IMPLEMENTATION OF A SUPERDIRECTIVE BEAMFORMER BASED ON AN EXTENDED MODAL SUBSPACE DECOMPOSITION

Martin Eichler and Arild Lacroix

Institute of Applied Physics, J. W. Goethe University, Frankfurt am Main
eichler@iap.uni-frankfurt.de

Abstract: Microphone arrays in combination with a suitable beamforming algorithm provide means to spatially filter signals from a given sound field. One very general beamforming approach is the method of Modal Subspace Decomposition (MSD) which regards the array and the beamformer as an operator applied to a coefficient vector. Based on this operator, an eigenvalue problem is considered the solution of which yields a set of eigen-beampatterns which serve as a basis for the series expansion of the desired beampattern (directivity). From the series expansion coefficients, the optimum filter coefficients for the beamformer filters are calculated. In an earlier work, an extension to MSD was presented which uses a certain class of IIR filters instead of FIR filters, thus making the algorithm superdirective. Hence, even with very compact arrays of arbitrary geometry, it is possible to achieve directivity at low frequencies where the wavelength is large in comparison to the array size. In this contribution, the underlying theory of this approach is outlined and measurement results obtained from a real-time implementation in a real acoustic environment are discussed.

1 Introduction

The underlying concept of modal subspace decomposition beamforming (MSD) is based on a filter-and-sum beamformer. In its original form, the beampattern B (directivity) is established by weighting and summing a given number of sound field samples taken at unique positions in space and relative time [1]. All weighting coefficients are arranged into a coefficient vector $\underline{b}$ and the connection between coefficient vector and beampattern is established by an operator $\hat{A}$. A given desired beampattern can be approximated by a linear combination of the so-called eigen-beampatterns of the beamformer, which are eigenfunctions of $\hat{A}\hat{A}^{\dagger}$. The spatial arrangement of the microphones is arbitrary; also, two- and three-dimensional array geometries are covered as well. This general approach also includes the FIR-filter-and-sum beamformer, which is obtained by taking M samples per microphone, each delayed by a time interval $\Delta t = m \cdot T_S$ with $m = 0, \ldots, M-1$. The discrete-time digital weighting filter of each microphone is then an FIR filter of $(M-1)$th order operating at sample rate $f_S = 1/T_S$. However, in spite of its above versatility and flexibility, this approach fails to achieve directivity with compact sensor arrays at low frequencies [2]. Raising the filter order improves the behaviour, yet mainly the numeric expense is increased while true superdirectivity is far from being obtained. Thus, an extension of the known MSD algorithm has been developed which successfully combines superdirectivity and the versatile operator approach [2].

2 Superdirective MSD Beamforming

Let us consider a filter-and-sum beamformer having N sensors located at the positions $\underline{r}_n$ ($n = 0, \ldots, N-1$). The microphones are weighted through individual weighting filters described by

complex-valued transfer functions $H_n(\omega)$ depending on the frequency f (with $\omega = 2\pi f$). We now assume that the sound field consists of planar waves having the wave vectors $\underline{k} \in K$, with $K \subset \mathfrak{K}$ being the considered *finite* range from the wave number vector space $\mathfrak{K}$. Given a source amplitude distribution $S(\underline{k})$, the sound pressure at the location $\underline{r}$ and time t can be written as

$$p(\underline{r}, t) \;=\; \iint_K S(\underline{k}) \cdot e^{i(\omega(\underline{k})t - \underline{k}\cdot\underline{r})}\, d^2k,$$

where $\omega(\underline{k}) = kc$ with $k := \|\underline{k}\|$ and c being the sound velocity. The output signal $y(t)$ of the filter-and-sum beamformer can now be expressed as

$$\begin{aligned} y(t) \;&=\; \iint_K \sum_{n=0}^{N-1} S(\underline{k}) \cdot H_n(\omega(\underline{k})) \cdot e^{i(\omega(\underline{k})t - \underline{k}\cdot\underline{r}_n)}\, d^2k \\ &=\; \iint_K S(\underline{k}) \cdot B(\underline{k}) \cdot e^{i\omega(\underline{k})t}\, d^2k, \end{aligned}$$

where

$$B(\underline{k}) \;=\; \sum_{n=0}^{N-1} H_n(kc) \cdot e^{-i\underline{k}\cdot\underline{r}_n} \tag{1}$$

is the so-called *beampattern function* describing the beamformer sensitivity as a function of angle and frequency. The above expressions consider the case of two spatial dimensions (2-D). The extension to three spatial dimensions (3-D) can be done by simply changing all double integrals to triple ones ($\iint \ldots d^2k \rightarrow \iiint \ldots d^3k$).

We will now introduce the *explicit form* of the weighting filters through which superdirectivity is achieved. In the original MSD approach [1], the weighting filters (in case of fixed sensor positions) have the form

$$H_n(\omega) \;=\; \sum_{m=0}^{M-1} h_{n,m} \cdot e^{-i\omega\Delta t_{n,m}},$$

which are digital FIR filters for $\Delta t_{n,m} = m \cdot T_S$ as stated above. The beamforming algorithm then calculates the coefficients $h_{n,m}$, thus creating an optimal linear combination of the individual microphone signals and their delayed versions, which approximates the given desired beampattern $B_{desired}(\underline{k})$ in an optimal way [1]. However, in order to achieve superdirectivity, we choose a fundamentally different form for the weighting filters, motivated by the following considerations:

1. The superdirectivity we aim for is an intrinsic property of multipole arrays and differential arrays [3, 4]. Their approximately frequency-invariant behaviour at low frequencies f (i.e. for $f \rightarrow 0$) is achieved through a magnitude compensation by means of *time-integration filters*: For a multipole of μ-th order, a μ-fold integration must be applied [3].

2. Combining a set of multipoles of different orders that share a common set of microphones eventually results in a mixture or *linear combination* of multiple integrated versions (of different integration order) of each microphone signal.

3. Maximum flexibility of the beamformer can be achieved by introducing individual coefficients for *all* integration orders of *each* microphone signal separately (in contrast to multipole beamforming where fixed linear combinations of sensors are used [3]).

The above points lead to the following weighting filter functions where *integration over time* expressed by the integration filter function $1/i\omega$ is used instead of time delay:

$$H_n(\omega) \;=\; \sum_{l=0}^{L-1} j_{n,l} \cdot \left(\frac{ac}{i\omega}\right)^l. \tag{2}$$

Here, $L-1$ is the *maximum integration order*; the coefficients $j_{n,l}$ describe the weighting of the l-fold integrated version of the signal from the nth microphone (with $n = 0, \ldots, N-1$ and $l = 0, \ldots, L-1$). a is a geometry factor having the unit m^{-1} which is introduced in order to make the term $(ac/i\omega)$ dimensionless. With (2) and $\omega = kc$, the beampattern (1) takes the form

$$B(\underline{k}) \;=\; \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} j_{n,l} \cdot \left(\frac{a}{ik}\right)^l \cdot e^{-i\underline{k}\cdot\underline{r}_n}, \tag{3}$$

comprising $V = NL$ summands in total. We now define the following vectors of length V:

$$\begin{aligned} \underline{w} &= (w_0, \ldots, w_{V-1})^T, \\ \hat{A}(\underline{k}) &= (A_0(\underline{k}), \ldots, A_{V-1}(\underline{k})) \end{aligned} \tag{4}$$

with ^{T} denoting the transpose, having the vector elements

$$\begin{aligned} w_{\nu=nL+l} &= j_{n,l}, \\ A_{\nu=nL+l}(\underline{k}) &= \left(\frac{a}{ik}\right)^l \cdot e^{-i\underline{k}\cdot\underline{r}_n}. \end{aligned} \tag{5}$$

These definitions allow us to rewrite the beampattern (3) in the more compact form

$$B(\underline{k}) \;=\; \sum_{\nu=0}^{V-1} A_\nu(\underline{k}) \cdot w_\nu \;=\; \hat{A}\,\underline{w}. \tag{6}$$

The beampattern shape is entirely controlled by the coefficient vector $\underline{w}$. The vector $\hat{A}$ is called the *beamforming operator* because it maps the coefficient vector $\underline{w}$ to the beampattern B. Fig. 1 shows the signal flow diagram of the so-defined beamformer.

2.1 Matrix Elements and Eigen-Beampatterns

The MSD method uses an eigenvalue problem to obtain the filter coefficients w_ν. We consider the V-dimensional complex vector space $\mathfrak{V} = \mathbb{C}^V$ which is spanned by all possible coefficient vectors. Similarly, we define $\mathfrak{F}$ as the function space of all beampatterns $B(\underline{k})$. It consists of two subspaces: If we define $\mathfrak{B} := \{B \,|\, \exists \underline{b} \in \mathfrak{V} : \hat{A}\underline{b} \equiv B\}$ as the function space of all beampatterns that can be achieved or created according to (6), and $\overline{\mathfrak{B}} := \{B \,|\, \hat{A}\underline{b} \not\equiv B \;\; \forall \underline{b} \in \mathfrak{V}\}$ as the function space of those which cannot, then we can write $\mathfrak{F} = \mathfrak{B} \cup \overline{\mathfrak{B}}$. The operator $\hat{A}$ can now be regarded as a linear map $\hat{A} : \mathfrak{V} \to \mathfrak{F}$ mapping each vector $\underline{w} \in \mathfrak{V}$ to a beampattern $B \in \mathfrak{B}$. For $\underline{w}, \underline{w}' \in \mathfrak{V}$ and $B, B' \in \mathfrak{F}$, we define the inner products

$$\langle \underline{w} | \underline{w}' \rangle_{\mathfrak{V}} \;=\; \sum_{\nu=0}^{V-1} w_\nu \cdot {w'_\nu}^*, \tag{7}$$

$$\langle B(\underline{k}) | B'(\underline{k}) \rangle_{\mathfrak{F}} \;=\; \iint_K B(\underline{k}) \cdot B'^*(\underline{k})\; d^2k \tag{8}$$

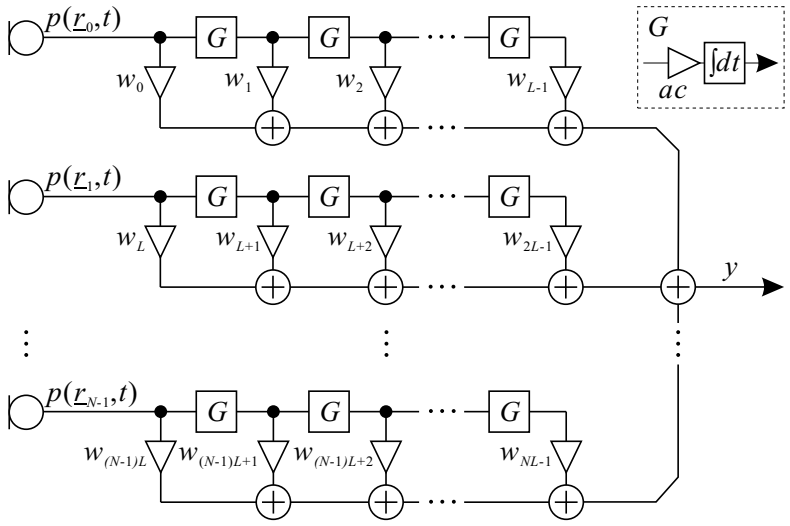

Figure 1 - Signal flow diagram of the proposed beamformer. Each of the N sensor signals is integrated 0 to $L-1$ times by means of the filters $G(\omega) = ac/i\omega$.

where * denotes the complex conjugate. Now, an *adjoint operator* $\hat{A}^\dagger : \mathfrak{F} \to \mathfrak{V}$ can be found that maps a beampattern B to a vector $\underline{w} = \hat{A}^\dagger B$ and satisfies the relation $\langle \hat{A}^\dagger B | \underline{w}' \rangle_\mathfrak{V} = \langle B | \hat{A} \underline{w}' \rangle_\mathfrak{F}$ for any $B \in \mathfrak{F}$ and any $\underline{w}' \in \mathfrak{V}$ [2]. Its explicit form can be derived from said relation using the definitions (7) and (8). One obtains

$$\hat{A}^\dagger = \begin{pmatrix} \iint_K d^2k \; A_0^*(\underline{k}) \cdot \square \\ \vdots \\ \iint_K d^2k \; A_{V-1}^*(\underline{k}) \cdot \square \end{pmatrix}, \tag{9}$$

where the placeholders $\square$ indicate that the integrands of $\iint_K d^2k$ must contain the respective operand which is multiplied from the right side. Now we consider $\boldsymbol{Z} = \hat{A}^\dagger \hat{A}$. From (4) and (9), it follows that $\boldsymbol{Z}$ is a $V \times V$ matrix having the matrix elements $Z_{\nu\nu'} = \iint_K A_\nu^* \cdot A_{\nu'} \, d^2k$ (with $\nu, \nu' = 0 \ldots V-1$). For evaluation of the matrix elements, the integration interval K must be defined. Since waves from all directions must be considered, it is reasonable to allow wave vectors $\underline{k}$ of all angles, with the wave number $k = \|\underline{k}\|$ inside a given range: $k \in [k_1; k_2]$. Thus, K is a ring-shaped area around the origin of $\mathfrak{K}$ in the 2-D case, and a spherical shell in the 3-D case, in each case bounded by the inner radius k_1 and the outer radius k_2. The matrix elements can now be evaluated by explicit angular integration in polar coordinates and spherical coordinates respectively: For 2-D, one obtains (with $n, n' = 0, \ldots, N-1$ and $l, l' = 0, \ldots, L-1$)

$$Z^{\text{2-D}}_{\nu=nL+l,\; \nu'=n'L+l'} \;=\; 2\pi(-1)^l \int_{k_1}^{k_2} \left(\frac{a}{ik}\right)^{l+l'} \mathrm{J}_0(k\|\underline{r}_n - \underline{r}_{n'}\|)\, k \, dk, \tag{10}$$

where $\mathrm{J}_0(x)$ is the 0^{th}-order Bessel function of the first kind. For 3-D, the matrix elements evaluate to

$$Z^{\text{3-D}}_{\nu=nL+l,\; \nu'=n'L+l'} \;=\; 4\pi(-1)^l \int_{k_1}^{k_2} \left(\frac{a}{ik}\right)^{l+l'} \mathrm{si}(k\|\underline{r}_n - \underline{r}_{n'}\|)\, k^2 \, dk, \tag{11}$$

$\mathrm{si}(x) = \sin(x)/x$ being the unnormalized sinc function [2]. It can be seen from (10) and (11) that $\boldsymbol{Z}$ is self-adjoint, i.e. $Z_{\nu\nu'} = Z^*_{\nu'\nu}$. Thus, its eigenvalues λ_ν are real and its normalized eigenvectors $\underline{u}_\nu$ form a complete orthonormal base of $\mathfrak{V}$, such that $\langle \underline{u}_\nu | \underline{u}_{\nu'} \rangle_{\mathfrak{V}} = \delta_{\nu\nu'}$ holds (δ denoting the Kronecker delta). Further, the functions $U_\nu(\underline{k}) = \hat{A}\underline{u}_\nu/\sqrt{\lambda_\nu}$ are eigenfunctions of the operator $\hat{A}\hat{A}^\dagger$ (i.e. $\hat{A}\hat{A}^\dagger U_\nu = \lambda_\nu U_\nu$) and form an orthonormal base of $\mathfrak{B}$, satisfying the orthogonality relation $\langle U_\nu | U_{\nu'} \rangle_{\mathfrak{F}} = \delta_{\nu\nu'}$. They are also called *eigen-beampatterns* and are orthogonal with respect to the given wave number interval $[k_1; k_2]$ which corresponds to a frequency interval $[f_1; f_2]$ with $f_{1,2} = k_{1,2} \cdot c/2\pi$. This interval is the *design frequency band* for which the set of eigen-beampatterns is valid.

2.2 Algorithm

A given desired beam pattern $B_{desired}$ can now be projected to $\mathfrak{B}$ and be approximated by a linear combination of the eigen-beampattern:

$$\begin{aligned} B_{approx}(\underline{k}) &= \hat{A}\,\underline{b}\,, \\ \underline{b} &= \sum_{\nu=0}^{V-1} \frac{1}{\sqrt{\lambda_\nu}} \cdot \langle B_{desired}(\underline{k}) | U_\nu(\underline{k}) \rangle_{\mathfrak{F}} \cdot \underline{u}_\nu. \end{aligned} \tag{12}$$

B_{approx} is an approximation if $B_{desired} \in \overline{\mathfrak{B}}$. Otherwise ($B_{desired} \in \mathfrak{B}$), it holds that $B_{approx} \equiv B_{desired}$. In the first case, the approximation is optimal in the sense that the mean square error is minimal. The beamforming algorithm can be summarized as follows:

1. For a given number N of sensors at the locations $\underline{r}_n$, calculate the matrix elements according to (10) or (11). The value of a must be decided beforehand.
2. Calculate the normalized eigenvectors $\underline{u}_\nu$ and eigenvalues λ_ν of $\boldsymbol{Z}$ and evaluate all eigen-beampatterns $U_\nu(\underline{k})$.
3. Given a desired beampattern $B_{desired}$, calculate the coefficient vector $\underline{b}$ according to (12).

Steps 1. and 2. are numerically expensive, but have to be carried out only once. Only step 3. needs to be repeated when the desired beampattern is changed. In step 2., one should confirm that the eigen-beampatterns actually satisfy the orthogonality relation $\langle U_\nu | U_{\nu'} \rangle_{\mathfrak{F}} = \delta_{\nu\nu'}$ with the deployed numerical approximation of the integral over K. Any eigen-beampattern violating this relation should be excluded from the series expansion (i.e. from the sum in (12)). Typically, this can be the case for some of the eigen-beampatterns corresponding to the smallest eigenvalues. Whether or not eigen-beampatterns are excluded also affects the choice of a: It can be shown analytically that the value of a is arbitrary and does not influence the resulting beam pattern B_{approx} iff *all* eigen-beampatterns can be used. This is because the beamformer then targets the entire function space $\mathfrak{B}$. Otherwise, only a subspace of $\mathfrak{B}$ is covered which depends on the choice of a. In this case, a does have an effect on the beampattern obtained.

2.3 Multi-Band Design for Broadband Beamforming

The fact that the eigen-beampatterns of the beamformer are calculated for a given design band $[f_1; f_2]$ is an interesting property of the MSD algorithm, because the coefficients obtained are thus optimized for that frequency range. This property can be advantageous in broadband beamforming: Let us assume that the desired bandwidth is very large and the design band $[f_1; f_2]$ is chosen such that it covers it completely. It may now happen that the eigen-beampatterns calculated with respect to this band are not flexible enough to provide a satisfactory realization of

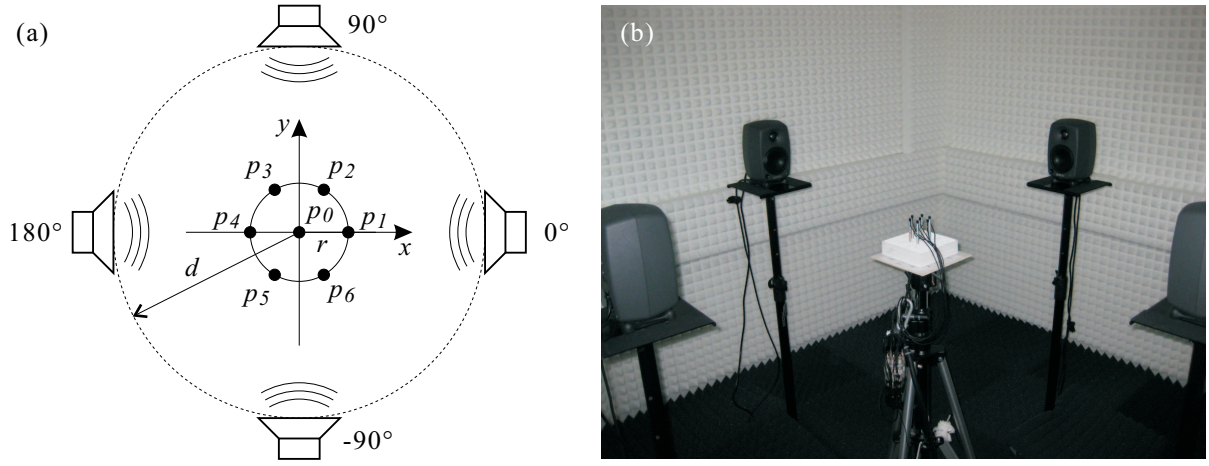

Figure 2 - Measurement setup with seven-sensor hexagonal microphone array and four sound sources at $0°$, $\pm 90°$ and $180°$; (a) scheme (not to scale), (b) actual setup. The radius of the array is $r = 3\,\text{cm}$, the distance between the sound sources and the array center is $d \approx 75\,\text{cm}$.

the desired beam pattern. In such a case, one may split the desired bandwidth into two or more subbands and use multiple beamformers, one for each band, and merge their output signals into one using suitable bandpass filters.

2.4 Scaling Property

Another interesting property of the beamforming operator (4) arises due to the geometry factor a: Imagine we have a coefficient vector $\underline{b}$ creating the beampattern $B = \hat{A}\,\underline{b}$ with sensors at the positions $\underline{r}_n$ and $\hat{A}$ using a geometry factor a having a given value (see eq. (5)). It is now easy to show that for the beampattern B' of an array enlarged by a factor $\rho \in \mathbb{R}^+$, having its sensors at the positions $\underline{r}'_n = \rho \cdot \underline{r}_n$, it holds that $B'(\underline{k}) \equiv B(\rho \cdot \underline{k})$ if $B' = \hat{A}'\,\underline{b}$ with the operator $\hat{A}'$ using the geometry factor $a' = a/\rho$ instead of a. This means that enlarging the array by a factor ρ results in scaling its beampattern by a factor $1/\rho$ along the frequency axis if the beamforming operator is adapted accordingly and the same coefficient vector is used. It can also be shown that for a desired beampattern $B'_{desired}(\underline{k}) \equiv B_{desired}(\rho \cdot \underline{k})$ the same coefficients $\underline{b}$ are found if the design frequency band is scaled as well ($[f'_1; f'_2] = [f_1/\rho; f_2/\rho]$).

3 Beamformer Implementation & Example Measurements

In order to implement the integration over time directly through the factor $1/i\omega$, and in order to avoid complex-valued signal processing in time domain, a frequency-domain implementation was chosen for the beamformer. Also, as the algorithm purely relies on phase information and assumes input signals of identical level (using $e^{-i\underline{k}\cdot\underline{r}}$ terms for plane waves), an FFT equalizer was added to each input channel in order to compensate the individual characteristics of the microphones which can be run in adaptive or static mode (keeping the magnitude corrections learned previously in adaptive mode). The example array geometry chosen comprises seven sensors ($N = 7$) in a hexagonal alignment with one center sensor (fig. 2(a)). The maximum integration order used was 3 (i.e. $L = 4$). The setup was realized with an array diameter of $6\,\text{cm}$, using microphones of type *Beyerdynamic MCE 60* (diameter: $\approx 7\,\text{mm}$). The array was exposed to sound fields generated by four louspeakers of type *GENELEC 8080A* (fig. 2(b)). By

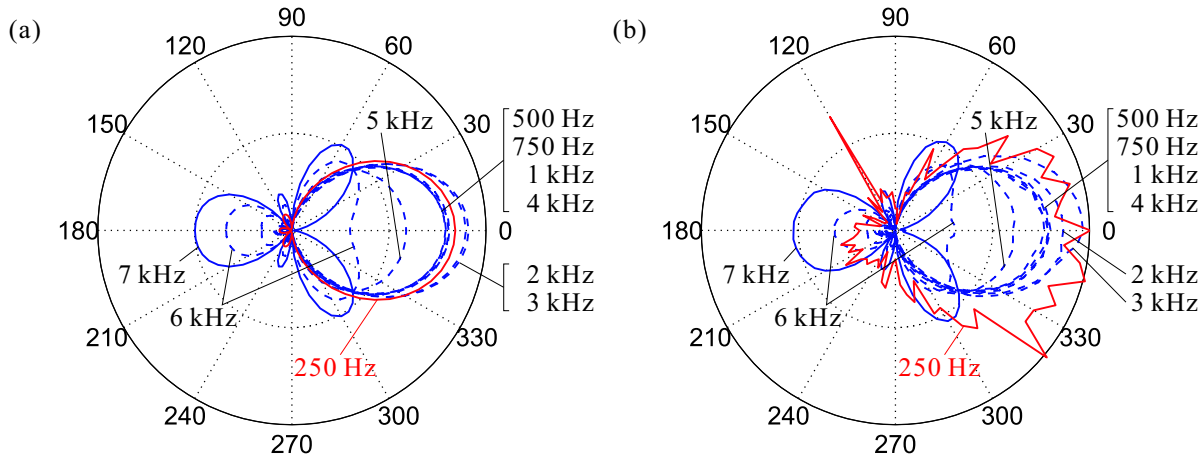

Figure 3 - Angular response of the example microphone array over $360°$ for the given desired beampattern (see text); (a) theoretical curves, (b) measured curves averaged over frequency bands 300 Hz wide. The indicated frequency values are the band center frequencies.

applying white noise from only one loadspeaker and rotating the array in $5°$ steps, the frequency response at each angle was measured; experiments to distinguish male and female speech played from two or four loudspeakers simultaneously were made as well. The desired beampattern used for the beamformer design was of frequency-invariant, triangular shape being 1 at $0°$, decreasing linearly to zero on both sides of $0°$, reaching zero at $\pm 79°$ and remaining zero for larger angles. Two beamformer designs were tested: (A) a single-band design integrating over the band $[25\,\text{Hz}; 4\,\text{kHz}]$, (B) a multi-band design using the three bands $[25\,\text{Hz}; 2\,\text{kHz}]$, $[2\,\text{kHz}; 4\,\text{kHz}]$ and $[4\,\text{kHz}; 6\,\text{kHz}]$. Design (B) almost fully exploits the bandwidth which can be used with the given geometry without aliasing effects growing too strong.

Fig. 3 shows the beampattern of the beamformer in case of design (A). The spectral data gathered by rotating the array was averaged over bands 300 Hz wide. It can be seen that the theoretical behaviour is matched very well with the exception of the band centered at 250 Hz. This is very likely due to the behaviour of the measurement chamber which exhibits resonances around 100 Hz. Fig. 4 compares the theoretical and the measured frequency response of the beamformer at selected angles ($0°$, $\pm 90°$, $180°$) for both design (A) and (B). The theoretical plots ((a)+(c)) were obtained by simulating ideal free-field conditions using fractional delay filters (FD-filters, [5]) and applying white noise. They exactly match the results obtained by direct evaluation of the algorithm in MATLAB, confirming that the beamformer implementation works properly. Hence, the differences between simulated and measured frequency responses must be caused by the physical measurement conditions such as room resonances, microphone alignment deviations, finite microphone size etc. Further, in both designs (A) and (B), it turned out that a couple of eigen-beampatterns with the smallest eigenvalues – though being orthogonal (see section 2.2) – may corrupt the beamformer behaviour. Thus, for (A) the first five, for (B) the first ten eigen-beampatterns (out of a total of $NL = 28$) had to be excluded.

4 Summary

A beamforming algorithm is proposed that achieves directivity down to very low frequencies and works for arbitrary array geometries. The superdirectivity is achieved by time-integration

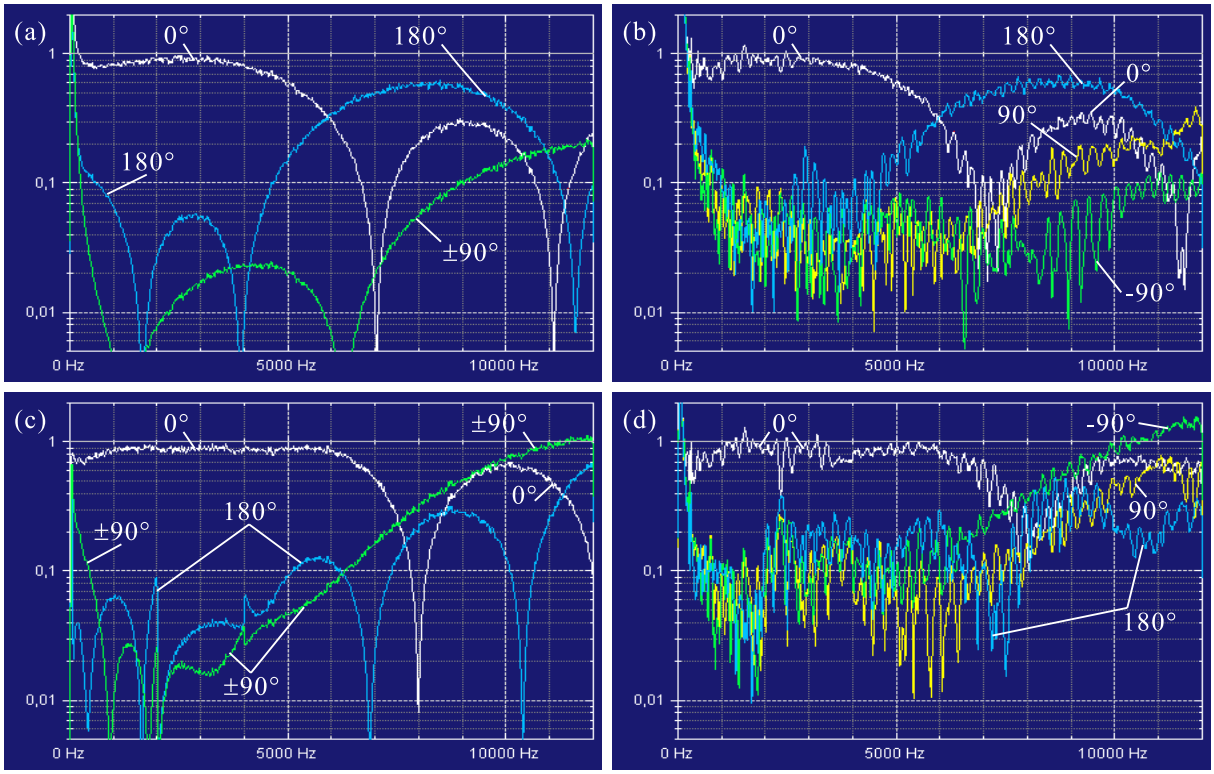

Figure 4 - Magnitude of the beamformer frequency response in selected directions (0°, ±90° and 180°); (a)+(b) single-band design, (c)+(d) three-band design; (a)+(c) theoretical curves (from simulation, see text), (b)+(d) measured curves.

filtering in analogy to multipole beamforming. Though the experiment setup to some extent compromises the shown measurement results (especially at low frequencies), it is shown that the algorithm itself is effective. Also, multi-band design can be used successfully to broaden the beamforming bandwidth.

References

[1] M. I. Y. Williams, T. D. Abhayapala, R. A. Kennedy, "Generalized Broadband Beamforming Using a Modal Subspace Decomposition", *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 68291, 9 pages, doi:10.1155/2007/68291, 2007.

[2] M. Eichler, A. Lacroix, „Superdirective Beamforming Using an Extended Modal Subspace Decomposition", *Proceedings NAG/DAGA 2009*, ISBN 978-3-9808659-6-8, pp. 133–136, Rotterdam, Netherlands, March 2009.

[3] M. Eichler, A. Lacroix, „Broadband Superdirective Beamforming Using Multipole Superposition", *in Proc. EUSIPCO 2008*, CD-ROM, Lausanne, Swiss, Aug. 2008.

[4] M. Ihle, "Differential Microphone Arrays for Spectral Subtraction", *Eighth International Workshop on Acoustic Echo and Noise Control*, *IWAENC 2003*, Kyoto, Japan, pp. 259–262, Sept. 2003.

[5] M. Eichler, A. Lacroix, "Maximally Flat FIR and IIR Fractional Delay Filters With Expanded Bandwidth", *in Proc. EUSIPCO 2007*, Poznań, Poland, pp.1038–1042, Sept. 2007.