

## **Contributors to 3D Perception of Sound**

Durand R. Begault and Elizabeth M. Wenzel

A study was conducted to evaluate the role of individualized head-related transfer functions (HRTFs) for localization accuracy and the experience of realism using auditory displays. Realistic auditory displays that emulate free field auditory sound sources are required for efficient and effective design of virtual reality systems for real-world applications, including teleconferencing and interacting with virtual systems. A number of studies have evaluated humans' abilities to localize aural signals delivered via headphones in a virtual space. The results suggest that optimal localization performance and perceived realism result from inclusion of factors that emulate those typical of everyday spatial hearing experiences (e.g., head position effects, a realistically diffuse field or reflections, and head-related transfer function specific to the listener). Our current work described here directly compares the relative contributions of these factors to localization accuracy and perceived realism using realistic, meaningful acoustic stimuli.

A laboratory study was performed at Ames using headphone-delivered virtual speech stimuli, rendered via HRTF-based acoustic auralization software and hardware. The relative advantages of individualized vs. generic HRTFs were evaluated because they have been shown to improve localization accuracy and externalization while reducing front-back reversal errors. Virtual speech was presented with: (1) No reflections; (2) A few HRTF-filtered "early reflections;" or (3) "Full auralization" reverberation. This factor has been shown to influence whether or not a headphone-delivered sound is perceived as being external to the listener, but has not been studied in the presence of the other factors under consideration. On half of the trials, the sounds were updated in real time to reflect the listener's head position based on output from a head-tracker. Nine naïve male and female volunteers participated in the study after being instructed how to make azimuth, elevation, distance and realism judgments using the interactive, self-paced software.

The results were evaluated to defined system requirements for improved virtual acoustic simulation. Overall, it appears that auditory signals that include reverberation yield lower azimuth errors and higher externalization rates (here, by a ratio of about 2:1), but at the sacrifice of elevation accuracy. The presence of only 80 msec of early reflections, as opposed to a full auralization, was sufficient to achieve the benefit. Head-tracking did not significantly increase externalization rates, nor did it yield more accurate judgments of azimuth.

It was surprising that the presence of reverberation caused the accuracy of azimuth estimation to improve by approximately 5 degrees. By contrast, the inclusion of head tracking significantly reduced reversal rates, also by a ratio of about 2:1, a phenomenon explained by the differential integration of interaural cues over time, but did not improve localization accuracy or externalization. These findings contrast with previous results indicating that head movements enhance source externalization. Except for the interaction of head tracking and HRTF for azimuth error, there is no clear advantage to including individualized HRTFs for improving localization accuracy, externalization or reversal rates, within a virtual acoustic display of speech. For speech, unlike earlier results for non-speech signals, no dramatic improvement in reversals was found with individualized HRTFs. The fact that individualized HRTFs did not significantly increase azimuth accuracy was not surprising given the spectral characteristics of speech.

The results taken as a whole suggest that a complete, high-fidelity emulation of all free-field sound characteristics is not required for an effective spatialized auditory display.