Acoustics'08
Paris
June 29-July 4, 2008
www.acoustics08-paris.org

Diffuseness and intensity analysis of spatial impulse responses

T. Lokki

Helsinki University of Technology, P.O.Box 5400, 02015 TKK, Finland
tapio.lokki@tkk.fi

Spatial impulse responses, meaning responses measured with a microphone grid, were measured in six concert halls. The microphone array consisted of 12 omnidirectional microphones, enabling a construction of 3 intensity pairs (in x, y, and z directions) with 1 cm spacing and 3 intensity pairs with 10 cm spacing. In each hall impulse responses were measured with at least 3 loudspeaker and 4 microphone positions. Responses were analyzed with directional audio coding methodology, which enables analysis of diffuseness and instantaneous intensity as a function of time and frequency. In other words, with this analysis it is possible to analyze the directions of early reflections and to estimate diffuseness of sound field in a measurement position. Preliminary results indicate that diffuseness is quite similar in different positions in one hall, but it varies more between halls. The directions of early reflections are hard to visualize, however some example illustrations are shown to get an idea about the possibilities of such an analysis technique.

1 Introduction

Traditionally room impulse responses are measured with omnidirectional microphones or with two channel microphones, e.g., dummy heads. However, with current technology it is possible to use more advanced microphone techniques to measure spatial impulse responses. Multichannel impulse responses enable to analysis of spatial information, e.g., the spatial distribution of reflections. In this paper an energy-based analysis method, which is developed for spatial audio recording and reproduction [1, 2], is applied to study the diffuseness and intensity in concert halls. Diffuseness could tell interesting information about the diffuse characteristics of the sound field. On the other hand intensity tells the direction of energy flow in the measurement positions.

Impulse responses were measured in six halls with an omnidirectional dodecahedron loudspeaker as a sound source. Spatial sound was captured with a custom microphone array consisting of 12 Sennheiser KE 4-211-2 omnidirectional electret microphone capsules arranged as two pairs on each Cartesian coordinate axis [3], see Fig. 1. The considered concert halls have different shapes and sizes and they have quite different acoustics as seen in Table 1.

1.1 Related work

The idea of analysing spatial impulse responses has been proposed many times. For example, Essert proposed the measurement and analysis methods for such responses [4]. However, only a few papers are published about analysis of spatial responses of concert halls. The intensity analysis has been presented earlier by Merimaa et al. [5] and by Omoto and Ushida [6]. However, their visualizations are different than presented in this paper. Another visualization attempt is presented by Fukushima et al. [7]. They converted reflections to image sources and plotted them on top of the photograph of the real space. Okubo et al. [8] proposed novel directional room acoustic parameters with spatial responses measured with several omni microphones.

Measurements of spatial impulse responses with a spherical microphone array have been done by Gover et al. [9], Park and Rafaely [10], and Rafaely et al. [11]. They all managed to find early reflections from responses measured in real rooms.

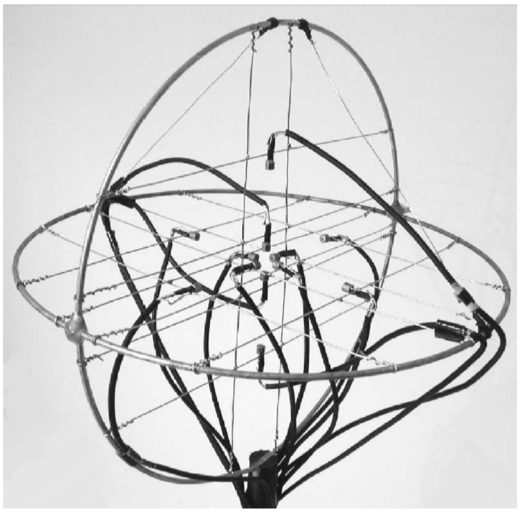

Figure 1: The custom 12 microphone grid [3].

2 Energy-based analysis

The diffuseness and the direction estimates of the energy flow is performed with an energy-based analysis. Such analysis has been successfully applied in multichannel audio reproduction, namely in spatial impulse response rendering [1, 2] as well as in directional audio coding [13]. The details of the energy analysis as well as deeper theory can be found in [5, 14]. Here, it is briefly explained how the analysis is performed for the spatial impulse responses measured with the custom 12 microphone grid.

The analysis is done in time-frequency blocks. The impulse response is divided into time frames which are again divided into frequency bands. Such analysis can be done, e.g., with STFT. For analysis both sound pressure and particle velocity of the signal is needed. The sound pressure can be obtained by averaging the six inner microphone signals ($p_i(t)$):

$$p(t) \approx \frac{1}{6}\sum_{i=1}^{6} p_i(t) \tag{1}$$

The particle velocity can be approximated at the point halfway between a microphone pair [15, 14]

$$u_x(t) \approx \frac{1}{\rho_0 d}\int_{-\infty}^{t} [p_1(\tau) - p_2(\tau)]d\tau \tag{2}$$

where d is distance between microphones and ρ_0 is mean density of air. Respectively, particle velocities in y and z directions are computed.

Hall	Attr.	125	250	500	1k	2k	4k
Tampere 2000 seats	EDT	1.5	1.5	1.6	1.6	1.6	1.4
	T30	2.0	1.9	1.9	1.8	1.7	1.6
	G	8.5	1.2	1.3	2.3	0.7	2.4
	LF	12.6	16.8	18.2	19.2	21.2	21.9
Helsinki 1700 seats	EDT	1.3	1.5	1.7	1.8	2.0	1.8
	T30	1.8	1.8	1.9	1.9	2.0	1.8
	G	7.3	0.6	3.6	4.2	3.5	5.8
	LF	6.4	11.6	14.1	14.2	17.8	22.2
Lahti 1500 seats	EDT	2.2	2.2	2.2	2.3	2.1	1.9
	T30	2.4	2.4	2.5	2.5	2.3	2.1
	G	10.6	2.8	3.8	5.4	3.6	3.9
	LF	18.1	22.6	23.0	26.7	29.9	36.5
Kuopio 1060 seats	EDT	1.5	1.6	1.6	1.7	1.4	1.2
	T30	1.5	1.7	1.7	1.8	1.6	1.4
	G	9.5	3.8	4.5	5.8	4.9	5.1
	LF	15.9	18.9	20.7	23.1	26.7	31.2
Pori 700 seats	EDT	2.5	2.2	2.2	2.2	1.9	1.5
	T30	2.5	2.3	2.2	2.2	1.9	1.6
	G	12.3	5.7	6.7	7.6	5.9	6.6
	LF	21.4	21.8	24.0	26.6	28.6	30.6
Järvenpää 570 seats	EDT	1.5	1.5	2.0	2.0	1.9	1.6
	T30	1.6	1.5	1.9	2.0	1.9	1.7
	G	8.6	4.4	6.2	8.4	6.5	7.3
	LF	12.9	14.9	19.3	20.6	25.5	26.0

Table 1: Averaged acoustical attributes [12] of six halls.

With sound pressure and particle velocity, the single-sided frequency distribution of the active intensity in an analysis window can be written as

$$\mathbf{I}_a(\omega) = 2\mathrm{Re}\{P^*(\omega)\mathbf{U}(\omega)\} \tag{3}$$

where $P(\omega)$ and $\mathbf{U}(\omega)$ are the Fourier transforms of the sound pressure and the particle velocity in a time window, and * denotes complex conjugate. Since the grid has microphone pairs at three orthogonal axis (x, y, and z), it is possible to compute the direction of the active intensity in the 3-D space. The 1 cm and 10 cm pairs are used to cover a frequency range from 100 Hz to 7 kHz [3] (the cross-over between 1 cm and 10 cm pairs is 800-1000 Hz).

The diffusity of the sound field can be defined as a ratio between intensity and energy density, thus, the diffuseness estimate can be written as

$$\psi(\omega) = 1 - \frac{||\mathbf{I}_a(\omega)||}{E(\omega)c} = 1 - \frac{2Z_0\,||\mathrm{Re}\{P^*(\omega)\mathbf{U}(\omega)\}||}{|P(\omega)|^2 + Z_0^2|\mathbf{U}(\omega)|^2} \tag{4}$$

where $||.||$ denotes the norm of a vector, $E(\omega)$ is energy density, Z_0 is the acoustics impedance of air, defined as $Z_0 = \rho_0 c$, and c is speed of sound.

3 Results

Spatial room impulse responses are inherently multidimensional, having components in three coordinate directions in time and frequency. In addition, based on the proposed energy-based analysis the time-frequency blocks can be further devided into directional and diffuse components. Therefore, it is hard to visualize such multidimensional data. In this paper, first the diffuseness estimates are visualized as a time-frequency presentation and intensity data is visualized from the human perception point of view.

The analysis is performed for 12 responses measured in all six concert halls. The three source and four receiver positions were similarly chosen in each hall, see Fig. 3, although they were not exactly the same positions since the dimensions of the halls vary.

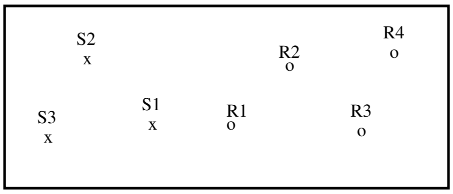

Figure 3: Source (S) and receiver (R) positions.

3.1 Diffuseness in measured halls

As mentioned earlier the proposed analysis is performed in the time-frequency blocks. The measured responses were first divided into time frames with 1024 samples long hanning window with 75% overlap in consecutive windows (fs = 48 kHz, 1024 samples = 21.3 ms). Then an FFT was taken from each time block to divide data into time-frequency tiles. Then, the diffuseness estimates were computed for each single block with Eq. (4).

Each time-frequency tile has a diffuseness value between 0 and 1, even if the tile contains only noise. To be able to interpret the data, the diffuseness estimates are set to 0 or 1. The 0 value is obtained if the tile has less than -30 dB energy compared to direct sound and the diffuseness estimate $\psi(\omega, t)$ is less than 0.4. Correspondingly, a tile is given value 1 if the tile has enough energy and $\psi(\omega, t) \geq 0.4$. Time-frequency visualizations of first 300 ms are illustrated in Fig. 2 where “0 tiles” are black and “1 tiles” are white.

Visualizations of diffuseness estimates in Fig. 2 show how fast diffuseness is developed after the direct sound. In Tampere it takes longer time as in Kuopio or Pori which is quite natural since Tampere hall is much bigger than other halls. Interestingly it also seems that the diffuseness plots are quite similar in all responses measured in one hall, but the overall shape and density of white tiles vary more between halls. This suggest that the amount of diffuse energy in one hall is not very position dependent.

To have more insight to the density of diffuse (=white) tiles at each time moment, the number of tiles were calculated. In other words “diffuseness curves in a function of time” were computed, see Fig. 4. All 12 responses in one hall are plotted with different colors and the black thick line is the average of all 12 curves. There are differences between halls, e.g., in the raise times of these “diffuseness curves” as well as in overall levels of diffuseness.

3.2 Intensity in measured halls

The intensity of sound in each time-frequency tile was analysed with Eq. (3). Intensity information has at least 5 dimensions (x, y, z, time, and frequency), thus, it is also very hard to visualize with static 2-D plots. Since early reflections are the most interesting from spatial sound point of view, we visualize the first 100 ms of analyzed data from 2 responses from each hall in Figs. 5 and 6. The intensity value is shown with the darkness of the marks in dB scale on one narrow frequency band.

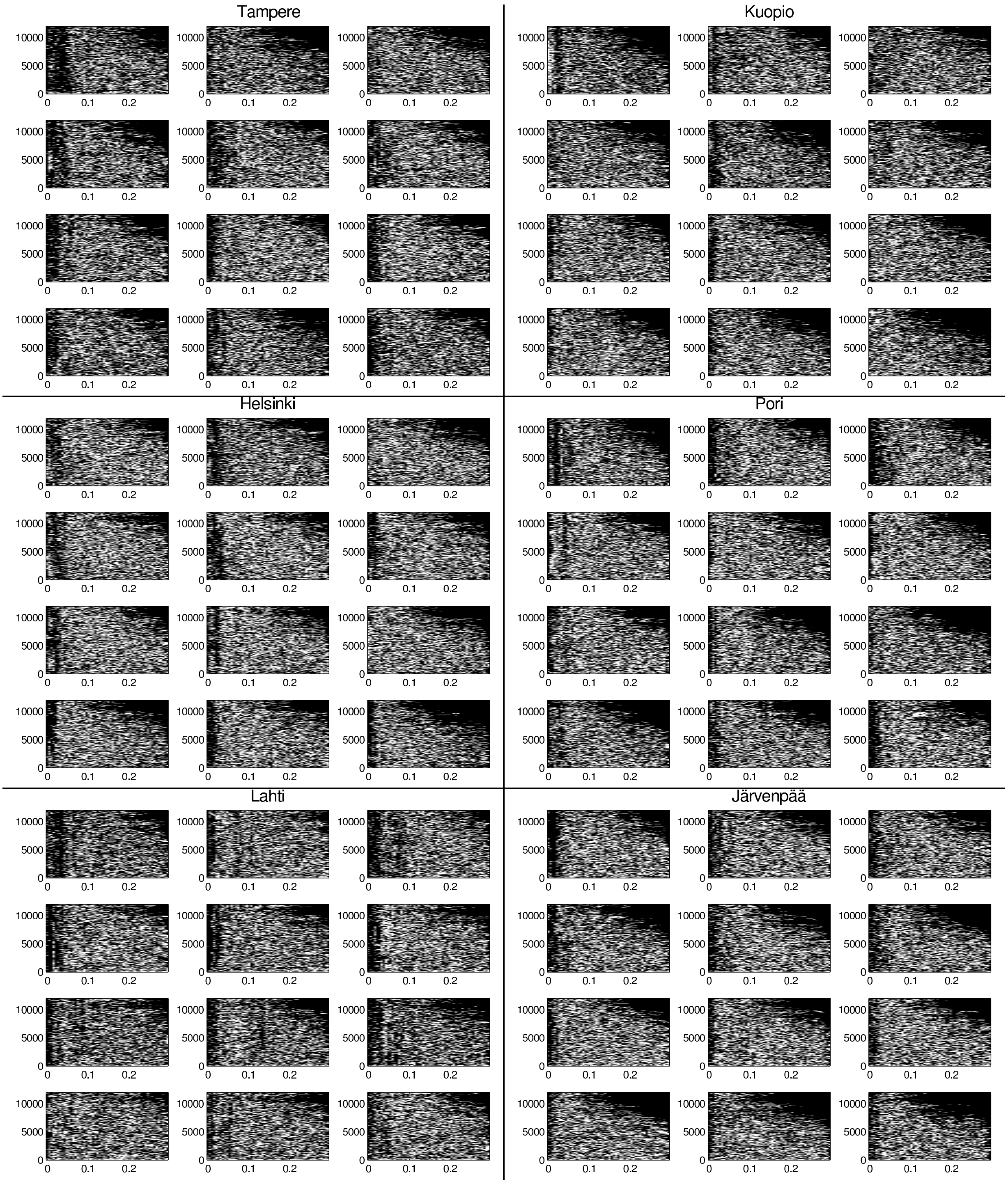

Figure 2: Diffusiness estimates of all six halls. Each hall has 12 responses; 3 sources (columns) and 4 receivers (rows). If the time-frequency tile is white it is considered to have significant diffuse energy; SPL ≥ -30 dB (direct sound normalized to 0 dB) and $\psi(\omega, t) \geq 0.4$. In each plot x-axis is time [s] and y-axis is frequency [Hz].

The direction of intensity vector is mapped on the unit circle and looked from right and left. In other words, the plots show the incoming angle of sound, mapped to the cone-of-confusion circles (at intervals of 10°) of both ears. The 100 ms is divided into 128-samples frames (=2.67 ms), and the FFT size for each frame was 1024.

Figure 5 shows the direct sound (normalized to 0 dB) coming to the right ear. Regarding early reflections, there is big differences between halls. For example, in Helsinki hall only a few early reflections are coming from side. This supports the low LF value at 250 Hz, see Table 1. In Fig. 6 it is seen that Pori and Lahti have reflections coming from side at 1 kHz, as proposed by high LF values.

4 Conclusion

An energy-based analysis of spatial impulse responses, measured in six concert halls is presented. The proposed method enables to extract intensities and diffuseness in time-frequency blocks. The analysis results in multidimensional data which is hard to visualize. Here, only a few visualizations are illustrated, but many other techniques should be studied in the future. Analysis of spatial impulse responses could also be done with different

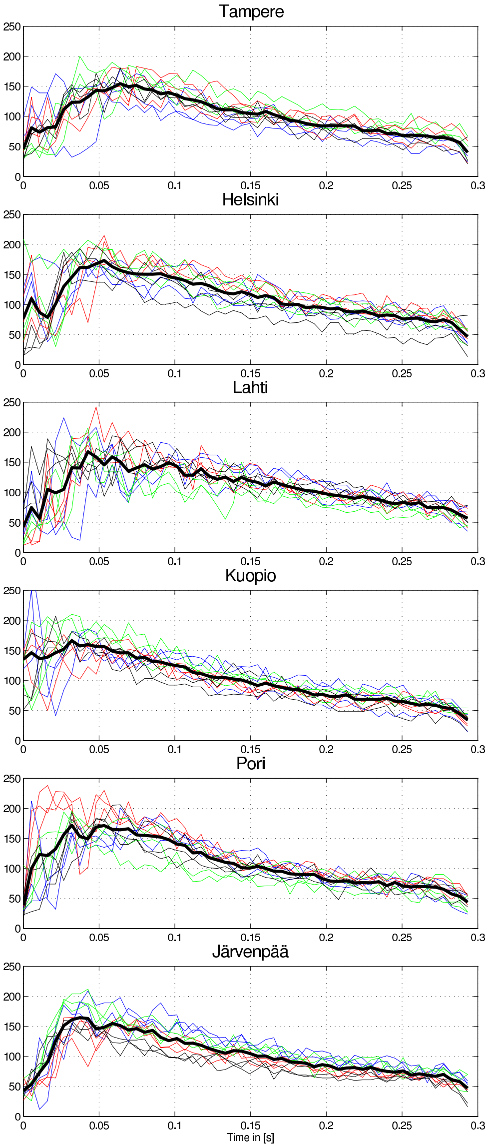

Figure 4: "Wide band diffuseness" of all six halls as a function of time.

time and frequency resolutions. The proposed method enables also investigations of continuos signals, not only impulse responses. This work remains as a future work.

Acknowledgments

This work has been supported by the Academy of Finland (project no. 119092). In addition, I'd like to thank Dr. Juha Merimaa for a few matlab scripts.

References

[1] J. Merimaa and V. Pulkki. Spatial impulse response rendering I: Analysis and synthesis. *J. Audio Eng. Soc.*, 53(12):1115–1127, 2005.

[2] V. Pulkki and J. Merimaa. Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests. *J. Audio Eng. Soc.*, 54(1):3–20, 2006.

[3] T. Peltonen, T. Lokki, B. Gouatarbes, J. Merimaa, and M. Karjalainen. A system for multi-channel and binaural room response measurements. In *the 110th Audio Engineering Society (AES) Convention*, Amsterdam, the Netherlands, May 12-15 2001. preprint no. 5289.

[4] R. Essert. Progress in concert hall design – developing and awareness of spatial sound and learning how to control it. *EBU Technical Review*, pages 31–39, Winter 1997.

[5] J. Merimaa, T. Lokki, T. Peltonen, and M. Karjalainen. Measurement, analysis, and visualization of directional room responses. In *the 111th Audio Engineering Society (AES) Convention*, New York, USA, September 21-24 2001. preprint no. 5449.

[6] A. Omoto and H. Uchida. Evaluation method of artificial acoustical environment: Visualization of sound intensity. *Journal of Physiological Anthropology and Applied Human Science*, 23(6):249–253, 2004.

[7] Y. Fukushima, H. Suzuki, and A. Omoto. Visualization of reflected sound in enclosed space by sound intensity measurement. *Acoustical Science and Technology*, 27(3):187–189, 2006.

[8] H. Okubo, M. Otani, R. Ikezawa, S. Komiyama, and K. Nakabayashi. A system for measuring the directional room acoustical parameters. *Applied Acoustics*, 62:203–215, 2001.

[9] B.N. Gover, J.G. Ryan, and M.R. Stinson. Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array. *J. Acoust. Soc. Am.*, 116(4):2138–2148, 2004.

[10] M. Park and B. Rafaely. Sound-field analysis by plane-wave decomposition using spherical microphone array. *J. Acoust. Soc. Am.*, 118(5):3094–3103, 2005.

[11] B. Rafaely, I. Balmages, and L. Eger. High-resolution plane-wave decomposition in an auditorium using a dual-radius scanning spherical microphone array. *J. Acoust. Soc. Am.*, 122(5):2661–2668, 2007.

[12] ISO Standard 3382. *Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters.* International Standards Organization, 2000.

[13] V. Pulkki. Spatial sound reproduction with directional audio coding. *J. Audio Eng. Soc.*, 55(6):503–516, June 2007.

[14] J. Merimaa. *Analysis, synthesis, and perception of spatial sound – Binaural localizaion modeling and multichannel loudspeaker reproduction.* PhD thesis, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, report 77, 2006.

[15] F. J. Fahy. *Sound Intensity.* E & FN SPON, 2nd edition, 1995.

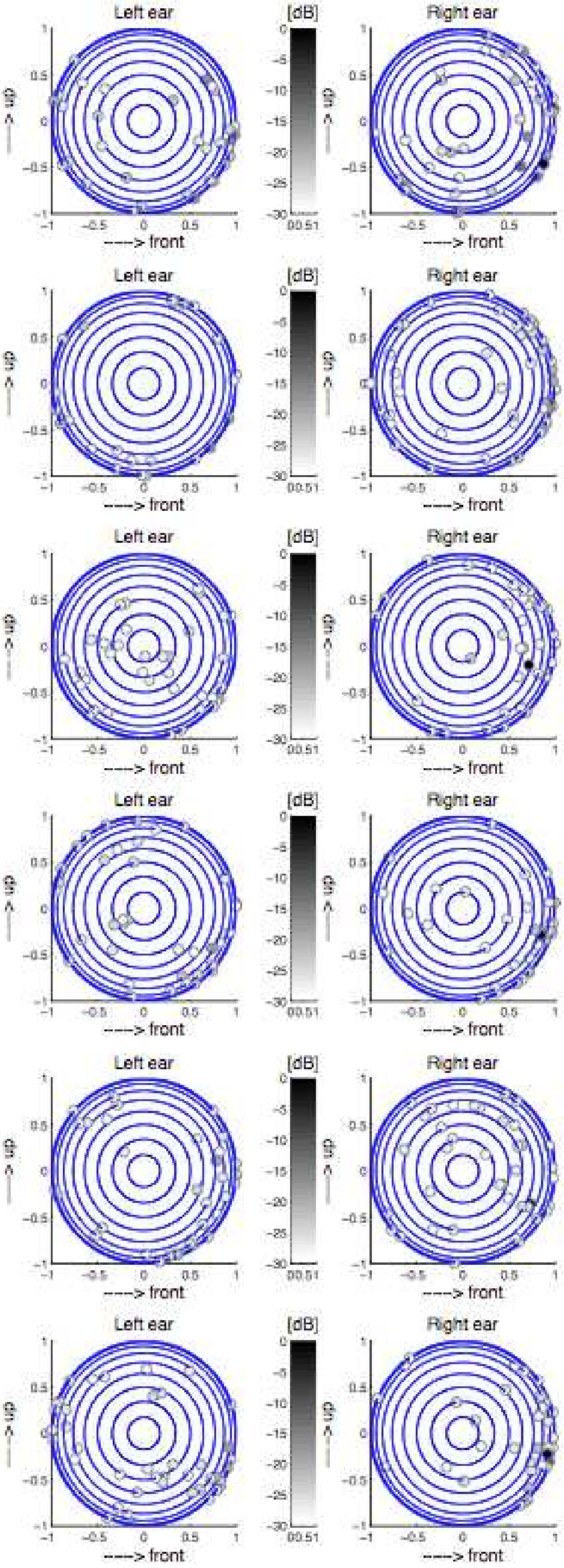

Figure 5: Intensity (O = 250 Hz) from source S2 to receiver R1 in each six halls (order of halls as in Fig. 4).

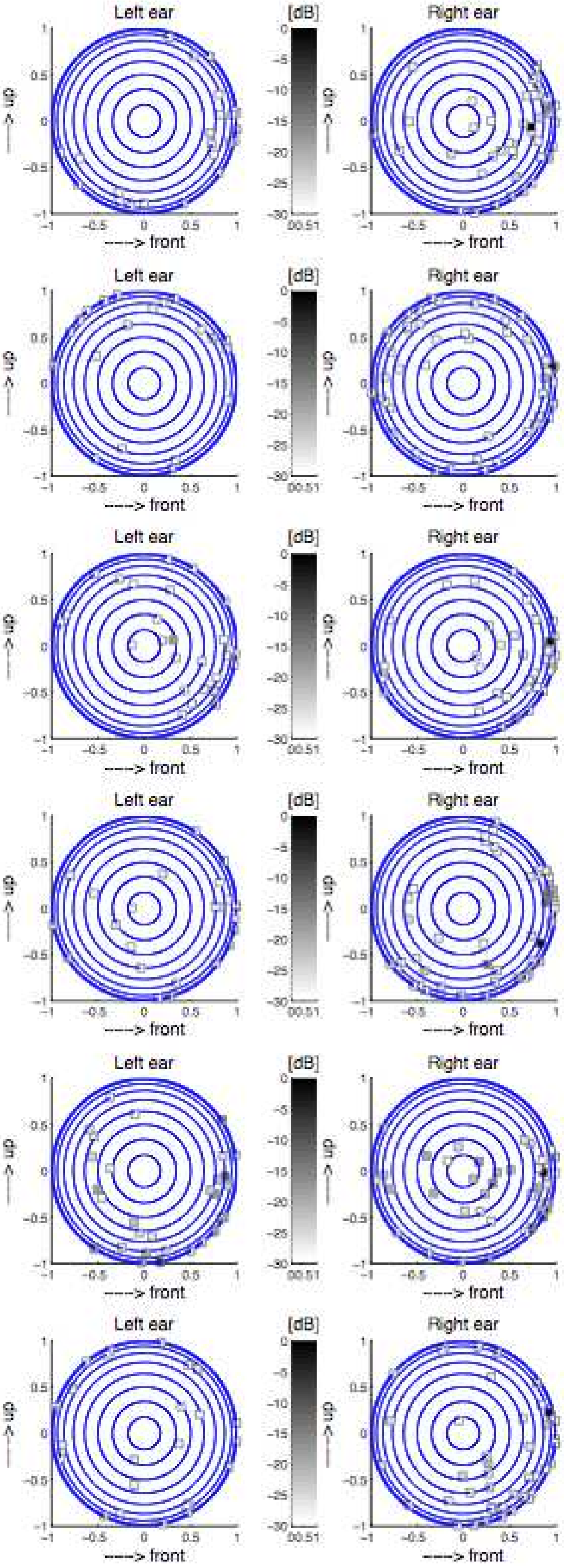

Figure 6: Intensity (□ = 1 kHz) from source S2 to receiver R3 in each six halls (order of halls as in Fig. 4).