A preliminary model for the synthesis of source spaciousness

Item Type	Meetings and Proceedings
Authors	Piggott, Darragh;Walker, Jacqueline
Citation	17th International Conference on Digital Audio Effects (DAFx-14);
Download date	2026-06-14 19:39:34
Item License	https://creativecommons.org/licenses/by-nc-sa/1.0/
Link to Item	https://hdl.handle.net/10344/6604

Proc. of the 17th Int. Conference on Digital Audio Effects (DAFx-14), Erlangen, Germany, September 1-5, 2014

A PRELIMINARY MODEL FOR THE SYNTHESIS OF SOURCE SPACIOUSNESS

Darragh Pigott,
CSIS,
University of Limerick,
Limerick, Ireland
`darragh.pigott@ul.ie`

Dr. Jacqueline Walker,
ECE,
University of Limerick,
Limerick, Ireland
`jacqueline.walker@ul.ie`

ABSTRACT

We present here a basic model for the synthesis of source spaciousness over loudspeaker arrays. This model is based on two experiments carried out to quantify the contribution of early reflections and reverberation to the perception of source spaciousness.

1. INTRODUCTION

The subject of spatial audio covers a vast and wide-ranging array of topics from psychology, acoustics, engineering, mathematics, and computer science. The varied contributions from these different fields make for a fascinating and challenging path towards understanding. One challenge that arises is the exact definition of any particular concept. Our primary concern is the synthesis of circumstances under which a certain perceptual attribute of a reproduced sound field arises in the listener. In particular we are interested in *source spaciousness* i.e. the perceived extent of a sound source in three dimensions.

Spaciousness has been the subject of experiments and studies in the past and there is much to learn from the work of [1]-[4]. One of the drawbacks of the term spaciousness is its use as an everyday term as a descriptor for the sense of space. The lack of a clear definition can lead to ambiguity in discussions about perceptual attributes such as source spaciousness. There are places in the literature where spaciousness is discussed but not defined, and others where a definition is offered which do not correspond to definitions found elsewhere. With this in mind we offer here a concise definition of source spaciousness to remove any possible ambiguity for the purposes of the experiments described below.

1.1. Definitions

In the scientific disciplines of acoustics and psychophysics there is a tendency to define spaciousness in terms of its physical correlates [5]. In some cases the term spaciousness is used as a synonym for Auditory Source Width (ASW) [6]. Griesinger opts for a more intuitive definition of spaciousness to mean the impression of a large and enveloping space [7].

Since we are using the definition to relate a concept to a group of potentially inexperienced listeners, we have opted for a more descriptive definition that describes the perceptual attributes of the sound as the three dimensional extent of the perceived source.

Source spaciousness is the perceived extent of a sound source in three dimensions. It can be expressed as a combination of source width, source depth, and source height. Width describes the extent of the perceived source from left to right, depth describes the source extent from front to back, and height is the extent from bottom to top.

This definition accommodates an extension of the sound source such that the boundaries of the source can expand to include the listener "within" the sound. Such a situation may lead to the need for terms such as source envelopment and source engulfment as special cases of listener envelopment (LEV) [2] and engulfment [8].

With the range of definitions used for the term spaciousness we have to tread carefully and state that we are referring to the work of others only in as much as it reflects on the work presented here, that is to say, we are using the definition of source spaciousness provided above even when we refer to the results of others who may themselves be using the term spaciousness to mean something else.

1.2. Past Experiments/Results

In the area of concert hall acoustics, source spaciousness is treated as a contributing component of an all-encompassing perceptual attribute referred to as Spatial Impression [1], [9]-[11]. The three dimensional nature of a sound image is described in [5] as the subjective effect of early reflections. *"As the lateral reflection level is increased, the source appears to broaden and the music gains body and fullness".*

The importance of the frequency content of early reflections to source spaciousness is reported in [5] to the degree that it contributes to the source broadening of the image, with the effect being most prominent around the 1kHz range. Blauert and Lindemann reported on the effect of various frequencies had on both the width and depth of a perceived auditory image [1]. Early reflections made up of primarily low frequencies were attributed to the cause of an increase in depth while the presence of higher frequencies resulted in the lateral expansion of the image.

Since the introduction of elevated speakers into the standard reproduction systems is a relatively new development, experiments covering the perception of height as a perceptual attribute are fewer in number relative to the number of experiments dealing with width and depth.

1.3. Research Question

We know from [1] and [5] that a certain amount of source spaciousness is determined by the presence of early lateral reflections and low frequency reverberation. We also know that the degree to which each dimension of source spaciousness is affected by a lateral reflection is dependent on the frequency content of the reflection(s). We present here a preliminary model to implement these ideas for a system that can synthesize and control the perception of source spaciousness. The purpose of the following experiments is to (a) quantify the contribution of early reflections and low frequency reverb to the perception of source spaciousness and (b) to quantify the contribution of unique frequency bands to the perception of width, depth and height independently.

2. EXPERIMENTAL SETUP

The experiment was carried out in the Spatialization and Auditory Display Environment (SpADE) at the University of Limerick. A description of the acoustic performance of that space can be found in [12]. Many of the features of the experiments are similar to those found in [1].

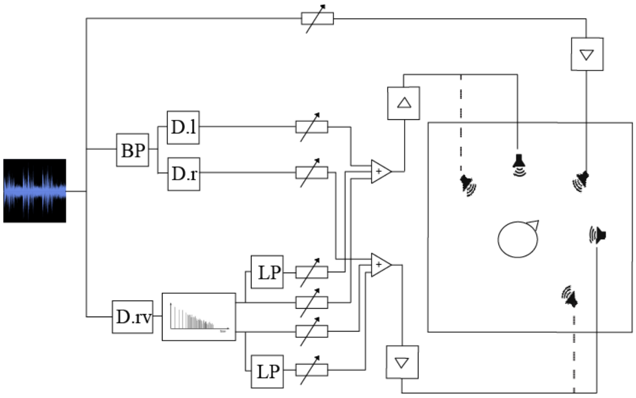

Figure 1: *schematic of the experimental setup*

2.1. Hardware & Software

The speaker setup consisted of 5 Genelec 8030 active near-field monitors positioned 2m from the listening position at angles of 0°, ±45° and ±90°. The direct sound was fed through the centre loudspeaker at 0° along with a reverb signal. The delayed lateral reflections were played back through the speakers to the side with the delayed reverb signal. The parameters of each test signal being examined in each part of the experiment are outlined below.

A reverb signal was created using an EMT 140ST with the reverb time set to 1.75s. This reverb signal was processed with a low pass filter and then mixed with the dry anechoic signal with a delay of 75ms.

Signal processing was applied to the source material in the Max/MSP audio environment. The DSP consisted of gain control, digital delays and 4th-order 24db/oct Chebyshev filters (low-pass and band pass). Each test signal was recorded to disk for use during the experiment to avoid any potential problems with running the signal processing “live”. The average Sound Pressure Level (SPL) at the listening area for each of the sound fields presented was 76dB ±2dB.

2.2. Test Signals

The test signals were generated from an anechoic recording of Glinka’s Overture, Russlan and Ludmilla, from the Denon Anechoic Orchestral Music Recording CD. The left channel was extracted from the stereo recording and used as source material for both experiments. The spectrum of the opening 15 seconds used for the experiment is shown in Fig 1. Each experiment consists of a direct signal played back from the front loud speaker, along with 2 simulated reflections played back over the left and right loud speakers with an applied delay of 20ms and 30ms respectively. In experiment 1 the speakers at angles ±45° were used while in experiment 2 the speakers at ±90° were used along with the frontal speaker.

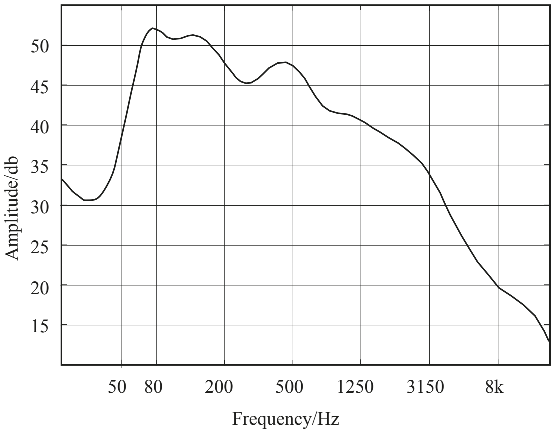

Figure 2: *spectrum of the test signal used*

2.3. Set 1

The parameters for the set of 15 signals in Set 1 are outlined below in Tables 1 & 2. The actual values for each variable were chosen on the basis of their inclusion in [1] where the emphasis was placed on “naturalness” for choosing the parameters outlined below. The sound fields were then arranged into pairs, making 105 pairs for comparison by the participants in the experiment.

Table 1: *variable values for experiment 1*

Cutoff frequency of low pass filtered reverb **fg**	Step +1: 900 Hz Step 0: 650 Hz Step -1: 400 Hz
Level of low pass filtered reverb relative to direct sound **NT**	Step +1: -12 db. Step 0: -14 db. Step -1: -18 db.
Level of early lateral reflections relative to direct sound **S**	Step +1: -3 db. Step 0: -5.6 db. Step -1: -13 db.

Table 2: *Variable values for each test signal in Experiment 1*

Test Signal		**A**	**B**	**C**	**D**	**E**	**F**	**G**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**
Parameter Settings	**S**	1	1	1	1	1	0	0	0	0	0	-1	-1	-1	-1	-1
	fg	1	-1	1	0	0	-1	1	0	-1	-1	0	0	1	-1	-1
	NT	1	0	-1	-1	1	0	1	-1	1	-1	1	0	0	0	-1

While listening to a pair, it was possible to switch between the sound fields freely, and repetition was allowed. In the first part of the experiment, the subjects were asked to compare the sound fields of the pair and make a judgment as to which was more spacious. Judgments of “no difference” were allowed. Their responses were submitted via a touch screen tablet device via OSC and saved in Max/MSP as a text file.

2.4. Set 2

The parameters for the filters applied to the simulated early reflections of part 2 of the experiment are outlined in Table 3. These sound fields were arranged in pairs resulting in 65 pairs for comparison. For each pair, the subject was asked to make a judgment as to which sound field was (a) wider (b) deeper, and (c) taller of the two. Their responses were in the form of a judgment plus a rating between 1-6 depending on the degree to which one was wider/deeper/taller than the other in each pair. Judgments of ‘no difference” were allowed and a rating of 0 was applied to all such responses. During playback it was possible to switch freely between the two sound fields of the pair and repetition was allowed.

Table 3: *variable values for experiment 2*

Test Signal	Bandwidth
1	50 Hz – 80 Hz
2	50 Hz – 200 Hz
3	50 Hz – 500 Hz
4	50 Hz – 1250 Hz
5	50 Hz – 3150 Hz
6	50 Hz – 8000 Hz
7	80 Hz – 20 kHz
8	200 Hz – 20kHz
9	500 Hz – 20 kHz
10	1250 Hz – 20 kHz
11	3150 Hz – 20 kHz
12	8000 Hz – 20 kHz

2.5. Test Subjects

There were 18 participants in total ranging in age between 19 - 28 years old. All were post-graduate students who were studying courses with a strong emphasis on audio and music. Each reported to have normal hearing.

2.6. Pre-Experiment Examples

Prior to the experiment a brief training session was carried out where each subject was presented with several example sound fields with varying degrees of source spaciousness. The definition of source spaciousness was defined as described above and subjects were allowed to make their own judgments of source spaciousness of the example sound fields.

3. RESULTS

3.1. Experiment 1

The participant’s responses to part 1 of the experiment were recorded as source spaciousness scores. For each pair under consideration, a 1 was assigned to the sound field judged to be more spacious and a -1 assigned to the sound field judged less spacious. In cases where the elements of the pair were considered to be equally spacious, a value of zero was assigned to both. Using this scoring scheme we can construct a ranking of spaciousness from the data, see Figure 3. The ranking clearly shows the test signals grouped into 3 clusters, each representing a different value for the variable S: level of early reflections.

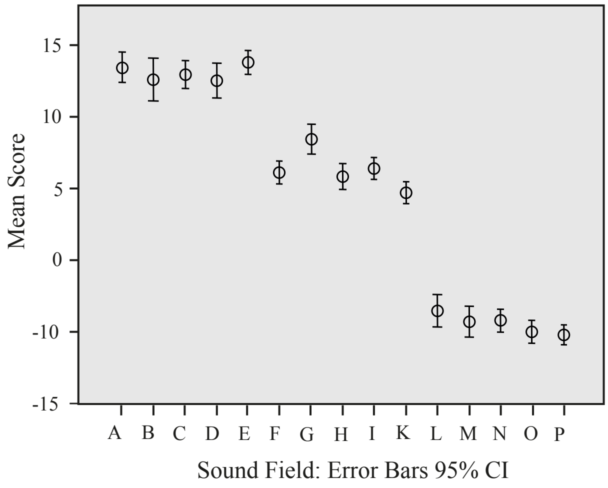

Figure 3: *spaciousness scores for experiment 1*

The existence of a strong relationship between the variable S and the source spaciousness score can be established by visually

Table 4: *Regression model for source spaciousness*

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	12.780	.994		12.851	.000
	Reflections Level	1.668	.031	.944	54.005	.000
	Reverb Filter Cutoff	.004	.001	.095	5.408	.000
	Reverb Level	.202	.053	.067	3.843	.000

a. Dependent Variable: Score

inspecting the chart in Figure 3 and the contents of Table 2. A standard multiple regression was performed between the spaciousness score for each sound field as the dependent variable and the variables S, fg and NT as the independent variables.

An analysis of the effect of the variables fg and NT on source spaciousness revealed little correlation between either variable and the score variable (correlation of .23 & .01 respectively). We also found that the contributions to the end score of the independent variables fg and NT were quite small (Standardized Coefficients of 0.09 and 0.06 respectively). It was proposed that the filter cut-off frequency had influence over the perceived source spaciousness only in as much as it affected the overall energy in the reverberation signal. A new variable was introduced that was the measured peak RMS level of the reverberation signal. Three level groups were identified, and the sound fields were given a new variable with value of -30db, -35db, or -45db according to the measured reverb level R.

This proved to slightly decrease the overall apparent contribution of the reverb signal to the perception of source spaciousness in the analysis. Although the difference is minor it leaves the question open as to whether there is an effect on source spaciousness by varying the frequency of a low cut filter applied to the reverb signal.

The number of cases submitted to analysis was 270, which is a sufficient amount to qualify as suitable for regression analysis [13]. No outliers were found with criteria for Mahalanobis distance set to $p < 0.001$.

Table 4 shows the unstandardized and standardized coefficients for the analysis along with the t value and significance levels. The R, R^2, and adjusted R^2 values for the model are 0.96, 0.92 and 0.92 respectively. This high value for R^2 signifies how dominant the level of early reflections is in determining source spaciousness.

As expected, the primary contributing variable for the spaciousness score is the level of the early reflections. The variation of the reverb signal does have and effect on the result but its significance is negligible in comparison to that of S. When we control for S we found that the effect of the reverb signal on the score was dependent on S. At extreme levels of S, the contribution was minimized, presumably because of the dominance of S. However the effect on the result caused by the reverb became more pronounced when S was in the middle of its range. This effect increased by a factor of 3 compared to its effect at the higher and lower values for S.

3.2. Experiment 2

The focus of the second experiment was on quantifying the contribution of various frequency bands to each of the three dimensions of source spaciousness i.e. source width, source depth and source height. Participants were asked to judge which test signal gave the impression of a wider, deeper and taller source. Compiling the scores in a similar way as we did in experiment 1, the ranking for each dimension is shown in Figures 4,5, & 6.

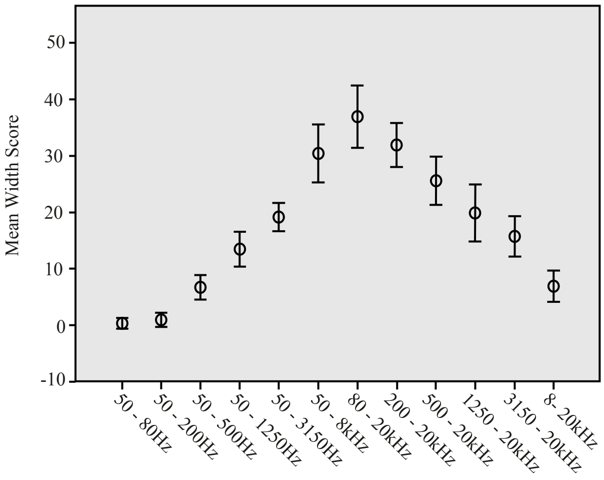

Figure 4: *width scores. Error Bars 95% CI.*

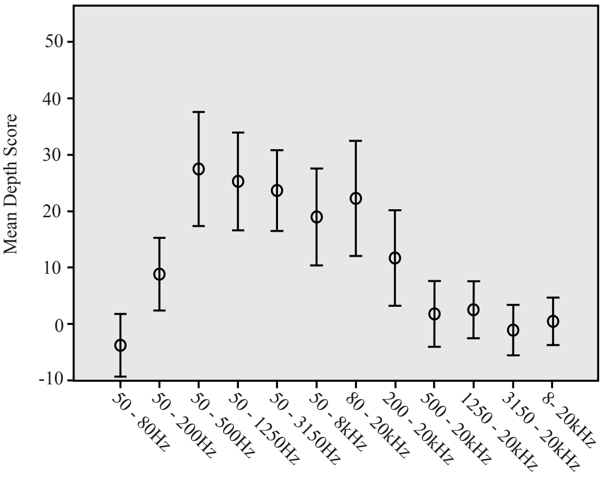

Figure 5: *depth scores: Error Bars 95% CI*

Table 5: *Regression model for width score*

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.583	.186		3.135	.002
	fb_200_500	1.819	.279	.294	6.518	.000
	fb_500_1250	1.500	.322	.242	4.654	.000
	fb_1250_3150	1.278	.322	.206	3.964	.000
	fb_3150_8000	2.056	.322	.332	6.377	.000
	fb_8000_20k	1.597	.279	.258	5.722	.000

a. Dependent Variable: Width Score

As there is little energy in the source material between 50 Hz and 80 Hz, we cannot conclude much about the effect of energy in that region on the source spaciousness. Looking at Figure 4 we can see that all frequency components contribute to the perceived width of the source. Figure 5 indicates that the depth of the perceived source is determined by frequencies below 500 Hz. The presence of energy at frequencies above 500 Hz adds nothing to the perception of depth and may in fact reduce the effect.

The perceived height of the source is determined according to the presence of frequencies above 1250 Hz. The frequencies between 1250 and 8000 Hz contribute the most to the perception of height in the experimental setup. When the higher frequencies (>1250) are present, the addition of any energy in the range below 1250Hz has little effect on the perception of height. However in the absence of energy in the upper range of the frequency spectrum, the lower frequencies may increase the height of the perceived source.

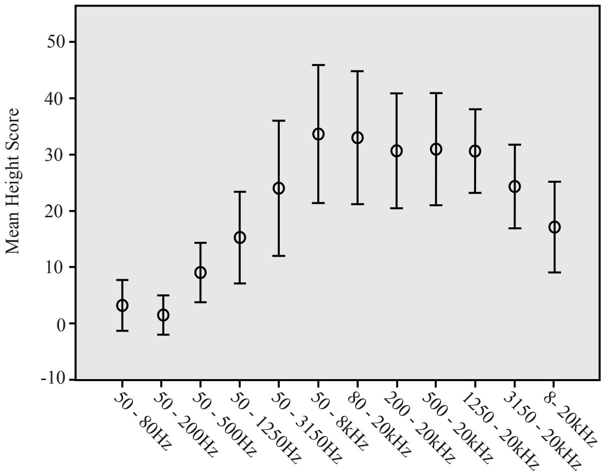

Figure 6: *height scores: Error Bars 95% CI.*

3.3. Models of Width, Depth, and Height.

The variables defining the test signals in experiment 2 were coded into non-overlapping frequency bands. If a frequency band is present as a reflection in a signal it is assigned a value of 1, otherwise it is 0. To determine the contributions from each frequency band to the perception of width depth and height we employed a standard multiple regression with the scores as the dependent variable and the frequency band variables of the early reflections as the independent variables.

After some exploratory analysis we found that the maximum number of independent variables contributing to the perceptual attribute source width is 5. With 216 cases submitted to the regression, the criterion for ratio of cases to independent variables is satisfied. No outliers were found.

The amount of variation in the width score is accounted for by the five frequency bands is shown in Table 5. The frequency range below 200Hz did not make any significant contribution to the width score. The model in Table 5 accounts for 81% of the variance in width score.

The variation in depth score is accounted for by the two frequency bands that make up the range between 80 Hz to 500 Hz. The coefficients for the depth regression are shown in table 6. The contribution of these frequency bands accounts for 39% of the total variation in depth score.

The results of the regression analysis with height as the dependent variable are summarized in Table 7.

We have found that the frequency content of the early reflections accompanying a direct signal have a significant influence on the perception of source spaciousness in terms of the width, depth and height of the perceived source. This confirms the results found in [1] and [5] although there is some disagreement over the exact frequency band which can be said to influence each of the dimensions.

4. SOURCE SPACIOUSNESS MODEL

Based on the results of the experiments presented above, we have devised an equation to represent a linear model of source spaciousness.

$$SS = \sum_{i=1}^{n} (\alpha_{w_i} + \alpha_{d_i} + \alpha_{h_i}) G_i + I_r \alpha_r G_r \qquad (1)$$

where G_i and G_r are the gain of the i^{th} frequency band of the simulated early reflections and the reverb signal respectively. α_{w_i}, α_{d_i}, and α_{h_i} are the regression coefficients from the linear approximations for perceived width, depth, and, height respectively. I_r is the scaling factor applied to the effect of the reverb due to the value of S. α_r is the reverb coefficient from the regression applied to the result of experiment 1.

Table 6: *Regression model for depth score*

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.352	.237		9.934	.000
	fb_80_200	1.889	.428	.317	4.410	.000
	fb_200_500	2.167	.428	.364	5.059	.000

a. Dependent Variable: Depth Score

Table 7: *Regression model for height score*

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.863	.244		7.635	.000
	fb_1250_3150	1.896	.423	.343	4.485	.000
	fb_3150_8000	1.250	.535	.226	2.338	.020
	fb_8000_20k	.590	.423	.107	1.396	.164

a. Dependent Variable: Height Score

The first term of (1) represents the contribution of early reflections while the second term accounts for the reverberation signal. Although we found there to be minimal effect of the reverb signal on the perception of source spaciousness, we kept this term in the equation to allow for potential future developments involving a reverb signal.

The overall content of the model is based on experiment 1 while the details of the filters applied to the early reflections to control for perceived width, height and depth independently is derived from the results of experiment 2. According to our findings, source spaciousness is a three dimensional spatial attribute that can be described in terms of width, depth, and height.

5. CONCLUSION AND FUTURE WORK

We have presented here a preliminary model for source spaciousness that is to serve as a starting point for the development of a more comprehensive study of this perceptual attribute. While changes in the width are well accounted for by the variables included in the experiments, the other two dimensions are less affected. Future experiments could potentially seek to get a more detailed picture at how the frequency spectrum of early reflections affects the perceptual attribute. The inclusion of elevated loudspeakers for the simulation of source height will also be investigated.

6. REFERENCES

[1] J. Blauert and W. Lindemann, “Auditory spaciousness: Some further psychoacoustic analyses”, *The Journal of the Acoustical Society of America*, vol. 80, p. 533, 1986.

[2] M. Morimoto, K. Iida, and K. Sakagami, “The role of reflections from behind the listener in spatial impression”, *Applied Acoustics*, vol. 62, no. 2, pp. 109–124, 2001.

[3] D. Griesinger, “General overview of spatial impression, envelopment, localization, and externalization”, *AES 15th International Conference: Audio, Acoustics & Small Spaces, 1998.*

[4] J. Berg and F. Rumsey, “Systematic evaluation of perceived spatial quality”, *AES 24th International Conference on Multi-channel Audio, 2003.*

[5] M. Barron and A. H. Marshall, “Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure”, *Journal of Sound and Vibration*, vol. 77, no. 2, pp. 211–232, 1981.

[6] D. L. Valente, S. A. Myrbeck, and J. Braasch, “Matching Perceived Auditory Width to the Visual Image of a Performing Ensemble in Contrasting Multi-Modal Environments”, *AES 127th Convention, New York, 2009.*

[7] D. Griesinger, “Spaciousness and envelopment in musical acoustics”, *Audio Engineering Society, Preprints,* 1996.

[8] R. Sazdov, G. Paine, and K. Stevens, “Perceptual Investigation into envelopment, spatial clarity, and engulfment in reproduced multi-channel audio”, *AES 31st International Conference, London, UK, 2007.*

[9] T. Hidaka, L. L. Beranek, and T. Okano, “Interaural cross-correlation, lateral fraction, and low and high-frequency sound levels as measures of acoustical quality in concert halls,” *The Journal of the Acoustical Society of America*, vol. 98, p. 988, 1995.

[10] M. Barron, “Late lateral energy fractions and the envelopment question in concert halls”, *Applied Acoustics*, vol. 62, no. 2, pp. 185–202, 2001.

[11] M. Barron, “The subjective effects of first reflections in concert halls--The need for lateral reflections”, *Journal of Sound and Vibration*, vol. 15, no. 4, pp. 475–494, 1971.

[12] M. Ronan, D. Pigott, and R. Sazdov, “Configuring SpADE: Practical Considerations Influencing the Design of a 3D Multi-Channel Listening Environment”, *2nd Convocation of ISSTC, Cork, 2012*.

[13] B. G. Tabachnick and L. S. Fidell, "Using Multivariate Statistics.", Allyn & Bacon, 2001.