Audio Engineering Society

Conference e-Brief

Presented at the International Conference on
Immersive and Interactive Audio
2019 March 27 – 29, York, UK

HMDiR: an HRTF dataset measured on a mannequin wearing XR devices

Andrea Genovese and Agnieszka Roginska

New York University - Music and Audio Research Lab

Correspondence should be addressed to Andrea Genovese (`genovese@nyu.edu`)

ABSTRACT

This document covers the release of the open-source HMDiR dataset (Head-Mounted-Display acoustic Impulse Responses) of HRTFs, useful to study the occlusion effect of wearing XR devices on the auditory perception. The data was collected for a previous publication in which the effect of wearing HMD gear on the HRTFs of a mannequin was described. This document covers in detail the measurement procedure, equipment, and specifications, including instructions on how to download the data files. The measurement library includes a free-head case (no HMD), two mixed reality headset cases, and three virtual reality headsets, chosen among those commercially available.

1 Introduction

The field of spatial audio makes use of *Head-Related-Transfer-Functions* (HRTFs), FIR filter pairs used to synthesize the perceptual localization of a sound source. These location-dependent filters are usually measured by recording a measurement signal using small electret microphone capsules placed in the ear canals of either a dummy plastic head (generic HRTFs) or a human subject (individual HRTFs) in an anechoic environment. Novel immersive technologies such as Augmented Reality (AR), Mixed Reality (MR) and Virtual Reality (VR) use HRTFs to simulate auditory objects superimposed within the current auditory environment, or within a virtual soundscape.

Several HRTF measurements databases have been publicly released, each with the intent of capturing a different effect or variation. Recent releases have been focused on the capture of inter-subject deviations [1, 2, 3], distance effect [4, 5], speaker configurations [6], room effect [7] and others.

The purpose of the HMDiR database (HMD impulse Responses), is to address the effect of Head-Mounted-Displays (HMD) occlusion on the acoustic path from source to ear. Due to their size, shape and material, HMDs often obstruct the direct path to the ear causing acoustic perturbations in the form of constructive reflections or dampening effects [8]. This is relevant for the application of Mixed Reality, where acoustic sources and virtual sources are desired to blend into a single coherent auditory scene. To a user wearing an HMD the acoustic character of real sound sources may be offset by the perturbations created by the rigid body. It may then prove beneficial to match the signal perturbations experienced by real sources on the virtual audio objects, or perhaps work on diminishing the perturbation effects. However these issues need to yet be explored in depth.

By measuring a set of HRTFs on a plastic head wearing such devices, it is possible to learn about observable differences to a "free-head" case and further investigate the perceptual impact of these perturbations. As of the time of publication of this e-brief no perceptual listening test has been run on this data. The data is therefore made public to the audio scientific community to both encourage such listening experiments, and to provide additional HRTF data to the current academic pool.

2 Measurement Procedure

The NYU Dolan ISO booth (Research Lab) was selected as measurement location due to its semi-anechoic isolation condition. The T_{60} reverberation time of this rectangular room of size 4.5×3.7×2.5 mt was measured to be 0.14 seconds. Six Genelec 8030A loudspeakers were arranged as a spiral array at six elevation levels starting from a low of -36° up to 54° in steps of 18°. The loudspeakers were placed at a radial distance of 1 meter to the binaural dummy head.

An automated rotated mount was implemented in order to achieve a high-resolution grid of measurement without the hand of an experimenter having to move equipment components. An *ARDUINO UNO* microcontroller, connected to an *Adafruit Motor Shield V2*, was interfaced to ScanIR and a compatible stepper motor [9]. The stepper motor resolution was of 1.8° steps for a total of 200 steps of azimuth rotation. By rotating the measurement head in relation to the spiral setup a full sphere of source positions can be achieved. A MOTU UltraLite MK3 was used as interface between the binaural recording head and the destination computer.

A Neumann KU100 dummy head was secured on top of the rotating mount. At each round of measurement, a different HMD device was fitted to the head (best ergonomic fit mount), a no-HMD case was also measured. A total of 1200 source positions per HMD case were measured (200 azimuths × 6 elevation levels).

All measurements were carried using the *ScanIR* MATLAB software tool [10, 11] which relies on the Psychtoolbox package [12]. A two-seconds logarithmic sinesweep [13] was chosen as measurement signal, using a sampling rate set at 96 kHz. To improve the SNR, each sinesweep was played twice by each of the six loudspeakers in succession, after which the rotating motor would move by a step to repeat the procedure at a new azimuth orientation.

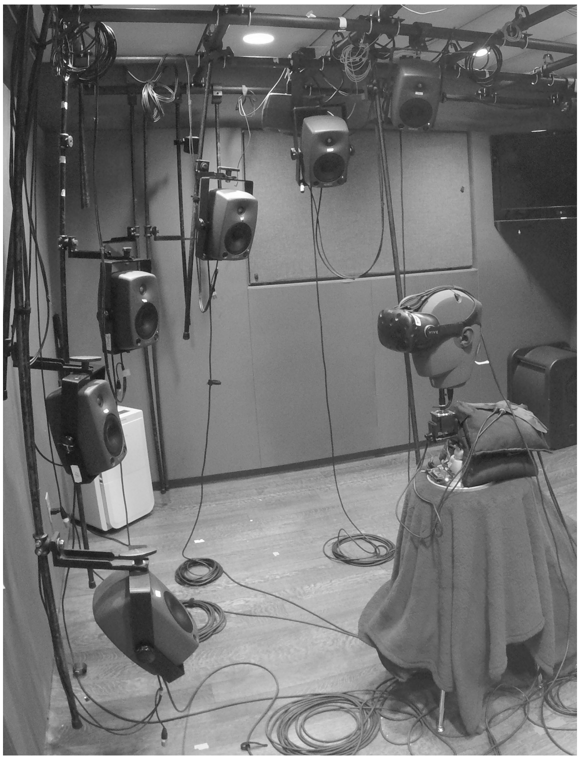

Fig. 1: Measurement setup spiral for the HTC Vive HMD measurement

3 Database Specifications

Post-processing All measurements are available as minimum phase time-domain filters in MATLAB (*.mat*) format, one per HMD case. Recordings were wrapped for correct angle alignment and formatted in two different versions (see below).To smooth occasional reflections, the time-domain IRs are truncated to 256 samples and scaled by the second half of a Tukey window with a 25% cosine taper amount [14].

HMDs Six Head-mounted-display cases were collected, all items were the latest HMD versions as of December 2017:

- Free-head (no HMD)
- Microsoft Hololens (2016 version)
- Metavision Meta-2
- Oculus Rift CV1
- HTC Vive
- Samsung GearVR SM-R322

Source Positions A total of 1200 angle positions per measured HMD have been collected:

- Six elevation levels: [$+54°$, $+36°$, $+18°$, $0°$, $-18°$, $-36°$]
- 200 azimuth positions from 0° to 358.2° in steps of 1.8°

Download The database is available on *Zenodo.org* (DOI 10.5281/zenodo.2558629) using the following link: `https://doi.org/10.5281/zenodo.2558629`. All data is publicly available under the Creative Commons Attribution 4.0 International license. Please cite the present publication for any use of this dataset. It is recommended to process the data in MATLAB.

Each HMD case is available in two *.mat* formats:

- **MARL Format:** complete HRIR individual-entry-based structure, includes metadata information for elevation, azimuth, distance, headset type, location, measurement signal and sample rate
- **Compact Format:** simplified matrix for only HRIR data. Packed as a 6×200 cell matrices with rows spanning elevations from $-36°$ to $+54°$ from top-to-bottom. Columns spanning azimuth positions from 0° to 358.2° (clockwise) left-to-right.

4 Acoustic Perturbations

In order to inspect the occlusion effects of the HMD rigid bodies, all measurements related to the mixed reality devices (Hololens and Metavision) have been analyzed and discussed in a previous publication [8]. In that study, objective signal differences between the free-head case and the other HMD cases were analyzed in order to identify regions of acoustic perturbations in terms of angle of incidence and frequency band, as well as an investigation of distortions of spatial cues.

While no particular ITD deviation was found among the different cases, spectral distortions (SD) were found to be potentially disrupting for listeners. Particularly for contralateral source locations (source located at the opposite hemisphere from the reference ear), SD levels as high as 10dB were found at the band between 3 to 6 kHz, along HRTF notch deviations around 7 kHz. Lower elevations of incidence were found to cause some energy boost due to reflections with the HMD rigid body, while higher elevations caused dampenings due to obstructed paths to the ears. For more details on the observed distortions please refer to the mentioned publication.

Virtual reality headsets had not been investigated in such detail given the different intended paradigm of use; while in mixed and augmented reality it is desired to blend the real with the virtual, in VR the user is completely shielded from the local environment. It is usually the case, in VR, that only virtual sources are reproduced, making the case for acoustic matching not necessary. However the VR HMDs were recorded for reference and included in the dataset for public use.

Perceptual studies that can validate the hypothesized effects of the observations, on this data, and their impact on source localization and source coloration are yet to be performed. It is still not yet clear whether acoustic matching would improve the auditory perception. However, similar studies found a decrease of localization accuracy and timbral quality on subjects wearing HMDs in the absence of visual cues, especially at contralateral angles and higher frequencies [15, 16], vertical localization was also affected.

5 Summary

This document accompanies the release of HMDiR, a new HRTF database measured on a dummy head wearing different commercially available HMDs used for mixed reality and virtual reality. Past work on this data has previously shown that wearing HMDs can affect the acoustic field around the head due to occlusion and refraction effects. The document focuses on the measurement procedure and data specifications to support and guide interested scientists. The open-source data, downloadable from Zenodo, can be used for further analysis on occlusion effects, perceptual listening tests or simply as HRTFs sets within applications.

6 Acknowledgements

The authors would like to thank the students who helped to collect the measurements described. In particular Gabriel Zalles for the help in sitting through long measurement hours and helping to develop the rotating motor tools used within the ScanIR software.

References

[1] Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C., "The cipic hrtf database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, pp. 99–102, IEEE, 2001.

[2] Warusfel, O., "Listen HRTF database," *online, IRCAM and AK, Available: http://recherche. ircam. fr/equipes/salles/listen/index. html*, 2003.

[3] Gupta, N., Barreto, A., Joshi, M., and Agudelo, J. C., "HRTF database at FIU DSP lab," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 169–172, IEEE, 2010.

[4] Qu, T., Xiao, Z., Gong, M., Huang, Y., Li, X., and Wu, X., "Distance dependent head-related transfer function database of KEMAR," in *Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on*, pp. 466–470, IEEE, 2008.

[5] Wierstorf, H., Geier, M., and Spors, S., "A free database of head related impulse response measurements in the horizontal plane with multiple distances," in *Audio Engineering Society Convention 130*, Audio Engineering Society, 2011.

[6] Kearney, G. and Doyle, T., "An HRTF database for virtual loudspeaker rendering," in *Audio Engineering Society Convention 139*, Audio Engineering Society, 2015.

[7] Jeub, M., Schafer, M., and Vary, P., "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *Digital Signal Processing, 2009 16th International Conference on*, pp. 1–5, IEEE, 2009.

[8] Genovese, A., Zalles, G., Reardon, G., and Roginska, A., "Acoustic perturbations in HRTFs measured on Mixed Reality Headsets," in *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, Audio Engineering Society, 2018.

[9] Zalles, G., Kamel, Y., Anderson, I., Lee, M. Y., Neil, C., Henry, M., Cappiello, S., Mydlarz, C., Baglione, M., and Roginska, A., "A Low-Cost, High-Quality MEMS Ambisonic Microphone," in *Audio Engineering Society Convention 143*, Audio Engineering Society, 2017.

[10] Boren, B. and Roginska, A., "Multichannel impulse response measurement in matlab," in *Audio Engineering Society Convention 131*, Audio Engineering Society, 2011.

[11] Vanasse, J., Genovese, A., and Roginska, A., "Multichannel impulse response measurement in matlab: an update on ScanIR," in *Audio Engineering Society Conference: AES 2019 Conference on Immersive and Interactive Audio*, Audio Engineering Society, 2019.

[12] Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., et al., "What's new in Psychtoolbox-3," *Perception*, 36(14), p. 1, 2007.

[13] Farina, A., "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio Engineering Society Convention 108*, Audio Engineering Society, 2000.

[14] Treeby, B. E., Pan, J., and Paurobally, R. M., "The effect of hair on auditory localization cues," *The Journal of the Acoustical Society of America*, 122(6), pp. 3586–3597, 2007.

[15] Ahrens, A., Lund, K. D., Marschall, M., and Dau, T., "Sound source localization with varying amount of visual information in virtual reality," *bioRxiv*, 2018, doi:10.1101/489484.

[16] Gupta, R., Ranjan, R., He, J., and Woon-Seng, G., "Investigation of effect of VR/AR headgear on Head related transfer functions for natural listening," in *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, Audio Engineering Society, 2018.