Audio Engineering Society

Convention e-Brief 448

Presented at the 144th Convention
2018 May 23 – 26, Milan, Italy

Ambisonics Directional Room Impulse Response as a new Convention of the Spatially Oriented Format for Acoustics

Andrés Pérez-López[1,2] and Julien De Muynke[1]

[1]*Multimedia Technologies Unit, Eurecat - Technology Centre of Catalonia, Barcelona, Spain*
[2]*Music Technology Group, Pompeu Fabra University, Barcelona, Spain*

Correspondence should be addressed to Andrés Pérez-López (`andres.perez@eurecat.org`)

ABSTRACT

Room Impulse Response (RIR) measurements are one of the most common ways to capture acoustic characteristics of a given space. When performed with microphone arrays, the RIRs inherently contain directional information. Due to the growing interest in Ambisonics and audio for Virtual Reality, new spherical microphone arrays recently hit the market. Accordingly, several databases of Directional RIRs (DRIRs) measured with such arrays, referred to as Ambisonics DRIRs, have been publicly released. However, there is no format consensus among databases. With the aim of improving interoperability, we propose an exchange format for Ambisonics DRIRs, as a new Spatially Oriented Format for Acoustics (SOFA) convention. As a use-case, some existing databases have been converted and released following our proposal.

1 Introduction

1.1 Directional Room Impulse Responsess

Impulse Responses (IRs) measurements constitute a compact way of representing the acoustic properties of a linear time-invariant system. When such measurements are performed in a specific room or enclosure, the so-called Room Impulse Responses (RIRs) are able to capture the intrinsic reverberation and acoustic characteristics of the enclosure, for which several methods have been developed these past years [1]. Furthermore, it is possible to account for different emitter/receiver positions in the measurement, usually performing the measurement with a microphone array. In that case, this kind of measurements is referred to as Directional Room Impulse Responses (DRIRs) [2]. DRIRs have a wide range of applications: auralization [2], room acoustics analysis [3, 4] and modelling [5], spatial audio synthesis [6], source separation and dereverberation [7], acoustic heritage preservation [8, 9], etc.

1.2 Ambisonics

Ambisonics is a sonic theory, developed in the 1970s by Gerzon [10], based on the spatial decomposition of a soundfield into a sequence of spherical harmonics. Spherical harmonics constitute a complete set of orthogonal functions along a spherical surface.. Since the spherical harmonics expansion has theoretically an infinite number of terms, in a practical context the spatial decomposition must be truncated at some order $L \in \mathbb{N}$. The value of L - referred to as the *Ambisonics Order* - will determine the spatial resolution of the decomposed soundfield.

Let B_{mn} be the *Ambisonics Component* [11] of order m and degree n, where $m \in \mathbb{N}, m \in [0, L]$ and $n \in \mathbb{N}, |n| \in [0, m]$. B_{mn} can be described as :

$$B_{mn} = SY_{mn}(\phi_s, \theta_s) \qquad (1)$$

where S is the audio signal located at the position described in a spherical surface by the azimuth-elevation tuple (ϕ_s, θ_s), and Y_{mn} is the real-valued Spherical Harmonic of order m and degree n evaluated at (ϕ_s, θ_s):

$$Y_{mn}(\phi_s, \theta_s) = \sqrt{(2 - \delta_{0n})\frac{(m-n)!}{(m+n)!}}P_{mn}(sin(\theta_s))f(n\phi_s) \qquad (2)$$

where:

$$f(n\phi_s) = \begin{cases} cos(n\phi_s), & \text{if } n > 0 \\ 1, & \text{if } n = 0 \\ sin(n\phi_s), & \text{if } n < 0 \end{cases} \qquad (3)$$

δ_{ij} is the Kronecker delta, and $P_{mn}(sin(\theta_s))$ are the associated Legendre functions, with ommited Condon-Shortley phase.

These equations describe the *Ambisonics* transformation using *SN3D* normalization and *ACN* channel ordering, which corresponds to the *Ambix* convention [12] (the current *de facto* standard for spatial audio production). Other conventions might be used by applying the corresponding transformations (see [13] for a more comprehensive explanation).

Equation 1 describes the *encoding* of a signal located at a determined angular position with respect to the listener, for a given order L. The resulting $(L+1)^2$ signals constitute the Ambisonics representation of the signal, also referred to as *B-Format* in audio production-oriented contexts.

1.3 Ambisonics Recording

It is also possible to capture Ambisonics audio scenes by using specific recording devices. Spherical microphone arrays (also known as *Ambisonics microphones*) are a type of microphone arrays in which the capsules are located around a spherical surface, presenting rotational symmetry. Such geometric arrangement allows the recorded signals to be transformed to the Ambisonics domain, by means of projection into the Spherical Harmonics basis, and further equalization with radial filters [11].

This process is known in the audio production domain as *A-B Conversion*. All major spherical microphone array manufacturers provide tools for achieving this transformation [14, 15, 16, 17, 18].

1.4 Ambisonics DRIRs

The intrinsic spatial capabilities of Ambisonics microphones might be applied to DRIR recording, as originally proposed by Gerzon in the context of acoustical heritage preservation [8]. In recent years, several datasets of Ambisonics DRIRs have been publicly released, such as the OpenAIR database [19] or the set of measurements performed in the scope of the S3A project [20, 21].

1.5 The SOFA Conventions

In general, the Ambisonics DRIR databases show a common approach for describing the measurements: given a specific room, usually represented as a folder, IR data consist of several multichannel audio files, with one audio channel per spherical harmonic, and one file per emitter/receiver combination. Furthermore, it is also usual to provide a *metadata* file, describing the different emitter and receiver positions, and eventually some information about the measurement setup, methodology, etc. Such files might be formatted as plain text or delimiter-separated value files.

Despite the common approach, it can be foreseen that each database generated by a different individual or institution might potentially have a different naming convention, folder structure, file format, and so on. This situation hinders data manipulation and exchange, and forces users to write ad-hoc parsers and algorithms for each specific database.

The *Spatially Oriented Format for Acoustics* (SOFA) is a file format designed for a consistent, standardized storage and manipulation of IR data [22]. The need for such standard arose from dealing with different databases of *Head-Related Transfer Functions* (HRTFs), in a similar manner as the one mentioned for Ambisonics DRIRs.

There are several SOFA conventions, each one addressing a particular type of IR measurement. In the case of Ambisonics DRIRs, given the data representation in existing databases, one could outline the following specificities:

- Presence of Ambisonics-related information (Ambisonics order, channel ordering and normalization)

- Audio stored in the Ambisonics (spherical harmonics) domain

- Data structure support for different combinations of source and receiver positions

However, none of the existing SOFA conventions meets those requisites. The potentially first candidate by name, *SingleRoomDRIR*, is limited to one source position per file. On the other hand, *MultiSpeakerBRIR* allows for multiple sound sources, but restricts the number of receivers (microphone capsules) to two, as expected in a binaural recording. *GeneralFIRE* is intended for "data which are too general to store in more specific conventions" [23].

2 Proposal

2.1 Considerations

The SOFA specification defines some criteria that must be fulfilled in order to propose a new convention [24]. These criteria are:

1. Data must exist.

2. Data can not be described by existing SOFA conventions.

3. Relevant information about the data must be available.

As we already mentioned, existing databases of Ambisonics DRIRs can be found in OpenAIR and S3A databases (criterium 1). Criterium 2 has been discussed in the previous paragraph. All available Ambisonics DRIR datasets are accompanied by explanations, pictures, diagrams and related information (criterium 3).

2.2 Specifications

We therefore propose a new SOFA convention, *AmbisonicsDRIR*, designed for DRIRs measured with spherical microphone arrays, and presented in the Ambisonics domain. In other words, we propose to store the instantaneous B_{mn} values for a given Ambisonics order L, provided that S is a unit impulse (Eq. 1).

The *Listener* - as defined by the SOFA specification - is embodied by the spherical microphone array. The different *Receivers* are the different Ambisonics Components B_{mn}, so that their number is fixed provided L (more precisely, by following the relationship $R = (L+1)^2$), and their positions are not applicable. Furthermore, there might be many different *Sources*, all of them omnidirectional and consisting of one only *Emitter*. In that sense, M represents the number of different *Source* positions.

The proposed convention is based on *GeneralFIRE*, with the following modifications:

- Mandatory field *GLOBAL:AmbisonicsOrder* (type *double*, dimension *I*, default *1*). Indicates the order of the Spherical Harmonic expansion.

- Mandatory field *DATA:ChannelOrdering* (type *attribute*, default *acn*). Describes the ordering of the different Ambisonics Channels. Must be one of: *acn* or *fuma*.

- Mandatory field *DATAChannelNormalization* (type *attribute*, default *sn3d*). Describes the Ambisonics normalization convention used in the data. Must be one of: *sn3d*, *n3d*, *fuma* or *maxn*.

- Fields *ReceiverPosition*, *ReceiverPosition:Type* and *ReceiverPosition:Units* are not needed, since it is assumed that *Receivers* represent the Ambisonics channels. Furthermore, the values of R (number of receivers) are defined by $R = (L+1)^2$. Accordingly, the values of *GLOBAL:AmbisonicsOrder* and R are valid only if they follow the given equation.
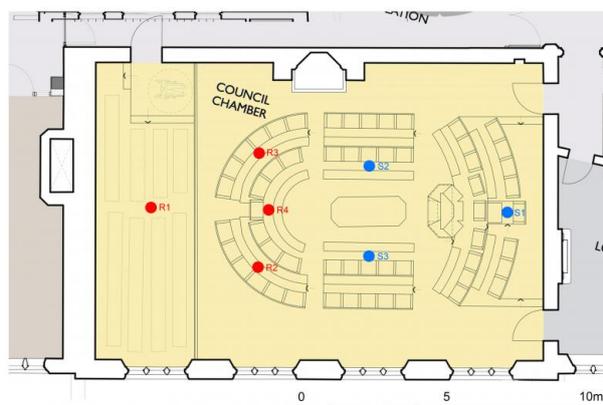
Fig. 1: Source and Listener position diagram for the "York Guildhall Council Chamber" Ambisonics DRIR (original diagram attribution to [28])

- Field *DATA:Delay* is not mandatory.

- Relevant information about the microphone model and brand, Ambisonics encoding methodology, software used, etc, is recommended to be added into the field *GLOBAL:Comment*.

3 Results

As a preliminary result, we have extended the current SOFA C++ [25] and Matlab [26] APIs to be fully compatible with the AmbisonicsDRIR convention.

Furthermore, as a use-case, a selection of existing Ambisonics DRIRs have been transcoded to the proposed convention: *Emmanuel Main Church* from S3A database [20], and *Heslington Church* and *York Guildhall Council Chamber* from the OpenAIRlib [21]. Figure 1 shows a schematic diagram of the different *Emitter/Receiver* positions at the *York Guildhall Council Chamber* recordings. The data, as well as the tools used to perform the conversion, are available online under an open source license [27].

Finally, we must remark that the *AmbisonicsDRIR* SOFA convention proposal is currently under discussion, and the described specifications are subject to change with upcoming versions.

4 Summary

This document addresses the lack of compatibility among different databases of Ambisonics DRIRs. The present proposal consists in defining a new SOFA convention, specifically designed for Ambisonics DRIRs. That way we contribute to improve the ease of data manipulation and database interoperability. Software implementations and tools for automatic conversion are provided.

5 Acknowledgements

References

[1] Stan, G.-B., Embrechts, J.-J., and Archambeau, D., "Comparison of different impulse response measurement techniques," *Journal of the Audio Engineering Society*, 50(4), pp. 249–262, 2002.

[2] Embrechts, J.-J., Lesoinne, S., and Werner, N., "Computation of Directional Impulse Responses in rooms for Bettter Auralization," in *Audio Engineering Society Convention 118*, Audio Engineering Society, 2005.

[3] Embrechts, J.-J., "Measurement of 3D room impulse responses with a spherical microphone array," in *Proceedings of the Euronoise 2015 Congress*, pp. 143–148, 2015.

[4] Clapp, S. W., Guthrie, A. E., Braasch, J., and Xiang, N., "Investigations of room acoustics with a spherical microphone array," in *Audio Engineering Society Convention 131*, Audio Engineering Society, 2011.

[5] Romblom, D., "Diffuse Field Modeling using Physically-Inspired Decorrelation Filters: Improvements to the Filter Design Method," *Journal of the Audio Engineering Society*, 65(11), pp. 943–953, 2017.

[6] Coleman, P., Franck, A., Menzies, D., and Jackson, P. J., "Object-based reverberation encoding from first-order Ambisonic RIRs," in *Audio Engineering Society Convention 142*, Audio Engineering Society, 2017.

[7] Baqué, M., Guérin, A., and Melon, M., "Separation of direct sounds from early reflections using the entropy rate bound minimization algorithm," in *Audio Engineering Society Conference: 60th International Conference: DREAMS (Dereverberation and Reverberation of Audio, Music, and Speech)*, Audio Engineering Society, 2016.

[8] Gerzon, M. A., "Recording concert hall acoustics for posterity," *Journal of the Audio Engineering Society*, 23(7), pp. 569–571, 1975.

[9] Murphy, D. T., "Multi-channel impulse response measurement, analysis and rendering in archaeological acoustics," in *Audio Engineering Society Convention 119*, Audio Engineering Society, 2005.

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

[11] Bertet, S., Daniel, J., and Moreau, S., "3D Sound Field Recording With Higher Order Ambisonics - Objective Measurements and Validation of a 4th Order Spherical Microphone," *Audio Engineering Society Convention 120*, pp. 1–24, 2006.

[12] Nachbar, C., Zotter, F., Deleflie, E., and Sontacchi, A., "Ambix-a suggested ambisonics format," in *Ambisonics Symposium, Lexington*, 2011.

[13] Carpentier, T., "Normalization schemes in Ambisonic: does it matter?" *142nd Convention Audio Engineering Society*, 0, 2017.

[14] Soundfield, "Soundfield Product Listing `http://www.soundfield.com/products`," Accessed on March 8th, 2018.

[15] Sennheiser, "AMBEO Plugins `https://en-us.sennheiser.com/ambeo-blueprints-downloads`," Accessed on March 8th, 2018.

[16] Zylia, "ZYLIA Ambisonics Converter `http://www.zylia.co/zylia-ambisonics-converter.html`," Accessed on March 8th, 2018.

[17] VVAudio.com, "VVTetraVST `https://www.vvaudio.com/products/VVTetraVST`," Accessed on March 8th, 2018.

[18] Acoustics, M., "EigenStudio `https://mhacoustics.com/download`," Accessed on March 8th, 2018.

[19] Murphy, D. T. and Shelley, S., "Openair: An interactive auralization web resource and database," in *Audio Engineering Society Convention 129*, Audio Engineering Society, 2010.

[20] Coleman, P., Remaggi, L., and Jackson, P., "S3A room impulse responses," 2015, doi:10.15126/surreydata.00808465.

[21] "The Open Acoustic Impulse Response Library `www.openairlib.net`," Accessed on Marcg 8th, 2018.

[22] Majdak, P. and Noisternig, M., "AES69-2015: AES standard for file exchange-Spatial acoustic data file format," *Audio Engineering Society.(March 2015)*, 2015.

[23] SOFA, "GeneralFIRE `https://www.sofaconventions.org/mediawiki/index.php/GeneralFIRE`," Accessed on February 13th, 2018.

[24] SOFA, "SOFA Conventions. `https://www.sofaconventions.org/mediawiki/index.php/SOFA_conventions`," Accessed on March 12th, 2018.

[25] Pérez-López, A., "Github, SOFA Cpp API `https://github.com/andresperezlopez/API_Cpp`," Accessed on February 13th, 2018.

[26] Muynke, J. D., "Github, SOFA Matlab/Octave API `https://github.com/jdemuynke/API_MO`," Accessed on February 13th, 2018.

[27] Pérez-López, A., "Github, AmbisonicsDRIRexamples. `https://github.com/andresperezlopez/AmbisonicsDRIRexamples`," Accessed on March 12th, 2018.

[28] of York Council, C., "Guildhall drawings `https://www.york.gov.uk/downloads/file/11337/guildhall_drawings`," Accessed on February 13th, 2018.