The newsletter of
The Acoustical Society of America

ECHOES

Volume 8, Number 2
Spring 1998

Head-Related Transfer Functions

by William M. Hartmann

Ever since the third edition of Lord Rayleigh's *The Theory of Sound* (1907), it has been known that a listener can make use of both interaural differences in level and interaural differences in timing to localize sound. If a sound source is to the left of the listener, the right ear is in the shadow of the head so that the signal level in the left ear is more intense. Also, with the source to the left, a signal arrives sooner at the left ear, creating an interaural time difference for each waveform feature.

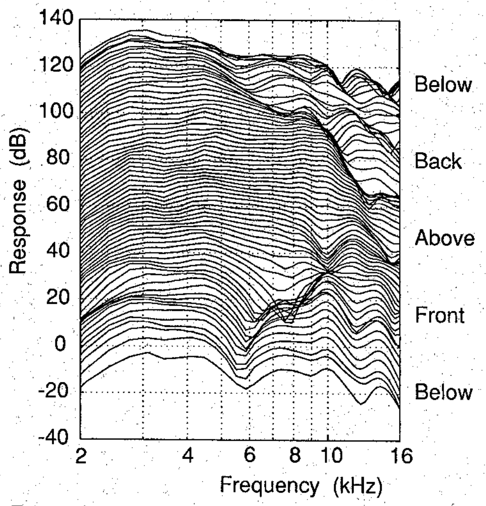

Figure 1: Frequency response curves for a KEMAR head for about eighty different source positions distributed around a circle that runs directly below the head, in front, directly overhead, in back, and below again. The KEMAR is the Knowles Electronics Manikin for Acoustic Research. It has silicone pinnae and microphones where the inner ears should be. (Courtesy of R.O. Duda)

Over the years, psychoacousticians have verified the roles of interaural differences in level and timing in great detail. But it has also been clear (to Rayleigh too) that there must be more. Interaural level and timing are not enough to disambiguate all different source locations. For example, locations in the median-sagittal plane (includes the points directly in front, directly in back, and overhead) are symmetrical with respect to the two ears so that interaural differences are small or absent. Yet listeners can readily distinguish between different sources in this plane. Further, when an experimenter presents signals by headphones using carefully controlled interaural differences in level and timing, listeners do not experience a sound that is correctly localized in space. Instead, the sound often seems to be within the head. The signal is said to be lateralized to the left or right. It is not externalized and localized. What is missing from the simple interaural-differences model of sound localization is the head-related transfer function.

The head-related transfer function (HRTF) is the frequency response, like the response of a filter, between a point in space, where a source might be, and an ear, attached to a head in the normal way. A HRTF is measured in an otherwise free field to avoid room effects, and the receiving microphone is normally a probe inserted well into the ear canal to capture the spectral details of the entire external ear. The frequency response of the external ear (pinna) is directional. In particular, it discriminates between source locations in the median-sagittal plane. The pinna is so important that HRTFs are sometimes called pinna response functions. However, its small size limits its directional response to the frequency range above 5 kHz. By contrast, psychoacoustical localization experiments exhibit some spectrally mediated directional effects at frequencies as low as 500 Hz. Therefore, the HRTF might be thought of as an *anatomical* transfer function.

HRTFs measured in the median-sagittal plane are shown in Figure 1. The curves show a broad maximum near the ear canal resonance, 2-4 kHz. Most significant is the *pinna notch* which occurs near 6 kHz when the source is below, and which occurs at ever higher frequencies as the source is elevated. When the source is overhead, the notch tends to disappear leading to a relatively flat frequency response.

The HRTFs in Figure 1 correspond well with psychoacoustical measurements of human localization ability. Increasing the frequency of a characteristic spectral feature like the notch is known to lead to a sense of increased elevation. The flat frequency response for locations overhead agrees with the psychoacoustical fact that listeners find it most difficult to distinguish between slightly different locations for

continued on page 6

...*Transfer Functions, continued from page 1*

locations directly overhead. The frequency response curves show an approximate, but not exact, symmetry between locations in front and in back. This agrees with the fact that front-to-back confusions are the most difficult to eliminate in attempts to synthesize precise source locations using headphones.

Psychoacousticians have now mastered the techniques of measuring individual HRTFs and reproducing their effects with headphones. The results can be dramatic. Sound images from headphones are no longer confined to the head or to an indeterminate space having a diffuse leftness or rightness. It is possible to use headphones to synthesize sound images that are externalized, compact, and correctly localized in azimuth, elevation, and (to some degree) distance.

The sense of reality can be made still more striking by including the dynamical effects of head motion. It is tricky to to this, and it requires high-speed dedicated digital convolvers, but it can be done. First, HRTFs are measured for left and right ears for source locations over the full 4p steradians of solid angle (a grid with 10-degree spacing will work). The functions are stored as impulse responses in digital memory ready for use. The listener wears a head tracker, which regularly informs the controlling computer of the pitch angle, yaw and roll of the head. The computer then calls up the appropriate left and right HRTFs, or interpolates among neighboring HRTFs, so that the location of the virtual sound source, created by headphones, is fixed as the head rotates. Listeners, it seems, are forever running subconscious tests on their auditory sensations. A virtual sound image that passes the head-motion test is perceived to be more convincing than an image that is untested.

Because of its ability to simulate sources of sound located in space, the HRTF technique opens the way to virtual auditory reality. The head-related impulse response combined with the response of a concert hall can create the impression of listening to a live performance in the hall. Computer modeling of the impulse response of the hall for different listener locations can, in principle, enable a listener to evaluate the quality of sound in the hall before it is built.

Simulating auditory images in space can lead to enhanced audio for entertainment or multimedia use. A helicopter on a video screen can be made to sound as though it is flying overhead by moving the pinna response features together with the screen image. In fact, a major television network has processed audio signals to do exactly this. Meanwhile, scientists at Wright Patterson Air Force Base have used HRTFs to provide vital information to military pilots, whose sensory channels are overloaded during combat. Localized auditory images have been used for targeting information and to indicate the location of an approaching missile.

Back in the lab, HRTFs are helping acousticians understand the basis of sound localization and the sensations of externalization and distance. HRTFs in the horizontal plane are shown in Figure 2. There is a lot of detail there, and in addition to the magnitude information in the figure, there is also phase information. One wonders whether all that detail is really necessary. Experiments show that considerable information reduction is possible with negligible effect on perception. Although physical measurements of phase show significant dispersion, it appears that phase shifts can be modeled as frequency-independent delays without adverse effect. Further, the frequency response curve for a single location can be modeled by a Fourier sum with as few as eight coefficients, at least for low signal levels. On the other hand, additional detail, in the form of a radial dependence, is required to synthesize images that are within a meter of the head.

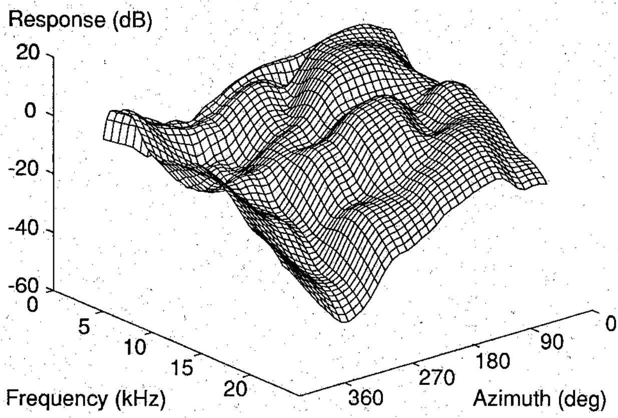

Figure 2: Frequency response curves for a KEMAR head for sources in the horizontal plane, distributed around the complete circle. The nose points at zero degrees, and the measuring ear is at 90 degrees. (Courtesy of R.O. Duda)

As would be expected for an anatomical feature like the HRTF, different listeners have different features. Also, perhaps unfortunately, individual differences in HRTFs are perceptually important. Experiments at the University of Wisconsin in which subjects listen through someone else's ears find increased confusion in localization, especially between sources in front and back. Images made with imperfect HRTFs can become diffused in space; although they are localized they are not compact. They often appear much closer to the head than they should be.

Widespread application of HRTF technology may

continued on next page page

...Transfer Functions, continued from previous page

depend on solving the individual difference problem. It is clear that some individual HRTFs are more successful than others. One commercially available device, intended to enhance television sound (the Auri from Virtual Listening Systems) gives the user a switch to choose among 15 representative HRTFs. Other techniques, some described at the ASA meeting at Penn State in June 1997, allow the listener to deform HRTFs to optimize localization. However, the space of possible HRTFs is so vast that an unrestricted search is bound to fail. Physical or psychoacoustical constraints must be judiciously placed to have any chance of getting close to the right answer. Perhaps one day we shall all routinely visit the HRTF-metrist who will fit us with individual virtual reality filters. Then we will be able to plug in to an expanded three-dimensional world of information and entertainment.

Acknowledgements: Information in this article was borrowed from work by Fred Wightman and Doris Kistler, Elizabeth Wenzel, Robert Butler, Richard Duda,Abhijit Kulkarni and Steven Colburn, Richard McKinley and Timothy Anderson, also Andrew Wittenberg and myself.

Bill Hartmann is a professor of physics at Michigan State University. His research is on the perception of sound, particularly sound localization and the perception of pitch. He is the author of the textbook "Signals, Sound, and Sensation" (AIP Press/Springer-Verlag, 1997) and is currently vice-president elect of the Acoustical Society.

World of Sound Calendars Available for Students at No Charge

If you are a teacher at a college, university or high school, call ASA at (516) 576-2360 to receive calendars for your students.