The Acoustics and Psychoacoustics of Workstation Audio Systems

Floyd E. Toole

Harman International Industries, Inc., 8500 Balboa Blvd., Northridge, CA 91329

Abstract: Workstations have become a common feature in both the professional and consumer worlds. The audio portion of programs created and reproduced on workstations include two-channel stereo, 5.1-channel matrixed and discrete, binaural and interactive binaural. The typical audio systems associated with computers cannot do justice to any of these formats, and the physical constraints limit the options. The paper reviews the technical requirements for such systems, describes what can be expected, and offers some advice on performing subjective evaluations. When properly done, the results can be surprisingly good.

WHAT IS EXPECTED OF A WORKSTATION AUDIO SYSTEM?

Most computer-based workstations employ loudspeakers that are tiny, not very good sounding, and floating somewhere on the desktop amidst a clutter of books, papers, printer and telephone. For simple auditory cues, and basic communication, this may be adequate, but if any of the more sophisticated auditory pleasures in our present and future are to be enjoyed, this will not do.

Workstations are now the production environment for all or much of the content in multimedia, games, film, TV and audio programming. They are also the environment in which such programs are frequently enjoyed by consumers. Good sound, high fidelity in the conventional sense, is now a requirement. But there is much more. The audio portions of these programs have moved beyond stereo, to include all of the multichannel formats, both matrixed for delivery through two channels (e.g. Dolby ProLogic) and digital discrete (e.g. Dolby Digital, MPEG2, DTS). In addition, binaural techniques are being employed to deliver a 3-D auditory experience in interactive games.

Multichannel sound tracks for films, TV and games require five loudspeakers, left, center and right front, and left and right surrounds at the sides. Add a subwoofer, and we have an amount of clutter that is not practical in a workstation. A further problem is the closeness of the loudspeakers. It would be more realistic if the sounds were perceived to be more distant. Consequently, the industry is adopting binaural simulation techniques that allow listeners to experience the illusion of surround sound, using only two real loudspeakers with acoustical crosstalk cancellation. Phantom images of five loudspeakers are created at the angular locations where real loudspeakers would normally be positioned, and with a perceived sense of distance which is not related to the distance of the real loudspeakers. As might be expected, the technique is also very effective in enhancing the reproduction of conventional stereo music [1].

EVALUATING BINAURAL WORKSTATION AUDIO SYSTEMS

Physically, a workstation can be the familiar desktop computer setup most of us work and play with, or it can be a large, densely-packed array of picture and sound editing apparatus in a professional production studio. Both present acoustical challenges, because of dimensional variations and acoustically reflecting surfaces and objects in the vicinity of the loudspeakers. Binaural reproduction methods require that there be a well-defined geometrical relationship between the two loudspeakers and the listener's head. Since the success of the illusion relies on the excellence of acoustical crosstalk cancellation, the integrity of the sound path between the loudspeakers and ears is important. Reflections can be problems. Both of these concerns argue for standardization of the workstation arrangement. The adjacent illustration is part of a suggestion proposed to the AES [2].
The rest of the concerns are related to the specific locations of the loudspeakers relative to the monitor, and to the proximity of walls, office dividers, etc.

The performance of a binaural audio system can only be evaluated by listening. Listening tests, to have validity, must be conducted with proper scientific rigor [3]. Part of that is the requirement that the products are hidden from view, and that rapid comparisons of different products are possible. With computer audio systems this requires, for example, a positional-substitution turntable, with the test systems arranged on its periphery behind an acoustically transparent, visually-opaque screen.

The nature of binaural sound reproduction techniques changes some of the ways in which a listening evaluation is done. Some of the factors are known, but more needs to be understood about their sensitivity to measurable quantities. For example:

1. A jury evaluation is essential. In conventional loudspeaker evaluations, differences among listeners tend to be related to their listening experience and their hearing thresholds [4]. With binaurally encoded or processed sounds it is also necessary to consider individual differences of anything related to the all-important head-related transfer functions (HRTF's). Simplifying assumptions are made about HRTF's in the binaural encoding of the directional information, and in the cancellation of the acoustical crosstalk. Consequently, the perceived directional and spatial illusions will not be identical for all listeners. No one person's opinion is absolute, except for that person.

2. Loudspeakers. Left-right matching of the loudspeakers is critical, otherwise the acoustical crosstalk-cancellation process cannot work.


5. Program. The spectrum and bandwidth of the test sounds can influence what is perceived.

6. Program. Steady-state sounds are more difficult to localize than sounds rich in transients.

7. Environment. Reflections and diffracted sounds from apparently innocent surfaces and objects can influence the localization of images, especially those at the sides.

Test sounds should be varied, including broadband music, game and movie effects, and pink noise steered to the phantom loudspeaker locations.

Listenors should be instructed to listen for incorrect locations of the images, ambiguously localized images, components of the sounds localized at the loudspeakers (there should be no "leakage"), and front-back reversals.

Listeners should also evaluate the nature of the "sweet spot". In any crosstalk-cancelled loudspeaker binaural system capable of precise directional localizations there must be a listening location where the auditory "picture" is in sharpest focus. In practice listeners will move around, so it is important that the sound illusion not collapse rapidly as the head is rotated left or right. There should not be any discomfort or anomalous effects associated with sideways movements. With large lateral movements, the illusion should gracefully degrade to something resembling off-axis stereo.

A SURPRISING CONCLUSION

The crosstalk-cancelled loudspeaker reproduction technique is antisocial, since only one listener can be properly entertained at a time. This limitation it shares with conventional stereo. However, the limitless number directional effects (or phantom loudspeakers) and the persuasively-real impressions of envelopment, distance and space, can rival, and in some ways surpass, conventional stereo or multichannel loudspeaker systems. In this respect, the close-listening situation of a workstation is not a disadvantage at all. But it must be done properly.

REFERENCES