Spatial Auditory Display: Comments on Shinn-Cunningham et al., ICAD 2001

BARBARA G. SHINN-CUNNINGHAM and TIMOTHY STREETER
Department of Cognitive and Neural Systems
Boston University

Spatial auditory displays have received a great deal of attention in the community investigating how to present information through sound. This short commentary discusses our 2001 ICAD paper (Shinn-Cunningham, Streeter, and Gyss), which explored whether it is possible to provide enhanced spatial auditory information in an auditory display. The discussion provides some historical context and discusses how work on representing information in spatial auditory displays has progressed over the last five years.

Categories and Subject Descriptors: [General Literature - General]: *Conference Proceedings*
General Terms: Experimentation, Human Factors, Performance
Additional Key Words and Phrases: Auditory Display, Spatial Hearing, Virtual Auditory Space

1. HISTORICAL CONTEXT

The next time you find yourself in a noisy, crowded environment like a cocktail party, plug one ear. Suddenly, your ability to sort out and understand the sounds in the environment collapses. This simple demonstration of the importance of spatial hearing to everyday behavior has motivated research in spatial auditory processing for decades.

Perhaps unsurprisingly, spatial auditory displays have received a great deal of attention in the ICAD community. Sound source location is one stimulus attribute that can be easily manipulated; thus, spatial information can be used to represent arbitrary information in an auditory display. In addition to being used directly to encode data in an auditory display, spatial cues also are important in allowing a listener to focus attention on a source of interest when there are multiple sound sources competing for auditory attention [Shinn-Cunningham et al. 2005]. Although it is theoretically easy to produce accurate spatial cues in an auditory display, the signal processing required to render natural spatial cues in real time (and the amount of care required to render realistic cues) is prohibitive even with current technologies. Given both the important role that spatial auditory information can play in conveying acoustic information to a listener and the practical difficulties encountered when trying to include realistic spatial cues in a display, spatial auditory perception and technologies for rendering virtual auditory space have both been well-represented areas of research at every ICAD conference held to date (e.g., see [Brungart et al. 2004; Carlile et al. 2002; Jin et al. 2003; Minnaar et al. 2001; Scarpaci et al. 2005; Wenzel et al. 2000]).

Even with a good virtual auditory display, the amount of spatial auditory information that a listener can extract is limited compared to other senses. For instance, auditory localization accuracy is orders of magnitude worse than visual spatial resolution. The study reprinted here, originally reported at ICAD 2001, was motivated by a desire to increase the amount of spatial information a listener could extract from a virtual auditory display. The original idea was to see if spatial resolution could be improved in a virtual auditory display by emphasizing spatial acoustic cues. The questions we were interested in were: 1) Can listeners learn to accommodate a new mapping between exocentric location and acoustic cues, so that they do not mislocalize sounds after training? and 2) Do such remappings lead to improved spatial resolution, or is there some other factor limiting performance?

2. RESEARCH PROCESS

The reprinted study was designed to test a model that accounted for results from previous experiments investigating remapped spatial cues. The model predicted that spatial performance is restricted by central memory constraints, not by a low-level sensory limitation on spatial auditory resolution. However, the model failed for the experiments reported: listeners actually achieved better-than-normal spatial resolution following training with the remapped auditory cues (unlike in any previous studies). These results were encouraging on the one hand, as they suggested a

This research was supported by the Air Force Office of Scientific Research (FA9550-04-1-0260) and the National Institutes of Health (R01 N00014-04-1-0131).
Authors' addresses: Barbara G. Shinn-Cunningham or Tim Sreeter, Department of Cognitive and Neural Systems, Boston University, 677 Beacon St., Boston, MA, 02215, USA; http://www.cns.bu.edu/~shinn; email: {shinn, timstr}@cns.bu.edu.

method for generating better-than-normal auditory spatial resolution. However, further investigations into how listeners cope with rearrangements of auditory space were needed to verify the unexpected results and to develop a computational model to account for the findings. Subsequent studies confirmed that better-than-normal spatial resolution can be achieved after training with an appropriate remapping [Streeter et al. 2001]. A computational model based on rapid adjustment of an internal mapping between spatial cue and perceived exocentric direction qualitatively accounted for these results [Streeter and Shinn-Cunningham 2002].

3. BODY OF WORK

Research in our laboratory has explored the human limitations in extracting and interpreting acoustic (often spatial) information, and how human abilities influence performance in an auditory display. At ICAD 1994, we discussed how much information a listener could pull out of different possible acoustic dimensions [Shinn-Cunningham and Durlach 1994]. In 2000, we demonstrated that perceptual learning was an important factor affecting directional hearing accuracy in the presence of reverberant energy [Shinn-Cunningham 2000]. At ICAD 2002, we presented results that showed that spatial auditory displays of distance could lead to improvements in a listener's ability to understand speech in the presence of competing sources [Shinn-Cunningham 2002]. In 2003, we presented studies showing that listeners are relatively insensitive to room location in a virtual auditory display [Shinn-Cunningham and Ram 2003], and that perception of competing speech sounds in natural (reverberant) settings improves when listening with two ears compared to one ear [Devore and Shinn-Cunningham 2003]. Most recently, we looked at how spatial auditory information affects a listener's ability to simultaneously monitor two different sound sources [Best et al. 2005; Shinn-Cunningham and Ihlefeld 2004].

4. RELATION TO THE FIELD OF AUDITORY DISPLAY

The paper reprinted here is but one example of how spatial auditory displays have been studied in the ICAD community. For example, in addition to developing technology for spatial auditory display, Wenzel and colleagues have considered everything from perception to performance [Miller et al. 2003; Wenzel 1994; Wenzel et al. 2000]. Carlile and his colleagues have developed and evaluated methods for rendering realistic spatial cues in a virtual auditory display [Best et al. 2003; Carlile et al. 2002; Jin et al. 2003]. Grohn and colleagues have studied navigation in immersive auditory displays [Grohn 2002; Grohn et al. 2003, 2004]. Brungart and his colleagues have focused on the influence of spatial cues (and other attributes) on speech understanding [Brungart et al. 2002; Brungart and Simpson 2003; Brungart et al. 2004].

5. FUTURE WORK

Although spatial hearing often has a large influence on auditory perception in a complex settting (like a cocktail party), the neural mechanisms underlying these effects are not fully understood. Existing computational models of the contribution of spatial hearing to detecting and understanding sounds in a complex environment predict that the efficacy of spatial hearing should degrade in the presence of noise and reverberation; however, if anything such complex listening environments make spatial hearing more rather than less important for understanding sound sources [Shinn-Cunningham et al. 2005]. Ongoing work (in my laboratory and others) is now focusing on teasing apart the many ways in which spatial acoustic cues influence perception, both through simple bottom-up processes and through other top-down, attentional mechanisms.

6. CONCLUDING THOUGHTS

The use of sound source location as a way of conveying information and improving information transfer to a listener is both promising and important. Research in spatial auditory perception and virtual auditory space technologies will no doubt continue to be an area of intense research activity both within and beyond the ICAD community.

7. REFERENCES

BEST, V., IHLEFELD, A., AND SHINN-CUNNINGHAM, B. 2005. The effect of spatial layout in a divided attention task. *Proceedings of the International Conference on Auditory Display*, Limerick, Ireland.

BEST, V., VAN SCHAIK, A., AND CARLILE, S. 2003. Two-point discrimination in auditory displays. *Proceedings of the International Conference on Auditory Display (ICAD 2003)*, Boston, MA, 17-20.

BRUNGART, D.S., ERICSON, M.A., AND SIMPSON, B.D. 2002. Design considerations for improving the effectiveness of multitalker speech displays. *Proceedings of the International Conference on Auditory Display (ICAD 2002)*, Kyoto, Japan.

BRUNGART, D.S., AND SIMPSON, B.D. 2003. Optimizing the spatial configuration of a seven-talker speech display. *Proceedings of the International Conference on Auditory Display (ICAD 2003)*, Boston, Massachusetts.

BRUNGART, D.S., SIMPSON, B.D., MCKINLEY, R., KORDIK, A.J., DALLMAN, R.C., AND OVENSHIRE, D.A. 2004. The interaction between head-tracker latency, source duration, and response time in the localization of virtual sound sources. *Proceedings of the International Conference on Auditory Display (ICAD 2004)*, Sydney, Australia.

CARLILE, S., JIN, C., AND LEUNG, J. 2002. Performance measures of the spatial fidelity of virtual auditory space: Effects of filter compression and spatial sampling. *Proceedings of the International Conference on Auditory Display (ICAD 2002)*, Kyoto, Japan.

DEVORE, S., AND SHINN-CUNNINGHAM, B.G. 2003. Perceptual consequences of including reverberation in spatial auditory displays. *Proceedings of the International Conference on Auditory Display (ICAD 2003)*, Boston, Massachusetts, 75-78.

GROHN, M. 2002. Localization of a moving virtual sound source in a virtual room: The effect of a distracting auditory stimulus. *Proceedings of the International Conference on Auditory Display (ICAD 2002)*, Kyoto, Japan.

GROHN, M., LOKKI, T., AND TAKALA, T. 2003. Comparison of auditory, visual, and audio-visual navigation in a 3-D space. *Proceedings of the International Conference on Auditory Display (ICAD 2003)*, Boston, Massachusetts.

GROHN, M., LOKKI, T., AND TAKALA, T. 2004. An orientation experiment using auditory artificial horizon. *Proceedings of the International Conference on Auditory Display (ICAD 2004)*, Sydney, Australia.

JIN, C., VAN SCHAIK, A., BEST, V., AND CARLILE, S. 2003. Perceptual spatial-audio coding. *Proceedings of the International Conference on Auditory Display (ICAD 2003)*, Boston, Massachusetts.

MILLER, J.D., ANDERSON, M.R., WENZEL, E.M., AND MCCLAIN, B.U. 2003. Latency measurement of a real-time virtual acoustic environment rendering system. *Proceedings of the International Conference on Auditory Display (ICAD 2003)*, Boston, MA.

MINNAAR, P., OLESEN, S.K., CHRISTENSEN, F., AND MOLLER, H. 2001. The importance of head movements for binaural room synthesis. *Proceedings of the International Conference on Auditory Display (ICAD 2001)*, Espoo, Finland.

SCARPACI, J.W., COLBURN, H.S., AND WHITE, J.A. 2005. A system for real-time virtual auditory space. *Proceedings of the International Conference on Auditory Display (ICAD 2005)*, Limerick, Ireland.

SHINN-CUNNINGHAM, B.G. 2000. Learning reverberation: Implications for spatial auditory displays. *Proceedings of the International Conference on Auditory Display (ICAD 2000)*, Atlanta, GA, 126-134.

SHINN-CUNNINGHAM, B.G. 2002. Speech intelligibility, spatial unmasking, and realism in reverberant spatial auditory displays. *Proceedings of the International Conference on Auditory Display (ICAD 2002)*, Kyoto, Japan, 183-186.

SHINN-CUNNINGHAM, B.G., AND DURLACH, N.I. 1994. Defining and redefining limits on human performance in auditory spatial displays. *Proceedings of the International Conference on Auditory Display (ICAD 1994)*, Santa Fe, NM, 67-78.

SHINN-CUNNINGHAM, B.G., AND IHLEFELD, A. 2004. Selective and divided attention: Extracting information from simultaneous sound sources. *Proceedings of the International Conference on Auditory Display (ICAD 2004)*, Sydney, Australia.

SHINN-CUNNINGHAM, B.G., IHLEFELD, A., SATYAVARTA, AND LARSON, E. 2005. Bottom-up and top-down influences on spatial unmasking. *Acta Acustica, submitted*.

SHINN-CUNNINGHAM, B.G., AND RAM, S. 2003. Identifying where you are in a room: Sensitivity to room acoustics. *Proceedings of the International Conference on Auditory Display (ICAD 2003)*, Boston, Massachusetts, 21-24.

STREETER, T., AND SHINN-CUNNINGHAM, B.G. 2002. A model of short-term perceptual adaptation to remapped spatial auditory cues. *Proceedings of the International Conference on Cognitive and Neural Systems*, Boston.

STREETER, T., SHINN-CUNNINGHAM, B.G., AND BRUGHERA, A. 2001. Short-term adaptation to novel combinations of acoustic spatial cues. *Journal of the Acoustical Society of America, 109,* 2376.

WENZEL, E.M. 1994. Spatial sound and sonification. In G. KRAMER (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interface* XVIII, SFI Studies in the Science of Complexity. Santa Fe, New Mexico: Addison-Wesley.

WENZEL, E.M., MILLER, J.D., AND ABEL, J.S. 2000. A software-based system for interactive sound synthesis. *Proceedings of the International Conference on Auditory Display (ICAD 2000)*, Atlanta, GA.