Calibration of Microphone Arrays with Arbitrary Geometries

Karim Helwani[1], Herbert Buchner[2], and Sascha Spors[1]

[1] *Quality and Usability Lab, Deutsche Telekom Laboratories,* [2] *Machine Learning Group, Technische Universität Berlin, 10587 Berlin, Germany*

Email:{karim.helwani, sascha.spors}@telekom.de, herbert.buchner@tu-berlin.de

Introduction

There is an increasing interest in extracting spatial information from acoustic sound fields, for instance for data-based rendering of virtual scenes. Analyzing an acoustic sound field with respect to suitable spatial components, e.g., plane waves, can be performed by evaluating signals captured by a number of spatially distinct microphones, i.e., microphone arrays. Typically, the signal acquisition using these arrays is achieved in two stages. The first stage is to compensate spatio-temporal deviations of the microphones. This stage is called calibration filter in this paper. The second stage is to apply beamforming filters which are specified depending on the geometry of the array. Beamforming can be understood as extraction of plane waves from the recorded sound field. High resolution acoustic scene analysis requires a high number of microphones. In this paper we present an integrated approach for non-adaptive determination of calibration and beamforming filters for arbitrarily shaped arrays.

State of the Art in 3-Dimensional Acoustic Scence Analysis

Theoretically, a pressure sensitive sphere will isotropically capture all plane waves constituting an acoustic sound field. Therefore, a sphere exhibits the ideal sensor shape for 3-dimensional acoustic scene analysis. In the next section we briefly review the functional principle of spherical microphone arrays.

Functionality Principles and Calibration of Spherical Microphone Arrays

It can be shown that the plane wave decomposition is equivalent to a spherical convolution [1]. The spherical convolution of a sound field with the impulse response of a sphere is diagonalized in the domain of spherical harmonics. Therefore, transforming the signals captured by a continuous distribution of microphones on the surface of a sphere into the spherical harmonics domain allows extracting plane wave coefficients by simple division. To adapt the spherical beamforming principle to the discrete case, the required continuous integrals for the transformation into the domain of the spherical harmonics are approximated by weighted summations, or quadratures [2]. In [3] a method was proposed for the computation of the quadratures based on minimizing the orthonormality error of the discrete spherical harmonics in a frequency selective manner. Thus, the calibration filters are given analytically. However, these methods assume a shape of an ideal sphere, accurate microphone positioning, and identical microphone characteristics. The usual calibration process aims at equalizing the individual microphones.

Novel Approach for Calibration Based on MIMO-Inverse Modeling

In the recent decades, techniques, such as wave field synthesis (WFS) and near-field compensated higher-order Ambisonics (NFC-HOA), were developed for reproducing specific sound fields, e.g., plane waves using circular or spherical loudspeaker arrays. Hence, we can use such systems to synthesize plane waves and measure the response of the microphone array on the plane wave excitation. The desired calibration filters aim at extracting decoupled plane waves with predefined angles of incidence. Hence, they can be modeled as an inverse of a MIMO system consisting of a loudspeaker array reproducing plane waves with predefined angles of incidence at the microphone array. Under the assumption of free field propagation and time invariance, a virtual loudspeaker array can be sequentially built up by rotating the microphone array around its principal axes and capturing the impulse responses from the loudspeaker to the microphones. We can assume that the filters reproducing the plane waves that we like to recover are collected in the MIMO system convolution matrix $\mathbf{D}$, as the synthesis operator for the loudspeaker signals, see Fig. 1. Hence, the condition for the inversion may be expressed in terms of an ideal overall system matrix $\mathbf{C}$, i.e.,

$$\mathbf{D}\cdot\mathbf{H}\cdot\mathbf{B} = \underbrace{\mathrm{Bdiag}\left\{[0,\dots,0,1,0,\dots,0]^{\mathrm{T}},\dots,[0,\dots,0,1,0,\dots,0]^{\mathrm{T}}\right\}}_{:=\mathbf{C}}.$$

Here, we denote the MIMO systems consisting of the loudspeaker and microphone array, and the desired calibration filters by the convolution matrices $\mathbf{H}$ and $\mathbf{B}$, respectively. The Bdiag$\{\cdot\}$ operator describes a block-diagonal matrix containing the listed vectors on the main diagonal. These target vectors represent pure modeling delays in the ideal case. This leads to the solution

$$\mathbf{B} = \mathbf{H}^{-1}\cdot\mathbf{D}^{-1}\cdot\mathbf{C}. \qquad (1)$$

Synthesis of Plane Waves with Circular Loudspeaker Arrays

Plane waves can be synthesized using WFS or NFC-HOA. In these techniques each loudspeaker is driven by

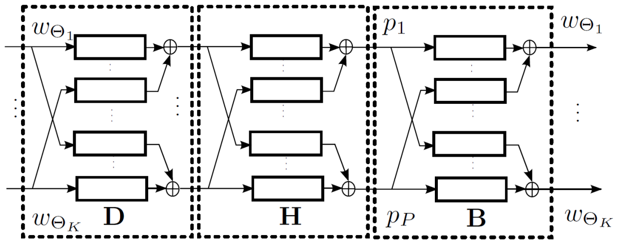

Figure 1: Block diagram of the proposed calibration approach. $w_{\Theta_{k=1...K}}$ denote plane waves with specified angles of incidence Θ, $p_{p=1...P}$ the pressure captured by a microphone of the array, Q is the number of the virtual loudspeakers, and P the number of microphones.

an individual driving function that can be analytically specified, depending on the array geometry [4], [5].
The frequency up to which an accurate synthesis is possible depends on the number of the loudspeakers of the virtual array which in turn influences the invertibility of the electro-acoustic MIMO system as we will discuss in the next section.
The sound field emitted by a loudspeakers can be approximated quite well by a plane wave for high frequencies and/or large distances. Hence, above the spatial aliasing frequency of the virtual loudspeaker array, plane waves can be approximated by single loudspeakers. Note, that applying the techniques of WFS or NFC-HOA on circular arrays offers the desired reproduction of plane waves only in the plane of the array. Moreover, an amplitude mismatch referred to as 2.5D synthesis [6] must be taken into account. Therefore, for relatively large microphone arrays better calibration results can be obtained, e.g., by building virtual line loudspeaker arrays by varying the height of the loudspeaker or virtual spherical loudspeaker arrays by rotating the microphone array around two axes.

Inversion of the Electro-Acoustic Transfer Function

Conditions for the invertibility of the electro-acoustic MIMO system can be obtained by considering the multiple-input/output inverse theorem (MINT) [7]. It states that the basic requirement for the MIMO system $\mathbf{H}$ in order to be invertible is that its convolution matrix $\mathbf{H}$ is of full rank. This condition corresponds to the requirement that the individual FIR acoustic impulse responses of the MIMO-system do not possess any common zeros in the z-domain. Another requirement for the invertibility of $\mathbf{H}$ is that the number of its rows equals the number of its columns.
An important conclusion of this consideration is that the MIMO mixing system can be inverted exactly even with a finite-length MIMO demixing system, as long as $P > Q$, i.e., the number of sensors is greater than the number of sources [7]. The resulting filter length $L_{\mathrm{opt,inv}}$ is given by the MINT in dependence of P and Q. Note, that P and Q must be chosen such that the resulting filter length $L_{\mathrm{opt,inv}}$ is an integer number in order to allow the matrix inversion.
In practice, a numerically stable approximate inverse model for $L < L_{\mathrm{opt,inv}}$ can be estimated by adaptive filtering techniques [8]. For our multichannel case, adaptive filtering techniques should be used that allow systematic spatio-temporal regularization, such as, transform-domain adaptive filtering [9].

Discussion and Conclusion

In this paper we presented an approach for calibrating microphone arrays with arbitrary geometry for beamforming. The computation of the calibration filters is based on the technique of inverse modeling of MIMO systems and exploits the techniques of synthesizing plane waves using virtual loudspeaker arrays. We have shown, based on MINT the theoretical limitations of using arbitrary shaped arrays for extracting plane waves. At low frequencies we synthesize plane waves using a virtual loudspeaker array, and for high frequencies we exploit the far field approximation of spherical waves as plane waves. The estimation of the calibration filter is proposed to be achieved in a spatio-temporal transform domain offering the ability of systematic regularization.

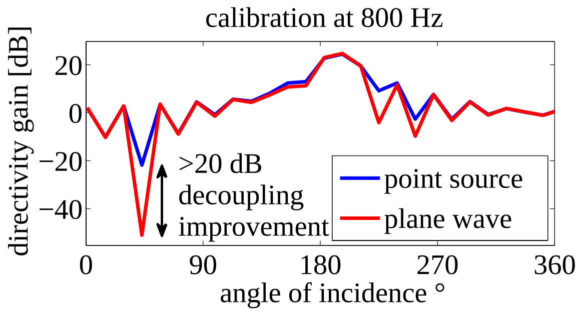

Figure 2: Directivity gain for plane waves arriving with an angle of incidence of 190° and a frequency of 800 Hz of a spherical array with 128 omnidirectional microphones randomly placed on a sphere surface and a radius of 7.5 cm . The red line is produced by calibrating the array using inversion of a MIMO system synthesizing plane waves.

The results in Fig. 2 show that especially at low frequencies, exploiting the prior knowledge regarding the synthesis of plane waves offers better sidelobe cancellation.

References

[1] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *JASA*, vol. 116, pp. 2149–2157, 2004.
[2] J. Meyer and G. Elko, "Spherical microphone arrays for 3D sound recording," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, pp. 67–89, 2004.
[3] Z. Li and R. Duraiswami, "Flexible and optimal design of spherical microphone arrays for beamforming," *Audio, Speech, and Language Processing, IEEE Trans. on*, vol. 15, pp. 702–714, Feb. 2007.
[4] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acustica*, vol. 94, no. 6, p. 988–999, 2008.
[5] A. Berkhout, D. De Vries, and P. Vogel, "Acoustic control by wave field synthesis," *JASA*, vol. 93, pp. 2764–2778, 1993.
[6] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," *Audio Eng. Soc. Conv. Paper*, vol. 124, 2008.
[7] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *Acoustics, Speech and Signal Processing, IEEE Trans. on*, vol. 36, pp. 145–152, Feb. 1988.
[8] S. Haykin, *Adaptive filter theory*. Prentice Hall, Inc., 1991.
[9] S. Spors, H. Buchner, and K. Helwani, "Block-based multichannel transform-domain adaptive filtering," in *(EUSIPCO)*, Aug. 2009.