The 21st International Congress on Sound and Vibration

13-17 July, 2014, Beijing/China

EFFICIENT REPRESENTATION OF HEAD-RELATED TRANSFER FUNCTIONS USING SPATIALLY ORIENTED FORMAT FOR ACOUSTICS

Majdak Piotr, Ziegelwanger Harald

Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria

e-mail: piotr@majdak.com

Wierstorf Hagen

Assessment of IP-based Applications, Technische Universität Berlin, Berlin, Germany

Noisternig Markus

IRCAM, Paris, France

Head-related transfer functions (HRTFs) describe the spatial filtering of the incoming sound by the head, pinna, and torso of a listener. Transformed to the time domain, HRTFs are referred to as head-related impulse responses (HRIRs). While usually measured under free-field conditions, HRIRs are referred to as binaural room impulse responses (BRIRs) when room effects are considered additionally. HRIRs and BRIRs have been measured by a number of laboratories and are usually stored using the labs' own file formats, which often complicates the exchange of data due to format incompatibilities. Recently, the "spatially oriented format for acoustics" (SOFA) has been proposed to unify the representation of such data. SOFA provides data compression, network transfer, file hierarchy, and possibility for a link to complex room geometry data. SOFA specifications allow to develop application programming interfaces (APIs) for various programming languages; current implementations include Matlab and Octave. Within SOFA, the different types of data are represented by different conventions, i.e., defined sets of variables and attributes. In this contribution, we introduce the general SOFA specifications and describe SOFA conventions allowing to store HRIRs. Currently, in the SOFA database, eight large databases have been unified offering HRIRs/BRIRs of over 200 listeners stored in the same format.

1. Introduction

Head-related transfer functions (HRTFs) describe the spatial filtering of the incoming sound as an effect of acoustic interaction between the sound field and the listener's torso, head, and pinna[1]. HRTFs are important for the binaural reproduction of sounds via headphones[2]. HRTFs have been measured by a number of laboratories and are typically stored in each lab's native file format[3]. While the different formats are of advantage for each lab, an exchange of such data is difficult due to incompatibilities between formats. In this work, we describe specifications of the spatially oriented format for acoustics (SOFA) which was designed to store HRTFs focusing on their interchangeability and extendability[4]. In particular, SOFA fulfills the following requirements:

- Description of a measurement setup with arbitrary geometry, i.e., not limited to special cases like a regular grid, or a constant distance;
- Self-describing data with a consistent definition, i.e., all the required information about the measurement setup must be provided as metadata in the file;
- Flexibility to describe data of multiple conditions (listeners, distances, etc) in a single file;
- Partial file and network support;
- Available as binary file with data compression for efficient storage and transfer.

SOFA aims at representing spatial data in a general way, allowing to store not only HRTFs but also more complex data, e.g., directional room impulse responses (DRIRs) measured with a multichannel microphone array excited by a loudspeaker array. In order to simplify the adaption of SOFA for various applications, SOFA supports the so-called conventions: sets of rules for the representation of specific type of data. In this study, we focus on the representation of HRTFs measured under assumed free-field conditions. We describe how to store such HRTFs in SOFA and explain the corresponding SOFA conventions: *SimpleFreeFieldHRIR*.

2. SOFA

Typical HRTF measurement setups have the following properties in common. In an anechoic chamber or in a room, excitation signals are generated by *emitters* and microphones are used as *receivers* to record the incoming signals (see Fig. 1). The measurement is repeated while varying the spatial position of the excitation *source* relative to the *listener*, which is done by varying the position of the listener, the sound source, or both in different dimensions.

Thus, in SOFA, an HRTF measurement setup is described by various objects and their relations. SOFA considers the following objects:

- *Receiver* is any acoustic sensor like a microphone. The number of receivers in not limited in SOFA and defines the size of the data matrix.
- *Listener* is the object incorporating all the receivers. For HRTFs, a listener can be a head or dummy-head microphone. For DRIRs, a listener represents the microphone-array structure such as a sphere or a frame. Incorporating the receivers in the listener as a single logical object is important because in measurements, usually the orientation and/or the position of the listener vary without substantial changes in the head-microphone relation. For example, in measurements done

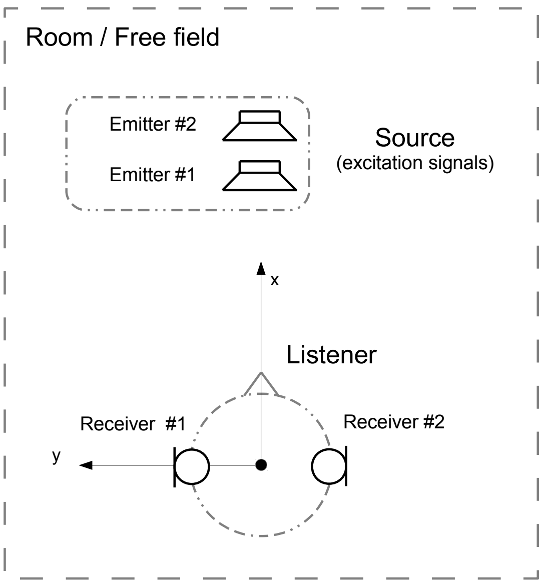

Figure 1: Typical HRTF measurement setup.

for multiple positions in a room, the position of the head varies and the relation between the head and the microphones does not change. Note that only one listener is considered.
- *Emitter* is any acoustic excitation used for the measurement. The number of emitters is not limited in SOFA. The contribution of the particular emitter is described by the metadata (see later).
- *Source* is the object incorporating all emitters. In SOFA, source might be a multi-driver loudspeaker (with the particular drivers as emitters), or a speaker array (with the particular speakers as emitters), or a choir (with the particular human as emitter), etc. Note that only one source is considered but the source may incorporate an unlimited number of emitters.
- *Room* is the volume enclosing the measurement setup. In the case of a free-field measurement, the room is not considered. An optional room description is considered for measurements performed in reverberant spaces, with a direct description of a simple shoebox, or with a link to a digital asset exchange file for a more complex description.

The information is stored in a numeric container, structured by the *measurement*. Measurement is a discrete sampled observation done at a specific time and under a specific condition. A measurement consists of data, e.g., a set of impulse responses, and is described by its corresponding dimensions and metadata. All measurements are stored in a single data structure, e.g., a matrix of IRs.

Further details on the general SOFA concept can be found in Ref #4, the currently valid specifications are provided at http://www.sofaconventions.org. Note that this document describes SOFA version 0.6.

3. SimpleFreeFieldHRIR

3.1 General

This conventions essentially defines the setup used to measure HRTFs in free field see Fig. 2. The measured HRTFs are represented as FIR filters, with a single HRTF set of a listener per file. We defined this conventions to describe data from databases like ARI[5], MIT KEMAR[6], LISTEN[7], CIPIC[3], SCUT[2], and others. Those measurements have been done under assumed free-field with a single excitation source assuming an omnidirectional loudspeaker.

In SimpleFreeFieldHRIR, the measured HRTFs are represented as FIR filters, with a single HRTF set of a listener per file. Human listeners or mannequins are considered and thus, the number of receivers is two. The default position of the listener is the origin of the coordinate system, facing in the direction of the y-axis, given in cartesian coordinates. The different HRTF directions are represented as different positions of the source. Thus, SourcePosition is given in spherical coordinates and varies corresponding to the different azimuth and elevation angles of the measurement directions and the distance of the source. Source consists of a single and omnidirectional emitter. Thus, SourceUp, and SourceView are not considered.

Note that usually, only the apparent azimuth and elevation angles are provided and this information is modeled as the variation of the source position. This conventions can, however, also be used to describe more complex data where the tilt of the head is varied or the listener is not exactly in the center of the measurement setup.

In the following, we describe the specifications of SimpleFreeFieldHRIR as the differences imposed to the general SOFA specifications.

3.2 Datatype and general attributes

The values of the following general attributes are fixed:
- SOFAConventions: “SimpleFreeFieldHRIR”
- SOFAConventionsVersion: “0.4”
- Datatype: “FIR”
- RoomType: “free field”

Thus, the datatype is fixed to “FIR”. Further, the following general attributes are required:

- DatabaseName: provides the name of the database. It is required in order to distinguish between databases and can be used for classification of the data. It must be unique across all databases available. Examples are "ARI", "LISTEN", "CIPIC", etc. DatabaseName is empty per default.
- ListenerShortName: provides the ID of the data set. It is required in order to distinguish between different data sets within a database. Thus, it must be unique within a database. Examples are "NH12", "003", "AK", etc. Default value for ListenerShortName is "unknown". Note that ListenerShortName is not a narrative short name of a listener.

3.3 Data

Datatype is "FIR", thus, Data.IR is a 3-dimensional matrix representing all measured impulse responses. The number of measurements M, the number of receivers R and the length of the impulse responses N define the size of the matrix dimensions, respectively.

3.4 Dimensions

The size of the dimension M and N in the variable Data.IR define the size of M and N for the rest of the data in the file, respectively. M and N can be arbitrary positive integers. The number of receivers R is fixed to 2. The number of emitters E is fixed to 1.

3.5 Objects

The listener is in the origin of the setup, facing in the direction of the x-axis. Thus, for default values, ListenerPosition = [0 0 0], ListenerView = [1 0 0], and ListenerUp = [0 0 1]. The coordinate type is Cartesian, thus ListenerPosition:Type = "cartesian", ListenerPosition:Units = "meter", ListenerView:Type = "cartesian" and ListenerView:Units = "meter"

Two omnidirectional receivers are considered placed on a head. Thus, ReceiverPosition = [0 -H 0; 0 +H 0] with H as the head radius, all given in Cartesian coordinates in meters. Thus, ReceiverPosition:Type = "cartesian" and ReceiverPosition:Units = "meter". Default for ReceiverPosition is H of 0.09 m.

Source position varies corresponding to the different azimuth and elevation angles of the measurement directions. Thus, SourcePosition is given in spherical coordinates. Per default, SourcePosition = [0 0 1] (azimuth 0°, elevation 0°, distance 1 m). SourceUp and SourceView are not considered.

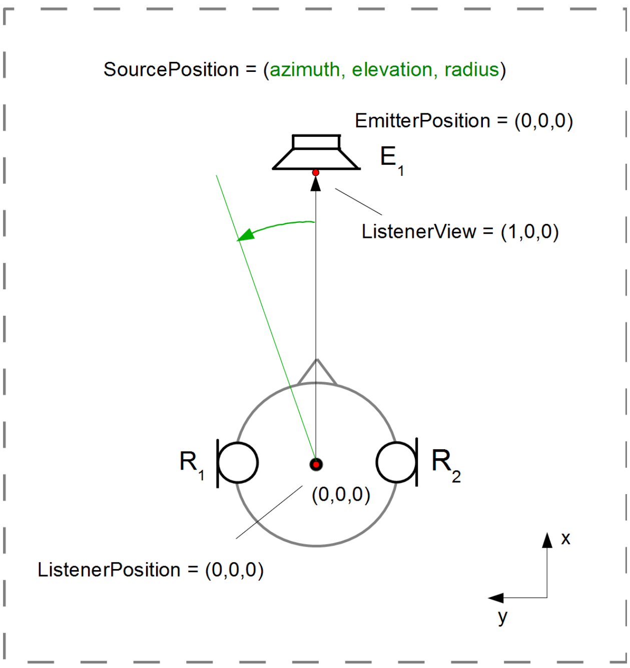

Figure 2: SimpleFreeFieldHRIR

Table 1: Summary of mandatory data in SimpleFreeFieldHRIR. See text for details.

Name	Value/Default	Dimensions	Type
GLOBAL:Conventions	**SOFA**		attribute
GLOBAL:Version	**0.6**		attribute
GLOBAL:SOFAConventions	**SimpleFreeFieldHRIR**		attribute
GLOBAL:SOFAConventionsVersion	**0.4**		attribute
GLOBAL:APIName			attribute
GLOBAL:APIVersion			attribute
GLOBAL:AuthorContact			attribute
GLOBAL:DataType	**FIR**		attribute
GLOBAL:License	No license provided, ask the author for permission		attribute
GLOBAL:Organization			attribute
GLOBAL:RoomType	**free field**		attribute
GLOBAL:DateCreated			attribute
GLOBAL:DateModified			attribute
GLOBAL:Title			attribute
ListenerPosition	[0 0 0]	IC, MC	double
ListenerPosition:Type	cartesian		attribute
ListenerPosition:Units	meter		attribute
ReceiverPosition	[0 -0.09 0; 0 0.09 0]	rCI, rCM	double
ReceiverPosition:Type	cartesian		attribute
ReceiverPosition:Units	meter		attribute
SourcePosition	[0 0 1]	IC, MC	double
SourcePosition:Type	spherical		attribute
SourcePosition:Units	degree, degree, meter		attribute
EmitterPosition	[0 0 0]	eCI, eCM	double
EmitterPosition:Type	cartesian		attribute
EmitterPosition:Units	meter		attribute
GLOBAL:DatabaseName			attribute
GLOBAL:ListenerShortName			attribute
ListenerUp	[0 0 1]	IC, MC	double
ListenerView	[1 0 0]	IC, MC	double
ListenerView:Type	cartesian		attribute
ListenerView:Units	meter		attribute
Data.IR	[0 0]	mRn	double
Data.SamplingRate	48000	I	double
Data.SamplingRate:Units	hertz		attribute
Data.Delay	[0 0]	IR, MR	double

The Source object consists of a single omnidirectional emitter in its origin. Thus, EmitterPosition = [0 0 0], and the Emitter's coordinate type is irrelevant.

3.6 Summary

Table 1 shows the summary of all data used in SimpleFreeFieldHRIR. The column “Name” lists all the data names. The prefix "GLOBAL:" indicates that the data is a global attribute. Attributes of a variable are described as "X:Y" where Y is the attribute of the variable X. All other data are variables, which type and allowed dimensions are given in the column “Type” and “Dimensions”, respectively. The dimensions are given according to the SOFA specifications. Lower case indicates that variable size in that dimension determines the dimension size in the file. Upper case indicates that variable must be of that dimension (or one of these dimensions). The column “Value/Default” shows in bold font, the values fixed by the conventions and in standard font, the default values to be subject of changes.

4. Conclusions

SOFA was particularly designed for efficient interchangeability and extendability of spatially oriented data like HRTFs. Thus, SOFA aims at implementation of the concept of reproducible research as far as possible. To this end, not only the specifications are free and open but SOFA is also the basis for the “AES-X212 HRTF file format standardization” project aiming at creating a standard by the AES. SOFA conventions are available at http://www.sofaconventions.org. SOFA API for Matlab and Octave is freely available at http://sourceforge.net/projects/sofacoustics/.

Many HRTFs are stored in SOFA already now. An example of a repository currently containing HRTFs and DRIRs from eight databases and (in total) of more than 200 listeners can be found at http://sofacoustics.org/data/database. SOFA is also the format for storing HRTFs in the Auditory Modeling Toolbox, available from http://amtoolbox.sourceforge.net/.

SOFA aims at encouraging researchers and developers of binaural-acoustics applications to offer open access for a wide exchange of spatially oriented data.

5. Acknowledgments

This study was supported by the Austrian Science Fund (FWF, P 24124-N13) and the German Research Foundation (DFG RA 2044/1-1).

REFERENCES

[1] Møller, H., Sørensen, M. F., Hammershøi, D. & Jensen, C. B. Head-related transfer functions of human subjects. *J Audio Eng Soc* **43**, 300–321 (1995).

[2] Xie, B. *Head-related transfer function and virtual auditory display*. (J. Ross Publishing: Plantatation, FL, 2013).

[3] Algazi, V. R., Duda, R. O., Thompson, D. M. & Avendano, C. The CIPIC HRTF database. (2001).

[4] Majdak, P., Carpentier, T., Nicol, R., Roginska, A., Suzuki, Y., Watanabe, K., *et al.* Spatially Oriented Format for Acoustics: A Data Exchange Format Representing Head-Related Transfer Functions. *Proc. 134th Conv. Audio Eng. Soc. AES* Convention Paper 8880 (2013).

[5] Majdak, P., Goupell, M. J. & Laback, B. 3-D localization of virtual sound sources: effects of visual environment, pointing method, and training. *Atten Percept Psychophys* **72**, 454–69 (2010).

[6] Gardner, W. G. & Martin, K. D. HRTF measurements of a KEMAR. *J Acoust Soc Am* **97**, 3907–3908 (1995).

[7] Warusfel, O. *LISTEN HRTF database*. (2003). [Online] available: http://recherche.ircam.fr/equipes/salles/listen/