Audio Engineering Society

Convention e-Brief 47

Presented at the 142nd Convention
2017 May 20–23, Berlin, Germany

An impulse response dataset for dynamic data-based auralisation of advanced sound systems

Chris Pike[1,2] and Michael Romanov[3,4]

[1]*BBC Research & Development*
[2]*Audio Lab, University of York*
[3]*University of Music and Performing Arts Graz*
[4]*Graz University of Technology*

Correspondence should be addressed to Chris Pike (`chris.pike@bbc.co.uk`)

ABSTRACT

This engineering brief presents a freely-available binaural room impulse response (BRIR) dataset measured on a multichannel loudspeaker system. The 32-loudspeaker array includes all loudspeaker layouts specified in Recommendation ITU-R BS.2051. Measurements were carried out in an ITU-R BS.1116-compliant listening room using a Neumann KU100 dummy head microphone. BRIRs were measured at 2° steps of rotation of the dummy head. The dataset can be used for dynamic data-based auralisation of multichannel loudspeaker signals, such as those generated by the so-called advanced sound systems described in ITU-R BS.2051, i.e. systems that can render surround sound with height signals from channel-based, object-based and/or scene-based content representations. The dataset is made freely-available in the SOFA file format.

1 Introduction

In development and evaluation of spatial audio systems for loudspeaker reproduction, it is often useful to auralise the loudspeaker signals on headphones, using binaural techniques to create a virtual acoustic environment. Loudspeaker systems for spatial audio (surround with height) are expensive and often site-specific. Dynamic data-based auralisation of loudspeaker signals uses binaural impulse responses measured in a reverberant environment, so-called binaural room impulse responses (BRIRs). This process is also called binaural room scanning [1]. It requires BRIR measurements are available for multiple different orientations of the head and use of a head-tracking system in real-time rendering. It can create plausible simulations of real loudspeaker systems [2] and so can be a valuable tool in development of spatial audio systems for loudspeakers, allowing assessment of systems or loudspeaker signals without requiring access to a real loudspeaker setup [3].

Recommendation ITU-R BS.2051 [4] describes a so-called “advanced sound system”, which can handle channel-based, object-based and scene-based (higher-order ambisonics) representations of audio content and can generate signals for a set of 8 loudspeaker layouts, many of which include loudspeakers at different heights.

The listening room at BBC R&D in Salford [5] has 32 full-range loudspeakers arranged to cover the layouts described in ITU-R BS.2051, as well as layouts with loudspeakers arranged in regular octagons and squares in each vertical layer. This room meets the acoustic criteria specified in Recommendation ITU-R BS.1116 [6].

This engineering brief presents a dataset of BRIRs measured in this listening room and therefore suitable for dynamic data-based auralisation of advanced sound systems. Subjective assessments such as [7], where rendering of object-based audio content to multiple loudspeaker layouts was evaluated, could therefore be carried out over headphones.

2 Measurement Details

A dataset of BRIRs was measured in the BBC R&D listening room [5] using Genelec 8030B loudspeakers and a Neumann KU100 dummy head microphone, which was rotated in 2° steps about the vertical axis. Additionally omnidirectional room impulse responses were measured with an Earthworks M30 microphone. The exponential sinusoidal sweep method [8] was used to measure BRIRs. Swept-sine signals of length 2^{18} samples were used, ranging from 20 Hz to 24 kHz, at a sampling rate of 48 kHz. An RME Micstacy preamplifier was used.

The loudspeaker positions are given in Table 1 and depicted in Figures 1 and 2. The coordinate system used is that of AES69:2015 [9]. Table 1 also indicates which loudspeakers are used for each of the layouts in ITU-R BS.2051. Note that layout G uses screen-related speakers, at the left and right edges of a video display. It is proposed that speakers at $\phi = \pm 45°$ might optionally be used to represent these speakers.

The loudspeaker sound pressure levels were aligned at the centre of the array to 70 dBA within ±0.1 dB using a band-limited pink noise signal (20 Hz-20 kHz). The loudspeaker driving signals were pre-equalised using an IOSONO Core system, to adjust for differences in the magnitude response of the direct sound and ensure time-alignment at the central listening position.

The dummy head microphone was placed on a computer-controlled rotational mount [10]. It was placed at the centre of the loudspeaker array, with the capsules at a height of 1.18 m from the floor. The initial microphone orientation was tested by evaluating the interaural time difference for the front centre loudspeaker. Custom software for controlling the mount and measuring the impulse responses was written in the Python language. The latency of the measurement system has been compensated. The impulse responses were truncated to 2^{14} samples.

3 Results

Figures 3 and 4 shows the energy time curve and energy decay relief for an example BRIR measurement. It can clearly be seen that the noise floor has been reached before the impulse response has been truncated. The achieved signal-to-noise ratio is approximately 60-85 dB in the frequency range 200 Hz to 18 kHz. Below 200 Hz there is more background noise present, users may wish to apply a high-pass filter or alternatively apply an anechoic head model at low frequencies [11]. Figure 5 shows the first 20 ms of the BRIRs for the front centre loudspeaker at each rotation of the dummy head microphone, from which the smoothly changing time-of-arrival of the direct sound can be observed, as well as the first two reflections.

4 Tools for Application

To allow easy application of the dataset, configuration files have been created for freely available open-source spatial audio software tools.

4.1 SoundScape Renderer Scenes

Scene configurations have been created for the Soundscape Renderer [12] to allow dynamic auralisation of loudspeaker signals in the ITU-R BS.2051 layouts.

4.2 AmbiX Configuration Files

Configurations for the AmbiX plug-ins [13] have also been created for rendering higher-order ambisonics to the ITU-R BS.2051 loudspeaker layouts and auralising the results. Ambisonics decoder design was performed according to [14]. These plug-ins can be used within a digital audio workstation and a head tracking system such as [15] can be used to rotate the ambisonic sound field to achieve a dynamic auralisation.

4.3 Dramatic Scenes

The S3A Object-Based Audio Drama Dataset[1] [16] provides object-based dramatic scenes using the Audio Definition Model [17]. These scenes have also been rendered to the ITU-R BS.2051 loudspeaker layouts using a VBAP-based panning algorithm, as well as to higher-order ambisonics. These files can be used as demonstration material for auralisation.

[1] http://epubs.surrey.ac.uk/811764/

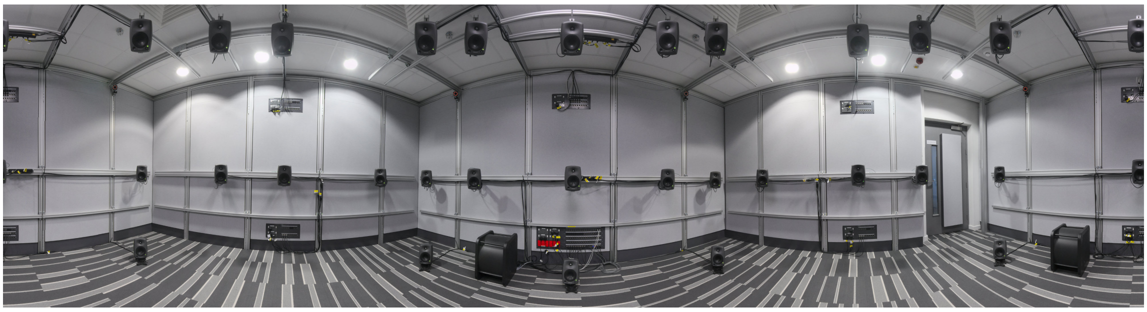

Fig. 1: 360° view inside the listening room, showing the current loudspeaker layout, excluding the top centre loudspeaker. The positive *x* axis (see Figure 2) points towards the centre of the image.

Table 1: Loudspeaker positions relative to the central listening position and ITU-R BS.2051 loudspeaker layouts

Speaker Label	**Azimuth** ϕ (°)	**Elevation** θ (°)	**Radius r** (m)	**A 0+2+0**	**B 0+5+0**	**C 2+5+0**	**D 4+5+0**	**E 4+5+1**	**F 3+7+0**	**G 4+9+0**	**H 9+10+3**
B+045	45	-22	2.59								x
B+000	0	-26	2.20								x
B-045	-45	-22	2.59								x
B-135	-135	-22	2.59								
B+135	135	-22	2.59								
M+045 [M+SC]	45	0	3.01							o	
M+030	30	0	2.37	x	x	x	x	x	x	x	x
M+000	0	0	1.99		x	x	x	x	x	x	x
M-030	-30	0	2.37	x	x	x	x	x	x	x	x
M-045 [M-SC]	-45	0	3.01							o	
M-060	-60	0	2.71								x
M-090	-90	0	2.28						x	x	x
M-110	-110	0	2.46		x	x	x	x			
M-135	-135	0	3.01						x	x	x
M+180	180	0	1.99								x
M+135	135	0	3.01						x	x	x
M-110	-110	0	2.46		x	x	x	x			
M+090	90	0	2.28						x	x	x
M+060	60	0	2.71								x
U+045	45	40	1.91						x	x	x
U+030	30	40	1.91			x	x	x			
U+000	0	40	1.91								x
U-030	-30	40	1.91			x	x	x			
U-045	-45	40	1.91						x	x	x
U-090	-90	40	1.87								x
U-110	-110	40	1.91				x	x		x	
U-135	-110	40	1.91								x
U+180	180	40	1.91						x		x
U+135	135	40	1.91								x
U+110	110	40	1.91				x	x		x	
U+090	90	40	1.86								x
T+000	0	40	1.30								x

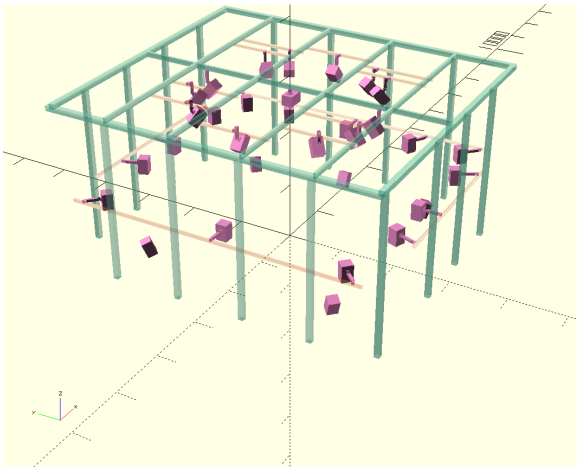

Fig. 2: Loudspeaker mounting system used in the BBC R&D listening room with loudspeaker positions measured (shown in purple)

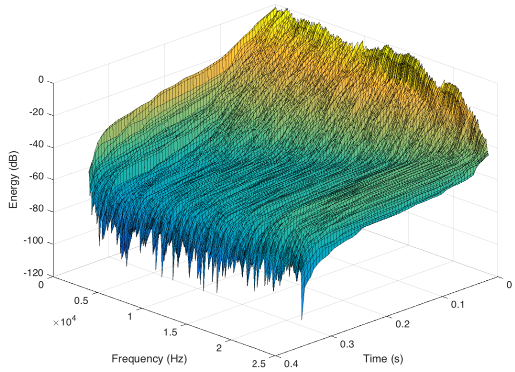

Fig. 4: Energy decay relief for an example BRIR

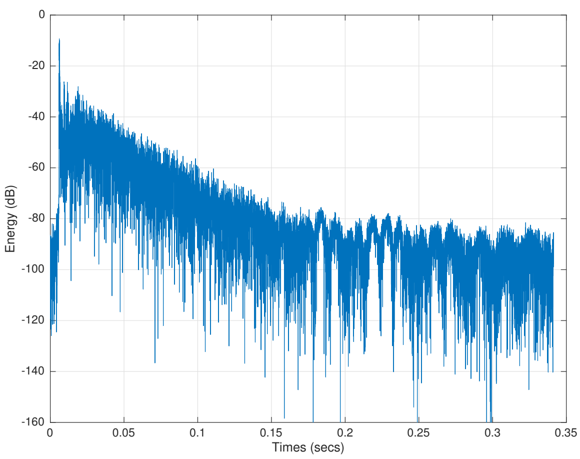

Fig. 3: Energy time curve for an example BRIR

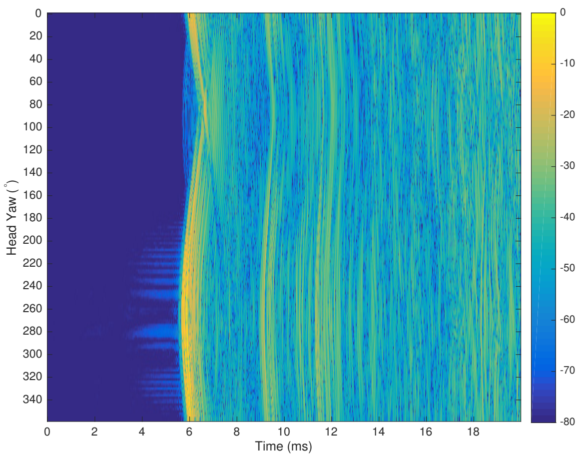

Fig. 5: BRIR energy (dB) over time with head rotation

5 Free Dataset

The impulse responses are made freely available in the Spatially Oriented Format for Acoustics (SOFA)[2] format, standardised in AES69:2015 [9]. A separate SOFA file has been created for each of the loudspeaker layouts, using the *MultiSpeakerBRIR* convention. The data is subject to the *Creative Commons Attribution-NonCommercial-ShareAlike 4.0* license[3]. The dataset of impulse responses and configuration files for AmbiX and SoundScape Renderer is made available at: `https://github.com/bbc/bbcrd-brirs`.

This work was supported by the BBC as part of the BBC Audio Research Partnership.

References

[1] Mackensen, P., Felderhof, U., Theile, G., Horbach, U., and Pellegrini, R. S., "Binaural room scanning—A new tool for acoustic and psychoacoustic research," *The Journal of the Acoustical Society of America*, 105(2), 1999.

[2] Lindau, A. and Weinzierl, S., "Assessing the Plausibility of Virtual Acoustic Environments," *Acta Acustica united with Acustica*, 98(5), 2012.

[3] Gedemer, L. A. and Welti, T., "Validation of the Binaural Room Scanning Method for Cinema Audio Research," in *135th AES Convention*, New York, 2013.

[4] ITU Radiocommunication Sector, "Recommendation BS.2051 - Advanced sound system for programme production," 2014.

[5] Nixon, T., Bonney, A., and Melchior, F., "A Reference Listening Room for 3D Audio Research," in *International Conference on Spatial Audio*, 2015.

[6] ITU Radiocommunication Sector, "Recommendation ITU-R BS.1116-3 - Methods for the subjective assessment of small impairments in audio systems," 2015.

[7] Zacharov, N., Pike, C., Melchior, F., and Worch, T., "Next generation audio system assessment using the multiple stimulus ideal profile method," in *International Conference on Quality of Multimedia Experience*, 2016.

[2] `http://www.sofaconventions.org`

[3] `http://creativecommons.org/licenses/by-nc-sa/4.0/`

[8] Müller, S. and Massarani, P., "Transfer Function Measurement with Sweeps," *Journal of the Audio Engineering Society*, 49(6), pp. 443–471, 2001.

[9] AES69:2015, "AES standard for file exchange - Spatial acoustic data file format," 2015.

[10] Shotton, M., Pike, C., and Melchior, F., "A Motorised Telescope Mount as a Computer-Controlled Rotational Platform for Dummy Head Measurements," in *136th AES Convention*, Berlin, 2014.

[11] Xie, B., "On the low frequency characteristics of head-related transfer function," *Chinese Journal of Acoustics*, 28(2), 2009.

[12] Geier, M. and Spors, S., "Spatial Audio with the SoundScape Renderer," in *VDT International Convention*, 2012.

[13] Kronlachner, M., "Plug-in Suite for Mastering the Production and Playback in Surround Sound and Ambisonics," in *AES Student Design Competition, 136th AES Convention*, Berlin, 2014.

[14] Romanov, M., Frank, M., Zotter, F., and Nixon, T., "Manipulations improving amplitude panning on small standard loudspeaker arrangements for surround with height," in *29th Tonmeistertagung*, Köln, 2016.

[15] Romanov, M., Berghold, P., Rudrich, D., Zaunschirm, M., Frank, M., and Zotter, F., "Implementation and Evaluation of a Low-cost Head-tracker for Binaural Synthesis," in *142nd AES Convention*, Berlin, 2017.

[16] Woodcock, J., Pike, C., Melchior, F., Coleman, P., Franck, A., and Hilton, A., "Presenting the S3A Object-Based Audio Drama dataset," in *140th AES Convention*, Paris, France, 2016.

[17] ITU Radiocommunication Sector, "Recommendation BS.2076 - Audio Definition Model," 2015.