Microphone Array Impulse Response (MAIR) Library and Renderer

Hyunkook Lee and Connor Millns
h.lee@hud.ac.uk connor.millns@hud.ac.uk
Applied Psychoacoustics Lab, University of Huddersfield, UK

1 Introduction

The Multichannel Array Impulse Responses (MAIR) library is an extensive database of room impulse responses captured for 13 different source positions in a reverberant concert hall, using numerous microphone arrays from 2ch stereo to 9ch 3D with height channels. MAIR is available for free under the CC-BY-4.0 license, and it was created with an aim to support researchers, sound engineers and students in spatial audio research and content creation. For example, MAIR can be used for creating realistically sounding multichannel stimuli for spatial audio experiments by convolving the impulse responses with multiple dry sources. Using the accompanied renderer, it also allows one to compare various microphone arrays, which can be also useful for spatial ear training. Furthermore, the renderer can process audio stream from a DAW session and therefore a virtual microphone array recording can be made flexibly and easily.

This document describes the impulse response capture method, microphone techniques used for the library, the structure of the library and the user instruction of the rendering tool.

2 Impulse Response Acquisition

Microphone array impulse responses (MAIRs) were captured at St. Paul's concert hall in Huddersfield, UK (Fig. 1), which is a church-converted hall with a high ceiling and the average T30 of 2.2s. The MAIRs were sampled at 44.1 kHz/32 bits using the exponential sine sweep method provided in the HAART software [1], which allows impulse response measurements for multiple sources and receivers with various measurement methods. The MAIRs were obtained for one or two configurations individually so that the microphones could be placed accurately at the intended positions.

Fig. 1. St. Paul's concert hall at the University of Huddersfield, where MAIRs were captured.

2.1 Microphone Setup

A total of 13 loudspeakers (nine Genelec 8040As and four Genelec 1029As) were arranged in front of the microphone array as depicted in Fig. 2. All loudspeakers were placed on loudspeaker stands and tilted so that they were on-axis with the microphones. The first row loudspeakers were placed on the stage floor, whereas the second row ones were on the platform. The height of the acoustic centres of the first row loudspeakers was 1.3m, whereas that of the second row loudspeakers was 2.1m.

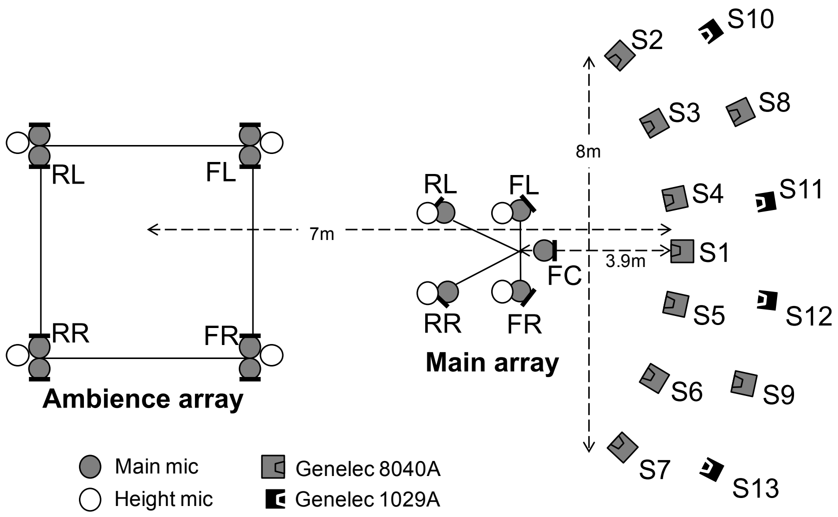

Fig. 2. Physical setup for the stage-audience scenario.

2.2 Microphone Setup

Table 1 summarises a total of 39 microphone configurations that were used for the stage-audience scenario. All of the microphones used for the measurements were Neumann digital microphones (kk184 cardioid, kk185 hyper-cardioid, kk133 omni, kk120 figure-of-eight) powered by Neumann DMI-8, except for the dummy head (Neumann KU100) and FOA (Sennheiser Ambeo) amplified by Merging HAPI.

The main microphone arrays included 2ch arrays (XY, AB, ORTF, dummy head – Neuamnn KU100), 3ch arrays (Decca Tree, Fukada Tree [2]), 4ch arrays (ESMA25 (IRT-Cross) [3], ESMA50 [4], First order Ambisonics – Sennheiser AMBEO) and 5ch arrays (OCT [3], ICA [5,6], PCMA [7], 2L-inspired [8]). They were raised 2.6m from the stage floor. The distance between the centre loudspeaker and the base point of the arrays (i.e., the middle point between the front left and front right microphones) was 3.9m.

For the height microphone layer, 1m x 1m square arrangements of four cardioid and omni patterns was used and they were placed at 0m and 1m above the main layer. The cardioids faced either towards the opposite direction from the stage for a maximum suppression of the direct sounds or directly upwards as suggested by Lee and Gribben [9]. Placing the height layer at 0m vertical spacing from the main layer makes the PCMA-3D configuration. Using omnis at 1m height from the OCT main layer was proposed for the OCT-9 by Theile and Wittek [10] and a similar approach is also used for the 2L Cube by Lindberg [8]. Additional height RIRs were captured also with pairs of omnis and upward-facing cardioids with 2m apart from each other and raised at 1m and 2m heights from the main layer.

The basic arrangement for the ambience arrays was the 2m x 2m Hamasaki Square [11], but for the current measurements the polar patterns were varied among figure-of-eight, cardioid and omni. The 4-channel ambience array with each polar pattern was raised at five different heights (2.6m, 3.1m, 3.6m, 4.1m and 4.6m), thus covering 0 to 2m vertical spacings between lower and upper layers for capturing ambience in 3D (i.e., Hamasaki Cube [12]). The centre point of the lowest square was 7m away from the centre loudspeaker, which was around the critical distance for the loudspeaker.

	Array	Channel	Polar pattern	Horizontal mic spacing	Front-Rear spacing	Mic angle from the base point	Vertical spacing from the main layer
Main arrays	OCT-inspired	5	FL/FR: Hypercardioid FC/RL/RR: Cardioid	Base-FC: 8cm FL-FR: 70cm RL-RR: 90cm	40cm	FL/FR: ±90° RL/RR: 180°	N/A
	ICA-5		Cardioid	Base-FC: 17.5cm FL-FR: 35cm	50cm	FL/FR: ±90° RL/RR: ±150°	
	PCMA1		FC/FL/FR: Cardioid RL/RR: Cardioid	Base-FC: 25cm FL-FR: 100cm RL-RR: 100cm	100cm	FL/FR: ±25° RL/RR: 180°	
	PCMA2		FC/FL/FR: Cardioid RL/RR: Hypercardioid			FL/FR: ±25° RL/RR: ±90°	
	PCMA3		FC/FL/FR: Cardioid RL/RR: Figure-of-8			FL/FR: ±25° RL/RR: ±90°	
	2L-inspired		Omni	Base-C: 25cm FL-FR: 100cm RL-RR: 100cm	100cm	FL/FR: ±45° RL/RR: ±45°	
	ESMA50	4	Cardioid	FL-FR: 50cm RL-RR: 50cm	50cm	FL/FR: ±45° RL/RR: ±45°	
	ESMA25 (IRT-Cross)		Cardioid	FL-FR: 25cm RL-RR: 25cm	25cm	FL/FR: ±45° RL/RR: ±45°	
	Ambisonics (1st order)		N/A	N/A	N/A	N/A	
	Decca Tree	3	Omni	Base-FC: 100cm FL-FR: 200cm		FL/FR: ±90°	
	Fukada Tree		Cardioid	Base-FC: 100cm FL-FR: 200cm		RL/RR: ±90°	
	XY130	2	Cardioid	FL-FR: 0cm		FL/FR: ±65°	
	AB50		Omni	FL-FR: 50cm		FL/FR: 0°	
	ORTF		Cardioid	FL-FR: 17cm		FL/FR: ±55°	
	Dummy head		N/A	N/A		N/A	
	Height01	4	Cardioid	FLh-FRh: 100cm RLh-RRh: 100cm	100cm	Off-axis to the source	0cm
	Height02					Directly upwards	
	Height03					Directly off-axis to the source	100cm
	Height04					Directly upwards	
	Height05	2	Cardioid	FLh-FRh: 200cm	N/A	Directly upwards	100cm
	Height06						200cm
	Height07	4	Omni	FLh-FRh: 100cm RLh-RRh: 100cm	100cm	Directly upwards	100cm
	Height08	2	Omni	FLh-FRh: 200cm	N/A	Directly upwards	100cm
	Height09						200cm
Ambience arrays	Hamasaki01	4	Figure-of-8	FL-FR: 200cm RL-RR: 200cm	200cm	FL/FR: ±90° RL/RR: ±90°	0cm
	Hamasaki02						50cm
	Hamasaki03						100cm
	Hamasaki04						150cm
	Hamasaki05						200cm
	Hamasaki06		Cardioid			FL/FR: 180° RL/RR: 180°	0cm
	Hamasaki07						50cm
	Hamasaki08						100cm
	Hamasaki09						150cm
	Hamasaki10						200cm
	Hamasaki11		Omni			Directly upwards	0cm
	Hamasaki12						50cm
	Hamasaki13						100cm
	Hamasaki14						150cm
	Hamasaki15						200cm

Table 1. Microphone configurations used for the MAIR library.

3 MAIR Library

The MAIR package can be downloaded free from www.github.com/APL-Huddersfield. It comprises of four file folders: MainArrays, MainHeightArrays, AmbienceArrays and SADIE_BRIRs. Each of the microphone array folders contains sub-folders for impulse responses captured with different microphone configurations described in Table 1. The naming convention for each impulse response is as follows.

Array Type_Array Name_Mic ID_Speaker ID_Output Channel.wav

For example, "Main_PCMA_M01_S04_FL.wav" is the impulse response for Speaker 4 captured by Microphone 1 of a main mic array called PCMA, which is supposed to be routed to the Front Left output channel. The ID for each speaker position can be found in Fig. 1.

These MAIRs can be convolved with dry sound sources manually by the user using suitable convolution plugin on a DAW or Matlab. However, the accompanied MAIR renderer provides convenient and easy-to-use graphical user interface and efficient real-time convolution engine.

The SADIE_BRIRs folder currently contains head related room impulse responses (HRIRs) for Auro-3D 9ch loudspeaker configuration. The database used for these HRIRs is SADIE by University of York [13]. These HRIRs are used for the binauralisation of the Auro-3D loudspeaker reproduction within the MAIR renderer.

4 User Instruction for the MAIR Renderer

This section provides instructions for using the MAIR renderer. Fig. 2 shows a screenshot of the renderer. Note that the current version only works on Mac OS. A Windows version will be provided in the next update soon.

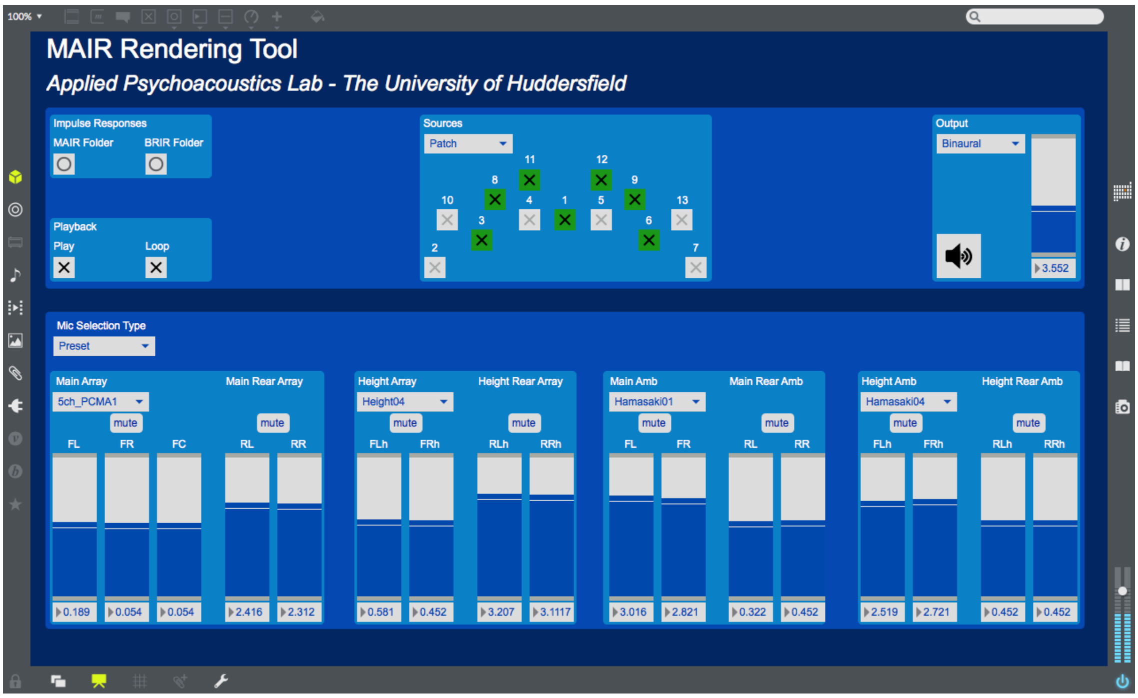

Fig. 2. MAIR renderer that allows users to convolve the MAIR with multiple sources and also to mix various microphone configurations.

4.1 Quick Start Guide

1. Download and install Max 7
2. Load *MAIR_RenderingTool_v1.mxf*
3. Press the 'MAIR Folder' button in the patch and find the *MAIR_Lib2017* folder on your Mac.
4. Press the 'BRIR Folder' button in the patch and find your BRIR folder (e.g. SADIE_BRIRs) folder on your Mac.
5. Load a sound source into the patch.
6. Select a Microphone array.
7. Choose output type (binaural or speaker).
8. Turn on the Audio DSP in Max.

4.2 Impulse Response Section

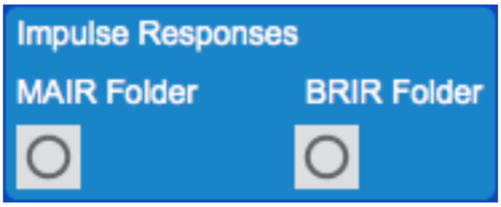

Fig. 3. Impulse response loading section of the MAIR renderer.

This section of the patch deals with loading the MAIR Library and the BRIRs for binauralisation. Once the 'MAIR Folder' button is pressed a dialogue box opens, in which you must locate the main folder for the MAIR Library (the folder that holds all the mic array subfolders). Fig. 4 shows an example of this.

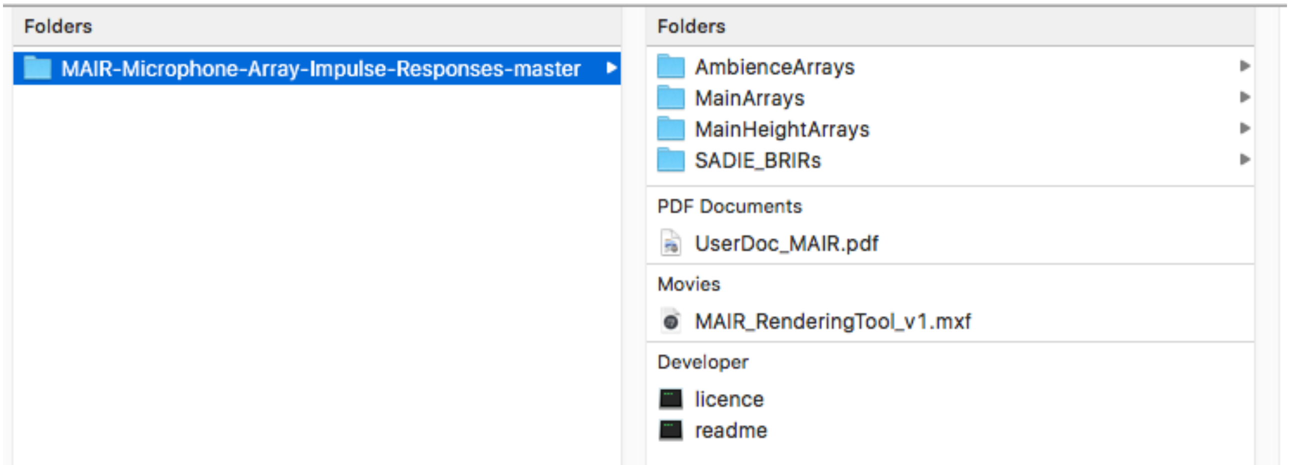

Fig. 4. The main folder for MAIR needs to be loaded first.

The 'BRIR Folder' button works in the same way. You are asked to find the folder which contains the BRIRs you wish to use. You can use other BRIRs than the ones provided, so long as the order and positions of the impulses are the same as what is shown in Table 2. For instance, the names of your BRIRs can start with numbers that correspond to specific output channels, e.g. 01_azi_30_ele_0.wav for FL, 01_azi_330_ele_0.wav for FR and 01_azi_0_ele_0.wav for FC. The BRIRs also need to be in stereo and have a sampling frequency of 44.1kHz.

Order number	Output channel for BRIR
1	Front Left (FL)
2	Front Right (FR)
3	Front Centre (FC)

4	Rear Left (RL)
5	Rear Right (RR)
6	Front Left Height (FLh)
7	Front Right Height (FRh)
8	Rear Left Height (RLh)
9	Rear Right Height (RRh)

Table 2. The required order of BRIRs for output channels.

4.3 Playback Section

This section will control the playback of the sound sources loaded directly into the rendering tool.

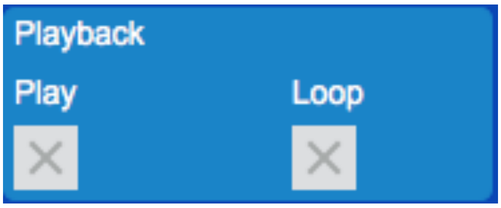

Fig. 3. Playback section of the MAIR renderer.

NOTE: Once a new sound source has been loaded, you must repress the play button to restart playback.

4.4 Source Section

In this section, you can load any .wav and .au files with a sampling frequency of 44.1kHz to any of the 13 source positions shown in Fig. 4 when the 'Patch' mode is selected in the dropdown menu. The source positions shown on the GUI correspond to the source positions shown in Fig. 2. Once one of source positions is clicked, a dialogue box opens which allows you to find your chosen audio file. To clear the audio file, simply click the source position it was loaded into. This will however mean you must reload the audio file if you want to hear it again in that position.

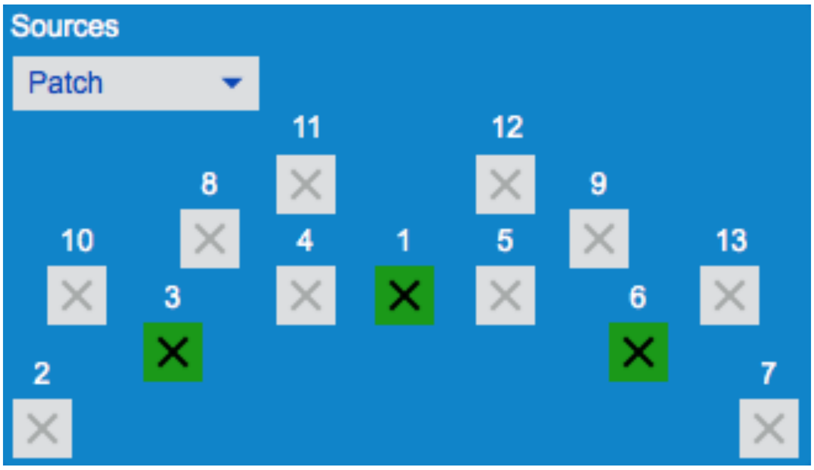

Fig. 4. 'Patch' mode in source section.

Another option is to stream audio to the rendering tool through a DAW. To do this you must use a virtual audio interface (e.g. Soundflower for Mac) to send the audio into the inputs of the rendering tool, which are inputs 1-13.

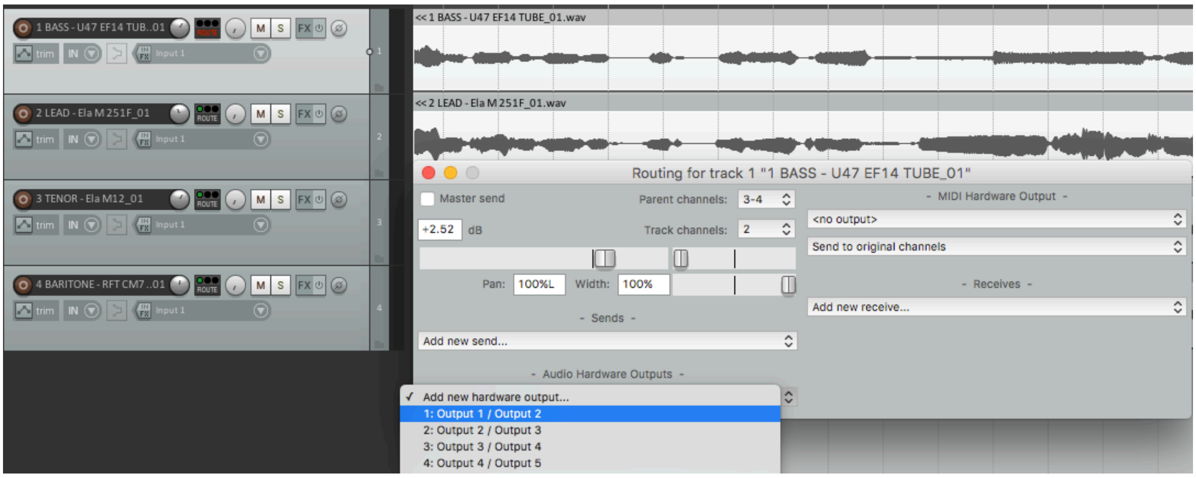

Fig. 4. Example of output channel selection on Reaper

An example of sending audio to the rendering tool through a DAW is shown in Fig. 5. As the tool accepts audio from inputs 1-13, all that needs to be done is to change the DAW's audio device to the virtual audio interface and send the audio through one of the hardware outputs from 1-13. The virtual audio interface also needs to be set as the rendering tool's input device in Max 7. The output channel number will correspond with the source position in the tool.

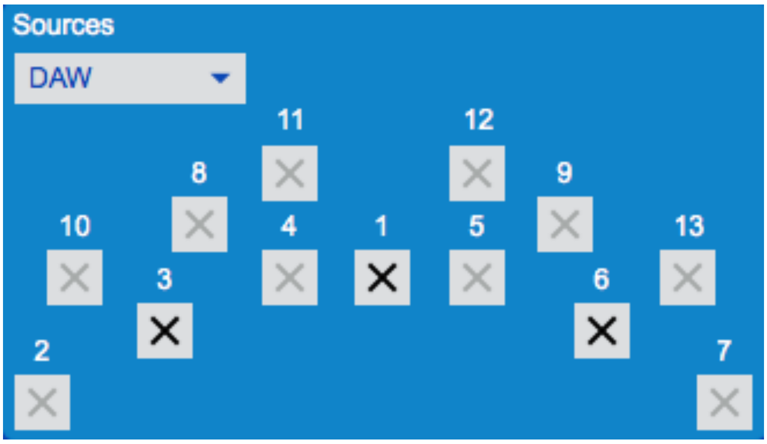

Fig. 6. DAW mode in source selection section.

When the Source section's dropdown menu is set DAW the function of the source positions changes. The toggles on the source positions now have a mute and unmute function. By default, all positions are muted at first so once audio has been sent by the DAW, you must unmute the chosen source position, e.g. if you have sent audio to output channel 1, 3 and 6 on the DAW, you must unmute source positions 1, 3 and 6 on the MAIR renderer.

4.5 Mixer Section

This section allows you select which microphone array is used to capture the sound source. The 'Mic Selection Type' has two options in the dropdown menu: 'Preset' and 'Custom'. 'Preset' will allow you to select complete microphone arrays (e.g., all 5 microphones of 5ch arrays like OCT or PCMA), whereas 'Custom' allows you to change what microphone array is used for the front pair/triplet and which microphone array is used for the back pair, e.g., combining the front triplet of PCMA with the rear pair of OCT, as shown in Fig. 7.

The mute function allows you to silence the audio for the front and back of an array individually. This mute can also help reduce CPU load as it mutes some DSP functions in Max 7.

The sliders control the level for the virtual microphone signals. They have up to +6 dB of gain.

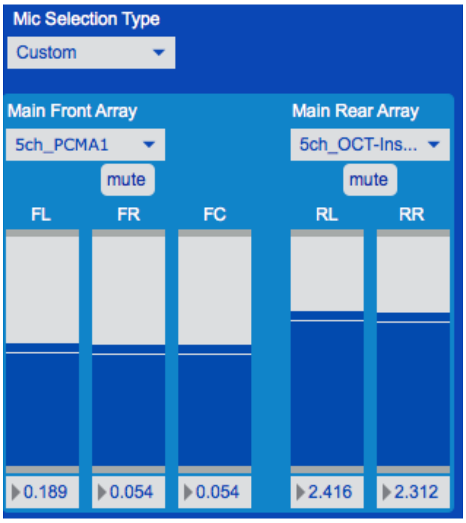

Fig. 7. Mic selection type and mixer section.

4.6 Output Section

The Output section controls the level of the output signal and whether the signal is sent straight to a multichannel loudspeaker array or binauralised for headphone use. The dropdown menu will set the output type and output level slider has a maximum of +12 dB. The speaker icon turns the DSP in Max 7 on and off.

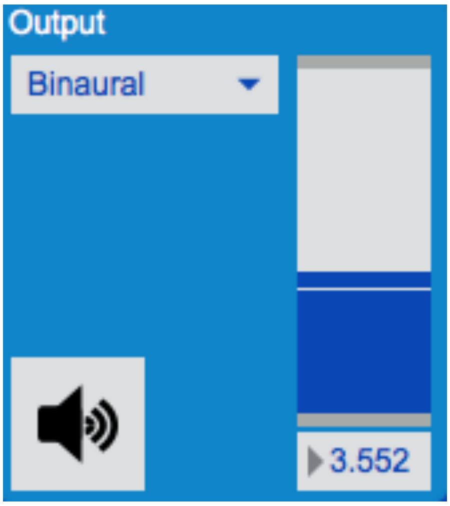

Fig. 8. Output section of the MAIR renderer.

When using the loudspeaker for playback, the output audio signals are mapped to the channels as shown in Table 3. If your loudspeaker setup uses different output channels to the one shown in Table 3, you can use the I/O Mapping system within Max to route them correctly for your system.

Output Channel	Speaker
1	Front Left (FL)
2	Front Right (FR)
3	Front Centre (FC)

4	NOT USED – No LFE
5	Rear Left (RL)
6	Rear Right (RR)
7	Front Left Height (FLh)
8	Front Right Height (FRh)
9	Rear Left Height (RLh)
10	Rear Right Height (RRh)

Table 3. Loudspeaker output channel mapping.

Feedback and Bug Report

All feedbacks are welcome. To report any bugs or ask any query about the rendering tool, please contact Connor Millns (connor.millns@hud.ac.uk) or Hyunkook Lee (h.lee@hud.ac.uk).

If you would like to be updated for MAIR or other research news by the Applied Psychoacoustics Lab, please sign up here → https://goo.gl/forms/CkqgkdePwOTR8XCI2

Acknowledgement

All microphones used for this project were supported by Sennheiser/Neumann UK. The authors would like to thank Andrew Lillywhite of Sennheiser/Neumann UK and Prof. Rupert Till of the University of Huddersfield for arranging the loan of the microphones. The authors are also grateful to Maksims Mironovs and Callum Eaton of the Applied Psychoacoustics Lab and Matt Jefferson of Fraunhofer IIS for contributing to the acquisition of the impulse responses.

References

[1] D. Johnson, A. Harker and H. Lee, “HAART: A new impulse response toolbox for spatial audio research”, *presented at the 138th Convention of the Audio Engineering Society* (2015), engineering brief 190.

[2] A. Fukada, K. Tsujimoto and S. Akita, “Microphone techniques for ambient sound on a music recording,” presented at the 103rd Convention of the Audio Engineering Society (1997), convention paper 4540.

[3] G. Theile, “Multichannel natural recording based on psychoacoustic principles,” In *Proceedings of the Audio Engineering Society 19th International Conference*, pp.201-229, 2001.

[4] H. Lee, “Capturing and rendering 360° VR audio using cardioid microphones,” *Proceedings of the Audio Engineering Society International Conference on Virtual and Augmented Reality*, 2016.

[5] U. Herrmann and V. Henkels, “Main microphone techniques for the 3/2- stereo-standard”, www.hhton.de.

[6] M. Williams, “Microphone arrays for natural multiphony,” *presented at the 91st Convention of the Audio Engineering Society* (1991), convention paper 3157.

[7] H. Lee, “A New Microphone Technique for Effec- tive Perspective Control,” presented at the 130th Convention of the Audio Engineering Society (2011), convention paper 8337.

[8] M. Lindberg, http://www.2l.no/artikler/2L-VDT.pdf

[9] H. Lee and C. Gribben, “Effect of vertical microphone layer spacing for a 3D microphone array,” *J. Audio Eng. Soc.*, vol. 62, pp. 870–884 (2014 Dec.).

[10] G. Theile and H. Wittek, “Principles in Surround Recordings with Height,” *presented at the 130th Convention of the Audio Engineering Society* (2011 May), convention paper 8403.

[11] K. Hamasaki, “Multichannel Recording Techniques for Reproducing Adequate Spatial Impression,” *Proc. of the Audio Engineering Society 24th International Conference: Multichannel Audio—The New Reality* (2003 June).

[12] K. Hamasaki and W. Van Baelen, “Natural sound recording of an orchestra with three-dimensional sound”, *presented at the 138th Convention of the Audio Engineering Society* (2015), convention paper 9348.

[13] G. Kearney, “SADIE HRIR database”, https://www.york.ac.uk/sadie-project/binaural.html