Individualized HRTFs for playing VR videos with Ambisonics spatial audio on HMDs

M. Binelli[1], D. Pinardi[1], T. Nili[2] and A. Farina[1]

[1] University of Parma, Department of Engineering and Architecture (Parco Area delle Scienze 181/A, 43124 Parma, Italy) – [2] ASK Industries S.p.A. (Via Togliatti 46/A, 42020 Montecavolo, Italy)

Abstract

Current audio/video head-mounted rendering systems for virtual and augmented reality rely on a binaural approach combined with Ambisonics technology. These head-tracking systems employ generic HRTFs commonly measured with a dummy head in anechoic room.
Here, we describe a new solution designed to play 360 video with spatial audio, developed for PC and standalone platforms and built from existing open source software.
The developed tools can **load HRTF** sets from a standard audio file chosen in an existing database or from an ad-hoc measurement. The capability to **switch HRTF** sets while playing files has been added.

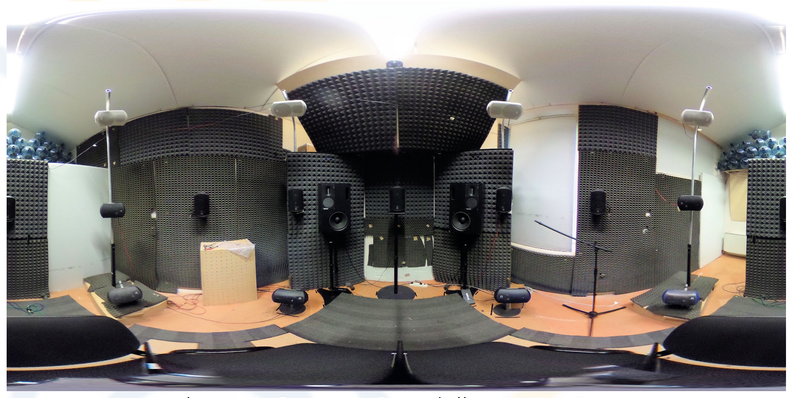
Fast HRTF measurement

The experimental measurements were carried out at Casa della Musica, Parma (Italy), in an Ambisonics Room equipped with:

- 16 "Turbosound Impact 50" loudspeakers arranged in a 4-8-4 configuration
- 2 studio monitors "Genelec S30D" employed as subwoofers (50-120 Hz)

Loudspeakers have been pre-equalized with the Kirkeby inversion method, for flattening the spectrum and aligning the phase in the measurement position.
The test signal (Exponential Sine Sweep of 10s), played sequentially by each loudspeaker, is overlapped by a factor 0.8, so the total measurement time is **just 42s**.

The reference HRTF set has been measured with a Neumann KU100 dummy head; human HRTFs were measured wearing EXOFIELD® microphones.



Ambisonics Room at Casa della Musica, Parma

Equipment

Vive Cinema (PC solution)
+ Open-source
+ Multi-HMDs (HTC Vive, Oculus Rift, Samsung Odyssey)
+ Up to 3rd order Ambisonics

Unity3D (standalone solution)
+ Easy-to-start
+ Well supported by community
+ Multi-target (Facebook Oculus Go, Samsung Gear VR, etc.)
– Limited to 1st order Ambisonics
– Cannot extract audio from a video container

Google Resonance Audio (standalone solution)
+ Open-source
+ Multi-environment (Unity, Unreal Engine and others)
+ Ready for 3rd order Ambisonics

JVCKENWOOD EXOFIELD®
+ In-ear microphones
+ Currently best solution for individualized HRTFs measurement

Neumann KU100
+ Worldwide reference dummy head
+ High-quality, pre-eq microphones
– Not really In-Ear microphones, impeding to equalize with in-ear headphones

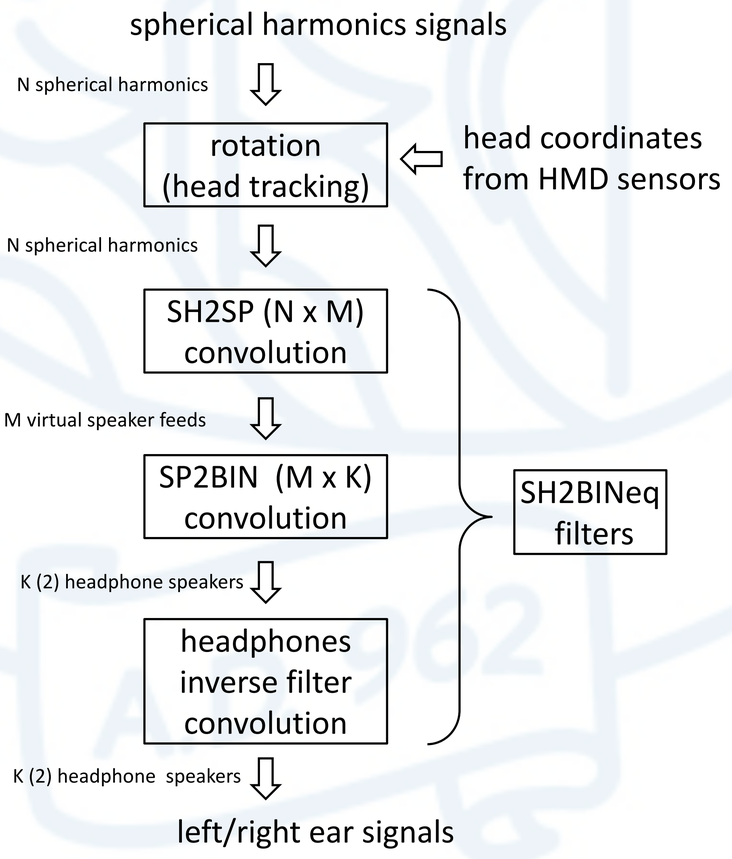
Oculus GO (Android based)
+ Good quality standalone HMD
+ Possibility to use external high quality headphones
+ Cheap and portable
– Android limitations

Samsung Odyssey
+ Currently top quality HMD
– Expensive, powerful PC required (not standalone)
– Unsatisfactory, not-replaceable headphones

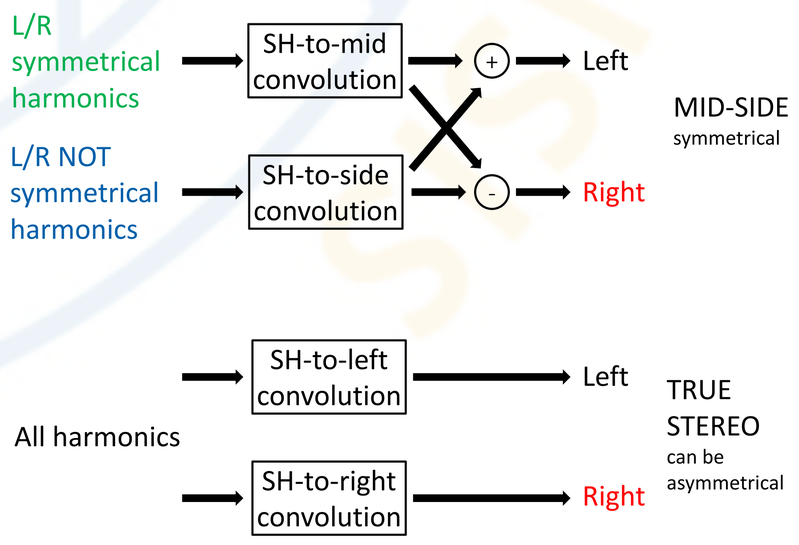
Headphones Inverse Filters

For avoiding coloration due to transducers employed, an equalization procedure is mandatory. This is obtained placing the headphones over the head (either dummy or human, in this case wearing the EXOFIELD® microphones). The impulse responses are measured with an exponential sine sweep and an inverse filter for each ear is generated by using the Kirkeby method.
This method allows to remove the effect both of the headphones and of the microphones, providing a great improvement, particularly with the integrated, low-quality headphones embedded with the HMD.

Signal Flow and Results



The desktop solution based on VIVE Cinema has been recompiled with a **TRUE-STEREO** convolution instead of the native solution, which is called **MID-SIDE** convolution. The last is based on the assumption that HRTF are perfectly symmetrical, which in general is not true.



The «Spherical2BinauralEqualized» (**SH2BINeq**) filters are stored in multichannel wav files and loaded by our software. They are obtained by convolving a decoding matrix (**SH2SP**) provided by an external Ambisonics decoder with the measured HRTF set (**SP2BIN**) and the headphone inverse filters.

Due to computational limits the suboptimal **MID-SIDE** convolution has been kept for the standalone solution based on Resonance Audio engine and running on Gear VR and Oculus Go devices, due to the limited computational power available on these devices. This limitation will be removed in a future development.

Contacts

Angelo Farina, Professor
e-mail: angelo.farina@unipr.it
Marco Binelli, Ph.D.
e-mail: marco.binelli@unipr.it
Download link
http://www.angelofarina.it/Public/ViveCinema

Acknowledgements

This work was supported by the Italian Ministry of Economic Development (MISE), FUND FOR THE SUSTAINABLE GROWTH (F.C.S) under grant agreement (CUP) B48I15000130008, project VASM ("Vehicle Active Sound Management").
The authors want to express their profound gratitude to JVCKENWOOD Corporation for making available the EXOFIELD® technology, which allowed to measure quickly and reliably human HRTFs and to perform accurate, individualized headphones equalization.
Special thanks to Hisako Shinbara, JVCKENWOOD Corporation, Tokyo, Japan.