Audio Engineering Society
# Convention Paper 5449
Presented at the 111th Convention
2001 September 21–24 New York, NY, USA



## Measurement, Analysis, and Visualization of Directional Room Responses

Juha Merimaa[1], Tapio Lokki[2], Timo Peltonen[3] and Matti Karjalainen[1]

[1]Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing,
P.O.Box 3000, FIN-02015 HUT, Finland
[2]Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory
P.O.Box 5400, FIN-02015 HUT, Finland
[3]Akukon Oy Consulting Engineers, Kornetintie 4 A, FIN-00380 Helsinki, Finland
`juha.merimaa@hut.fi, tapio.lokki@hut.fi, timo.peltonen@akukon.fi,`
`matti.karjalainen@hut.fi`

### ABSTRACT
Room impulse responses are inherently multidimensional, including components in three coordinate directions, each one further being described as a time-frequency representation. Suc h 5-dimensional data is difficult to visualize and interpret. We propose methods that apply 3-D microphone arrays, directional analysis of measured room responses, and visualization of data, yielding useful information about the time-frequency-direction properties of the responses. The applicability of the methods is demonstrated with three different cases of real measurements.

### INTRODUCTION
A room impulse response, measured from a source to a receiver position, is inherently multidimensional. Traditionally, the evolution of an omnidirectional sound pressure response in a single point has been studied as a function of time and frequency. However, dividing the response further into directional components can reveal much more information about the actual propagation of sound in the room, as well as about its perceptual aspects. In this paper we propose methods that are based on 3-D microphone arrays, directional analysis of the measured responses, and visualization of such data in a way that yields maximal information about the time-frequency-direction properties of the response.

The measurement of directional room responses is made with a special 3-D microphone probe which basically consists of two intensity probes in each x-, y-, and z-coordinate directions and is constructed of small electret capsules. The responses are analyzed either with a uniform or an auditorily motivated time-frequency resolution.

The analysis results in a significant amount of 5-dimensional data that is hard to visualize and interpret. Based on measured x/y/z-intensity components, intensity vectors (magnitude and direction) can be plotted in a spectrogram-like map, one vector for each time-frequency bin, illustrating the directional evolution of the field in time and frequency. Additionally, a pressure-related time-frequency spectrogram can be overlaid with the vectors, in gray levels or colors, illustrating for example a perceptually motivated spectrogram with no directional information. One such map can be used to illustrate the horizontal information and another one can be added for the elevation information.

This technique is a part of a Matlab visualization toolbox for directional room responses developed by the authors, and it includes several other possibilities to analyze and represent room acoustical data. Traditional parameters and presentations are also available, some of them in 3-D versions, such as energy-time plots in desired directions.

The paper starts with a discussion on measurements of directional room responses and sound intensity. This is followed by descriptions of the visualization method and the auditorily motivated time-frequency analysis. Finally, the applicability of the methods is demonstrated with three different cases of real measurements.

## DIRECTIONAL SOUND PRESSURE COMPONENTS

Existing literature on room acoustics discusses mainly omnidirectional measurements with the exception of some special directional parameters. Directional room responses can be measured with either directional microphones or arrays of microphones. However, an array of omnidirectional microphones has some distinct advantages compared to directional microphones. Omnidirectional capsules can be made smaller and they usually behave more like ideal transducers. Further, if the omnidirectional signals are stored at the measurement time, it is possible to afterwards create varying directivity patterns based on a single measurement.

Typical directivity patterns can be formed with an array of two or more closely spaced omnidirectional microphones and some equalization to compensate for the resulting non-flat magnitude response. For example the difference of two microphone signals gives a dipole pattern and adding an appropriate delay to one of the signals changes the pattern to a cardioid. Okubo et al. [1] have also proposed a method that uses a product of cardioid and dipole signals to achieve a directivity pattern more suitable for some directional room acoustics measurements.

Various directional sound pressure responses can be used to plot traditional impulse responses, energy-time-curves or spectrograms that give information about the directional properties of the room responses. With larger microphone arrays it is also possible to form directivity patterns with very narrow beams and thus good spatial resolution. However, groups of similar plots for several different directions are not very visual or easy to interpret. Sound intensity as a vector quantity can solve some of the visualization problems in the method we are proposing in this paper.

## SOUND INTENSITY

Sound intensity [2] describes the propagation of energy in a sound field. Instantaneous intensity vector is defined as the product of instantaneous sound pressure $p(t)$ and particle velocity $\mathbf{u}(t)$

$$\mathbf{I}(t) = p(t)\mathbf{u}(t) \tag{1}$$

Based on the linearized fluid momentum equation, particle velocity in the direction $n$ can be written in the form

$$u_n(t) = -\frac{1}{\rho_0}\int_{-\infty}^{t} \frac{\partial p(\tau)}{\partial n} \mathrm{d}\tau \tag{2}$$

where $\rho_0$ is the mean density of air. The pressure gradient can be approximated with a finite difference of the signals of two closely spaced omnidirectional microphones aligned such that a vector connecting the microphones points to the direction $n$. On the other hand, an estimate for the sound pressure at the same point halfway between the microphone pair is given by the average of the two microphone signals. Hence, an approximation for the instantaneous intensity is given by

$$I_n(t) \approx \frac{1}{2\rho_0 d}\left[p_1(t) + p_2(t)\right]\int_{-\infty}^{t}\left[p_1(\tau) - p_2(\tau)\right]\mathrm{d}\tau \tag{3}$$

where $p_1(t)$ and $p_2(t)$ are the two microphone signals and $d$ is the distance between the microphones. Furthermore, a 3-D intensity vector can be calculated by utilizing three concentric pairs of microphones aligned such that the vectors connecting each pair span a 3-D space.

Equation (3) can be directly applied in intensity measurements. If discrete-time signal processing is used, the integral is simply replaced with a summation of sound pressure differences. Typically the result is further averaged in time to yield the so called active intensity component, i.e., the component corresponding to the net transport of sound energy during the averaging period.

What is of interest in the frequency distribution of sound intensity is the contribution of sound pressure and particle velocity at the same frequency or frequency band. The most straightforward way to calculate this is to feed the microphone pair signals through a filterbank before applying equation (3). The filterbank approach also provides means to use an arbitrary frequency resolution. If, however, uniform resolution is desired, there is a computationally more efficient way to determine the frequency distribution. It can be shown [2, 3] that the frequency distribution of active intensity is given by

$$I_n(\omega) \approx -\frac{j}{\rho_0 \omega d} Im\left\{G_{p1p2}(\omega)\right\} \tag{4}$$

where $\omega$ is the angular frequency, $j$ is the imaginary unit, and the single-sided cross-spectral density of the microphone signals $G_{p1p2}$ is given by

$$G_{p1p2}(\omega) = 2P_1^*(\omega)P_2(\omega) \tag{5}$$

where $^*$ denotes the complex conjugate and $P_1(\omega)$ and $P_2(\omega)$ are the Fourier transforms of the microphone signals $p_1(t)$ and $p_2(t)$, respectively.

The finite difference approximations of both sound pressure and particle velocity at the point halfway between a microphone pair cause frequency dependent systematic error to the intensity approximation. The actual error depends on the acoustical field being measured [2]. Chung [4] has proposed a practical upper limit of

$$kd = 1 \Leftrightarrow f_L = \frac{c}{2\pi d} \tag{6}$$

where $k$ is the wave number, $f_L$ is the limiting frequency and $c$ is the speed of sound. The practical lower frequency limit, on the other hand, depends mainly on the phase errors of the measurement system. One method to alleviate the frequency range problems is to use a multi-microphone measurement technique proposed in [5]. However, our choice has been to use a microphone probe with two microphone pairs with different spacings at each coordinate direction in order to cover a wider frequency range.

**VISUALIZATION OF DIRECTIONAL ROOM RESPONSES**
The difference between the approach taken in this study and the traditional physical theory is that our sound field presentations are listener-centered, compared to plotting the sound pressure and intensity as functions of spatial position. This can also be considered as the difference between physical approach and signal modeling approach.

The visualization method proposed in this paper is based on a combination of two overlaid plots describing different aspects of the sound field being measured. Active intensity is represented with vectors (quiver plot) on top of a sound pressure related spectrogram. Both of them are analyzed with either one of two alternative time-frequency resolutions; a uniform resolution useful in physical analysis of the responses, and an auditory resolution related to the properties of human hearing. One combination plot is used to illustrate the horizontal information and another one the elevation information in the median plane.

The directions of the vectors represent the direction of the mean flow of sound energy at a given time and frequency band. The length of a vector is scaled proportionally to the logarithmic magnitude of the active intensity exceeding an adjustable threshold. More specifically, the range of the magnitudes of the vectors that are plotted can be set relative to the peak value. With this control, the amount of details shown can be adjusted. With a small range, for example, strong discrete reflections are clearly highlighted, while most details of the sound field will be hidden.

For the sound pressure related spectrogram, similar thresholding is applied. The spectrogram represents the logarithmic magnitude of an omnidirectional pressure signal. This serves two purposes. First, human ear is sensitive to sound pressure rather than intensity, and thus the spectrogram provides valuable additional information about the perceptual aspects of room responses. Furthermore, the differences between sound pressure and intensity describe the diffuseness of the sound field. In a completely diffuse field the active intensity should be zero, wheras a sound pressure signal exists.

**AUDITORILY MOTIVATED ANALYSIS**
Traditionally, uniform, octave, or one-third octave band frequency resolution is used in analysis. However, these are not optimal from a perceptual point of view. The frequency resolution of human hearing is a complex phenomenon which depends on many factors, such as frequency, signal bandwidth, and signal level. Despite the fact that our ear is very accurate in single frequency analysis, broadband signals are analyzed using quite sparse frequency resolution. Nowadays, Equivalent Rectangular Bandwidth (ERB) scale [6, 7] is considered the most accurate one for auditory research, while the Bark scale [8] is often used in engineering applications.

The time resolution of human hearing is an even more complex phenomenon. In some cases monaural time resolution is as good as about 1–2 ms at high frequencies and a little worse at lower frequencies. On the other hand, the temporal integration time constant and the post-masking effect after a noise masker are over 100 ms, even 200 ms. Spatial hearing complicates the situation further, since the precedence effect [9] causes the sound following a sharp onset to be suppressed in sound source localization while the sound still affects the perception of timbre and space. A complete and detailed model for the time resolution is not known. In this study we have tried to find an integrating window that roughly simulates the temporal resolution of human hearing.

The analysis method applied in this paper is based on an auditorily motivated method, presented earlier in [10]. A block diagram of the method is depicted in Fig. 1. The microphone input signals are fed into a gammatone filterbank which divides them into 32 ERB bands ranging from 100 Hz to 9.2 kHz. After this, the instantaneous intensity and the squared sound pressure signal for each band are formed. Taking the square of the pressure resembles the half-wave rectification done by the hair cells in human hearing. A sliding window is used to simulate the time resolution of the ear and to form the time-averaged active intensity.

The resulting directional intensity components are used to form a quiver plot for both horizontal and vertical planes. The directions of the vectors are determined separately and their magnitudes, as well as the magnitude of the omnidirectional sound pressure, are presented on a logarithmic scale. Since intensity is already an energy related quantity, the logarithm of its squared magnitude must be multiplied by five to get a decibel representation.

Although this analysis method is far from a full-scale auditory model, it however respects the frequency and time resolution of human hearing better than a uniform or one-third octave frequency band analysis.

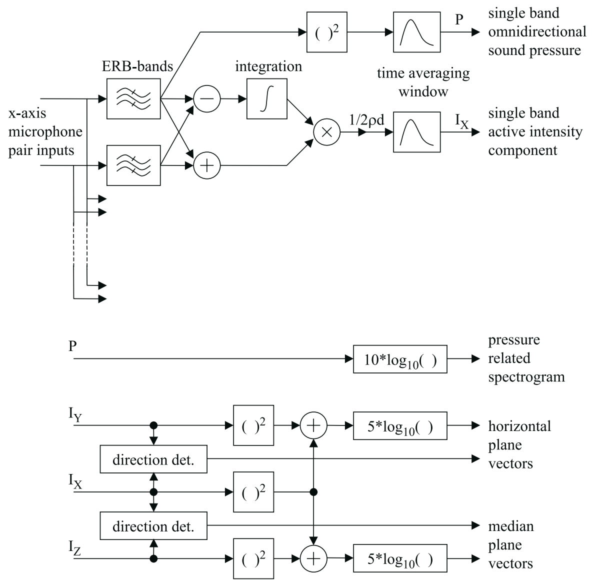


Fig. 1: *A block diagram of the auditorily motivated analysis method.*

**EXAMPLES**
In this section, three cases of directional room response visualization are demonstrated. In the first case, the response is measured in an anechoic chamber, including the direct sound and a single mirror-like reflection from an added surface. It shows the basic functionality of the system in a regular case. The second case is a listening room with rectangular geometry but having surfaces with diffuse reflections. The third case is a concert hall where the development of the early sound field is analyzed and visualized. In all three cases the method is able to visualize important information of the acoustic environment under study.

**Measurement system**
The measurements of directional room responses were made with a special 3-D microphone probe which is basically x-, y-, and z-oriented intensity probes constructed of small electret capsules. In each dimension two microphone spacings are used to cover the frequency range of 100 Hz to 8 kHz with good signal to noise ratio and directional resolution [11].

The responses were acquired using the IRMA measurement system

[11, 12], which enables simultaneous acquisition of multichannel impulse responses. The system hardware consists of a pc computer, a multichannel sound card, and an external AD converter. The measurement and analysis software runs in the Matlab environment. Impulse responses are measured using the MLS method [13].

The omnidirectional sound source consists of 12 wideband speaker elements mounted in a dodecahedron-shaped frame of 30 cm in diameter. The unit complies with the ISO-3382 [14] standard specifications for response magnitude and directivity, and provides a usable frequency range of 100 Hz to 8 kHz.

The performance of the measurement system can be enhanced with some simple postprocessing of the measured responses. In our system, the average magnitude response of the omnidirectional sound source is compensated with a minimum phase inverse filter to be flat at the frequency range of interest. Total compensation of all the errors caused by the source is impossible, since both its magnitude and phase responses are direction dependent at high frequencies. However, in most room acoustical measurements, these errors are negligible.

The microphones of the 3-D probe give fairly good results. A simple gain compensation was applied to the magnitude response of each capsule. With careful phase compensation the applicable frequency range could be extended to a little lower frequencies. However, no phase compensation was required for the frequency range provided by the omnidirectional sound source.

**Anechoic chamber**

To illustrate the performance of the proposed measurement, analysis, and visualization method, simplified measurements were made in an anechoic chamber with a hard floor. Fig. 2 shows the measurement setup. The paths of the direct sound and the floor reflection are also illustrated.

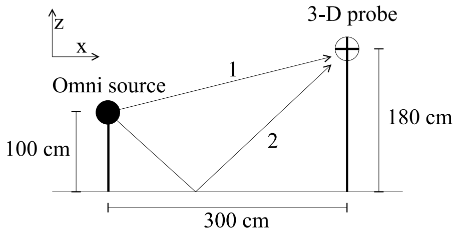


Fig. 2: *The measurement setup in an anechoic chamber with a hard floor.*

The omnidirectional sound source and the 3-D microphone probe were separated horizontally by 3 m and vertically by 80 cm. The probe was directed such that the sound source was in the median plane. Thus there should be no intensity component in the direction of y-axis. Furthermore, the angle between the direct sound and the horizontal plane should be about 15° and the angle between the floor reflection and the horizontal plane about 43°. The time difference between the two sound events should be a little less than 3 ms.

Fig. 3 shows the visualization of the results analyzed with a uniform time-frequency resolution. Measurements were done with a sampling rate of 48 kHz. 128-point FFT of hanning-windowed 50 % overlapping sections of the microphone signals was utilized. This yields a time resolution of roughly 1.3 ms. A magnitude range of 25 dB was used for the plot. The left-hand part of the figure shows the response in the horizontal plane such that the arrows describe the direction of propagation of the sound. Both the direct sound and the floor reflection are travelling to the direction of the positive x-axis. The right-hand part shows a similar plot of the median plane, in which the direct sound and the floor reflection have different directions. The non-flat magnitude characteristics of the two discrete events is caused by the direction dependence of the omnidirectional source at high frequencies.

The same response analyzed on an auditorily motivated time-frequency resolution is shown in Fig. 4. In order to facilitate comparison with Fig. 3 the same time step size of 1.3 ms is used although the actual time resolution is about 2 ms at its best. As can be seen, the floor reflection is merged with the direct sound. From a perceptual point of view this is a more correct representation of the response. However, for physical interpretation the uniform time-frequency resolution clearly provides more precise information.

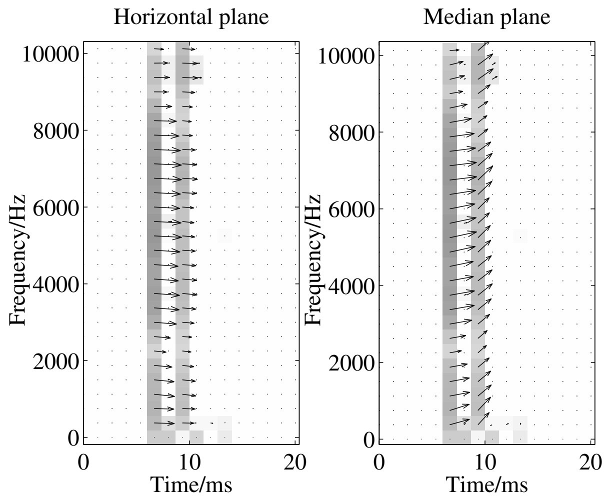


Fig. 3: *An anechoic response with a floor reflection, analyzed with a uniform time-frequency resolution.*

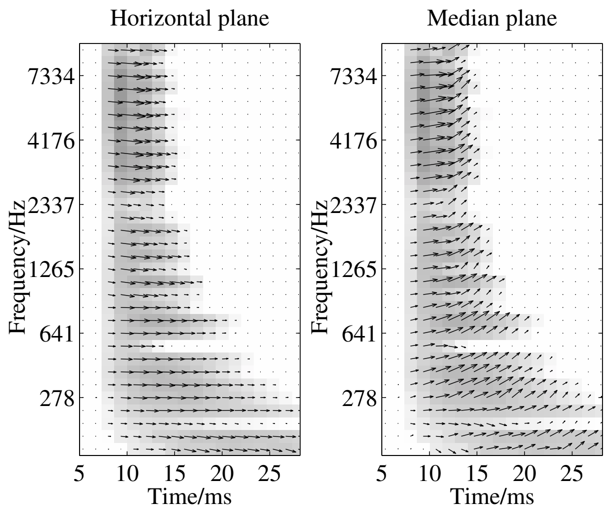


Fig. 4: *An anechoic response with a floor reflection, analyzed with an auditorily motivated time-frequency resolution.*

The spreading in time of the energy of the two impulses at low frequencies is caused by the sharper frequency resolution applied. The filters also introduce some delay to the signals. For visualization purposes the delays have been partially compensated so that an impulse appers to begin approximately at the same time in each frequency band.

**Listening room**
The second example shows directional room responses measured in the listening room of the Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing [15]. The room and the measurement setup are depicted in Fig. 5. Also the paths for the direct sound from the omnidirectional source to the 3-D microphone probe and the first three room reflections are shown.

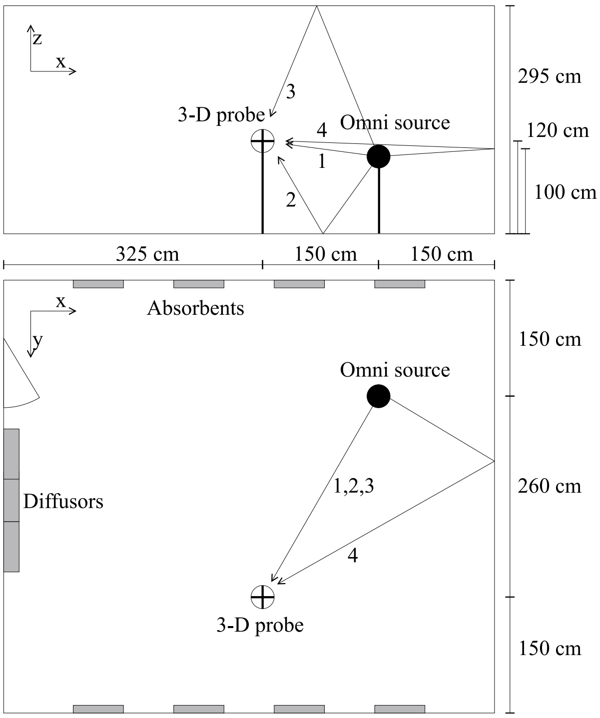


Fig. 5: *Layout of the listening room of the HUT Acoustics Laboratory. Top: sideview. Bottom: floor plan. First four paths from the omnidirectional sound source to the 3-D microphone probe have been drawn in the figure.*

Fig. 6 presents the directional responses visualized with a uniform time-frequency resolution. As in the previous example, the measurements have been done with a sampling rate of 48 kHz and similar analysis methods are applied. In this case, a shorter FFT window of the length of 64 samples is used in order to be able to better separate the first reflections. The magnitude range is limited to 25 dB. After 15 ms, the sound field can be seen to become more and more diffuse when a growing number of reflections starts to arrive from different directions at the same time. Some discrete reflections localized in frequency can still be seen.

Fig. 7 shows the auditorily motivated analysis of the listening room responses. Note that a different time scale is used. In this case, the discrete reflections are a harder to see, because they get merged with each other. However, the analysis shows some strong narrow band intensity components at frequencies below 250 Hz that cannot be seen with the uniform time-frequency analysis. Furthermore, the directions of these components vary significantly at adjacent frequency bands. The figure also shows how the response changes into diffuse reverberation.

**Concert hall**
The last example presents directional responses measured in a large Finnish concert hall located in Tampere. The omnidirectional sound source was located on the stage and the 3-D probe in the rear part of the parquet. The measurements were done with a sampling rate of 48 kHz. The results of a uniform time-frequency analysis are shown in Fig. 8 and those of an auditorily motivated resolution analysis in Fig. 9. The uniform analysis has been done with 128-point FFT of hanning-windowed sections of the responses. Again, both plots use a time step of 64 samples and a magnitude range of 25 dB.

In this case, both analysis methods are able to separate out the strongest discrete reflections. However, the frequency distributions of these reflections look quite different on different plots. For example, the first reflection that seems to be coming from the right side of the hall is very weak at frequencies below 2 kHz, which is very prominent in the auditorily motivated plot.

Compared to the previous examples these responses decay much slower. The direct sound arrives in about 50 ms and 120 ms after that no notable attenuation has yet occured except for the highest frequencies. By investigating several measured directional responses, they were also found to vary considerably as a function of the spatial positions of the probe and the sound source. A more detailed analysis of this kind of a concert hall with complex geometry and several irregular reflecting surfaces is beyond the scope of this paper.

## CONCLUSIONS AND FUTURE WORK

In this paper, methods to measure, analyze, and visualize directional room responses were proposed. The measurements are done with a specifically constructed 3-D microphone array and a multi-channel measurement system. Resulting responses are analyzed using either a uniform time-frequency resolution or a resolution simulating the properties of human hearing. Finally, the responses are visualized with two 2-D plots consisting of intensity vectors laid on top of a sound pressure related spectrogram.

The application of the methods is demonstrated with three cases of measurements of different acoustical spaces. In all three cases, the analysis and visualization methods are able to provide illustrative information about the directional properties of the responses.

Future work includes development of auditory models that better reflect the human perception of spatial sound, as well as evaluation of subjective importance of different properties of directional room responses.

## ACKNOWLEDGEMENTS

This study is a part of the VÄRE technology program, project TAKU (Control of Closed Space Acoustics), funded by Tekes (Finnish National Technology Agency).

J. Merimaa is supported by the Graduate School in Electronics, Telecommunications and Automation, and T. Lokki is supported by the Helsinki Graduate School in Computer Science.

## REFERENCES

[1] H. Okubo, M. Otani, R. Ikezawa, S. Komiyama, K. Nakabayashi, “A system for measuring the directional room acoustical parameters,” *Applied Acoustics* 2001, Vol. 62, Pp. 203–215.

[2] F.J. Fahy, “*Sound Intensity,*” Elsevier Science Publishers Ltd, Essex, England, 1989.

[3] F.J. Fahy, “Measurement of acoustic intensity using cross-spectral density of two microphone signals,” *J. Acoust. Soc. Am.* 1977, Vol. 62. Pp. 1057–1059.

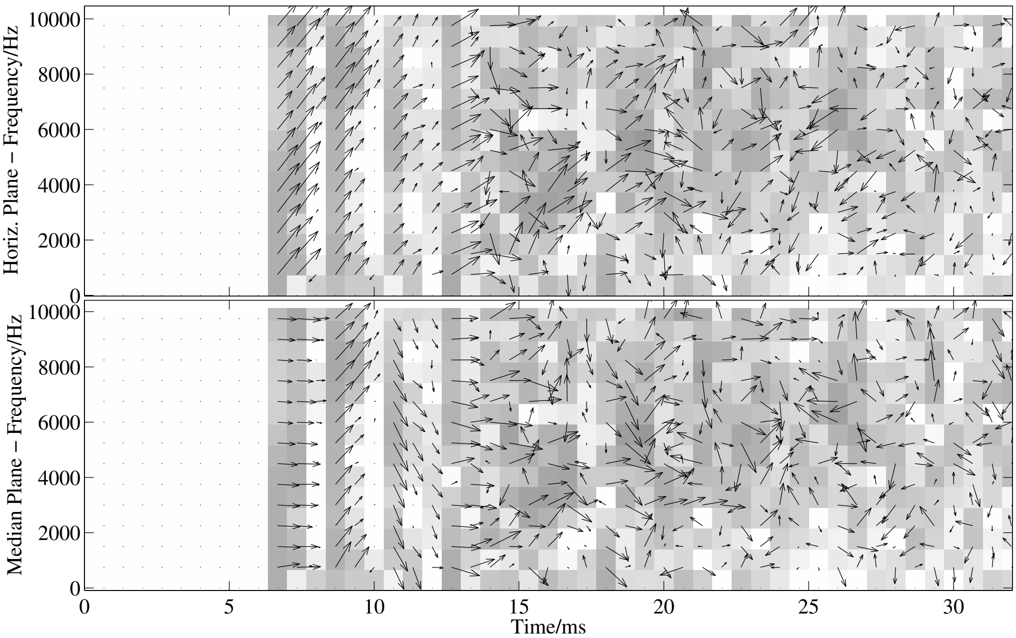


Fig. 6: *Directional room responses measured in the listening room of the HUT Acoustics laboratory, analyzed with a uniform time-frequency resolution.*


[4] J. Y. Chung, "Cross-spectral method of measuring acoustic intensity without error caused by instrument phase mismatch," *J. Acoust. Soc. Am.* 1978, Vol. 64, No. 6. Pp. 1613–1616.

[5] H. Kuttruff and A. Schmitz, "Measurement of Sound Intensity by Means of Multi-Microphone Probes," *Acustica* 1994, Vol. 80. Pp. 388–396.

[6] B.C.J. Moore, R.W. Peters, and B.R. Glasberg, "Auditory filter shapes at low center frequencies," *J. Acoust. Soc. Am.* 1990, Vol. 88, Pp. 132–140.

[7] B.C.J. Moore and B.R. Glasberg, "A revision of Zwicker's loudness model," *ACUSTICA united with acta acustica* 1996, Vol. 82, Pp. 335–345.

[8] E. Zwicker and H. Fastl, *Psychoacoustics — Facts and Models*. Springer-Verlag, Berlin, 1990.

[9] W.A. Yost and G. Gourevitch, "*Directional Hearing*," Springer-Verlag, New York, 1987.

[10] T. Lokki and M. Karjalainen, "An auditorily motivated analysis method for room impulse responses," in *Proc. COST-G6 Conference on Digital Audio Effects (DAFx-00)*, pp. 55–60, Verona, Italy, Dec. 7-9 2000.

[11] T. Peltonen, T. Lokki, B. Goutarbès, J. Merimaa and M. Karjalainen, "A System for Multi-Channel and Binaural Room Response Measurements," in *Proc. AES 110th Convention*, Amsterdam, The Netherlands 2001.

[12] T. Peltonen, "*A Multichannel Measurement System for Room Acoustics Analysis,*" M. Sc. Thesis, Helsinki University of Technology, Department of Electrical and Communications Engineering, Laboratory of Acoustics and Audio Signal Processing, 2000.

[13] D.D. Rife and J. Vanderkooy, "Measurement with Maximum-Length Sequences," *J. Audio Eng. Soc.*, 1989, Vol. 37, No. 6, Pp. 419–444.

[14] ISO 3382:1997. "*Acoustics — Measurement of the reverberation time of rooms with reference to other acoustical parameters*," International Standards Organization, Geneve.

[15] A. Järvinen, "*Kuunteluhuoneen suunnittelu ja mallinnus,*" M. Sc. Thesis, Helsinki University of Technology, Department of Electrical and Communications Engineering, Laboratory of Acoustics and Audio Signal Processing, 1999. (In Finnish).

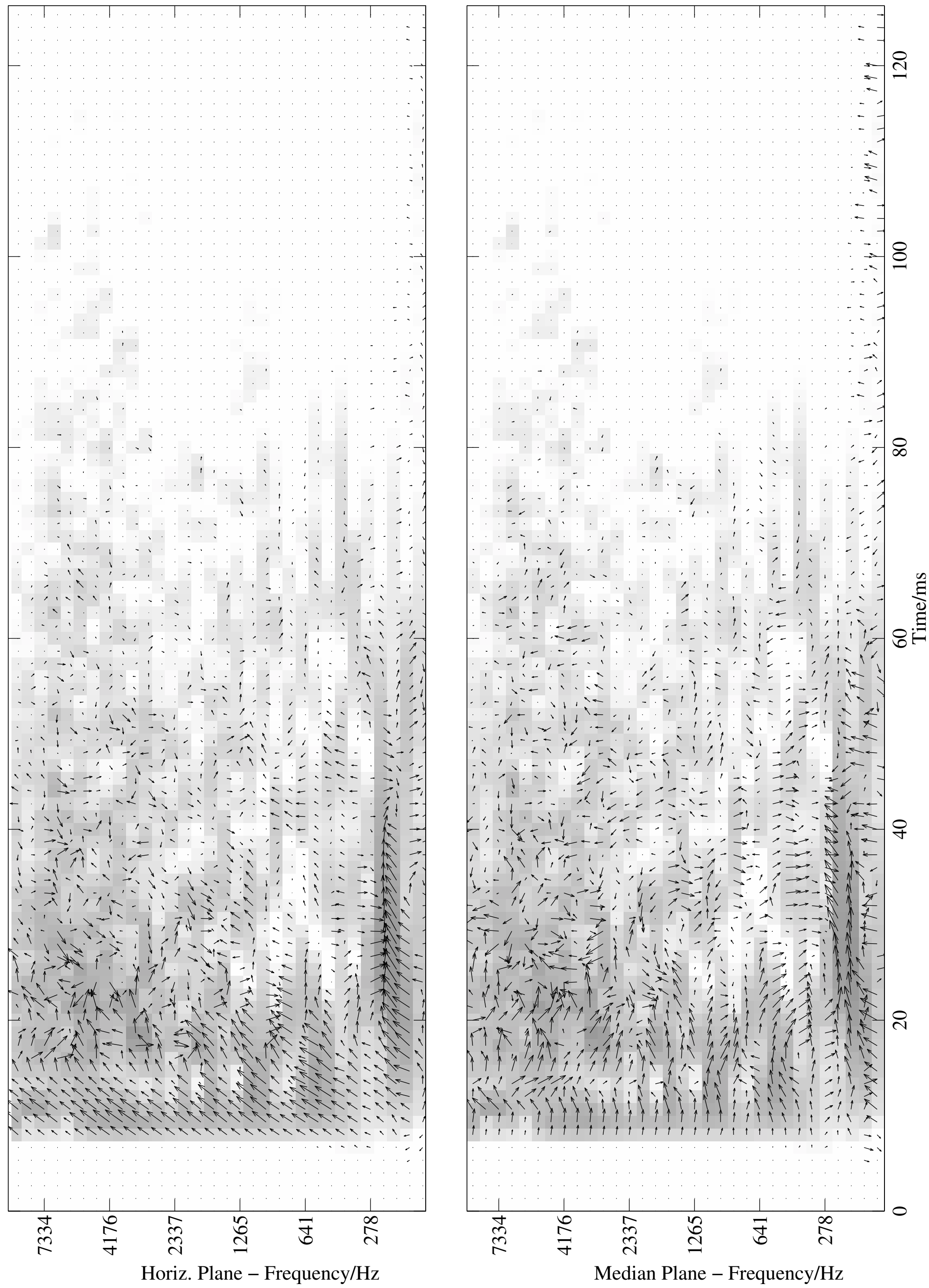


Fig. 7: *Directional room responses measured in the listening room of the HUT Acoustics laboratory, analyzed with an auditorily motivated time-frequency resolution.*

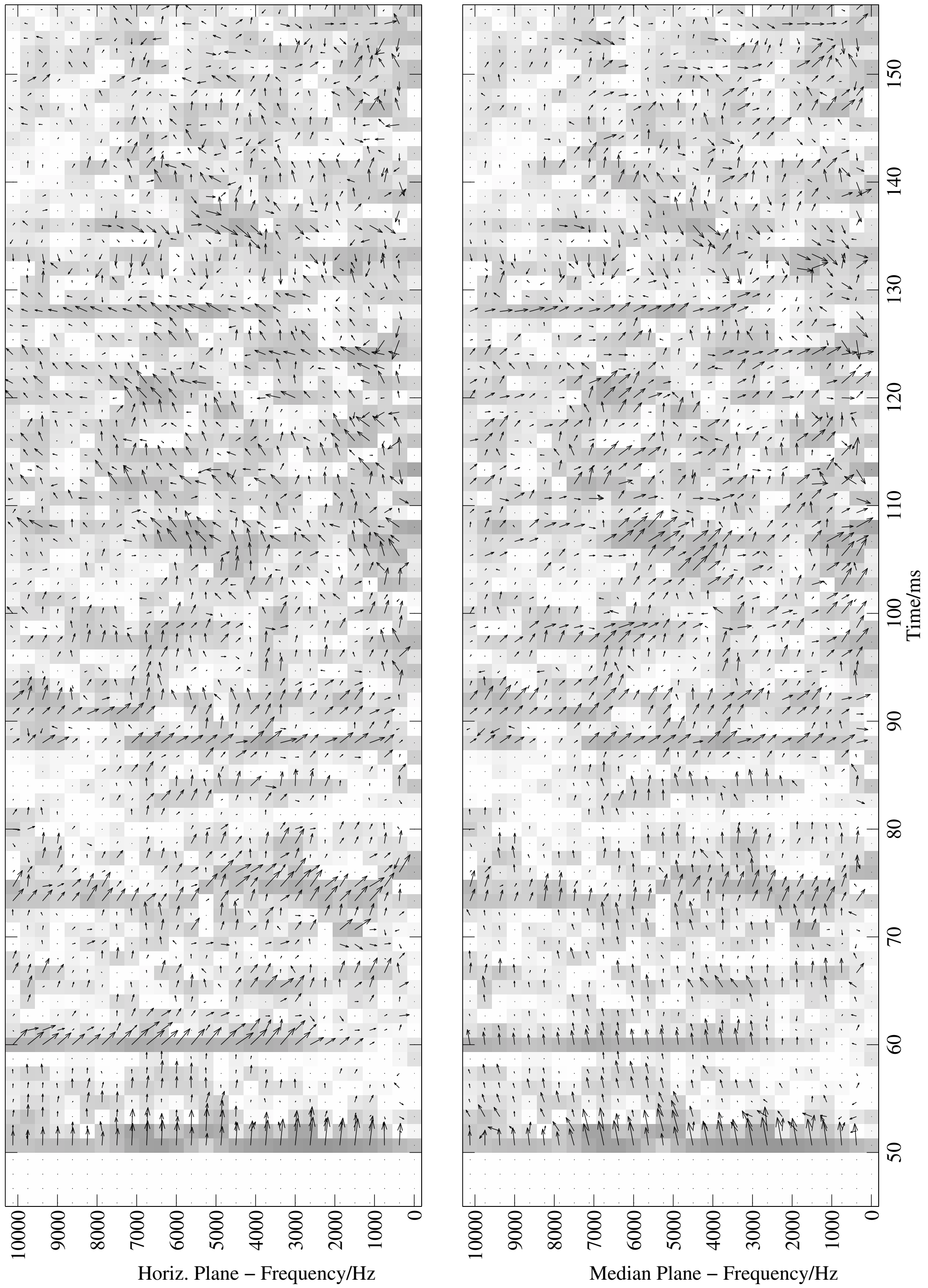


Fig. 8: *Directional room responses measured at the concert hall of Tampere-talo, analyzed with a uniform time-frequency resolution.*

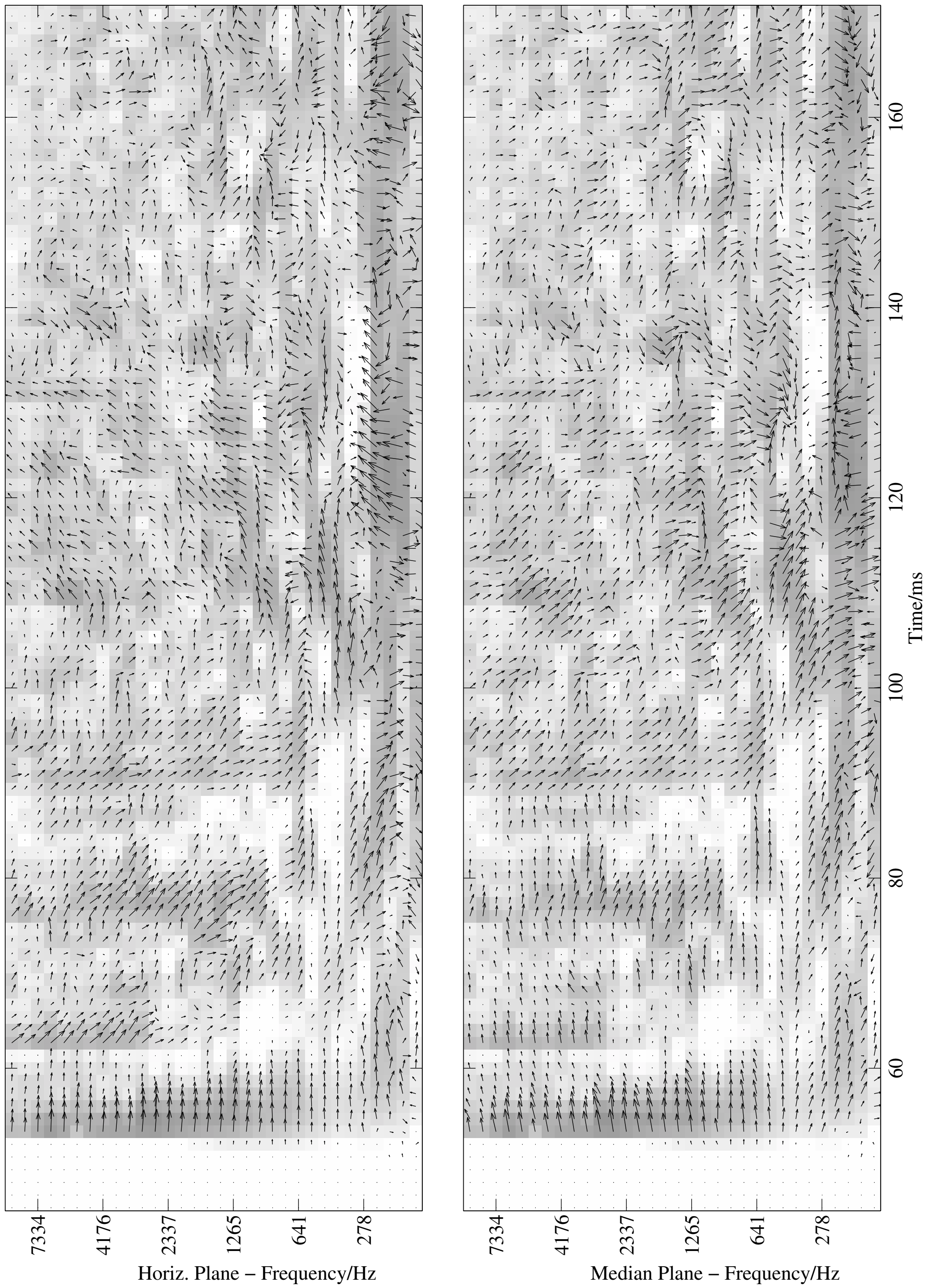


Fig. 9: *Directional room responses measured at the concert hall of Tampere-talo, analyzed with an auditorily motivated time-frequency resolution.*