AN IMMERSIVE 3D AUDIO HEADSET FOR VIRTUAL AND AUGMENTED REALITY

Nguyen Duy Hai[1], Santi Peksi[1], Rishabh Ranjan[3], Jianjun He[2], Boon Siang Tan[4], Rishabh Gupta[1], Woon-Seng Gan[1];

[1]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore; [2]Maxim Integrated, San Jose, United States; [3]Immerzen Labs Pte. Ltd., Singapore; [4]Sivantos Pte. Ltd, Singapore

DEMO 2.1

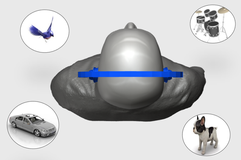
Motivation

Individualized Head Related Transfer Function (iHRTF)

"One's unique audio imprint"

With generic HRTF

Reproduce wrong or mismatched audio cues where brain could not relate to the real situation, resulting the sound to be perceived inside the head or front-back confusions.



With iHRTF

Reproduce the exact audio cues that trick a person's brain into a "being there" auditory illusion with accurate directional perception and externalized sound.

"Existing HRTF individualization techniques such a ***acoustical*** measurements in anechoic chamber, ***anthropometric*** measurements, or ***listening*** and evaluation either require **tedious** measurement, **training** or result in **degraded** performance"

This motivates us to develop a technique that:

Real-time processing

Allow unconstrained continuous head movements with direction recorded automatically.

Accurate measurement

Measure HRTFs with low error and good consistency as conventional static method.

Immersive/Interactive 3D sound

Personalized HRTFs create the most immersive and interactive 3D sound used in multimedia, VR/AR applications.

Easy setup for in-situ measurement

Binaural microphone, loud speaker, soundcard and VR/AR headset. That's all you need for the in-situ HRTF measurement.

High Resolution

Obtain HRTF at fine resolution up to 1 degree, providing smoother perception for dynamic source rendering.

Natural Sound Playback

Using the patented (US 9,357,282), 3D audio headphones to enhance the sensation of "being there".

Reference and Acknowledgment

References

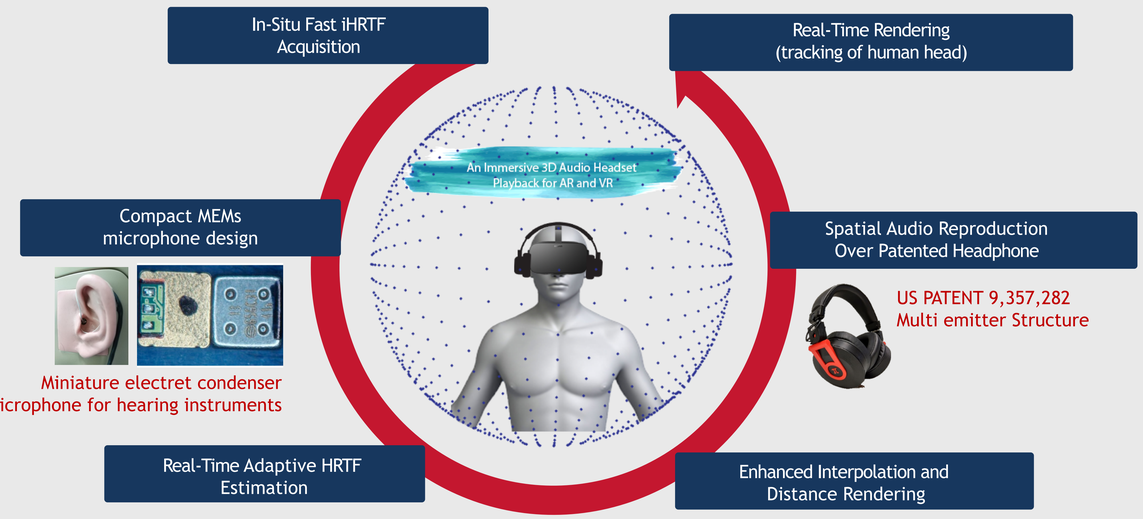
1. J. He, R. Ranjan, and W. S. Gan, "Fast continuous HRTF acquisition with unconstrained movements of human subjects," in Proc. ICASSP, Shanghai, China, Mar. 2016, pp. 321-325.
2. R. Ranjan, J. He, and W. S. Gan, "Fast Continuous Acquisition of HRTF for Human Subjects with Unconstrained Random Head Movements in Azimuth and Elevation " AES Conference on Headphone Technology, Aalborg, Denmark, 2016 Aug 24-26.
3. K. Sunder, J. He, E. L. Tan, and W. S. Gan, "Natural sound rendering for headphones: Integration of signal processing techniques," IEEE Sig. Process. Mag., vol. 32, no. 2, Mar 2015, pp. 100-113.
4. J. He, E. L. Tan, and W. S. Gan, "Primary-ambient extraction using ambient spectrum estimation for immersive spatial audio reproduction," IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 9, pp. 1430-1443, Sept. 2015.
5. R. Ranjan, and W. S. Gan. "Natural listening over headphones in augmented reality using adaptive filtering techniques." IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP) 23.11 (2015): 1988-2002.
6. K. Sunder, E. L. Tan, and W. S. Gan, "Individualization of binaural synthesis using frontal projection headphones," J. Audio Eng. Soc., vol. 61, no. 12, pp. 989-1000, Dec. 2013.

Acknowledgement

We would like to thank Facebook, Inc. and Sivantos Pte. Ltd. for their generous donation of Oculus Rift and the binaural microphones, respectively.

Scan Me for More Information

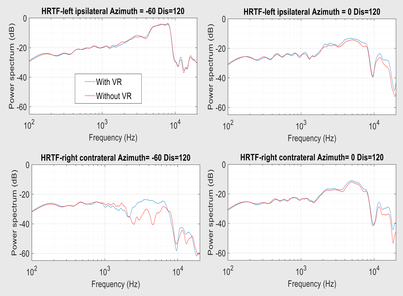
Integration of Acquisition and Rendering



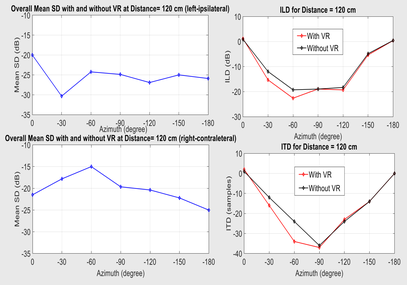
Results

Comparison between standalone Head Tracker and VR Headgear

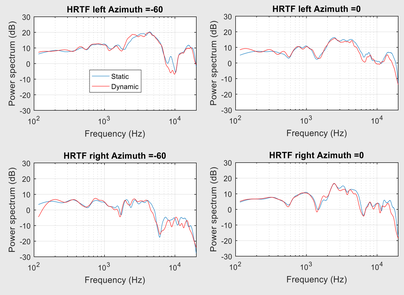
- **HRIR/HRTF for HATO aligned, az=0,-60, el=0 With and Without VR**



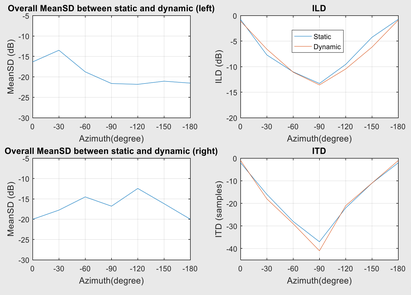
- **Overall Mean SD, ILD, ITD at 120cm, variable az With and Without VR**



- **HRIR/HRTF for HATO aligned, az=0,-60, el=0 Static and Dynamic**

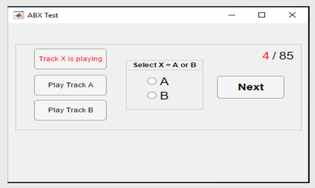


- **Overall Mean SD, ILD, ITD at 120cm, variable az Static and Dynamic**

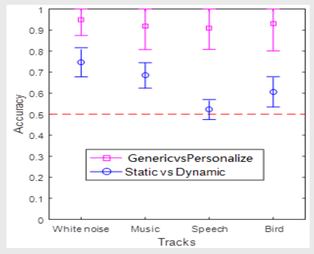


Subjective Performance

ABX test further confirmed our objective results. **95% identification accuracy** for Generic vs Individualized; and **50-70% identification accuracy for Static vs Dynamic**, which indicates that our measured HRTF is almost in-differentiable from conventional static methods.



Subjective Evaluation ABX



1- Easily differentiate
0.5 - Guessing level

Specification

ACQUISITION

Hardware

Microphone	Miniature electret condenser microphone for hearing instruments Deviation between left and right < 1dB Sensitivity @1khZ : 32(+/- 3) : dB, re.1V per Pascal Size: 3.55 x 3.55 x 1.27mm
Microphone placement	1-2mm into the ear canal
Head Tracker Speed	50 Hz

Software

No of Grid in Frontal Directions	Azimuth [-60:5:60] Elevation [-30:10:30] Distance 1 m
Excitation Signal	White Noise
Length of acquired impulse response	600 samples @ 48 kHz
Recording Duration	90-180 secs
Plugins	• Real-Time HRIR Estimation • Adaptive NLMS with variable step size • HRIR Final Selection based on MSE • Real-time Visual Feedback of movements • Automatic saving of audio/head movement/HRIR data

RENDERING

Hardware

Patented Headphone	**US PATENT 9,357,282** "Listening device and accompanying signal processing method"

Software

Rendering Update Rate	50 Hz
Interpolation	To 1 degree, Linear using Triangulation method
Head Tracking Speed	90 Hz
Playback Scenes	• Single Static Sound Source • Multiple Static Sound Sources • Moving Sound Sources

DATA ANALYTICS

HRIR Accuracy	MSE of ipsilateral and contralateral ear at each azimuth and elevation
Head movement pattern	6 Degree of Freedom (DoF) X Axis Pitch — X Axis Left/Right Y Axis Yaw — Y Axis Up/Down Z Axis Roll — Z Axis Front/Back
Degree of transition	Head and Torso Aligned (roll variation between head and torso < 2 degree) Head and Torso Not Aligned (roll variation between head and torso < 6 degree)