Measuring HRTFs in a Reflective Environment

Jonathan S. Abel and Scott Foster
Crystal River Engineering
490 California Avenue, Suite 200
Palo Alto, CA 94306 USA

Virtual audio displays control the apparent location of sound sources by applying left-ear and right-ear filters designed to mimic the acoustic properties of the human torso, head, and pinnae. These filters, called head-related transfer functions (HRTFs), attenuate and delay the signal as a function of frequency and direction of arrival. Though HRTFs have many features in common across a population, virtual audio displays are known to perform best when the HRTFs used are those of the listener.

Head-related transfer functions may be measured by analyzing ear canal signals received from known audio sources placed at various locations about the listener. Typically, the subject position is fixed, and an array of speakers is used to generate the needed HRTF directions. Swept sinusoids or maximum-length sequences are often used as test signals, and probe microphones make measurements near the ear drums. So that unwanted reflections do not corrupt the data, the subject and speaker array are placed in an anechoic chamber.

The requirement of an anechoic chamber is a barrier to widespread use of the current technology. In addition, since anechoic chambers are not truly anechoic, measurement accuracy suffers. Measurement accuracy is also degraded by the use of probe microphones which gather relatively little signal. To achieve a 60dB output signal-to-noise ratio, probe sequences several seconds long are commonly employed. Aside from exciting anechoic chamber modes, such long sequences invite subject head motion, thus smearing the measurement over a range of directions.

In this chapter, we describe a method for measuring HRTFs in a reflective environment. The basic idea is to constrain the speaker-subject geometry so that no reflections arrive until after the direct-path impulse response has decayed. Under these conditions, by measuring the impulse response of the entire system (which includes contributions from the environment as well as the direct path), the direct-path response is available via a simple windowing operation.

Given a raw direct-path response duration τ, the constraint essentially requires that no reflecting objects appear within the ellipse having the subject head and speaker as foci and having radius $c \cdot \tau$, c being the speed of sound. So that the speaker-subject geometry is not overly constrained, the raw direct path impulse response should be made as short as possible. By using blocked meatus microphones, rather than probe microphones, the ear canal resonance is eliminated from the raw response, and τ is shortened. In initial testing on human subjects, blocked meatus microphones were seen to generate raw responses on the order of a milisecond in length, whereas probe-type microphones yielded responses two to four times as long.

Blocked meatus microphones also provide greater sensitivity than that of probe-type microphones, and allow louder test sequences to be used without subject discomfort. Employing 0.1-second-long Golay code sequences, HRTFs with signal-to-noise ratios in excess of 70dB have been measured in office environments.