Invited paper, Special Session on Virtual Auditory Space, Proceedings of the First IEEE Pacific-Rim Conference on Multimedia, 13-15 December 2000, Sydney, Australia. 227-230.

DISTANCE CUES FOR VIRTUAL AUDITORY SPACE

B. G. Shinn-Cunningham

Boston University Hearing Research Center
Departments of Cognitive and Neural Systems and Biomedical Engineering
677 Beacon St., Boston, Massachusetts 02215, USA
Telephone: 617-353-5764 Fax: 617-353-7755 Email: shinn@cns.bu.edu

ABSTRACT

Recent experiments investigating the acoustic cues that underlie distance perception for nearby sources provide important insights into how distance cues can be incorporated into virtual auditory displays. Potential distance cues considered include overall level, reverberation, and interaural level differences. Perceptual research shows that the relative utility of these cues depends on the type(s) of sound source(s) and the kind of acoustic environment to be simulated as well as the range of distances to be encoded.

1. INTRODUCTION

While there has been great interest in the development of effective virtual auditory displays, most previous work has examined how to recreate accurate perceptions of source direction while paying scant attention to the simulation of source distance. Recent work in our laboratory has examined normal distance perception for sources nearby the listener (within one meter of the center of the head).

Focusing on distance perception of nearby sources is instructive for a number of reasons. First of all, most of the acoustic distance cues that arise for relatively distant sources are also present for nearby sources. Secondly, a given physical displacement often corresponds to a larger change in these common distance cues for nearby sources compared to more distant sources, making it possible to measure and evaluate the effectiveness of potential distance cues more easily. In addition, when sources are close to the listener, unique binaural distance cues arise which are not present for more distant sources [5, 10, 20]. Finally, many virtual reality simulations simulate objects near the user, so that creating realistic, compelling auditory simulations of nearby sources is of interest.

2. METHODS

In order to study how the signals at the ears change with source distance, both theoretical analyses and empirical studies were undertaken. A spherical head model [16] was employed to examine how the signals at the ears vary with source position in anechoic space. This approach has previously been used by a number of researchers to investigate how distance affects the signals reaching a listener s ears for nearby sources [5, 10, 20].

Empirical measurements of the transfer functions from source to the ear were made in a reverberant room (reverberation time T_{60} approximately equal to 450 ms) for sources at distances ranging from 15 cm to 1 m from the center of the head. These measurements were made using a Maximum Length Sequence (a standard technique for measuring acoustic transfer functions [21]). The resulting impulse responses characterize how both the direct and reverberant sound is filtered by the room and the head and ears of the listener. The total impulse response was time windowed (using a 10-ms long window with a $\cos^2$ onset and offset lasting 200 µs) to separate the direct from the reverberant portions of the transfer function. A longer (500-ms long) window with a similar $\cos^2$ ramp was used to separate the reverberant portion of the impulse response from the total measured impulse. This simple time-windowing enabled the direct and reverberant portions of the impulse response to be separately analyzed and compared.

3. RESULTS: DISTANCE CUES

The most important acoustic distance cues include overall sound level, interaural level differences (ILDs) in the sound reaching the two ears, and reverberation. While other potential cues have been identified (such as spectral changes with distance [8]), these three cues are arguably the most acoustically robust and perceptually salient.

3.1 Overall Level

The most obvious cue for source distance is a change in level with distance. For a sound wave traveling in free space, the root-mean-square (RMS) pressure level varies inversely with the square of the distance between receiver and source (r^2). Thus, when sources are relatively far from the head (and the interference of the head on the propagating wave can be ignored) doubling the source distance causes a 6-dB reduction in received energy. While this analysis is appropriate for distant sources, the interaction of the head with the sound wave is important for distances less than one meter. As a result, changes in distance depend on source direction for nearby sources.

Figure 1 shows how the RMS pressure (for a broadband source) reaching the right ear varies as a function of source distance (for various directions). The lines in the plot show theoretical results when the head is modeled as a rigid sphere. Symbols show how the measured RMS pressure reaching the ears varies with distance for sources in the median plane and along the interaural axis.

The agreement between empirical measures and spherical head predictions is very good. For both the theoretical and empirical results, the change in RMS pressure with change in log distance is constant at —20dB / $\log_{10}$(distance) for sources more than a meter from the listener (as expected). While the slope of the line relating sound level to distance is roughly constant for distant sources (independent of direction), the sound level at the ear varies with direction due to acoustic shadowing of the head.

In contrast, for nearby sources, sound level is no longer proportional to log distance. In fact, for a source approaching the head along the interaural axis (90ß), the level at the nearer ear grows three times faster with changes in distance than is observed for more distant sources. In contrast, as a source nears the head along the median plane, sound level changes very slowly with distance.

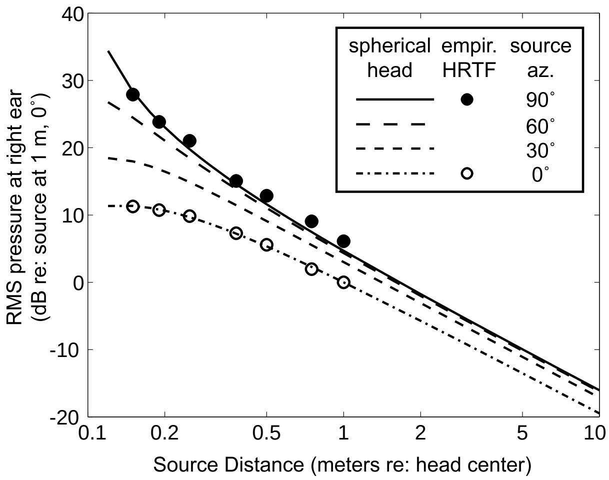

Figure 1: RMS pressure at the near ear as a function of source distance for sources at various directions in the horizontal plane. Lines show predictions for a spherical head model. Symbols show measured RMS pressure at the ear canal of a human subject.

These results show that the RMS pressure of the signals reaching the ears can provide distance information. However, since the absolute level of the direct sound varies both with distance and with the energy emitted from the source, the level at the ears can only provide *relative* distance information unless the listener has a priori knowledge about the source level.

3.2 Interaural Level Differences

It is well known that level differences in the signals at the two ears (ILDs) can arise at moderate to high frequencies due to the acoustic interference of the head (e.g., see [15]). Such head shadow effects provide directional spatial information but are essentially invariant with source distance. However, when sources are close to (within a meter of) the head and laterally displaced from the median plane, changes in source distance cause changes in ILD across all frequencies [5, 9]. To a first-order approximation, these two sources of ILD (i.e., the head shadow and the direction- and distance-dependent ILD that can arise for near sources) are additive [20]. Thus, the total ILD can be broken down into the direction- and frequency-dependent, distance-independent head-shadow component (that is normally simulated in VAS) and a distance- and direction-dependent, frequency-independent component (that is normally overlooked).

Figure 2 shows how ILD varies with source distance for sources in various directions. As in Figure 1, results are shown for both a rigid, spherical head and for empirical measures from a human subject.

For both measured and theoretical results, a change in distance for a nearby source causes significant changes in the overall ILD when the source is to the side of the listener. For sources on the interaural axis, the ILD ranges over 20 dB as distance ranges from 1 m to 15 cm. While the change in the ILD with distance is roughly equivalent for theoretical and empirical results, the ILDs observed in empirical measurements are larger than predicted by the spherical head model (by approximately 5 dB). This discrepancy is primarily due to the fact that the RMS pressure measured at the far (shadowed) ear was consistently less than predicted by the spherical-head model (note, for instance, the agreement in the predicted and measured levels at the right, near ear in Figure 1). While the reason that the model overestimates the sound level at the far ear is not clear, it may be a result of a mismatch between the assumed head size in the model and the actual head size for the subject whose results are shown in Figure 2.

For sources in the median plane, ILD is essentially zero, independent of source distance. Additionally, when the source is more than about a meter from the head, changes in distance cause no significant change in ILD; for relatively large distances, only the direction-dependent head shadow causes any level differences at the two ears.

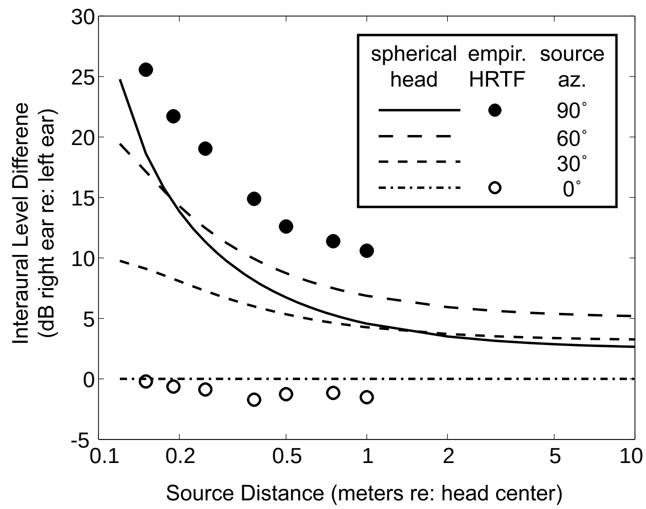

Figure 2: ILD (right re: left ear) as a function of source distance for sources at various directions in the horizontal plane. Lines show predictions for a spherical head model. Symbols show results measured at the ear canal of a human subject.

Results show that nearby sources give rise to unique binaural distance cues; however, the strength of these cues varies with direction. Unlike overall level, ILD distance cues are robust since ILD depends only on source location and is independent of the level of the emitted sound.

3.3 Reverberation

In most listening situations, sound emitted from a source reaches the ears directly as well as indirectly, after reflecting off of walls, floors, and other objects. For sources relatively far from the listener, the direct sound level at the ears varies as in the free field (i.e., proportional to $1/r^2$). However, for distant sources, the level of reverberation reaching the ears is roughly independent of the location (i.e., distance and direction) of the source relative to the listener. Thus, the level of the direct sound relative to the reverberant sound varies inversely with the square of the source distance (independent of the level emitted by the source) for sources beyond a meter from the head.

As with overall level and ILD, the direct to reverberant RMS pressure ratio shows a different pattern for near compared to far sources. As shown in Figure 1, the direct RMS pressure does not vary linearly with r^2 for sources within a meter of the head. In addition, the reverberant energy reaching the ears varies with source location for nearby sources. In particular, the first-arriving (and often largest) echoes reaching the ears are often due to floor and ceiling reflections. For nearby sources, the distance traveled by these reflections varies with the source distance. As a result, the floor/ceiling echoes tend to be larger when the source is near the listener, increasing the reverberation level. Of course, in some situations the head and body may absorb this floor/ceiling echo. For instance, when a source nears the head along the interaural axis, the floor/ceiling echo no longer dominates the reverberation and the reverberation level varies less with source location.

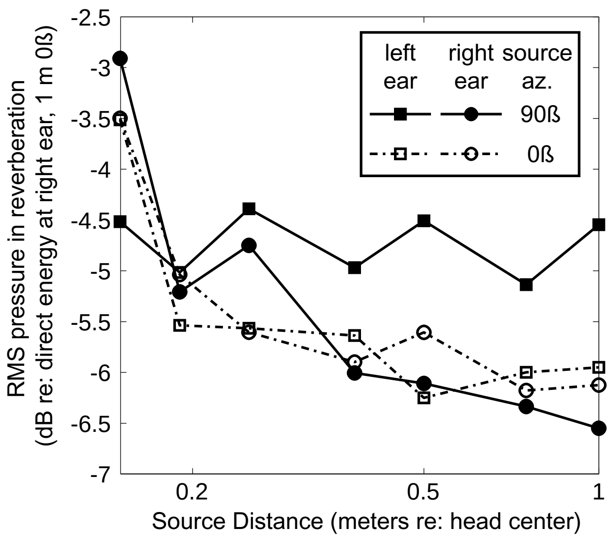

Figure 3: RMS pressure level of the reverberation measured at left (square) and right (circle) ears as a function of distance for sources straight ahead (dashed lines) and 90ß to the right (solid lines).

These effects are seen in Figure 3, which plots the RMS pressure of the reverberation reaching both left and right ears of a human listener sitting in the center of the rectangular reverberant room. For sources in the median plane, the left and right ears receive approximately the same reverberant energy, but the energy level varies with source distance. Similar results are obtained for the energy reaching the right ear when the source is positioned at 90ß azimuth. However, when the source is positioned laterally, the reverberation at the far ear is essentially independent of source distance.

Many researchers have argued that the direct to reverberant energy ratio provides a reliable distance cue [3, 12-14]. For distant sources, this ratio varies with the direct sound energy as $1/r^2$. However, since both the direct sound and the reverberant energy levels vary with distance and direction for nearby sources, the direct to reverberant energy will also vary with both distance and direction.

Figure 4 plots the ratio of the RMS pressures for the direct and reverberant sound. The ratio is shown for both the left and right ears in order to demonstrate how this ratio may encode distance information at both near and far ears.

It is clear that the direct to reverberant pressure ratio at the near ear is a potent distance cue for sources to the side of the head; the pressure ratio varies nearly linearly with distance with a slope of —25 dB /$\log_{10}$(distance). For lateral sources, the signal at the far ear also contains some distance information: the ratio varies from 1 dB to -7 dB for sources between 15 cm and 1 m. For sources in the median plane, the pressure ratio spans a range of approximately 10 dB as the source ranges from 15 cm to 1 m (slightly larger than the range observed in the far, shadowed ear for lateral sources). From this analysis, it appears that the near ear provides the strongest reverberant cue for distance, and that the salience of the reverberant distance cue varies directly with the laterality of the source.

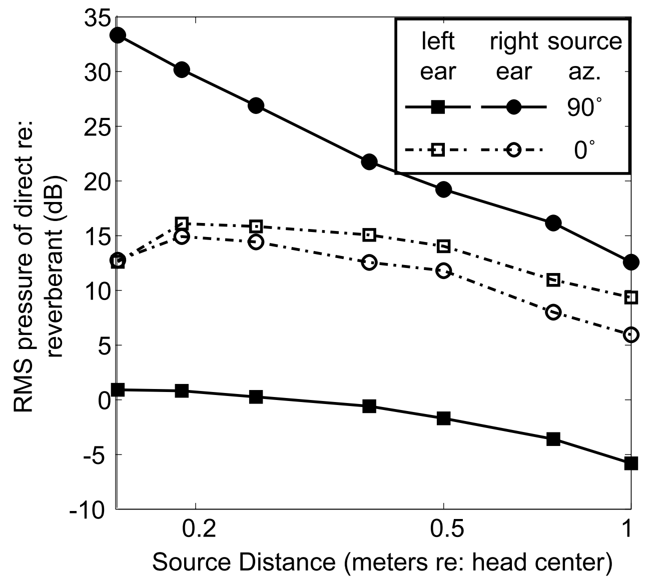

Figure 4: Ratio of direct to reverberant RMS pressure level at left (square) and right (circle) ears as a function of distance for sources straight ahead (dashed lines) and 90ß to the right (solid lines).

4. PERCEPTUAL SALIENCE OF CUES

The previous section examined three acoustic features of the signals at the ear that change with source distance. Results of various perceptual studies investigating the salience of these potential distance cues suggest that each can be useful in different listening situations.

There is ample evidence that overall level provides potent distance information. For a familiar source (such as speech) where the listener knows the level of the emitted sound, overall level provides an absolute distance cue [11, 13, 14]. When a listener has no such a priori expectations, overall level cannot be used to judge absolute source distance; however, it does provide a relative cue [1, 7].

A recent study by Brungart [4] demonstrates that ILD cues provide reliable distance information for nearby, lateral sources in anechoic space. Random broadband noise samples were used, so the listener had no a priori expectations about the absolute levels produced by the source. Stimulus level was set by first roughly equating the level at the head and then randomly roving level by an additional 15 dB. This processing removed any relative distance cues arising from sound level. While subjects were good at judging source distance for sources along the interaural axis, performance degraded to chance as the source direction approached the median plane.

Reverberation has long been known to provide distance information for relatively distant sources [2, 6, 12]. We recently replicated Brungart s experiment in a reverberant room. Even for nearby sources (where the reverberation is relatively quiet and ILD cues may already provide distance information), reverberation yields drastic improvements in distance perception [17]. In fact, we observe reasonably good distance performance for sources in all directions (including along the median plane, where ILD cues are nonexistent and reverberant cues are relatively weak; see Figure 4), although performance is best for lateral sources.

While Brungart s results show that ILD cues can be useful in a true anechoic setting [4], recent results in our own laboratory call into question whether anechoic simulations employing ILD cues yield robust distance percepts [19]. Individually-measured transfer functions were used to simulate anechoic and reverberant listening conditions for both medial and lateral sources under headphones. Listeners judged source distance for both binaural and monaural listening conditions using roving-level noise stimuli. Distance performance for all anechoic conditions was at chance levels. In the reverberant conditions, subjects were able to extract distance information reliably. Surprisingly, reverberant monaural and binaural conditions yielded essentially equivalent performance. These results suggest that ILD cues (which are not available in monaural listening conditions) do not contribute to the perception of source distance when reverberation is present. Further, even when large ILDs are present in anechoic headphone simulations, they do not lead to robust distance percepts.

5. DISCUSSION

These results suggest that creating ILDs that vary with nearby source distance is not critical for simulating source distance in VAS. In contrast, overall level and reverberation cues provide compelling percepts of source distance in real-world listening conditions and under headphones.

Incorporating overall level effects into a VAS is straightforward and requires little computational power. In fact, both overall level and ILD cues can be approximated simply by adjusting the overall signal level at the left and right ears appropriately for the given source location [20]. In contrast, incorporating realistic reverberation into a display requires extensive computation [18]. Results discussed above suggest that binaural attributes in the reverberation are not critical to the perception of distance [19]. This may imply that shortcuts can be taken in modeling reverberation, since binaural differences in the reverberation have at best a minor impact on distance perception. Unfortunately, few studies address how the brain computes distance from reverberation and which attributes of reverberation are critical to perception.

6. ACKNOWLEDGEMENTS

AFOSR grant F49620-98-1-0108 supported this work. Tara Brown and Noro Kop o assisted with data collection.

7. REFERENCES

[1] Begault D. R., "Preferred sound intensity increase for sensation of half distance." *Percept Motor Skills*, 72: 1019-1029, 1991.

[2] Begault D. R., "Perceptual effects of synthetic reverberation on three-dimensional audio systems." *J Audio Eng Soc*, 40: 895-904, 1992.

[3] Bronkhorst A. W. and Houtgast T., "Auditory distance perception in rooms." *Nature*, 397: 517-520, 1999.

[4] Brungart D. S. and Durlach N. I., "Auditory localization of nearbv sources II: Localization of a broadband source in the near field." *J Acoust Soc Am*, 106: 1956-1968, 1999.

[5] Brungart D. S. and Rabinowitz W. M., "Auditory localization of nearby sources I: Head-related transfer functions." *J Acoust Soc Am*, 106: 1465-1479, 1999.

[6] Butler R., Levy E., and Neff W., "Apparent distance of sounds recorded in echoic and anechoic chambers." *J Exp Psych: Human Percept Performance*, 6, 1980.

[7] Coleman P. D., "Failure to localize the source distance of an unfamiliar sound." *J Acoust Soc Ama*, 34: 345-346, 1962.

[8] Coleman P. D., "Dual role of frequency spectrum in determination of auditory distance." *J Acoust Soc Am*, 44: 631-632, 1968.

[9] Duda R. O. and Martens W. L., "Range-dependence of the HRTF for a spherical head," *IEEE ASSP Workshop on Applications of Digital Signal Processing to Audio and Acoustics*, 1997.

[10] Duda R. O. and Martens W. L., "Range dependence of the response of a spherical head model." *J Acoust Soc Am*, 104: 3048-3058, 1998.

[11] Gardner M. B., "Distance estimation of 0-degree or apparent 0-degree-oriented speech signals in anechoic space." *J Acoust Soc Am*, 45: 47-53, 1969.

[12] Mershon D. H., Ballenger W. L., Little A. D., McMurtry P. L., and Buchanan J. L., "Effects of room reflectance and background noise on perceived auditory distance." *Perception*, 18: 403-416, 1989.

[13] Mershon D. H. and Bowers J. N., "Absolute and relative cues for the auditory perception of egocentric distance." *Perception*, 8: 311-322, 1979.

[14] Mershon D. H. and King L. E., "Intensity and reverberation as factors in auditory perception of egocentric distance." *Percept Psychophysics*, 18: 409-415, 1975.

[15] Mills A. W., "On the minimum audible angle." *J Acoust Soc Am*, 30: 237-246, 1958.

[16] Rabinowitz W. R., Maxwell J., Shao Y., and Wei M., "Sound localization cues for a magnified head: Implications from sound diffraction about a rigid sphere." *Presence*, 2: 125-129, 1993.

[17] Shinn-Cunningham B.G., "Learning reverberation: Implications for spatial auditory displays," *International Conference on Auditory Displays*, Atlanta, GA, 126-134, 2000.

[18] Shinn-Cunningham B. G., Lehnert H., Kramer G., Wenzel E. M., and Durlach N. I., "Auditory Displays," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. Gilkey and T. Anderson, Eds. New York: Erlbaum, 1997, pp. 611-663.

[19] Shinn-Cunningham B. G., Santarelli S., and Kop o N., "Distance perception of nearby sources in reverberant and anechoic listening conditions: Binaural vs. monaural cues," *23rd mid-Winter ARO meeting*, St. Petersburg Beach, FL, 88, 2000.

[20] Shinn-Cunningham B. G., Santarelli S., and Kop o N., "Tori of confusion: Binaural localization cues for sources within reach of a listener." *J Acoust Soc Am,* 107: 1627-1636, 2000.

[21] Vanderkooy J., "Aspects of MLS measuring systems." *J Audio Eng Soc*, 42: 219-231, 1994.