The Effect of Inter-channel Cross-correlation Coefficient on Perceived Diffuseness

M. P. Cousins[1], S. Bleeck[1], F. M. Fazi[1]

[1] *University of Southampton, UK, Email: mpc1r13@soton.ac.uk*

Abstract

Reproducing diffuse sound fields such as reverberation, audience noise or rain using loudspeakers is an important part of creating surrounding and enveloping spatial audio. This paper presents the results of two listening tests that investigate how the inter-channel correlation between loudspeakers affects the perceived diffuseness. In the first experiment inter-channel cross-correlation coefficient (ICC) was varied for pink noise stimuli reproduced using 4 different loudspeaker layouts and at two listening positions. In a second experiment, 1/3 octave noise stimuli were investigated in a similar manner using variable ICC and two listening positions. The ICC was found to have a large effect on the perceived diffuseness in both experiments and for all frequency bands (125 Hz, 1 kHz and 8 kHz) where correlation between loudspeakers would reduce the perceived diffuseness. Although frequency was found to be a significant factor, the ICC and listener position had larger effects with although moving off-centre only reduced the perceived diffuseness when fewer loudspeakers were used. Off-centre appears more robust to small amounts of correlation between loudspeakers as the loudspeakers signals no longer arrived at the listener at the same time.

Introduction

This paper presents the results of two listening tests investigating how the correlation between loudspeakers affects the perception of diffuseness. The theoretical diffuse sound field is one created by an infinite number of uncorrelated plane waves coming from all directions. Sound fields such as the late part of reverberation, audience noise, or rain are often considered diffuse as the sound field approximates this theoretical diffuse sound field. These sound fields play a vital role in creating enveloping[1], engulfing[2] and surrounding auditory experiences. However, to recreate the exact sound field using loudspeakers is impractical or impossible. Individual loudspeakers would be required for each direction of rain drop or acoustic reflection and the data requirements are huge with each loudspeaker requiring a separate uncorrelated signal. However in reproduced audio, it is only important that the signals are perceived as diffuse rather than an accurate recreation of the entire sound field. The experiments in this paper show how the perception of the diffuseness of a sound field will change as the correlation between the loudspeaker signals is increased.

The first experiment uses full bandwidth pink noise with variable cross-correlation coefficient between loudspeakers. Four loudspeaker layouts and two listening positions were used to investigate how this inter-channel cross-correlation coefficient (ICC) varies with the number of loudspeakers but also across the listening area. From this experiment the ICC was found to be highly significant and important to high diffuseness perception. However the summation of signals that are correlated between loudspeakers will depend on both the frequency of the signals and the listening position. For this reason, the second experiment used three narrow 1/3 octave band noise signals with centre frequencies of 125 Hz, 1 kHz and 8 kHz to investigate any dependence on frequency.

Experiment 1: Inter-channel Cross-correlation Coefficient

Other authors [3] and papers [4] have shown the number of loudspeakers to be a highly significant factor for creating diffuse and enveloping sound fields. However the signals with which the loudspeakers are driven will also determine the overall perception of diffuseness or envelopment. Experiment 1 looks at how the correlation between loudspeakers affects the perceived diffuseness using wide bandwidth pink noise.

Stimuli

Stimuli with variable inter-channel correlation were created by mixing a set of signals that are uncorrelated between all loudspeakers with a common signal that is the same for all loudspeakers. Therefore the signal for the n'th loudspeaker is given by,

$$y_n(t) = \frac{1}{\sqrt{N}}\left(\sqrt{(1-a)}x_n(t) + \sqrt{a}c(t)\right), \qquad (1)$$

where N is the number of loudspeakers, $x_n(t)$ is a pink noise signal that is uncorrelated between loudspeakers and $c(t)$ is a pink noise signal that is common for all loudspeakers. The variable a is defined as the Inter-channel Cross-correlation Coefficient (ICC) and is equal to the cross-correlation coefficient between any two loudspeakers. It holds that,

$$R_{ij} = \langle y_i(t)y_j(t)\rangle = \begin{bmatrix} 1 & a & a & & a \\ a & 1 & a & \ddots & a \\ a & a & 1 & \ddots & a \\ & \ddots & \ddots & \ddots & a \\ a & a & a & a & 1 \end{bmatrix}, \qquad (2)$$

where R_{ij} is the cross correlation coefficient matrix, y_i

		Floor		Lower					Head-height						Upper						Ceiling		
	Elevation°	-56	-56	-20	-17	-24	-17	-20	0	0	0	0	0	0	27	26	24	32	24	27	52	52	90
	Azimuth°	±45	±135	0	±45	±90	±135	180	0	±30	±60	±90	±120	180	0	±30	±45	±90	±135	180	±45	±135	0
Loudspeaker Layout	Stereo*									✓													
	5.0								✓	✓			✓										
	9.0								✓	✓			✓			✓			✓				
	0/12/0								✓	✓	✓	✓	✓	✓									
	12/12/13	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓

Table 1: Positions of loudspeakers for the four test layouts. *Uncorrelated stereo was only used as a hidden anchor.

and y_j are the signals for the i-th and j-th loudspeakers respectively. $\langle . \rangle$ denotes averaging and a is the ICC (the cross-correlation coefficient between any two different loudspeakers). Therefore, this method of varying the correlation between loudspeakers allows a single value—the ICC—to represent the similarity between all the various combinations of loudspeaker pairs.

The four chosen loudspeaker layouts were standardised 5.0 and 9.0(5+4) layouts [5, 6] as well as two layouts of loudspeakers evenly distributed in azimuth around the listener and labelled as 0/12/0 and 12/12/13. The labels of the latter two layouts relate to the number of loudspeakers below/at/above head-height. These four layouts were rated at 50, 60, 68 and 78 respectively on an arbitrary scale of perceived diffuseness in a previous experiment [4] ranging from moderate perceived diffuseness to high perceived diffuseness when using uncorrelated pink noise. The layouts also feature two two-dimensional layouts and two three-dimensional layouts. There are also two layouts with more loudspeakers in front of the listener and two layouts evenly distributed in azimuth around the listener. The positions of the loudspeakers in each layout are shown in table 1.

The six values chosen for the ICC were 0 (all uncorrelated), 0.2, 0.4, 0.6, 0.8 and 1 (all loudspeakers have identical signals).

Finally, for the two listening positions, the on-centre listening position was in the middle of the room and the point of time and frequency alignment. The off-centre listening position was 80 cm to the right of centre. this equated to half the distance to the loudspeaker at ±90° and is similar to a large sofa in a domestic listening scenario. However, to reduce the length of the test, only the most and least diffuse layouts (5.0 and 12/12/13) were included for the off-centre position.

These variables led to 36 test stimuli (24 on-centre and 12 off-centre).

The Listening Room

The listening tests were both performed in the Audio Lab at the University of Southampton with R_{60} of 0.12 s ±0.02 s in 1/3 octave bands between 125 Hz and 8 kHz. The 39 Kef HS3001SE loudspeakers mounted to the walls, floor and ceiling were equalised in 1/3 octave bands from 95 Hz to 20 kHz. Below 95 Hz a -24 dB per octave roll off followed the approximate frequency response of the loudspeakers. Digital delay compensated for the differing propagation delays between the loudspeakers and the central listening position to the nearest sample (the sampling frequency was 48 kHz throughout).

Loudness Alignment

The correlation between loudspeakers will also affect the loudness of the stimuli. Uncorrelated signals add incoherently (by power) whereas correlated signals add coherently (dependent on amplitude and phase). To avoid bias from loudness differences, the loudness was aligned subjectively using a preliminary listening test.

The alignment gains were found using an adjustment task in which listeners were asked to match the loudness of a test stimulus to the loudness of a reference stimulus. The stimulus 12/12/13 with ICC=0 was used as the reference due to consistency of sound pressure level across the whole listening area. Two seconds of the reference were followed by two seconds of the test stimulus. Two buttons allowed either +1 dB or -1 dB of gain to be applied to the test stimulus before the reference and test stimuli were relayed. This process was repeated until the listener was happy the loudness of the two stimuli was identical. The listener could then click the next button to load a different test stimulus until all stimuli had been adjusted relative to the reference.

Each stimulus was adjusted twice by each of the 3 listeners. The mean values are given in table 2. The loudness alignment gains from the on-centre listening position were also used off-centre.

		ICC					
		0	0.2	0.4	0.6	0.8	1
Loudspeaker Layout	Stereo*	0.667					
	5.0	-0.333	-0.167	1	0.333	0.5	0.5
	9.0	-0.167	0.167	-0.167	0	-0.333	-0.833
	0/12/0	-0.5	-0.667	-0.333	-1	-1	-1
	12/12/13	0.167	-0.333	-0.667	-1.5	-1.5	-2.167

Table 2: The loudness alignment gains for each combination of loudspeaker layout and ICC. * Stereo only used as hidden anchor.

In addition to the adjusted loudnesses, ±2 dB of uniformly random gain variation randomised any residual loudness difference bias.

Stimuli presentation

A MUSHRA based methodology [7] was used to elicit the perceived diffuseness for all stimuli. Every page included the stimulus 12/12/13 with ICC=0 as a reference and uncorrelated stereo as a hidden anchor.

Listeners rated the stimuli on a scale of 0 to 100 in terms of perceived diffuseness. The description of diffuseness given to listeners for both experiments was based on the description of envelopment from [8] with additional advice considering the likely phasiness of the signals. This was:

> Diffuseness is the degree of being

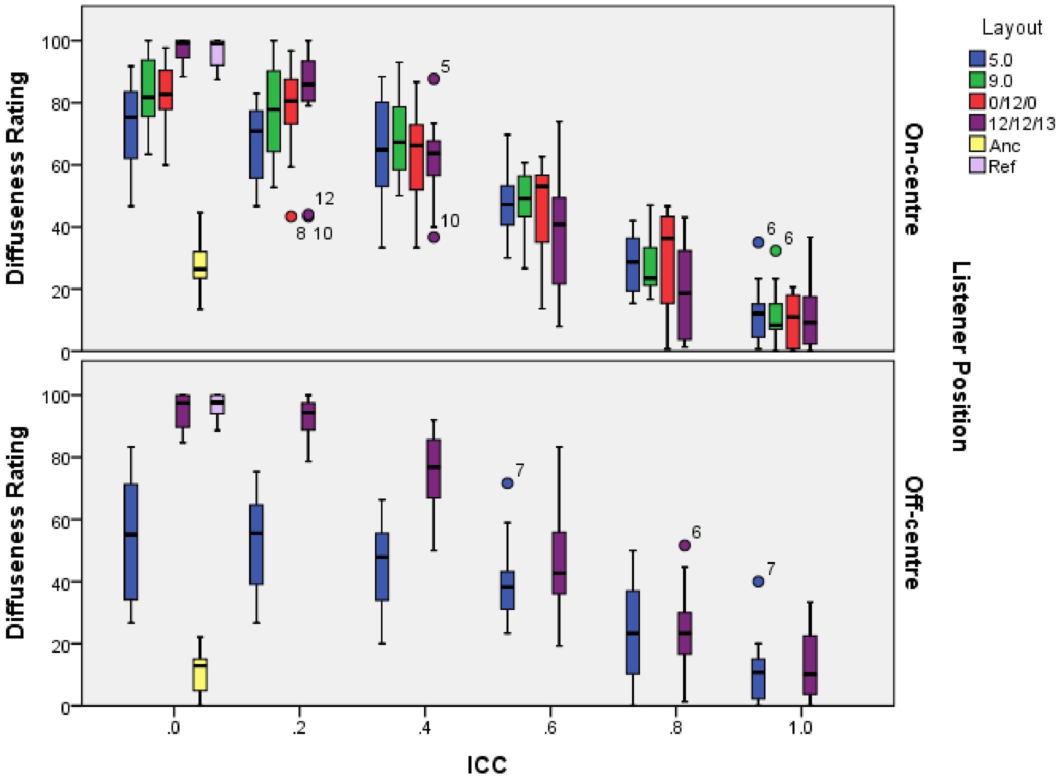

Figure 1: Box plots of mean perceived diffuseness ratings from each listener. Outliers are labelled with Listener IDs.

> surrounded/enveloped by the sound field. This may be heard when standing and listening to the rain hitting the pavement; applause in a concert hall; atmosphere or air conditioning (room tone). Being able to localise the source of the sound will decrease diffuseness. Holes (an absence of sound from a certain directions) would normally reduce diffuseness. Feeling the sound inside your head or as moving/unstable sources would also usually be less diffuse.

Each stimulus was rated 3 times by each of the 12 postgraduate/undergraduate listeners.

Results

All listeners were consistent between repeats and congruent with each other and so no listeners were excluded from analysis. The three repeats for each stimulus were averaged to give a single value for each stimulus for each listener and the results are plotted in figure 1. The mean ratings for each stimulus are plotted in figure 2.

ANalysis Of VAriance (ANOVA), shown in table 3, shows loudspeaker layout and ICC to be statistically significant. Listener position is not significant on its own although the interaction between layout and listener position is significant as is the interaction between the ICC and layout. ICC has the largest effect and layout the second largest effect on the rated diffuseness.

Discussion

In this section the features of the results shown in figure 2 are separated and discussed.

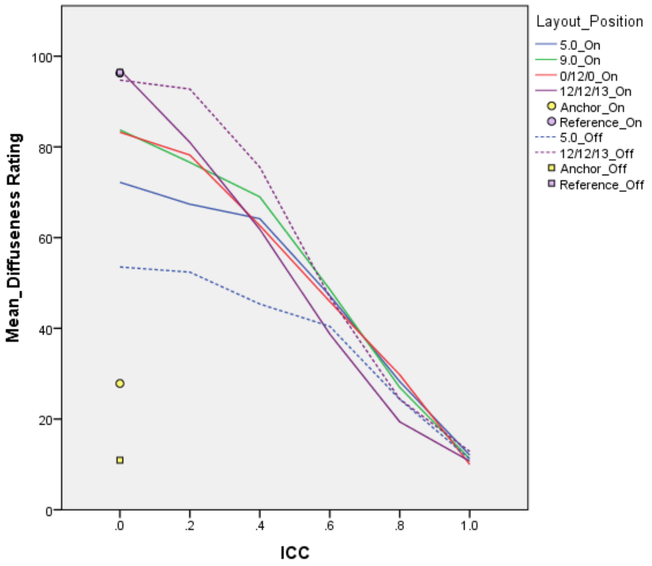

Figure 2: Mean diffuseness ratings for all stimuli.

Effect of ICC and Layout

The perceived diffuseness rating is strongly dependent on the ICC. At low ICCs, the perceived diffuseness is high and is dependent on the layout of the loudspeakers with more loudspeakers perceived as more diffuse. When the ICC is 0.4 or above the differences between layouts is very low with the perceived diffuseness independent of the number of loudspeakers. All arrangements at ICC of 1 give a perceived diffuseness less than uncorrelated stereo. Interestingly, on-centre and at medium ICCs, 12/12/13 is less diffuse than the other loudspeaker layouts despite having the most loudspeakers and being most diffuse at

Source	Type III Sum of Squares	df	Mean Square	F	Significance
Corrected	387162.207	39	9927.236	59.315	.000
Intercept	698385.725	1	698385.725	4172.820	.000
Loudspeaker Layout	93002.997	5	18600.599	**111.138**	**.000**
Listener Position	308.054	1	308.054	1.841	.176
ICC	242253.728	5	48450.746	**289.491**	**.000**
Loudspeaker Layout * Listener Position	6149.968	3	2049.989	**12.249**	**.000**
Loudspeaker Layout * ICC	15636.186	15	1042.412	**6.228**	**.000**
Listener Position * ICC	1111.106	5	222.221	1.328	.251
Loudspeaker Layout * Listener Position * ICC	1804.614	5	360.923	2.156	.058
Error	73640.780	440	167.365		
Total	1699112.285	480			

Table 3: ANOVA for all stimuli from Experiment 1. R Squared = .840 (Adjusted R Squared = .826)

low ICCs although this is not statistically significant.

Moving Off-centre, ICC=0

With fewer loudspeakers (in the case of stereo and 5.0), as the listener moves off-centre, the perceived diffuseness decreases as the listener becomes more aware of the loudspeaker(s) they are approaching. In stereo these loudspeakers are obviously turned off and in 5.0 the rear right loudspeaker becomes easy to localise. For the layout 12/12/13, which is more homogeneous, the perceived diffuseness remains more similar as the listener moves off-centre. The explicit reference that is available in both listening positions normalises the top of the scale for both listening positions. However, the same trend has been reported in [4] without explicit reference.

Off-centre, ICC=0.2

When off-centre there is very little difference between ICC=0 and ICC=0.2. It is suggested that on-centre, where the signals are time aligned, it is possible to detect the small addition of the correlated component. However when off-centre, the signals are not time aligned and the differences between ICC=0 and ICC=0.2 are harder to detect. This can be seen as a robustness to small amounts of correlation when moving off-centre. However, the robustness off-centre to ICC changes does not apply at high ICCs where the dependence of the perceived diffuseness on both layout, listener position is once again negligible with all layouts and listener positions rated equally not diffuse.

Summary

Some of the results of this experiment are exactly as expected. When the loudspeaker signals are uncorrelated it is the quantity and arrangement of loudspeakers that determines the perceived diffuseness and increasing the ICC reduces the perceived diffuseness. What is interesting is the degree to which the ICC plays a role. With a high ICC the effect of both the layout of the loudspeakers and the listener position is negligible. Highly correlated signals are perceived less diffuse than uncorrelated stereo regardless of the number of loudspeakers.

The effect of ICC is also interesting off-centre where the perceived diffuseness decreases for layouts with few loudspeakers at low ICC but increases for layouts with many loudspeakers and low ICC. At high ICC there appears no effect of moving off-centre as once again all layouts are perceived as equally not diffuse. These results show that in order to achieve high perceptual diffuseness, many uncorrelated loudspeakers are required. Although, when the listener is off-centre, the signals do not all arrive at the same time and we believe that this means that an ICC value of 0.2 is less degrading off-centre than it would be on-centre.

Experiment 2: Frequency Dependency

The first experiment looked at full bandwidth signals however the interaction between the coherent components of the stimuli (in any stimulus without an ICC of 0) is complex. The signals form an interference pattern that is a function of both frequency and position. In addition, the limited spacing of the human ears leads to a high correlation between the ear signals at low frequencies regardless of the ICC. For these reasons, in the second experiment, frequency was added as a variable. It was predicted that for low frequencies the ICC would have less effect on the perceived diffuseness where the long wavelength relative to the interaural distance would make differentiating conditions more difficult. Also it was predicted that moving off-centre would have the same effect as decorrelating the signals and therefore reduce the effect of increasing the ICC. This would be more noticeable at high frequencies where the highly detailed time structure is effectively decorrelated with shorter time shifts than at low frequencies. It was these hypotheses for high and low frequencies that the second experiment was designed to test.

Stimuli

Narrow 1/3 octave bandwidth noise signals were generated for three centre frequencies (125 Hz, 1 kHz and 8 kHz). The low frequency band was as low as possible considering the frequency response of loudspeakers. The high frequency band was chosen to allow a high modal density when off-centre. The bandwidth was wide enough to be perceived as noise like for all centre frequencies without exciting much more than a single critical bandwidth.

The off-centre listening position was changed from the first part of the experiment to be 1 m to the right of centre. This was done to hopefully increase any effects of moving off-centre by increasing the time between arriving wavefronts.

The layout 12/12/13 was chosen as the loudspeaker arrangement for the second experiment as this had a smaller difference in perceived diffuseness when moving off-centre. The reproduction system was identical to that in the first experiment.

The ICC values also the same as for the previous experiment (0, 0.2, 0.4, 0.6, 0.8 and 1) and were generated in the same way, using a summation between a component uncorrelated between loudspeakers and a component common to all loudspeakers (equation 1).

As with experiment 1, subjective loudness alignment was used to avoid the bias of loudness. In this case there were 3 listeners who sat 2 repeats. The alignment gains for each of the stimuli are given in table 4. Also in this case additional random gain of ±2 dB was added.

Listener Position	Frequency	ICC					
		0	0.2	0.4	0.6	0.8	1
On-centre	125 Hz	2.125	1	0.875	0.625	0.375	-0.125
	1 kHz	0	0	0.25	0.5	0.625	0.875
	8 kHz	0.375	0	0.375	0	0.125	0.125
Off-centre	125 Hz	2.625	2	2.125	2.125	2.25	2.5
	1 kHz	0	0	0.125	0.5	0.25	0.375
	8 kHz	0.25	0	0.375	0.375	0.375	1.125

Table 4: Alignment gains in dB for Experiment 2.

Listener Response Methodology

The first experiment showed a large effect of the ICC on the perceived diffuseness. The second experiment was concerned with whether the dependence on ICC is also consistent with frequency. It was decided that comparing different frequency bands with different ICCs would be too complex a task and so the test was reduced to compare a stimulus with another stimulus of the same frequency. However if a MUSHRA type user interface is used and comparisons are only made between stimuli of the same frequency then listeners can view the whole range of perceived diffusenesses for that particular frequency on a single page and there is a risk that range equalisation bias [9] might normalise the differences between frequencies and hide any differences in perceived diffuseness. The alternative chosen was to use an A/B type experiment [10] with only two stimuli to compare at once and both at the same frequency. A slider was used to elicit the difference in perceived diffuseness between a reference stimulus (with ICC=0) and a test stimulus (with variable ICC but the same centre-frequency). The disadvantage of only comparing stimuli with stimuli of the same frequency is that there is then no data to compare the relative diffusenesses of the different frequency bands. Therefore additional comparisons were added that compare the references for each frequency band which were all uncorrelated. Because the references are completely uncorrelated the task is simplified relative to comparing both frequency and ICC differences at the same time. All the comparisons are shown in the table 5.

To perform these comparisons, a user interface was used that featured two buttons labelled A and B which would play 2 s bursts of two stimuli (with 10 ms fade in and out to avoid clicking). One stimulus would be ICC=0 as the reference and the other would be the test stimulus with either an ICC between 0 and 1 but with the same centre-frequency or with an ICC of 0 but a different centre-frequency. The listener would then decide which of the two was most diffuse and then move the horizontal slider in that direction depending on how much more diffuse they perceived that stimulus. The slider was labelled "A is much more diffuse than B" (slider fully left), "A and B are equally diffuse"(slider in the centre) or "B is much more diffuse than A"(slider fully right).

			Reference Stimulus		
	Frequency		125 Hz	1 kHz	8 kHz
		ICC	0	0	0
Test Stimulus	125 Hz	0	✓		
		0.2	✓		
		0.4	✓		
		0.6	✓		
		0.8	✓		
		1	✓		
	1 kHz	0	✓A	✓	
		0.2		✓	
		0.4		✓	
		0.6		✓	
		0.8		✓	
		1		✓	
	8 kHz	0	✓B	✓C	✓
		0.2			✓
		0.4			✓
		0.6			✓
		0.8			✓
		1			✓

Table 5: Comparisons used in experiment 2. All stimuli use the 12/12/13 loudspeaker arrangement. ✓ABC are the comparisons between different frequencies (inter-frequency).

Each comparison was completed three times by 9 post-graduate listeners at the University of Southampton.

Results

In this experiment there is neither a common reference for all the comparisons nor comparisons between every possible stimulus pair. Each frequency and listening position is rated relative to its own reference and then, additionally, there are comparisons between the references in both of the listening positions. The stimuli are therefore shown in two sections. The first section has each listening position and frequency rated relative to its own reference (intra-frequency comparisons) and separately the comparisons between the frequency references (inter-frequency comparisons).

Firstly the data was screened using the standard deviation between repeats of the same stimulus as a metric of listener consistency. Listener 8 was found to be very inconsistent between repeats. They also often moved the slider when the test stimulus was identical to the reference stimulus. For these reasons, listener 8 was removed from the remaining analysis. The three repeats

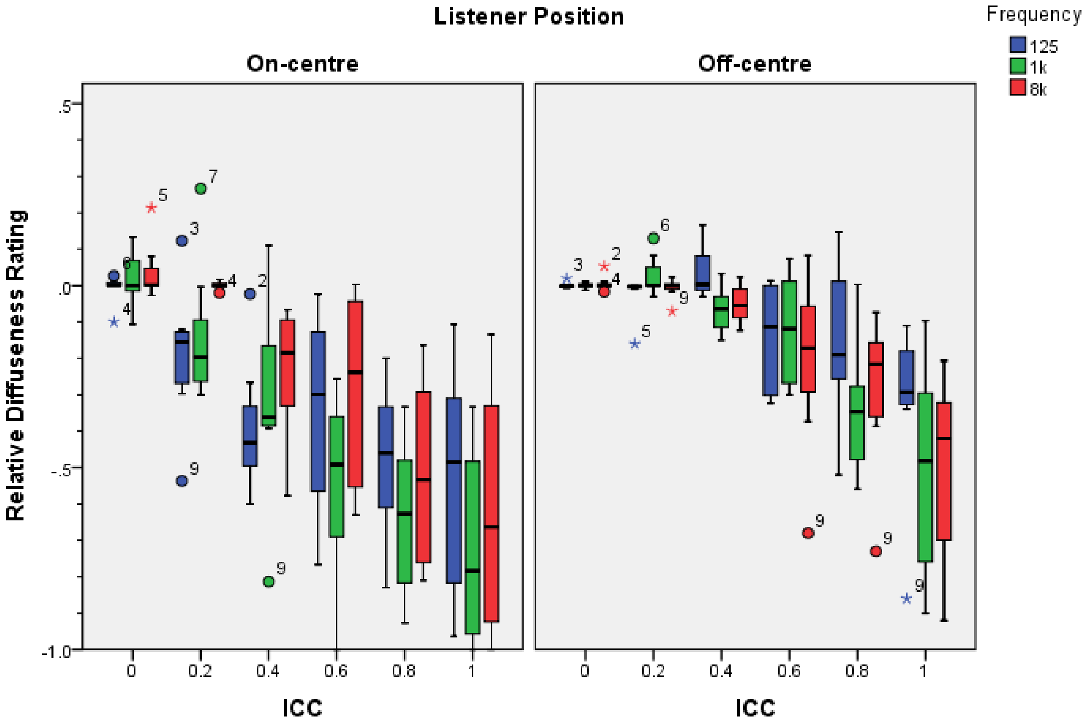

Figure 3: Box plots of mean ratings from each listener for intra-frequency stimuli. A rating of zero equates to the test ICC being equally diffuse as the ICC=0 for the same frequency. Outliers are labelled with listener IDs.

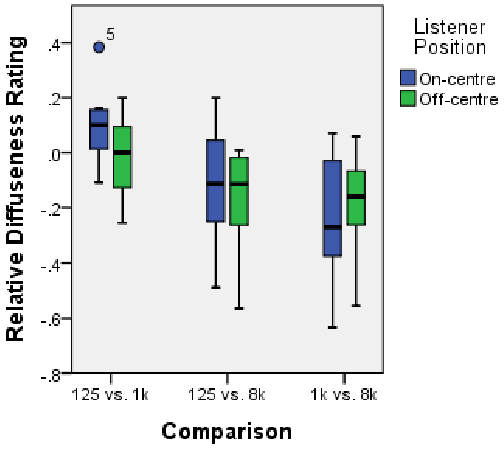

Figure 4: Box plots of mean ratings from each listener. In each case value is the diffuseness of the higher frequency relative to the lower frequency. i.e. positive values mean the higher frequency is more diffuse and negative values mean the higher frequency is less diffuse.

from the remaining 8 listeners were then averaged to give a single value as the rating for that particular comparison for each listener. These results are shown in the two box plots in figures 3 and 4. A value of zero represents that both stimuli were equally diffuse, a negative value indicates the reference was more diffuse and a positive value indicates the test stimulus was more diffuse. In the case of inter-frequency stimuli, the reference is arbitrarily the lower frequency.

Listener 9 has a few outliers and tended to use more of the bottom of the scale, but in general the listeners were fairly congruent between each other and so only listener 8 was excluded from further analysis.

ANOVA for the intra-frequency stimuli (table 6) shows that the listener position and the ICC have the largest significant effects. Frequency is significant although it has a smaller effect. The interaction between the position and frequency is significant as is the interaction between position and ICC.

ANOVA for the inter-frequency stimuli (table 7) reveals that, whilst the frequencies that are being compared is a significant factor, the fit to the data is poor indicating the variation due to other factors (such as differences between listeners and noise) are large in comparison to the differences in diffuseness between frequencies.

To view the trends in the intra-frequency comparisons, the mean ratings are taken and shown in figure 5 and discussed in the next section.

The data from the inter-frequency comparisons is less intuitive to understand. Therefore, to visualise the data from the inter-frequency comparisons, the mean ratings were combined to give a perceived diffuseness for each frequency arbitrarily normalised to the diffuseness at 1 kHz. The inter-frequency comparisons are labelled as A (125 vs. 1k), B (125 vs. 8k) and C (1k vs. 8k) in table 5. Therefore the perceived diffuseness of the different frequencies — with reference to the diffuseness of 1 kHz — can be calculated for 125 Hz by either $-A$ or by $C - B$, and the for 8 kHz by either C or $-A + B$. The two alternatives were averaged so that $Diffuseness_{125} = (-A + C - B)/2$ and $Diffuseness_{8k} = (C - A + B)/2$ where A, B and C are the mean diffuseness ratings for the inter-frequency comparisons. These relative frequency diffusenesses, that have been calculated from the inter-frequency comparisons, are plotted in figure 6

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	14.714a	35	.420	11.665	.000
Intercept	16.378	1	16.378	454.466	.000
Listener Position	2.353	1	2.353	**65.298**	**.000**
Frequency	.255	2	.128	**3.540**	**.030**
ICC	10.334	5	2.067	**57.349**	**.000**
Listener Position * Frequency	.254	2	.127	**3.524**	**.031**
Listener Position * ICC	.810	5	.162	**4.496**	**.001**
Frequency * ICC	.473	10	.047	1.311	.224
Listener Position * Frequency * ICC	.235	10	.023	.651	.769
Error	9.082	252	.036		
Total	40.174	288			
Corrected Total	23.795	287			

Table 6: ANOVA for all intra-frequency comparisons. R Squared = .618 (Adjusted R Squared = .565)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.632a	5	.126	3.358	.012
Intercept	.509	1	.509	13.523	.001
Comparison_Pair	.555	2	.277	**7.367**	**.002**
Listener Position	.017	1	.017	.461	.501
Comparison_Pair * Listener Position	.060	2	.030	.797	.457
Error	1.582	42	.038		
Total	2.723	48			
Corrected Total	2.214	47			

Table 7: ANOVA for inter-frequency comparisons Where comparison pair relates to the differences in frequency. R Squared = .286 (Adjusted R Squared = .201)

and discussed in the following section.

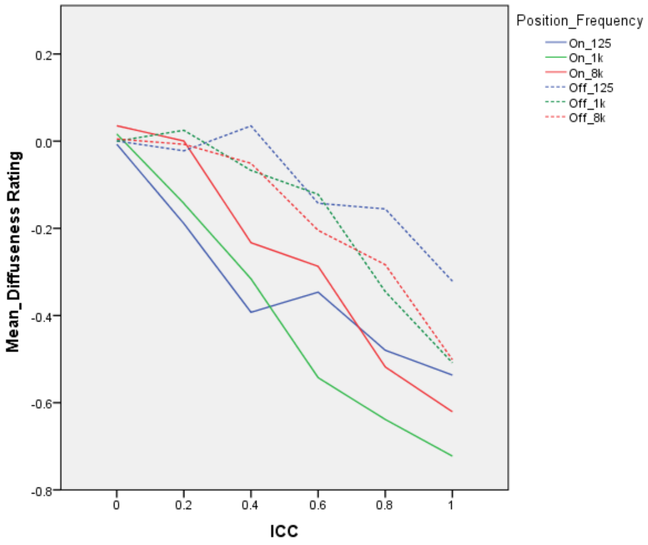

Figure 5: Mean diffuseness ratings for all intra-frequency stimuli comparisons.

Discussion

Intra-frequency Comparisons

As in experiment 1, the ICC has a large effect with high ICC less diffuse than the low ICC reference.

As in experiment 1, when off-centre, low ICCs have less effect on the perceived diffuseness than when on-centre. On-centre, 8 kHz appears to demonstrate a similar trend with ICC=0.2 seemingly indistinguishable from ICC=0. It is possible that the short wavelength of 8 kHz makes even the on-centre position exhibit the same trends seen off-centre at lower frequencies.

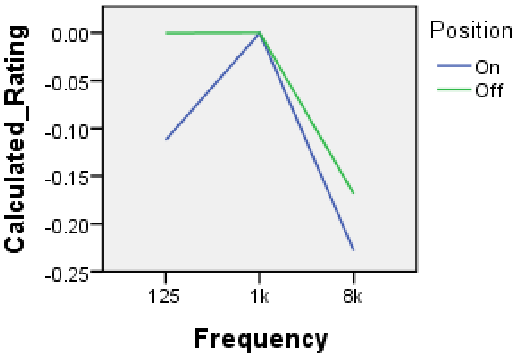

Figure 6: Diffuseness of each frequency band calculated with reference to the perceived diffuseness at 1 kHz.

The stimuli at 125 Hz also show a strong significant dependence on the ICC despite the high correlation between ear signals at this low frequency. This was unexpected and deserves further investigation.

In experiment 2 the listener position is also highly significant and the lines no longer converge at high ICC. It is possible that the MUSHRA type listening test of experiment 1 exhibited some range equalisation bias between the two different listening positions [9] encouraging the use of the bottom of the scale when off-centre and thereby hiding any differences.

The variation on-centre is greater than off-centre with larger interquartile range and standard deviation. This could be due to the fact that small variations in listener position have a larger effect on-centre than off-centre.

Alternatively, it may be harder for listeners to be consistent when the difference in perceived diffuseness between the reference and the test is large and whilst using an A/B type experiment (as opposed to the MUSHRA method from experiment 1).

Inter-frequency Comparisons

Looking at the calculated diffusenesses of the different frequencies (figure 6) we see that on-centre 1 kHz was more diffuse than the 125 Hz or 8 kHz bands. Off-centre, 1 kHz was equally diffuse as 125 Hz but more diffuse than 8 kHz. Whilst the differences between frequencies are significant, the differences are small especially when compared to the effect of the ICC.

Summary

Even when combining the differences between the references (the calculated inter-frequency reference diffusenesses) applied to the intra-frequency comparisons, the difference in diffuseness between frequencies is not obvious. It is therefore suggested that to ensure high perceived diffuseness, low correlation should be used at all frequencies (at least 125 Hz–8 kHz) and this is especially important on-centre. As a side note these results also highlight the limitations of metrics such as Interaural Cross-Correlation Coefficient (IACC) for measuring low frequency envelopment where the perception of variable diffuseness is strong to a much lower frequency than the the IACC is effective.

Conclusions

Two listening tests were conducted to investigate the relative effects of the loudspeaker layout, the inter-channel cross-correlation coefficient, the listener position and the centre-frequency of a stimulus. All these factors were shown to be significant.

Loudspeaker layout is only important at low ICC values. When the ICC is high, the diffuseness appears independent of the loudspeaker arrangement. At the off-centre listening position, the loudspeaker layout is even more relevant with fewer loudspeakers perceived as less diffuse. However, off-centre as the system is slightly more robust to low ICCs as the signals do not all arrive at the same time.

The centre-frequency of band-limited noise was found to be a significant factor although with a very small effect. This is interesting because high perceived diffuseness is dependent on a low ICC even when the wavelength is very long relative to the ear spacing (at low frequency) and equally when the wavelength is very short in comparison to the time of arrival differences between the loudspeakers (at high frequency when off-centre).

Therefore to generate a perceptually diffuse sound field it is important to maintain low inter-channel correlation from at least 125 Hz to 8 kHz. Whilst off-centre the system may be a little more robust to low amounts of inter-channel correlation, any low amount of correlation on-centre will reduce the perceived diffuseness.

Acknowledgements

This work has been carried out thanks to funding from the EPSRC and BBC R&D. The author would also like to thank Andrew Mason for his support and advice.

References

[1] Francis Rumsey. ‘Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm’. Journal of the Audio Engineering Society 50, no. 9 (2002): 651–666.

[2] Garth Paine, Robert Sazdov, and Kate Stevens. ‘Perceptual Investigation into Envelopment, Spatial Clarity, and Engulfment in Reproduced Multi-Channel Audio’. In Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio. Audio Engineering Society, 2007.

[3] Koichiro Hiyama, Setsu Komiyama, and Kimio Hamasaki. ‘The Minimum Number of Loudspeakers and Its Arrangement for Reproducing the Spatial Impression of Diffuse Sound Field’. In Audio Engineering Society Convention 113. Audio Engineering Society, 2002.

[4] Michael Cousins, Filippo Fazi, Stefan Bleeck, and Frank Melchior. ‘Subjective Diffuseness in Layer-Based Loudspeaker Systems with Height’. New York: Audio Engineering Society, 2015.

[5] ITU-R, B. S. ‘Rec. ITU-R BS.775-2: 2007 Multichannel Stereophonic Sound System with and without Accompanying Picture’. Int. Telecommun. Union, Geneva, Switzerland, 2007.

[6] ITU-R, B. S. ‘Rec. ITU-R BS.2051-0: 2014 Advanced Sound System for Programme Production’. Int. Telecommun. Union, Geneva, Switzerland, 2014.

[7] ITU-R, B. S. ‘Rec. ITU-R BS.1534-2:2014 Method for the Subjective Assessment of Intermediate Quality Level of Audio Systems’. Int. Telecommun. Union, Geneva, Switzerland, 2014.

[8] Nick Zacharov, Torben Pedersen, and Chris Pike. ‘A Common Lexicon for Spatial Sound Quality Assessment-Latest Developments’. In Quality of Multimedia Experience (QoMEX), 2016 Eighth International Conference On, 1–6. IEEE, 2016.

[9] Slawomir Zielinski, Philip Hardisty, Christopher Hummersone, and Francis Rumsey. ‘Potential Biases in MUSHRA Listening Tests’. In Audio Engineering Society Convention 123. Audio Engineering Society, 2007.

[10] ITU-R, B. S. ‘Rec. ITU-R BS.1116-1:1997 Method for the Subjective Assessment of Small Impairments in Audio Systems Including Multi-Channel Sound Systems.’ Int. Telecommun. Union, Geneva, Switzerland, 1997.