7.4.7 HRTF and BRIR sets

7.4.7.1 HRTF and BRIR latency

In case of multi-channel input with LFE, LFE input is added to the binaural output using a delay line and a gain:

$$O^{l,r} = O^{l,r} + G_{lfe} * delay(I_{lfe}, binraural\_latency\_s)$$

I_{lfe} is the LFE input signal.

The default values:

- $G_{lfe} = 10^{5.5/20} \approx 1.88364911$
- The $binraural\_latency\_s$ includes the HRTF/BRIR group delay and the processing delay. The default values are:
  - For Fastconv binaural renderer:
    - HOA to binaural filters, $binraural\_latency_s = 0.000020834\text{s}$
    - MC to binaural filters (HRIR), $binraural_{latency_s} = 0.000020834s$
    - MC to binaural filters (BRIR), $binraural\_latency_s = 0.000937500s$
  - For TD binaural binaural renderer:
    - HRIR, $binraural\_latency_s = 0.000020834$
  - For Crend binaural renderer:
    - HOA to binaural filters, $binraural\_latency_s = 0.000020833s$
    - MC to binaural filters (HRIR), $binraural\_latency_s = 0.000666667s$
    - MC to binaural filters (BRIR), $binraural\_latency_s = 0.000145834s$

During conversion HRTF/BRIR conversion process the binaural filter shall be estimated.

7.4.7.2 Parametrization of Binaural renderers using binary file

Head related filters for the binaural rendering may be provided to the decoder or the renderer by using dynamic loading of external binary file. The way to generate such a binary file from a set of SOFA file is described in [9].

The decoder program should be called with option -hrtf <binary_file>. This option can be used with the output configurations BINAURAL, BINAURAL_ROOM_IR and BINAURAL_ROOM_REVERB.

A binary file has a specific container format with a header and a sequence of entries.

The header of a binary file is defined according to table 7.4-2 as follows:

Table 7.4-2: Binary file header

Offset	Format	Length (in bytes)	Description
0	string	8	File identifier: “IVASHRTF”
8	integer	4	Size of file in bytes (header of file included)
12	integer	2	Number of entries (HR filters)
14	integer	4	Max size of raw data (HR filter in binary format)

Every entry contains a header followed by the related raw data which is the binary representation of the HR filter.

The header of each entry is defined as given in table 7.4-3:

Table 7.4-3: Entry headers

Offset	Format	Length (in bytes)	Description
0	integer	4	Renderer type The renderer type is defined according to the enumeration RENDERER_TYPE among the following values : - HRTF_READER_RENDERER_BINAURAL_FASTCONV - HRTF_READER_RENDERER_BINAURAL_FASTCONV_ROOM - HRTF_READER_RENDERER_BINAURAL_PARAMETRIC - HRTF_READER_RENDERER_BINAURAL_OBJECTS_TD - HRTF_READER_RENDERER_BINAURAL_MIXER_CONV - HRTF_READER_RENDERER_BINAURAL_MIXER_CONV_ROOM - HRTF_READER_RENDERER_REVERB_ALL
4	integer	4	Input audio configuration The input audio configuration is defined according to the enumeration BINAURAL_INPUT_AUDIO_CONFIG among the following values : - BINAURAL_INPUT_AUDIO_CONFIG_COMBINED - BINAURAL_INPUT_AUDIO_CONFIG_HOA3 - BINAURAL_INPUT_AUDIO_CONFIG_HOA2 - BINAURAL_INPUT_AUDIO_CONFIG_FOA - BINAURAL_INPUT_AUDIO_CONFIG_UNDEFINED
8	integer	4	Sampling frequency (16000, 32000, 48000)
12	integer	4	Raw data size in bytes

The format of the raw data depends on the rendering and the HR filters are represented in floating point.

Note:

- The HR filters for the renderer types HRTF_RENDERER_BINAURAL_PARAMETRIC , HRTF_RENDERER_BINAURAL_FASTCONV and HRTF_RENDERER_BINAURAL_FASTCONV_ROOM are fully defined at 48kHz.
- For the renderer type HRTF_RENDERER_BINAURAL_OBJECTS_TD the input audio configuration is always BINAURAL_INPUT_AUDIO_CONFIG_UNDEFINED.
- renderer type HRTF_READER_RENDERER_REVERB_ALL should be associated with HRTF_READER_RENDERER_BINAURAL_OBJECTS_TD and/or HRTF_READER_RENDERER_BINAURAL_CREND to specify the binaural reverberation parameters jointly with new HRIR parameters. They shall be computed on the same HRIR set.
- The binary file does not have to contain all data (HR filter) for all renderers. The following minimal configurations are accepted or any combination of those :

HRTF_READER_RENDERER_ BINAURAL_FASTCONV	BINAURAL_INPUT_AUDIO_CONFIG_ COMBINED	Contains data for combined HRIR

HRTF_READER_RENDERER_BINAURAL_FASTCONV	BINAURAL_INPUT_AUDIO_CONFIG_COMBINED	Contains data for Combined HRIR
HRTF_READER_RENDERER_BINAURAL_FASTCONV	BINAURAL_INPUT_AUDIO_CONFIG_HOA3	Contains data for HOA3
HRTF_READER_RENDERER_BINAURAL_FASTCONV	BINAURAL_INPUT_AUDIO_CONFIG_HOA2	Contains data for HOA2
HRTF_READER_RENDERER_BINAURAL_FASTCONV	BINAURAL_INPUT_AUDIO_CONFIG_FOA	Contains data for FOA
HRTF_READER_RENDERER_BINAURAL_FASTCONV_ROOM	BINAURAL_INPUT_AUDIO_CONFIG_COMBINED	Contains data for combined BRIR
HRTF_READER_RENDERER_BINAURAL_PARAMETRIC	BINAURAL_INPUT_AUDIO_CONFIG_HOA3	Contains data for HOA3, HOA2, FOA and reverberation from BRIR
HRTF_READER_RENDERER_BINAURAL_OBJECTS_TD	BINAURAL_INPUT_AUDIO_CONFIG_UNDEFINED	Contains data for HRIR
HRTF_READER_RENDERER_BINAURAL_CREND	BINAURAL_INPUT_AUDIO_CONFIG_COMBINED	Contains data for combined HRIR
HRTF_READER_RENDERER_BINAURAL_CREND	BINAURAL_INPUT_AUDIO_CONFIG_HOA3	Contains data for HOA3
HRTF_READER_RENDERER_BINAURAL_CREND	BINAURAL_INPUT_AUDIO_CONFIG_HOA2	Contains data for HOA2
HRTF_READER_RENDERER_BINAURAL_CREND	BINAURAL_INPUT_AUDIO_CONFIG_FOA	Contains data for FOA
HRTF_READER_RENDERER_BINAURAL_CREND_ROOM	BINAURAL_INPUT_AUDIO_CONFIG_COMBINED	Contains data for combined BRIR only (BINAURAL_ROOM_IR)
HRTF_READER_RENDERER_REVERB_ALL	BINAURAL_INPUT_AUDIO_CONFIG_UNDEFINED	Contains data for HRIR with reverberation (BINAURAL_ROOM_REVERB)when TD renderer or mixerconv are used

Table 7.4-4: HR filters for binaural renderer Fastconv Impulse response binary entries:
`HRTF_READER_RENDERER_BINAURAL_FASTCONV`

Offset	Format	Length (in bytes)	Description
0	integer	4	Scaling factor for latency value
4	integer	4	Latency value*
8	integer	2	Number of Binaural convolution bands (Nb)
10	integer	2	Number of channels (Nc)
12	integer	2	Number of taps per filter (Nt)
14	integer	2	Scaling factor for filters taps
16	integers	4 * Nb * Nc * Nt	Left ear real taps values*
16 + 4 * Nb * Nc * Nt	integers	4 * Nb * Nc * Nt	Left ear imaginary taps values*
16 + 2 * 4 * Nb * Nc * Nt	integers	4 * Nb * Nc * Nt	Right ear real taps values*
16 + 3 * 4 * Nb * Nc * Nt	integers	4 * Nb * Nc * Nt	Right ear imaginary taps values*

Table 7.4-5: HR filters for binaural renderer Fastconv Room Impulse Response binary entries:
`HRTF_READER_RENDERER_BINAURAL_FASTCONV_ROOM`

Offset	Format	Length (in bytes)	Description
0	integer	4	Scaling factor for latency value
4	integer	4	Latency value*
8	integer	2	Number of Binaural convolution bands (Nb)
10	integer	2	Number of channels (Nc)
12	integer	2	Number of taps per filter (Nt)
14	integer	2	Scaling factor for filters taps
16	integers	4 * Nb * Nc * Nt	Left ear real taps values*
16 + 4 * Nb * Nc * Nt	integers	4 * Nb * Nc * Nt	Left ear imaginary taps values*
16 + 2 * 4 * Nb * Nc * Nt	integers	4 * Nb * Nc * Nt	Right ear real taps values*
16 + 3 * 4 * Nb * Nc * Nt	integers	4 * Nb * Nc * Nt	Right ear imaginary taps values*
16 + 4 * 4 * Nb * Nc * Nt	integer	2	CLDFB max number of channels (Nm)
18 + 4 * 4 * Nb * Nc * Nt	integer	2	Scaling factor for reverberation time values
20 + 4 * 4 * Nb * Nc * Nt	integers	2 * Nm	reverberation time values*
20 + 4 * 4 * Nb * Nc * Nt + 2 * Nm	integer	2	Scaling factor for energies corrections values
22 + 4 * 4 * Nb * Nc * Nt + 2 * Nm	integers	2 * Nm	Energies corrections values *

Table 7.4-6: HR filters for binaural renderer parametric:
`HRTF_READER_RENDERER_BINAURAL_PARAMETRIC`

Offset	Format	Length (in bytes)	Description
0	integer	2	Number of channels (Nc)
2	integer	2	Number of bins (Nb)
4	integer	2	Scaling factor for filters taps
6	integers	2 * 2 * Nc * Nb	Real taps values* one for each ear
6 + 2 * 2 * Nc * Nb	integers	2 * 2 * Nc * Nb	Imaginary taps values* one for each ear
6 + 2 * 2 * 2 * Nc * Nb	integer	2	CLDFB max number of channels (Nm)
8 + 2 * 2 * 2 * Nc * Nb	integer	2	Scaling factor for reverberation time values
8 + 2 * 2 * 2 * Nc * Nb	integers	2 * Nm	reverberation time values*
8 + 2 * 2 * 2 * Nc * Nb + 2 * Nm	integer	2	Scaling factor for energies corrections values
10 + 2 * 2 * 2 * Nc * Nb + 2 * Nm	integers	2 * Nm	Energies corrections values *
10 + 2 * 2 * 2 * Nc * Nb + 4 * Nm	integer	2	Scaling factor for early part energies corrections values
12 + 2 * 2 * 2 * Nc * Nb + 4 * Nm	integers	2 * Nm	Early part energies corrections values *

Table 7.4-7: HR filters for binaural renderer Crend entries:
`HRTF_READER_RENDERER_BINAURAL_MIXER_CONV` or
`HRTF_READER_RENDERER_BINAURAL_MIXER_CONV_ROOM`

Offset	Format	Length (in bytes)	Description
0	integer	4	Scaling factor for latency value
4	integer	4	Latency value*
8	integer	2	Number of HRIR/BRIR (Nc)
10	integer	2	Number of Binaural channels (Nb)
12	integer	2	Max number of block iterations (Ni)
14	integers	2 * Nc * Nb * Ni	Max frequency value for each block
14 + 2 * Nc * Nb * Ni	integer	2	Max number of block iterations for diffuse blocks
16 + 2 * Nc * Nb * Ni	integers	2 * Nb	Num of iterations for diffuse part (Nid)
16 + 2 * Nc * Nb * Ni + 2 * Nb	integers	2 * Nb * Nid	Max frequency value for each block of diffuse part
16 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid	integer	2	Max frequency value over all diffuse blocks
18 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid	integer	2	Scaling factor for inverse diffuse weight values
18 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid	integers	2 * Nc	Left ear inverse diffuse weight values*
18 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * Nc	integers	2 * Nc	Right ear inverse diffuse weight values*
18 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * 2 * Nc	integer	4	Max number of bins over all HRIR/BRIR for direct part (Nf)
22 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * 2 * Nc	integer	2	Scaling factor for filters taps
24 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * 2 * Nc	integers	4 * Nb * Nc * Nf	Direct part real taps values*
24 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * 2 * Nc + 4 * Nb * Nc * Nf	integers	4 * Nb * Nc * Nf	Direct part imaginary taps values*
24 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * 2 * Nc + 2 *4 * Nb * Nc * Nf	integer	4	Max number of bins over all HRIR/BRIR for diffuse part (Nfdiff)
28 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * 2 * Nc + 2 *4 * Nb * Nc * Nf	integers	4 * Nb * Nc * Nfdiff	Diffuse part Real taps values*
28 + 2 * Nc * Nb * Ni + 2 * Nb + 2 * Nb * Nid + 2 * 2 * Nc + 2 *4 * Nb * Nc * Nf + 4 * Nb * Nc * Nfdiff	integers	4 * Nb * Nc * Nfdiff	Diffuse part imaginary taps values*

Table 7.4-8: HR filters for reveberation renderer: `HRTF_READER_RENDERER_REVERB_ALL`

Offset	Format	Length (in bytes)	Description
0	integer	2	Scaling factor for energies and coherences values
2	integers	2 * Nfft[1]	Left energies values *
2 + 2 * Nfft[1]	integers	2 * Nfft[1]	Right energies values *
2 + 2 * 2 * Nfft[1]	integers	2 * Nfft[1]	Coherence values *

(1) Nfft = 129 if sample rate = 16 kHz, 257 otherwise

(*) : for floating point binaural renderer implementation, the floating point value $valf$ is computed as follow :
$valf = vali * 2^{-Q}$ with $vali$