Perception-based investigations on the monopole synthesis for reproduction of directional sound sources

Jakob Bergner[1,2], Johannes Redlich[3], Christoph Sladeczek[2], Sandra Brix[2]
[1] Technische Universität Ilmenau, 98693 Ilmenau, jakob.bergner@tu-ilmenau.de
[2] Fraunhofer Institute for Digital Media Technology IDMT, 98693 Ilmenau
[3] Technische Universität Berlin, 10623 Berlin

Abstract

The monopole synthesis provides an analytic approach to approximate the sound radiation of arbitrary sound sources by a distribution of monopole sources. This procedure allows to describe, represent and reproduce the sound radiation of sources with complex directivity by means of a set of simple monopoles. However, for an acceptable error estimation the physical requirements of both the underlying acoustical raw data (from measurements or simulations) and the distribution of the monopoles (quantity and positioning) are quite challenging when it comes to practical applications.

Therefore this work is motivated by the human auditory perception of directional sources. A processing chain is presented to execute the monopole synthesis for directional sources and subsequently conduct a comparison of physical approximation errors and perceptive artifacts. This can be used for formulating a qualitative statement about the feasibility of the monopole synthesis for the reproduction of directional sound sources. This is done by conducting the procedure: raw data acquisition, monopole decomposition, resynthesis, auralization, listening tests for the example of a highly directional loudspeaker. This work is part of and motivated by ongoing researches on plausible virtual acoustic environments.

Introduction

The virtualization of environments and processes has gained and is still gaining much attention and technological advance in various areas of application. Relevant fields with both academic and commercial purposes include *virtual reality (VR)*, *virtual engineering* and *computer-aided engineering (CAE)* as well as *digital product design*. After initially focusing on the mere visualization of products and environments, further influences and effects are being considered intensively lately. This includes the physical behavior of products due to certain environmental influences on the one side and the multimodal perception of virtual environments on the other side. The research on aspects of *virtual acoustic environments (VAE)* becomes relevant which can be organized in three parts [2]:

- sound source modeling
- description of and interaction with sound(ing) scenes
- integration in object-based reproduction system, e.g. binaural synthesis, wave field synthesis (WFS), higher-order ambisonics (HOA)
- psychophysical perception

This paper deals with a modeling approach of sound sources with direction-dependent sound radiation (complex directivity) which is motivated by above mentioned object-based audio reproduction systems. These systems provide well known possibilities of synthesizing simple sound sources due to their straightforward analytic description, namely plane waves and point sources. For auralization of sound sources with complex directivities new models have to be found. For doing this analytical approaches of the *spherical harmonic expansion* exist [1] [9] as well as practical simulations of the monopole synthesis [3] [8]. Latter approach is based on the *multiple radiator synthesis* or *equivalent source system*, defined in [6], [7] and is the main object of this paper.

The aim of this approach is to approximate the sound radiation of a complex structure by a finite distribution of equivalent weighted monopoles within the respective structure. The error results from spatial sampling of the monopole distribution, the finite number of equivalent sources and the according weights has to be minimized. This approach is discussed here by presenting a concept to oppose the resulting physical errors to the perceptual quality, derived by appropriate listening tests. It consists of a processing chain and identification of critical parameters influencing the aural perception. Special interest is applied to the auralization of directional sound sources, i.e. the presented approach requires a solution in time domain realized by a set of FIR-filters. The procedure is subsequently executed to reconstruct the directivity of a loudspeaker.

Monopole synthesis

The principle of the monopole synthesis represents the sound pressure $P(\mathbf{x_r}, \omega)$ at receiver positions $\mathbf{x_r} = [x_r, y_r, z_r]^T$ as superposition of M weighted distributed monopoles at the positions $\mathbf{x_m}$. This can be written as

$$P(\mathbf{x_r}, \omega) = \sum_{m=1}^{M} C_m(\omega) \cdot G(\mathbf{x_r} - \mathbf{x_m}, \omega) \qquad (1)$$

where ω denotes the angular frequency, $C_m(\omega)$ the weighting functions and $G(\mathbf{x_r} - \mathbf{x_m}, \omega)$ Green's free field

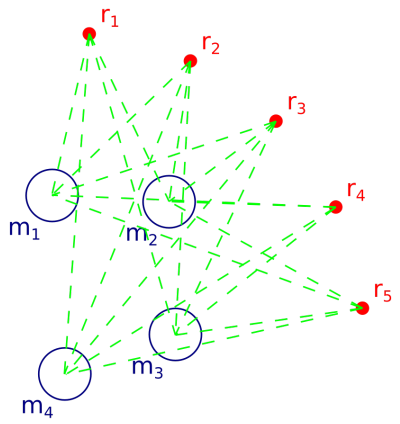

Figure 1: Principle of the monopole synthesis. Superposition of distributed monopoles (blue), described by the transfer paths (green) of the Green's function at specific receiver positions (red).

function

$$G(\mathbf{x_r} - \mathbf{x_m}, \omega) = \frac{1}{4\pi} \frac{\mathrm{e}^{-\mathrm{j}\frac{\omega}{c}|\mathbf{x_r}-\mathbf{x_m}|}}{|\mathbf{x_r} - \mathbf{x_m}|} \quad . \tag{2}$$

The underlying concept is depicted in Figure 1. To describe a directional sound source, the weights $C_m(\omega)$ must be determined for each monopole by solving the inverse problem of (1) with given $P(\mathbf{x_r}, \omega)$. To comprehend this problem, we can formulate the monopole synthesis (1) as matrix operation for each frequency

$$p = G \cdot c \quad . \tag{3}$$

To obtain the weighting vector c, the transfer path matrix G has to be inverted

$$G^{-1} \cdot p = c \quad . \tag{4}$$

Usually this latter operation is not feasible since the transfer path matrix is not square (overdetermined/underdetermined) or ill-conditioned. For that reason a regularization method must be applied to solve the minimization problem

$$\min \{||Gc - p||_n\} \tag{5}$$

where n denotes a suitable matrix norm. Since the overall aim is to auralize the reconstructed sound source, we need to find suitable time domain representations of the weights c. Hence, c must be restricted in both amplitude range and phase (or better group delay) which can be realized by regularizing the minimization (5) with additional constraints. The benefits of the regularization can deducted from Figures 2 and 3. An unbound minimization problem can lead to weights within an extreme magnitude range. For weights deducted by an exemplary Moore-Penrose pseudoinverse matrix in Figure 2 we find magnitudes within a range from -50 to 200 dB which is an evidence of an ill-conditioned problem [5]. With appropriate magnitudes constraints of a regularization method we can limit this range. Figure 3 shows weights

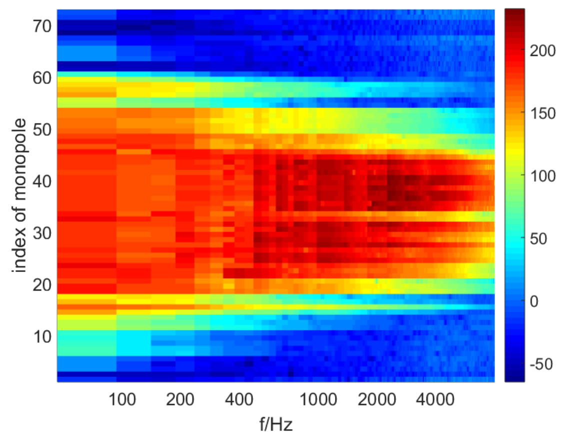

Figure 2: Magnitude $20\log_{10}|c|$ of exemplary frequency dependent weights derived by an unbound Moore-Penrose pseudoinverse matrix.

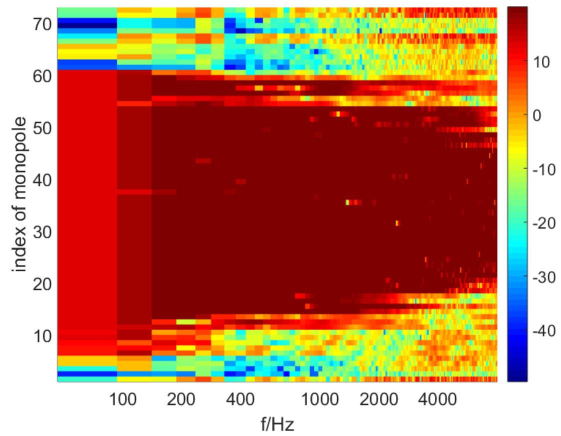

Figure 3: Magnitude $20\log_{10}|c|$ of exemplary frequency dependent weights derived by convex optimization subject to magnitude constraints.

for the same example but derived by convex optimization which is subject to certain constraints. The resulting magnitude range can be reduced to about 70 dB which is feasible for practical implementations.

Another way to find more suitable solutions of weights c is to form a better conditioned matrix G. To reduce linear dependencies within G, the transfer paths between monopole and receiver positions have to be analyzed and optimized. Since the latter are generally constant, this analysis is up to the distribution of monopoles. However the positioning of the monopole positions is subject to the spatial sampling theorem which states that the controllable frequency range depends on the smallest and largest distance between monopoles within a respective direction [4]. An exemplary monopole distribution can be found in Figure 4 where conical logarithmic spirals with respect to the three coordinate axes are depicted. The total number of monopoles M and the resulting spacing influences both the spatial sampling theorem and the conditioning of the transfer path matrix G.

So far the weights c of the monopoles exist in frequency domain. For auralization of the sound source we need

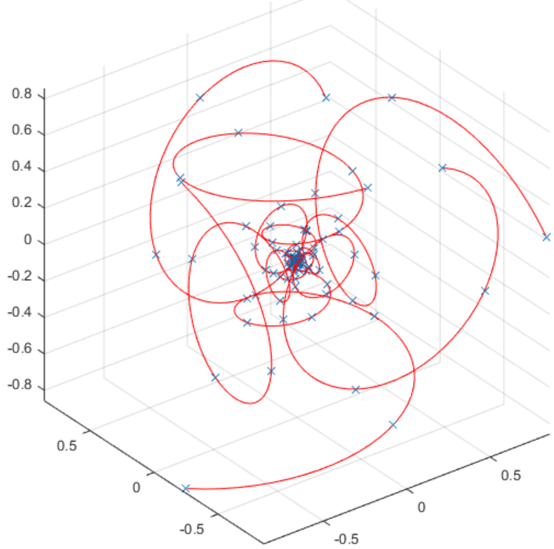

Figure 4: Exemplary distribution of monopoles (blue) alongside conical logarithmic spirals (red) oriented into the three respective coordinate axes.

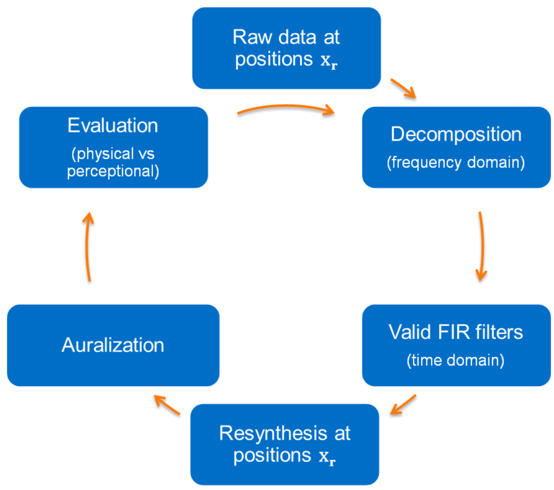

Figure 5: Processing chain of the monopole decomposition und subsequent synthesis of a measured directional sound source.

to find a valid way to transfer these weights into time domain, namely as a set of valid FIR filters h. It is not necessarily given that a straightforward inverse fft of c leads to valid, finite and causal FIR filters so a dedicated procedure might be required. This can be achieved by another minimization problem

$$\min\left\{||Fh - c||_n\right\} \tag{6}$$

where F denotes a Fourier transform matrix with elements $\mathrm{e}^{-\mathrm{j}\omega_k t_l}$.

Exemplary application

The previous explained strategies to decompose measured or simulated sound propagation of a sound source into a set of weighted monopoles is exemplary applied to a practical use case. For a general overview Figure 5 shows the processing chain to be conducted. The raw measurement data is obtained by a loudspeaker CA106 by *Kling&Freitag* whose directivity was measured on a 10° x 10° equiangular spherical grid. This example is a special case since the sound source under test can be described as an (almost) linear time-invariant system which can be defined by a set of directional impulse responses. The processing chain was applied to these measurements with frequency limits from 20 to 8000 Hz determining the monopole distribution and the minimization problem. The resulting sound field could be evaluated as deviation from the original data while comparing the frequency domain weights c and time domain filters h for different monopole distributions as shown in Figure 6. Three obvious findings can be withdrawn: (i) for frequencies below 1000 Hz the reconstruction works with deviations less than ±3 dB, above this frequency artifacts arise depending on the receiver position, (ii) the transform to the time domain introduces additional artifacts at frequencies above 4000 Hz and (iii) with an increasing

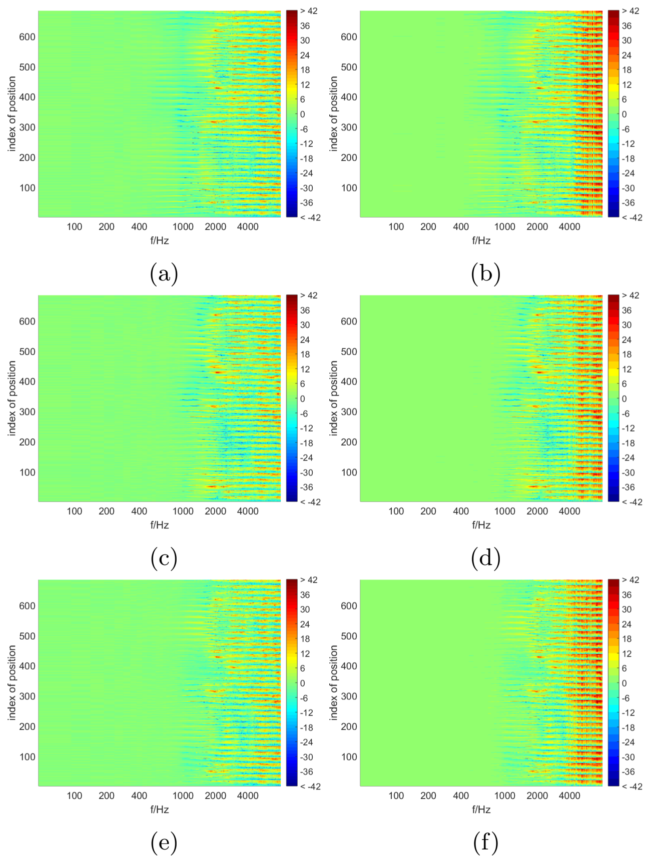

Figure 6: Deviation between reconstructed and original directivity $20\log_{10}(P_{\mathrm{recon}}(\mathbf{x_r},\omega)/P_{\mathrm{orig}}(\mathbf{x_r},\omega))$ for both frequency domain weights c (left) and time domain FIR filters h (right). Three monopole distributions are evaluated with $M = 49$ (top), $M = 73$ (middle) and $M = 97$ (bottom) monopoles.

number of monopoles M the main deviation contributions are shifted to higher frequency components. Since the data basis is too few at this point of the research, interpretation and cause analysis of the data is postponed to the future.

The final auralization of the reconstructed sound source was evaluated in an informal listening test. Differences between reconstructed and original directivities were observed however the distinction between the individual configurations of the monopole distribution were insignificant. More extensive perceptual investigations have to be done.

Conclusion

This work presented a processing chain for representing a directional sound source as a set of distributed, weighted monopoles motivated by the aim of a plausible auralization. The inverse problem of the monopole synthesis was discussed and relevant critical parameters were identified, namely

- regularization method of the inverse minimization problem Eq. (5)
- distribution of the equivalent monopoles (Figure 4)
- transformation of frequency domain weights to time domain filters Eq. (6)

We expect these parameters to be main influences on both physical and perceptual artifacts and thus essential to investigate individually. The presented conceptual work serves as basis for detailed further research. The following steps can be identified as

- developing systematic perceptual test routine,
- investigating the influence of the critical parameters individually,
- find a decision and calculation procedure for optimal sound source, decomposition
- extend and apply overall method to time-variant sound sources,
- conceptualize measurement methods for sound sources in non-anechoic environments.

Acknowledgment

This work was part of the project *Vista4F* funded by the *Thüringer Ministerium für Bildung Wissenschaft und Kultur (TMBWK)* under the "ProExzellenz" programme.

References

[1] J. Ahrens and S. Spors. Implementation of Directional Sources in Wave Field Synthesis. In *Applications of Signal Processing to Audio and Acoustics, 2007 IEEE Workshop on*, pages 66 –69, October 2007.

[2] J. Bergner, T. Clauß, A. Zhykhar, C. Sladeczek, and S. Brix. Application of wave field synthesis in virtual acoustic engineering. In *Proceedings of the 45th INTER-NOISE*, pages 3509–3516, Hamburg, Germany, March 2016.

[3] J. Escolano, J. J. López, and B. Pueo. Directive sources in acoustic discrete-time domain simulations based on directivity diagrams. *J. Acoust. Soc. Am.*, 121(6):EL256–62, June 2007.

[4] F. Giron. *Investigations About the Directivity of Sound Sources*. PhD thesis, Ruhr-Universität Bochum, 1996.

[5] P. C. Hansen. Regularization tools version 4.0 for matlab 7.3. *Numerical Algorithms*, 46(2):189–194, 2007.

[6] M. Ochmann. Die Multipolstrahlersynthese - ein effektives Verfahren zur Berechnung der Schallabstrahlung von schwingenden Strukturen beliebiger Oberflächengestalt. *Acustica*, 72(4):233–246, 1990.

[7] M. Ochmann. The source simulation technique for acoustic radiation problems. *Acustica*, 81:512–527, 1995.

[8] R. Opdan, K. Laztke, G. Behler, and M. Vorländer. Directivity as monopole decomposition for active noise reduction. In *42. Jahrestagung für Akustik, DAGA*, pages 461–463, Aachen, Germany, March 2016.

[9] A. Zhykhar, C. Sladeczek, and S. Brix. Synthesis of Directional Sound Sources with Complex Radiation Patterns Using a Planar Array of Loudspeakers. In *25th AES UK Conference / 4th Int. Symp. on Ambisonics and Sph. Acoustics*, York, UK, march 2012.