Evaluation of changes in stereophonic imagery during progressive optimization of loudspeaker configuration in critical listening rooms

David Spargo, William L. Martens, and Densil Cabrera

Faculty of Architecture, Design, and Planning, University of Sydney, Australia

ABSTRACT

While simple approaches to the configuration of loudspeakers in critical listening rooms are well known, there is no standardized practice for optimising loudspeaker positions that is based upon improvements in the perceived quality of the stereophonic imagery associated with adjustments to the loudspeaker positions. As the influence of the acoustical response of the reproduction space on stereophonic imagery is not so simple to predict, a pragmatic solution to this problem has been developed and is growing in popularity, particularly among recording engineers. This method utilising changes in stereophonic image quality that occur when moving either the listening position or loudspeaker positions in an effort to find an optimum configuration within a listening room. While the method gives good practical results, there is as yet no explanation as to how or why the method works, and no consensus regarding the characterisation of changes in auditory imagery to which listeners must attend. To provide a more scientific foundation for explaining this method, the steps involved in optimisation were documented for a test case, and the changes in the binaural responses during the adjustment of loudspeaker positions were analysed in relation to loudspeaker configurations that were preferred by a group of ten listeners.

INTRODUCTION

Critical listening rooms for 2-channel stereophonic sound production are intended to provide the listener with a reasonable degree of confidence that sound will be presented in a predictable and consistent way, allowing the construction of sound mixes that translate well to other rooms and channel formats, such as monophonic or multi-channel reproduction. While standards and design criteria help by providing guidance on factors such as reverberation time, which varies with room volume, the recommended loudspeaker and listener relationships are assumed more or less fixed regardless of other varying room conditions (see for example technical recommendations of the European Broadcasting Union, 1998, and the International Telecommunications Union, 1994).

Many studies in the past have focused on image or sound stage qualities in sound reproduction systems as influenced by the room properties, and conclude that the latter changes the former. Common themes in these studies are; rigid adherence to the traditional loudspeaker setup geometry (the well-known equilateral triangle formed between the listener and the left and right loudspeaker positions), and control of reflected sound, mainly lateral, to alter stereophonic imaging. For those interested in further information on the topic the authors suggest the recent book by Floyd Toole (Toole, 2008), which provides a rich synthesis of findings.

It is well-established that changes in room acoustic conditions result in changes in image quality from fixed loudspeaker positions (Olive et al., 2007). The corollary must also be true, but no well-established method exists for adapting loudspeaker configurations to fixed room acoustic conditions. Rather than follow a simple rule of thumb for such an important issue, it should be better to determine a robust method to optimise image quality in a given room through the positioning of the loudspeakers.

A practical technique for use by recording engineers has been suggested by Stavrou and outlined in his recent book (Stavrou, 2003). The method is designed to optimise image quality for critical listening in any room. However, there is as yet no physical or psychoacoustic explanation of how the method works. This study examines a method for quantifying the repeatability and the efficacy of Stavrou's method and begins the process of identifying the qualitative and quantitative elements that may lead to a prediction model.

METHOD

Subjects

Ten male listeners participated in this study. All had considerable experience in audio and acoustics, ranging from 21 to 51 years of age, and all reporting normal hearing.

Stimuli and Apparatus

A room with the dimensions 6.5 m x 4.8 m x 4 m was set up as a listening room. A pair of two-way bookshelf loudspeakers was selected with wide high frequency dispersion angle and 8” bass drivers. The loudspeakers were mounted on a long board centred on the room's length axis and positioned 1.6 m from the rear wall as shown in Figure 1. Hard plastic feet were attached to the loudspeakers so they would slide along the board's length. The board was marked with graduations in 5 cm steps symmetric about the centre and room centre line, 0 being the centre of the board. This experimental set-up follows closely that prescribed by Stavrou.

The two loudspeakers received the same signal for this test, which was a mix-down of the two-channel recording of “Penny Lane” by the Beatles. During the familiarization and test phases of this study, the entire track was played continuously from a CD source, and was repeated indefinitely until the task was complete.

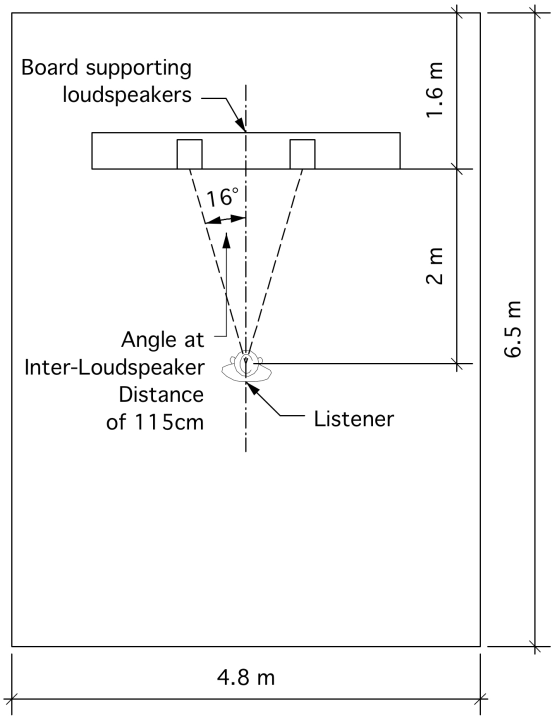

***Figure 1**. Listening room set-up, viewed form above, with a dotted line along the length axis of the room showing the 2-meter distance between listener and the plane containing the loudspeakers. Also illustrated is the offset angle from this plane formed when the Inter-Loudspeaker Distance is set to 115cm along the board supporting the loudspeakers (see text for details).*

Experimental Task

Subjects were seated upon a stool at the position indicated in Figure 1 facing the loudspeakers, with the level of their ears roughly matched to the height of the speakers as measured at the centre of the loudspeaker's 8" bass driver. Before beginning a listening session, they first received the following instructions (as described by Stavrou) about what to listen for during the task:

> Imagine a two-foot high ribbon that stretches between the speakers, supporting the plane of perception that sounds are projected upon. When the speakers are very close together, the ribbon is tight with a forward bulge in the middle. This bulge is instantly recognisable as extreme colouration. As your speakers move apart, the ribbon begins to straighten and the colouration disappears. But as they continue to move apart further, the ribbon begins to stretch, the centre thins out, gets a hole in it, and the sound loses depth and solidity. Continuing to pull them apart further still disrupts the overall acoustic impression even more – until it feels more like a cocktail party with subtle surround-sound effects. The ideal position is the one where the ribbon is dead straight, without being stretched or lumpy or full of holes.

After receiving these instructions, listeners were asked to attend to the variation in auditory imagery produced by the stereophonic loudspeaker pair in various configurations. At the beginning of each session, each listener was exposed to a sequence of positions that started with an extreme inter-loudspeaker distance of about 200 cm, and was gradually reduced in 10 cm steps until the distance between the speakers reached the minimum of about 25 cm. Then, the direction of the change in inter-loudspeaker distance was reversed, and gradually increased until the listener indicated that the image was beginning to "break up" or "degrade in quality."

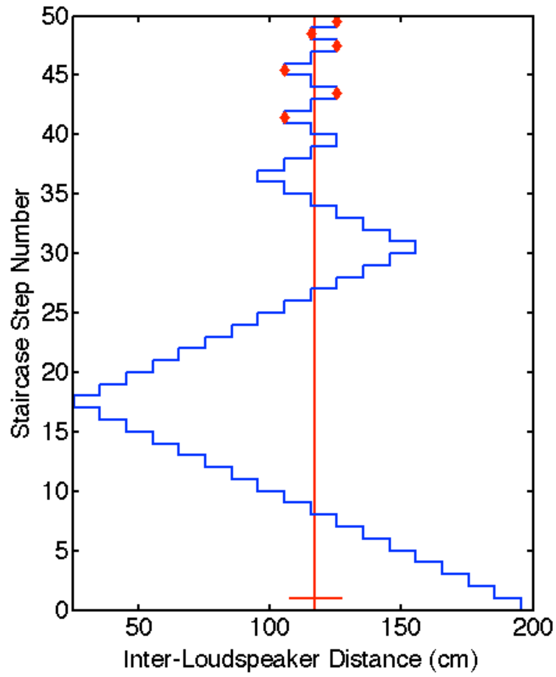

***Figure 2.** Example sequence of the inter-loudspeaker distance settings visited during a single staircase session tracking a listener's preferred auditory image in stereophonic loudspeaker reproduction. The staircase step number proceeds from the start of a listening session at the bottom of the plot toward the terminal step at the top of the plot (in the plotted case, this would be step number 49). The diamond-shaped symbols marking the extreme inter-loudspeaker distances visited during the final 6 sweeps near the optimum configuration that the staircase was intended to estimate. The solid vertical line shows the computed average level of 117.5 cm (based upon the final 6 turnarounds). The solid horizontal line that crosses the line drawn at the computed average level shows the standard deviation calculated for those final inter-loudspeaker distances for the last 6 turnarounds).*

These first two sequences of step-wise changes in inter-loudspeaker distance on auditory imagery are illustrated at the bottom of Figure 2 (swinging from right to left and then rightward).

Subsequent to this systematic examination of the influence of a wide range of inter-loudspeaker distances on reproduced auditory imagery, listeners began a more tightly controlled session in which an optimal configuration was tracked via a staircase method designed to converge upon that optimum. At each inter-loudspeaker distance visited, listeners were asked to indicate whether the optimum might be reached by continuing in the current direction of changing distance, or whether that direction of change should be reversed. Thus, while focussing on the character of the reproduced auditory imagery during the movement of the loudspeakers, the range of inter-loudspeaker distance visited was gradually reduced throughout the course of the staircase session. This convergence can be appreciated in the graphic depiction of the staircase track that swings through a narrower and narrower range as the staircase step number increases (towards the top of the plot). Once the staircase began to alternate between just two inter-loudspeaker distances, the listener was asked to make a judgment of which of the two seemed to be the more optimal. This was taken as a self-report of optimum inter-loudspeaker distance that corroborated the more objective measure of the optimum point that was based upon the common psychophysical staircase method described by Levitt (Levitt, 1971). Figure 2 shows how the staircase method provides an estimate of the optimum inter-loudspeaker distance for a given session (see figure caption). These experimental sessions typically took listeners only about 10 minutes to complete.

RESULTS

Listening Task

The results of the staircase tracking procedure for the ten listeners tested are illustrated in Figure 3. The optimum calculated inter-loudspeaker distance for the ten has been sampled into a histogram with an interval width of 10 cm. It can be seen that the distribution of optimum inter-loudspeaker distances is concentrated around a single preferred value, since five of the ten listeners converged on a point within the histogram bin centred on the value of 115 cm. The median value was indeed quite near the centre of this bin, at 115.6 cm (indicated by the vertical dashed line). So, at least for the single case tested, and for the single musical program employed, listeners are in good agreement.

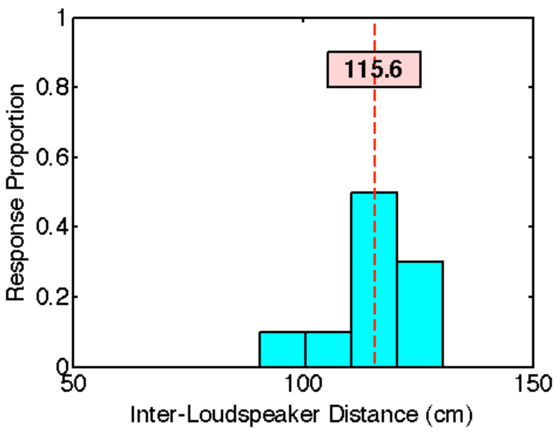

Figure 3. *Histogram showing the distribution of optimum inter-loudspeaker distances calculated for ten listeners tested. The ordinate shows the relative response proportion summed over the results for ten listeners within histogram bar width of 10 cm.*

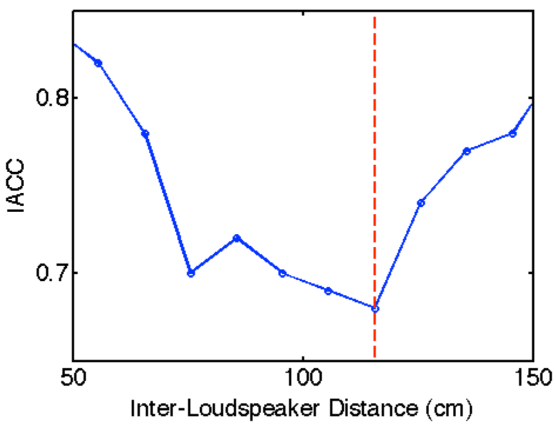

Figure 4. *Plot of IACC as a function of the inter-loudspeaker distances presented in the current study, calculated on measured broadband impulse responses. As in Figure 3, the vertical dashed line shows the optimum inter-loudspeaker distance that was calculated from the results of ten listeners.*

Prediction from Acoustical Analyses

Although a broader sampling of listening spaces would certainly provide more confidence in a prediction of optimized configuration in general, the measurements associated with loudspeaker configurations examined in the current study do show a strong relation with the parameter termed inter-aural cross correlation (IACC). The IACC values calculated on the measured binaural impulse responses at each of the tested inter-loudspeaker distances are plotted in Figure 4, and show a clear minimum near the optimum inter-loudspeaker distance that was calculated from the results of ten listeners. Low values of IACC are associated with listener preference for critical listening rooms and also predict greater perceived spaciousness (D'Antonio et al., 1989). It remains to be seen whether such a relation holds across a broad sampling of listening spaces.

DISCUSSION AND CONCLUSIONS

The current results provide preliminary confirmation of the effectiveness of a prescribed practice for optimizing loudspeaker positions that is based upon improvements in the perceived quality of the stereophonic imagery associated with shifts in the distance between loudspeakers. While the method was thought to give good practical results, there was systematic evaluation of the method’s effectiveness, and no explanation why the method works. On the basis of the findings reported here, it is possible to propose a model for predicting such results in other listening spaces just by examining measured changes in the binaural responses during the adjustment of loudspeaker positions. It is interesting to consider how different the selected loudspeaker configuration might be using the prescribed method rather than using the conventional rule of thumb that has not so often been questioned.

The convention for two-channel stereophonic loudspeaker geometry has been around since the mid 1930s. The original concept of the equilateral triangle formed between two loudspeakers and a listener was developed in a *free field* and was designed to reproduce the spaciousness of a concert experience. While this experience is undoubtedly enjoyable, particularly with a well-recorded acoustic performance, it presents problems in critical listening rooms where conditions are not free field and loudspeaker/room interactions tend to increase spaciousness. In these rooms accurate rather than enjoyable reproduction is desired.

Stavrou recognised this problem and developed his heuristic method based upon his substantial experience in critical listening and in applying this technique. Anecdotal evidence also suggests that loudspeakers should be closer together than the traditionally recommended equilateral triangle configuration to counter these interactions. Our results support the use of a smaller angle, at least for the reverberant room in which we tested.

REFERENCES

D'ANTONIO, P., KONNERT, J., BECKER, F. & BILELLO, C. (1989) Sound Intensity and Interaural Cross-Correlation Measurements Using Time-Delay Spectrometry. *Journal of The Audio Engineering Society,* 37**,** 1-15.

EUROPEAN BROADCASTING UNION (1998) Technical Document Tech 3276–Listening Conditions for the Assessment of Sound Programme Material: Monophonic and Two Channel Stereophonic, 2nd ed.

INTERNATIONAL TELECOMMUNICATIONS UNION - RADIOCOMMUNICATION ASSEMBLY (1994) ITU-R. Recommendation BS.775-1, Multichannel Stereophonic Sound Systems with and without Accompanying Picture.

LEVITT, H. (1971) Transformed up-down methods in psychoacoustics. *Journal of Acoustic Society of America,* 49**,** 467-477.

OLIVE, S. E., WELTI, T. & MARTENS, W. L. (2007) Listener Loudspeaker Preference Ratings Obtained In Situ Match those Obtained Via a Binaural Room Scanning Measurement and Playback System. *The 122nd Convention of the Audio Engineering Society*. Vienna, Austria.

STAVROU, M. P. (2003) *Mixing with your mind: Closely guarded secrets of sound balance engineering revealed*, Flux Research.

TOOLE, F. (2008) *Sound Reproduction Loudspeakers and Rooms*, Focal Press.