Twelfth International Congress
on Sound and Vibration

PERCEIVED ROOM SIZE AND SOURCE DISTANCE IN FIVE SIMULATED CONCERT AUDITORIA

Densil Cabrera[1], Andrea Azzali[2], Andrea Capra[2], Angelo Farina[2] and Paolo Martignon[2]

[1]School of Architecture, Design Science and Planning, University of Sydney
Sydney, NSW 2006, Australia
[2] Industrial Engineering Department, Università di Parma
Via delle Scienze, 43100 Parma, Italy
densil@arch.usyd.edu.au

Abstract
Room impulse responses (stereophonic and binaural) were recorded using identical procedures in five concert auditoria: the large, medium and small halls of Rome's Parco della Musica, Parma's Auditorium Paganini, and Kirishima's Miyama Conseru. A subjective test was conducted, using an anechoic recording of a piano accordion convolved with these impulse responses, and reproduced over four audio systems (binaural headphones, conventional stereophony, stereo dipole and double stereo dipole). The frequency responses of these playback systems were matched, and the playback level was realistic for the instrument and auditoria. Results indicate that the audio system can significantly affect distance estimation and realism judgments (with a non-significant effect on room size ratings). The situation (auditorium and seat position) significantly affects distance estimations and room size ratings (but not realism ratings). Room acoustical parameters were determined from for each situation, and relationships between these and the auditory distance and room size judgments were investigated. Auditory distance perception is primarily correlated (negatively) to the reproduced sound pressure level. Results for perceived room size are more complex, and demand further research for interpretation. Results tentatively indicate that among these non-individualized two-channel systems, the single stereo dipole system provides a plausible reproduction of the auditorium acoustical environment for the purpose of subjective assessment.

INTRODUCTION

In auditorium acoustics, the concept of auditory spatial impression has received great attention, since aspects of spatial impression are believed to be related to auditorium acoustical quality for music. The concept and components of auditory spatial impression vary somewhat between researchers, but apparent source width (ASW) and listener envelopment (LEV) are frequently studied. A subtlety with ASW and LEV is that, as subjective phenomena, they do not simply relate to room geometry, and their physical basis in the auditorium sound field remains a matter for research. However, the auditory impression of distance and room size, which could also be classified as aspects of 'auditory spatial impression', do have physical counterparts for comparison (namely physical source-receiver distance, and physical room size). Based on previous studies of auditory room size perception and auditory distance perception, it seems likely that physical room size and distance should correspond to listener perceptions [1, 2], but that room acoustical conditions will also significantly affect these [3] in music auditoria.

Studying auditory impression of multiple auditoria almost necessarily relies upon electro-acoustic simulations of these spaces, because these give the ability to instantly switch between listening situations even between the most remote auditoria. Simulations also have the advantage of greatly simplifying the stimulus presentation conditions, and so limiting the influence of uncontrolled variables on subjective responses. However, in many practical situations, simulation systems are limited to two non-individualized channels of audio. The present study investigates auditory distance and room size as it is conveyed by four such systems. Aspects of this project pertaining to audio system quality have been discussed previously [4]. This paper considers the results of the study especially as they pertain to auditorium acoustics.

PROCEDURE

This study exploits an archive of concert hall impulse responses (IRs) made by Farina and colleagues. As outlined by Farina and Ayalon [5], these impulse responses had been made for listening applications rather than merely to obtain numerical indices. An important characteristic of them was that the equipment was constant throughout, and the signal gain structure was documented. The auditoria selected for this project were the large, medium and small halls of Rome's Parco della Musica (abbreviated to RL, RM and RS), Parma's Auditorium Paganini (P), and Kirishima's Miyama Conseru (K), and these auditoria were chosen because of the consistency of the measurement method. Binaural, stereophonic (ORTF format) and first order ambisonic (B-format) IRs were recorded for various receiver positions within the audience areas of each hall, with the source on the stage. Only the binaural and stereophonic IRs were used for sound reproduction in the present study. Two receiver positions in each auditorium were selected for this study.

For this project, the direct sound component of the IRs was replaced with idealized versions – a single sample impulse for the ORTF IRs, and a measured 0° anechoic dummy head IR for the binaural IRs. The primary reason for this treatment of the direct sound was the problems with source directivity for dodecahedral loudspeakers in the high frequency range. The energy of these idealized direct sound IRs were matched to the originals at 500 Hz (where the loudspeaker is essentially omnidirectional), with a broadband 3 dB gain to approximate the effect of a weakly directional source. This process described further elsewhere [4].

An anechoic recording of a piano accordion was recorded for this project, along with a calibration tone for the recording microphone, allowing realistic playback levels to be approximated. The music was "La ballata di Michè", by Fabrizio de Andrè: a waltz, with a legato melody and articulated accompaniment. This was 45 s in duration, and it had an anechoic L_{eq} of 80 dBA at 1 m. This recording was convolved with the selected edited binaural and stereophonic IRs, preserving their relative gains.

Table 1 – Key features of the auditoria, auditorium impulse responses and stimuli, including source-receiver distance (*r*), maximum length, mid-length width, number of seats, stereophonic (*SPLS*) and binaural (*SPLB*) stimulus sound pressure levels (unweighted, measured from a dummy head in the listening room), reverberation time (*T30*), early decay time (*EDT*), clarity index (*C80*), speech transmission index (*STI*, assuming no background noise), bass ratio (*BR*), treble ratio (*TR*), inter-aural cross correlation coefficient (*IACC*), and lateral fraction (*LF*).

	r (m)	Length (m)	Width (m)	seats	*SPLS* (dB)	*SPLB* (dB)	*T30* (s)	*EDT* (s)	*C80* dB	*STI*	*BR*	*TR*	*IACC*	*LF*
RL1	21	56	32	2800	66	67	2.43	2.4	-0.1	0.49	1.04	0.61	0.15	0.32
RL2	30				65	65	2.44	2.2	-0.2	0.47	1.01	0.62	0.15	0.16
RM1	10	48	34	1200	71	69	1.89	1.5	1.3	0.53	1.09	0.58	0.18	0.36
RM2	31				66	67	2.04	1.8	-0.5	0.46	1.08	0.60	0.17	0.22
RS1	12	35	25	700	71	71	1.82	1.9	-0.6	0.47	1.16	0.67	0.14	0.29
RS2	24				69	70	1.87	1.8	-1.4	0.48	1.10	0.65	0.15	0.28
P1	13	48	17.5	780	72	72	2.29	2.3	-1.4	0.46	1.04	0.67	0.17	0.35
P2	22				68	70	2.34	2.2	-2.2	0.45	1.01	0.69	0.12	0.31
K1	8	45	30	770	73	74	1.95	1.8	-0.1	0.48	0.98	0.71	0.48	0.20
K2	24				74	71	1.95	1.7	0.0	0.54	0.99	0.72	0.45	0.13

A listening room was set up, with four audio playback systems: a conventional stereophonic array, stereo dipole [6] (which is a type of cross-talk cancellation, but with the loudspeaker pair having a narrow angle of separation in front of the listener), double stereo dipole (which adds a rear cross-talk canceling pair to front stereo dipole) and headphones. Thus, there was one stereophonic system, and three binaural systems. The system frequency responses and gains were carefully matched within the frequency range of the sound stimuli. The listening room and playback systems are described in more detail elsewhere [4].

In a subjective experiment, listeners rated the realism of each recording and the apparent room size, and also estimated the distance to the performer. The 30 subjects were experienced with listening to musical performances (the experiment was conducted in a music school), and were asked to imagine themselves in the auditoria, rather than in the darkened listening room. Each subject assessed five auditoria and two audio systems, with stimuli counterbalanced over the subject group. The experiment was conducted using purpose-written software which was presented on a screen above the stereo dipole loudspeaker pair, controlled with a wireless mouse. Subjects could switch between stimuli at any time, using buttons on the display. The three questions were presented in random order to the subjects, and they had to assess all ten stimuli on the first question before the second question became active – and likewise for the third question. When the headphones were used, subjects were instructed by the software to put on and remove the headphones when appropriate.

RESULTS

Auditory Distance Estimation

Analysis of variance (ANOVA) shows that the auditorium situation has a substantial effect on auditory distance (f=11.45, p<0.0001, df=9), with a weaker effect for the audio system (f=3.86, p=0.0099, df=3). Using logarithmic distance units (log base 10 of metres), the rms errors for the four audio systems are: stereophony = 0.23; headphones = 0.28; stereo dipole = 0.19; and double stereo dipole = 0.22. While this gives one indication that stereo dipole results are closest to veridical, the results do not compare the audio systems with *in situ* listening – which also could be expected to deviate somewhat from veridical [1]. Results are shown in Figure 1.

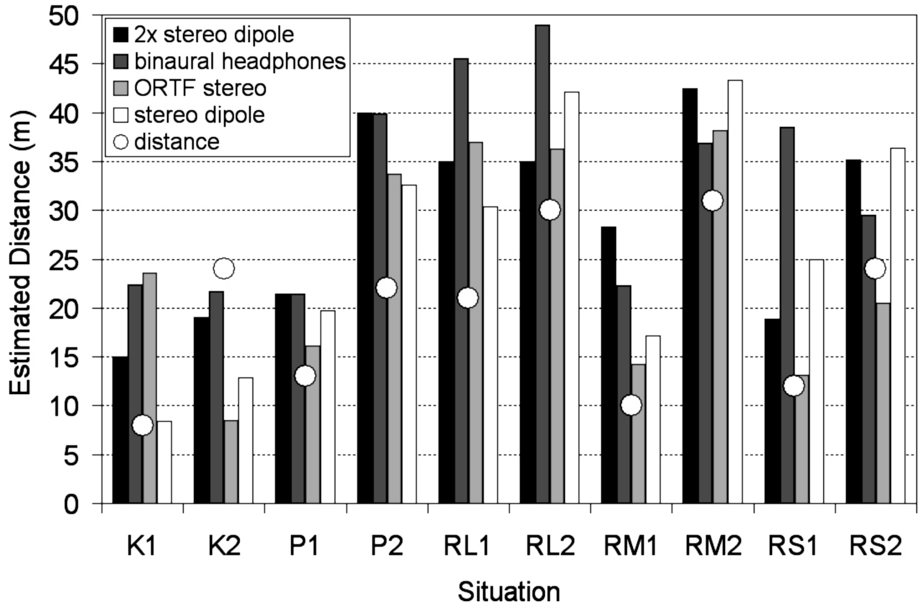

Figure 1 – Mean auditory distance estimates for the ten situations and the four audio systems, also showing the actual source-receiver distances.

Results suggest that auditory distance perception *within* an auditorium may, or may not, be strongly affected by the actual source-receiver distance. In RM, actual and estimated distance are closely coupled (especially for ORTF stereo and stereo dipole); in K, there is no apparent effect of actual distance on estimated distance; and in P, the effect appears to be exaggerated. In the case of K, Table 1 shows that position K2 has acoustical characteristics remarkably similar to K1. Sound pressure level varies little with distance (especially for the ORTF system), and minor changes in some parameters (*SPLS*, *EDT*, *C80*, *STI*) run counter to normal expectations for an increase in distance from K1 to K2. Hence, the unusual acoustical design of K creates a situation running counter to conventional expectations, such that ratings for all audio systems have underestimations for K2. P is also a distinctive auditorium, in that it is long and narrow – and as such its form is opposite to that of K. P is also distinctive for its long (unoccupied) reverberation time, and the associated low C80 values.

Cabrera *et al.* [7] have previously found that stimulus sound pressure level can better predict auditory distance estimates than actual source-receiver distance in auditoria (using binaural headphone reproduction of music). That is also the case in the present study for all audio systems, as shown in Table 2 and illustrated in Figure 2.

Table 2 – Correlations (*r*) between auditory distance estimates and sound pressure level or source-receiver distance for the four audio systems.

	2x Stereo dipole	Headphones	ORTF	Stereo dipole
SPL	-0.76	-0.79	-0.86	-0.82
Source-receiver distance	0.73	0.56	0.57	0.79

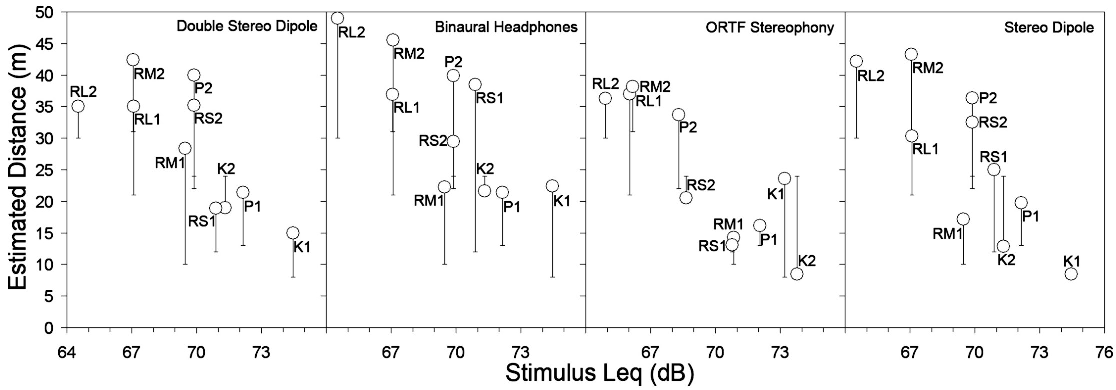

Figure 2 – Relationship between distance estimates and stimulus sound pressure level (unweighted binaural in the listening room) for the four audio systems. Error bars show deviation from actual source-receiver distance.

Auditory Room Size Ratings

Room size ratings are significantly affected by situation (*f*=6.89, *p*<0.0001, *df*=9), but not by the audio system (*p*=0.066). Nevertheless, there are some notable divergences between the audio systems, especially comparing ORTF stereophony with the three binaural systems. Ratings for ORTF stereophony are highly correlated with auditory distance estimates (r^2=0.90), whereas ratings for the binaural systems have lower correlations: 0.74 for double stereo dipole; 0.72 for headphones, and 0.34 for stereo dipole. As might be expected from this observation, the ORTF system results correlate best with sound pressure level (r=-0.73), whereas correlations are weaker for the binaural systems (-0.67, -0.54, and -0.43 for double stereo dipole, headphones, and stereo dipole respectively). Considering the mean results for the four audio systems, for three auditoria (K, RL and RS) the estimated room size is not affected by the source-receiver distance within each. However, P and RM both see an increase in room size ratings with source-receiver distance.

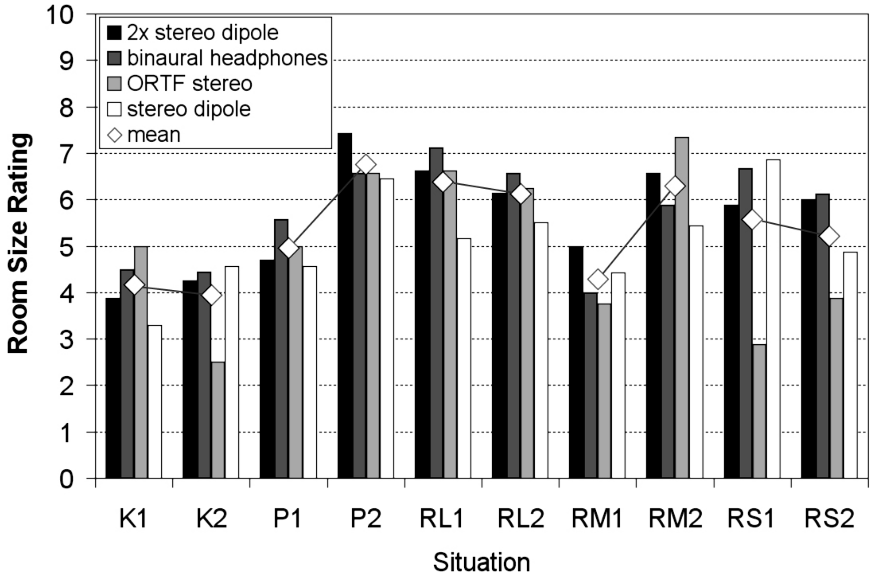

Figure 3 – Mean auditory room size ratings for the ten situations and the four audio systems.

IACC measurements for most situations are below 0.2. However, the measurements in K have *IACC* measurements close to 0.5. Correlations between room size ratings and acoustical parameters appear to show some influence of *IACC* for the binaural systems (but not ORTF stereophony), both with and without the K measurements – such that a lower *IACC* is associated with a larger room size rating (Table 3).

Table 3 – Correlations (r) between room size ratings and *IACC*.

	2x Stereo dipole	Headphones	ORTF	Stereo dipole
All Situations	-0.79	-0.69	-0.37	-0.69
Without K	-0.74	-0.74	-0.11	-0.79

Realism Ratings

ANOVA shows that realism is not affected by situation (p=0.29), but is affected by audio system (f=4.1, p=0.0068, df=3). The ORTF and stereo dipole systems receive greater realism ratings than the double stereo dipole and headphone systems. The fact that the listener's head was unrestrained is a likely explanation of double stereo dipole system's poor evaluation. Having the rear loudspeaker pair defeats one of the key advantages of stereo dipole – namely that it is robust in the face of listener head movements [8]. The rear loudspeaker pair reduces the 'sweet spot' through interference with the front pair. The low evaluation of the headphones is probably of more interest, because of the widespread use of headphones in non-individualized binaural reproduction and simulation. Such systems are generally unable to produce a frontally located auditory image, instead yielding images with poorly defined location, tending to be above or behind the listener [9]. Furthermore, the space rotates with the listener's head in headphone reproduction (without the use of head tracking).

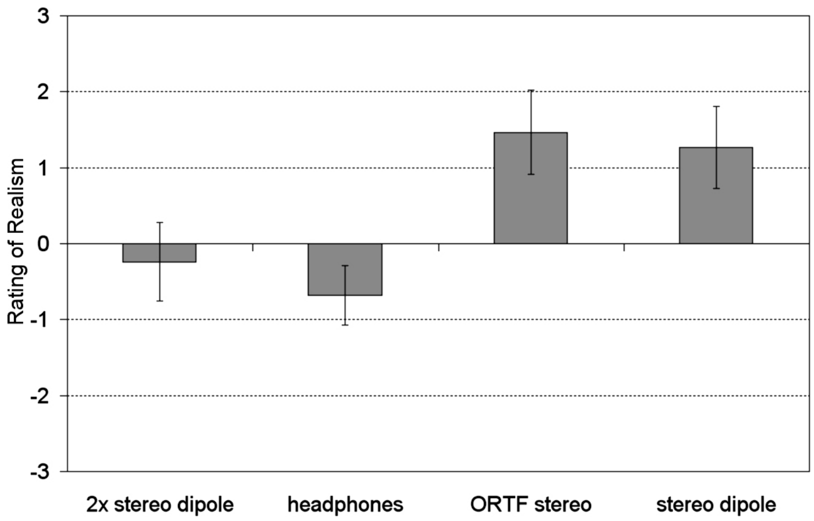

Figure 4 – Mean realism ratings for the four audio systems (±1 standard error).

CONCLUSIONS

This study appears to show that auditory distance perception in an auditorium is generally related to, but not necessarily closely coupled to, actual source-receiver distance. Strength factor (which is represented by stimulus sound pressure level in this study) can be a substantial influence on auditory distance estimates. The implication is that an auditorium can be designed for greater or lesser auditory intimacy, and that intimacy can be controlled within an auditorium through acoustic design (cf. [10], p.43).

Room size ratings in this study are not clearly related to the actual room size of the auditoria (except for the ORTF ratings, which are scarcely different to auditory distance estimates). Therefore it is not clear from the results whether room size is discernible through the sound of rooms such as these auditoria. As mentioned in the introduction, previous studies have shown that subjects can hear the size of rooms of substantially different size, but room acoustical conditions can have a larger effect than actual room size. The present results appear to show some influence of *IACC* on room size ratings.

All of the audio systems tested here yield approximate renderings of the situations, but greater accuracy is available through individualized binaural systems, high order ambisonics and wave field synthesis. However, two-channel non-individualized systems are much easier to work with than the more accurate alternatives, and will continue to be in widespread use for auditorium simulations for many years. The results of this study indicate that the stereo dipole system provides a more effective and realistic impression of the auditorium sound field than the other non-individualized two channel alternatives.

ACKNOWLEDGMENTS

The authors are grateful for the assistance of Alberto Amendola, Paolo Bilzi, ASK Industries, Casa della Musica, and Tommaso Dradi (piano accordion) in this research project.

REFERENCES

[1] P. Zahorik, "Assessing auditory distance perception using virtual acoustics," J. Acoust. Soc. Am., 111, 1832-1846 (2002)

[2] R. McGrath, T. Waldmann, M. Fernström, "Listening to rooms and objects," 16th Audio Eng. Soc. Int. Conf., Rovaniemi, Finland, 512-522 (1996)

[3] D.H. Mershon, W.L. Ballenger, A.D. Little, P.L. McMurtry, J.L. Buchanan, "Effects of Room Reflectance and Background Noise on Perceived Auditory Distance," Perception, 18, 403-416 (1989)

[4] P. Martingon, A. Azzali, D. Cabrera, A. Capra, A. Farina, "Reproduction of auditorium spatial impression with binaural and stereophonic sound systems," 118th Audio Eng. Soc. Conv., Barcelona, Spain (2005)

[5] A. Farina, R. Ayalon, "Recording concert hall acoustics for posterity," 24th Audio Eng. Soc. Conf., Banff, Canada (2003)

[6] O. Kirkeby, P. A. Nelson, H. Hamada, "The 'stereo dipole' – a virtual source imaging system using two closely spaced loudspeakers," J. Audio Eng. Soc., 46, 387-395 (1998)

[7] D. Cabrera, A. Nguyen, Y.-J. Choi, "Auditory versus visual spatial impression: a study of two auditoria," Int. Conf. on Auditory Display, Sydney, Australia (2004)

[8] T. Takeuchi, P. A. Nelson, O Kirkeby, H. Hamada, "Robustness of the performance of the 'stereo dipole' to misalignment of head position," 102nd Audio Eng. Soc. Conv., Munich, Germany (1997)

[9] D. Cabrera, D. Gilfillan, "Auditory distance perception of speech in the presence of noise," Int. Conf. on Auditory Display, Kyoto (2002)

[10] M. Barron, Auditorium Acoustics and Architectural Design (Spon, London, 1993)