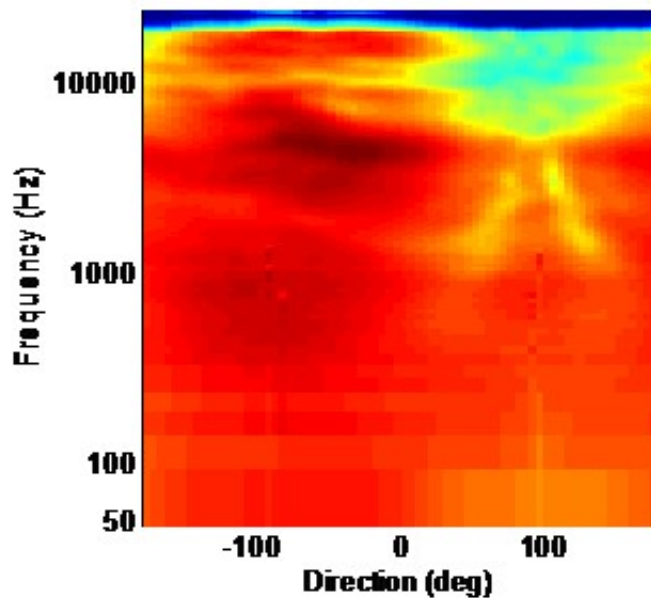# Researchers use head related transfer functions to personalize audio in mixed and virtual reality

June 26 2017

Credit: Microsoft

Move your head left, now right. The way you hear and interpret the sounds around you changes as you move. That's how sound in the real world works. Now imagine if it worked that way while you were listening to a recording of a concert or playing a video game in virtual reality.

During Acoustics '17 Boston, the third joint meeting of the Acoustical Society of America and the European Acoustics Association being held June 25-29, in Boston, Massachusetts, Ivan J. Tashev and Hannes Gamper, with Microsoft's Audio and Acoustics Research Group, will explain how they are using [head](#) related transfer functions (HRTF) to create an immersive [sound](#) environment.

HRTFs are the directivity patterns of the human ears and since every human is unique, capturing a person's HRTFs is the equivalent of fingerprinting. Everyone's ears are unique, so the imprint of one person's anatomy on a sound is completely different from that of another.

"Sound interacts not just with the area but the listener themselves," Gamper said. Sound doesn't just go straight into a listener's ear. It interacts with the shape of the listener's head, shoulders and ears. "All of the physical interactions with the sound create this unique acoustic fingerprint that the brain is trained to understand or decode," Gamper said. "This fingerprint is captured in the HRTF."

Using a database of HRTF measurements from 350 people, the largest such database of HRTF, Tashev, a senior project architect, and Gamper, a researcher, are working to create a more personalized, immersive audio environment. In mixed reality, where a digital image is imprinted in your actual environment, this technology will attach sound to that image. If there's a dragon in the corner of the room, you'll hear it as well as see it.

"In order for us to be able to create this illusion of the spatial sound that's somewhere in space we have to trick the brain," Gamper said.

There are three main components to Tashev and Gamper's work: spatial audio, head tracking and personalization of spatial hearing. Spatial audio is the concept of filtering audio as it comes from a certain dimension.

With head tracking, sound changes as the listener moves his head. Spatial hearing, the ability to localize the direction of sound, is unique for every person. The third is personalization of spatial hearing: the ability to localize the direction of sound.

"From those three components, our spatial audio system and spatial hearing filters are improving the perception of the user to be able to say, 'Yes, that sound came from there!'" Tashev explained.

"The whole goal of personalization is to make [the experience] more convincing, make it more accurate, and also eventually improve the quality of it," Tashev said.

For this personalization to work, the researchers need a lot of data. The 350 people in their HRTF database all had high resolution 3-D scans, which included detail around the ears, head and torso. For less data, but still a high amount of personalization, the researchers could ask for specific user metrics, such as head height, depth and width, and then search their database of 350 people for a close match. Eventually, they said, it might be possible to take a single image from a depth camera or an XBox Kinect and use models to map a template from a 3-D scan of that image.

"We care about impacting the life of people in a positive way," Tashev said.

 More information: www.microsoft.com/en-us/resear … oject/spatial-audio/

Provided by Acoustical Society of America