5. ACKNOWLEDGEMENTS

The authors acknowledge support from the following: D.R. Campbell; Defeating Deafness (246:PAI:DC). K. Palomäki; EC TMR SPHEAR, The Academy of Finland (1277811). G.J. Brown; EPSRC GR/R47400/01. The authors are grateful for permission to include in Roomsim the HRTF data from CIPIC at the University of California [7] and the MIT Media Laboratory [8].

6. REFERENCES

[1] Lokki, T., "Physically-based auralization" PhD Thesis, Publications in Telecommunications Software and Multimedia. Helsinki University of Technology, 2002.
[2] Allen, J. B. and Berkley, D. A., "Image method for efficiently simulating small-room acoustics", JASA 65(4), 943-950, 1979.
[3] Palomäki, K. J., Brown G. J., and Wang D. L., "A Binaural Processor for Missing Data Speech Recognition in the Presence of Noise and Small-Room Reverberation", submitted to Speech Communication, 2003.
[4] Peterson, P. M., "Simulating the response of multiple microphones to a single acoustic source in a reverberant room", JASA 80(5), 1527-1529, 1986.
[5] Gardner, W. G., and Martin, K. D., "HRTF measurements of a KEMAR dummy head microphone", In Standards in Computer Generated Music, Haus, G. and Pighi, I. eds., IEEE CS Tech. Com. on Computer Generated Music, 1996.
[6] Algazi, V. R. *et al.*, "The CIPIC HRTF Database", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2001.
[7] http://interface.cipic.ucdavis.edu, January 2004.
[8] http://xenia.media.mit.edu/~kdm/hrtf.html, January 2004.