Depth - Distance

A sense of a sonic object's *distance* is another dimension in spatial hearing which contributes to the complete localization process.

Distance can be separated into 2 different categories: *absolute* and *relative*. A sense of *absolute* distance implies that the listener can identify the range of a sonic object upon first hearing it, and without any other references. (For example, upon hearing an isolated sound, the listener would be able to estimate the distance). *Relative* distance implies comparing the sound to other sounds (or to itself) at varying distances (Mershon & Bowers 1979) (Nielsen 1993).

Distance (or depth) is a spatial attribute often overlooked in audio representations. This is partly because a true sense of depth and distance is difficult to achieve over loudspeakers. In many recordings, we may ask whether the object *really* sounds further away, or are we just too easily accepting of the intended audio illusion simply because (for example) there is more reverberation attached to that sound?[1]

In stereophonic reproduction, the depth of field is limited by the distance of the loudspeaker pair to the listener. That is, auditory images cannot be perceived closer than the loudspeaker positions[2] (Wöhr, Thiele, Goere & Persterer 1991) (Nielsen 1993) (Michelsen & Rubak 1997). This may also be related to the so-called "proximity-image effect" reported by Gardner (1968a) where test subjects consistently (and wrongly) perceived the sounds as coming from the nearest of a line of 5 loudspeakers .[3]

---

[1]Imagine this in the context of a recording reproduced through a single, tiny loudspeaker placed off to one side.
[2]An illusion of sounds being closer than the loudspeaker pair can be somewhat obtained through manipulation of the signal polarity, but this method is unreliable.
[3]Perhaps this is also related to the "ventriloquism effect" (Pick, Warren & Hay 1969) and (Begault 1994, p. 84).

It should also be considered that sounds which are assigned (i.e. panned) completely to one loudspeaker (left or right) seem slightly more forward-projected than the same sound as a phantom-image between the 2 loudspeakers. (This effect is even more pronounced with rear "stereo" surrounds and side images in multichannel reproduction).

Whereas cues for azimuth (horizontal) localization are predominantly binaural, distance cues are more often monaural. That is, we can judge the distance of a sonic event without necessarily relying on interaural (binaural) differences.

A brief review of the various distance cues is necessary at this point towards an understanding of the performance of the proposed system in this respect.

Distance cues can be separated into 4 different categories:

1. Loudness differences
2. Spectral differences
3. Direct-to-reverb ratio differences
4. Binaural differences

1. A basic acoustics principle states that the sound pressure level (as received by a stationary observer) will fall 6 dB for every doubling of distance. This predictable rate of attenuation is valid only for anechoic (or free-field) environments and is often referred to as the "inverse-square law". This change in intensity that manifests itself as a loudness cue is the most relied upon for distance judgements (Coleman 1963) (Ashmead, LeRoy & Odom 1990) even though this reliance may result in error.

It is a relative cue otherwise (to be an absolute distance cue) it would require a high level of familiarity with the sound source (Mershon & King 1975).

It is also a monaural cue since we can perceive loudness differences (between 2 sonic events) with one ear (or via one loudspeaker).

2. Changes in distance cause 2 types of spectral changes. At great distances there is a high frequency attenuation due to absorption properties of air. But, as a sound source moves closer, its loudness increases which causes us to perceive an accentuation of bass frequencies as is evident by the equal-loudness contours (a.k.a. Fletcher-Munson curves). This low-frequency boost can be further accentuated when using directional pressure-gradient microphones which cause an increase in bass as distance decreases (known as the "proximity effect").

This is a relative cue otherwise, simple alterations/distortions in the frequency response of the audio signal path can result in misjudgement of distance. It is also a monaural cue.

3. A change in distance causes a change in the relationship between the direct and its reflected sounds. With increasing distance, the intensity of the direct sound diminishes as the reverberant energy remains essentially constant (Mershon & King 1975) (Wöhr, Thiele, Goere & Persterer 1991).

As a listener (or microphone) is moved increasingly outside the room-radius[4], there is a corresponding sense of increasing distance (or separation) as the reflected energy begins to dominate the direct sound.

[4] "room radius" is line surrounding a sound source where direct and reflected energy are equal.

Changes in the relationship between the 1st-order reflections and direct sound occur with distance. That is, the level and time[5] both become more integrated and fused between the reflection and its direct sound source (Kendall, Martins & Wilde 1990). For example, the 1st-order floor reflection appears at a relatively lower amplitude and time delay when the receiver is up close to the direct sound source. As the source is moved away, the level of the floor reflection and direct sound become more equal and, the time of arrival difference is shortened causing a more fused image of the 2 soundfield components. This is related to the so-called precedence effect. As the time delay gap between the direct sound and first reflections widens, it becomes easier to segregate the 2 soundfield components.

Since most people spend a greater proportion of time indoors, it stands to reason that they are better accustomed to localization judgments within room enclosures. It can be equally considered that the inclusion of room reflections both aids and hinders distance judgments. It aids, by providing additional cues (of a different dimension) to loudness. It hinders by complicating and obscuring the direct wavefront's otherwise predictable minus-6-dB per distance doubling inverse-square law.

Room-related cues to distance are monaural and mostly relative. They can be absolute if the listener is highly familiar with the room environment.

4. It is still unclear whether binaural differences are used by the psychoacoustic system to discern distance - however there is objective evidence that changes in distance, cause changes in interaural signals (Hirsch 1968) (Greene 1968) (Molino 1973) (Cochran, Throop & Simpson 1967) (Mershon & Bowers 1979) (Nielsen 1993).

Binaural differences are particularly significant when the sound source is at very close proximity where the separation between the ears and the size of the head become significant in relation to the source distance. At greater distances, the relative interaural-time (ITD) and interaural amplitude (IAD)

[5] The (ITDG) Initial Time Delay Gap (Beranek 1996) becomes smaller as distance increases.

differences are negligible. It might be assumed that since sounds that arrive from ±90° result in the greatest ITD and IAD values, distance judgment would have the strongest possible cue available from binaural processing. However, the research shows conflicting reports on whether this factor of localization ability is enhanced dependent on direction angle (Cochran, Throop & Simpson 1968) (Gardner 1969) (Holt & Thurlow 1969).

However for sources arriving from around 0°, another form of binaural difference comes into play with changes in the Inter-Aural Cross Correlation (IACC). As the sound source recedes into the distance, the binaural differences are less and the direct sound signals arriving at both ears becomes more correlated. At very close proximity, the head itself will impose appreciable differences resulting in a lower IACC value. As well, IACC would lessen as the room reflections would seem to come from a wider angle in relation to the sound source and receiver positions.[6]

John, what makes a sound appear nearer or far?
Here's a list I've come up with as a general starting point.
From there, anyone can apply the relevant effects (EQ, levels, panning, compression, reverb) that would relate to any of these points.

The closer sound object:

1. is overall louder…
   louder also means that (according to the Fletcher-Munson curves) will result in perceiving relatively more bass.
2. has more high-frequencies (further away the sound suffers more HF absorption)…so, a far away sound is perceived as more midrange-y, having both less bass and high-frequencies…. think of this in actual practice, with for ex,. a quiet speaking voice, extremely close at your ear, it seems very bassy, full, with high frequencies, move that speaking voice 20 feet away and it seems to have mostly midrange frequency energy.

3. has more transient detail – transients are clearer, more defined….the room obscures the transients and smears them in time as the sound moves further away.

4. has more <u>dynamic contrast</u> (as a sound moves further into a room, the room reflections wash out the sound and seem to compress it more)…again, think of the wider dynamic range of for ex. a strummed acoustic guitar inches away, as opposed to listening to it from several feet away where it sounds more dynamically consistent, smoothed out.

5. has more direct sound level over its reflected sound (early reflections, reverb)

6. the time delay between the direct sound and its reflections is greater. A further away sound will have the direct sound and reflections more fused and jumbled together in time, less separate. Closer up, they appear to you more distinct and separate (due to the distance of the room surfaces and the extra time to travel back to the POV of the mic/listener in most situations).

7. closer sound appears w I d e r… due to exaggerated difference of sound between both ears that diminishes when the sound object is further away…imagine a piano up close, then how its image becomes much less wide as you move several feet away.

---

[6] This factor should be designed into the DSP algorithms of artificial reverberation software to enhance the engineer's ability to control any sound element's distance.

Notes:
All of the points 1-6 don't require listening in stereo, only #7…so they can be perceived under many less than perfect stereo-seat listening conditions

> Although in addition, I've found that panning an echo/delay or mono short-decay reverb directly in the same place as the direct sound gives it a feeling of depth, rather than say, a wide stereo echo/reverb effect, or one that is panned away from the direct sound…that sounds to me more spacious, than deep and far.

References

Ashmead, D., LeRoy, D., and Odom, R., (1990). "Perception of the relative distances of nearby sound sources", *Perception & Psychophysics* vol. 47, no. 4, pp. 326-331.

Begault, Durand R., (1994). *3D Sound for Virtual Reality and Multimedia.* Academic Press Inc., Chestnut Hill, Mass..

Beranek, Leo. (1996). *Concert and opera halls: how they sound.* Acoustical Society of America, Woodbury, New York.

Cochran, P., Throop, J., and Simpson, W., (1968). "Estimation of distance of a source of sound", *American Journal of Psychology*, vol. 81, p. 198.

Coleman, P., (1961). "Failure to localize the source distance of an unfamiliar sound", *Journal of the Acoustical Society of America*, vol. 34, no. 3, p. 345.

Coleman, P., (1963). "An analysis of cues to auditory depth perception in free space", *Psychological Bulletin*, vol. 60, no.3.

Gardner, M. B., (1968a). "Proximity image effect in sound localization", *Journal of the Acoustical Society of America*, vol. 43.

Gardner, M. B., (1968b). "Distance estimation of 0° or apparent 0°-oriented speech signals in anechoic space", *Journal of the Acoustical Society of America*, vol. 45, no. 1, p. 47.

Greene, D., (1968). Comments on "Perception of the range of a sound source of unknown strength", *Journal of the Acoustical Society of America*. vol. 44, no. 2, p. 634.

Hirsch, H., (1968). "Perception of the range of a sound source of unknown strength", *Journal of the Acoustical Society of America*. vol. 43, no. 2.

Holt, R., and Thurlow, W., (1969). "Subject orientation and judgement of distance of a sound source", *Journal of the Acoustical Society of America*. vol. 46, pp. 1584-1585

Kendall, G., Martens, W., and Wilde, M., (1990). "A spatial sound processor for loudspeaker and headphone reproduction", *Proceedings of the AES 8th International Conference*, Washington, D.C., May.

Mershon, D., and King, E., (1975). "Intensity and reverberation as factors in the auditory perception of distance", *Perception and Psychophysics*, vol. 18, no. 6. pp. 409-415.

Mershon, D., and Bowers, J., (1979). "Absolute and relative cues for the auditory perception of egocentric distance", *Perception*, vol. 8, pp. 311-322.

Michelsen, J., and Rubak, P., (1997). "Parameters of distance perception in stereo loudspeaker scenario", *presented at the 102nd Audio Engineering Society Convention*, Munich, Germany, (March), preprint # 4472.

Molino, J., (1973). "Perceiving the range of a sound source when the direction Is known", *Journal of the Acoustical Society of America*. vol. 53, no. 5.

Nielsen, S., (1993). "Auditory distance perception in different rooms", *Journal of the Audio Engineering Society*, vol. 41, no. 10, (October).

Pick, H.L., Warren, D.H., & Hay, J.(1969). "Sensory conflict in judgements of spatial direction". *Perception & Psychophysics, 6.*

Wöhr, M., Thiele, G., and Goeres, H.-J., (1991). "Room-related balancing technique: A method for optimizing recording quality", *Journal of the Audio Engineering Society*, Sept., vol. 39, no. 9.