SMALL MICROPHONE ARRAYS WITH OPTIMIZED DIRECTIVITY FOR SPEECH ENHANCEMENT

Matthias Dörbecker
Institute of Communication Systems and Data Processing,
Aachen University of Technology, Templergraben 55, 52056 Aachen, Germany
Phone: +49 241 806979, Fax: +49 241 8888186
E-Mail: matthias@ind.rwth-aachen.de http://www.ind.rwth-aachen.de

ABSTRACT

In many situations of digital speech communication (e.g. hands-free telephony or electronic hearing aids) the speech signal picked up by the microphone is disturbed by acoustic background noise. Therefore, adaptive filtering techniques which aim at the reduction of the disturbing noise are subject of current research activities (e.g. [1]). Although some of the already known adaptive techniques – especially concepts with two or more microphones – allow a significant reduction of the noise, most of the adaptive strategies result, particularly at low SNR, in a speech signal with an unnatural character due to time-variant distortions, and the occurrence of *musical noise*. An alternative approach, which does not affect the speech signal by time-variant distortions, is the application of a microphone array with a fixed directivity pattern aligned to the speaker's position, resulting in a suppression of spatially distributed noise sources. In this contribution it is shown that due to the proposed optimization even with only two microphones a reduction of diffuse noise sound up to 6 dB can be achieved.

1 INTRODUCTION

The structure of a linear array with $N = 4$ microphones is depicted in Fig. 1. The microphone signals $x_n(k)$ (with $1 \le n \le N$) are filtered with time invariant impulse responses $a_n(k)$, which are designed to obtain an output signal $\hat{s}(k)$ of maximum power for the direction of principal incidence. Conventionally, the filters $a_n(k)$ simply provide a compensation of the time delay between the microphone signals, which occurs for the direction of principal incidence. The resulting structure is often called *delay-and-sum-beamformer*. The destructive interference occurring for every other angle of incidence leads to a reduced output signal for directions outside the main lobe.

Although the delay-and-sum-beamformer is based on a distinct theory, this approach is not very appropriate for the application in microphone arrays because it results in a poor directivity. Therefore, a large number of microphones is required to obtain an appreciable reduction of spatially distributed noise sources. For this reason, this contribution deals with the application of the theory known from superdirective antennas to microphone arrays. The fundamentals of superdirective antennas have already been discussed in [2] and have been applied to microphone

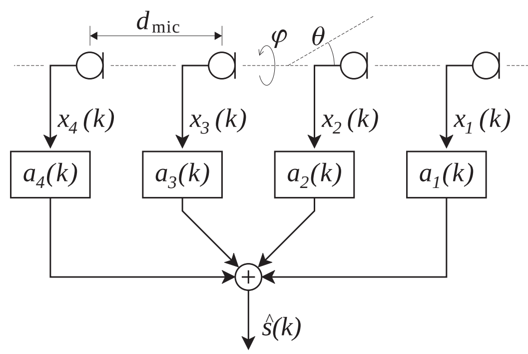

Figure 1: Structure of a linear array with $N = 4$ microphones

arrays in [3, 4, 5]. After a short review of superdirective arrays, it is shown in Section 2 that the superdirective design yields an excellent directivity even if a small number of microphones is used. A further improvement of the directivity can be obtained by the modified design proposed in Section 3. In conclusion, the simulations described in Section 4 confirm that significant speech enhancement can be achieved in real acoustic situations.

2 DESIGN OF SUPERDIRECTIVE ARRAYS

The characteristic of an array can be described by means of the power directivity pattern $\Psi(f,\theta,\varphi)$, which represents the power spectral density of the output signal $\hat{s}(k)$ as a function of the direction of incidence denoted by the angles θ and φ. In the following, we focus on linear arrays, i.e. the microphones have to be placed on a straight line as depicted in Fig. 1. Due to the symmetry of the array, the power directivity pattern is independent of the angle φ according to

$$\Psi(f,\theta) = \left|\sum_{n=1}^{N} A_n(f)\exp\left(j\beta\, d_{\rm mic}\cdot\left(\tfrac{N+1}{2} - n\right)\cos\theta\right)\right|^2, \tag{1}$$

where $A_n(f)$ denotes the transfer function of filter $a_n(k)$, $d_{\rm mic}$ the distance between two adjacent microphones, $\beta = 2\pi f/c$ the propagation factor, and c the speed of sound. Alternatively, the directivity can be measured by the gain $G(f)$, which is defined as the ratio of the power directivity pattern for the direction of principal incidence, $\Psi(f,\theta_0)$, relative to the power spectral density of the array's output signal in case of an omnidirectional incidence of sound. In terms of

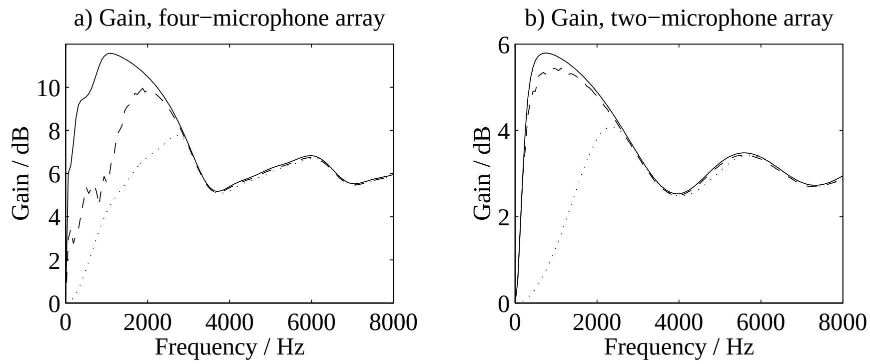

Figure 2: Gain of linear end-fire arrays with four and two omnidirectional microphones, respectively, and a microphone spacing of $d_{\mathrm{mic}} = 5\,\mathrm{cm}$.
········ delay-and-sum-beamformer
——— superdirective array, microphones with identical transfer functions
– – – superdir. array, microphones with non-identical transfer functions (see description in the text)

spherical polar coordinates the gain reads

$$G(f) \;=\; \frac{\Psi(f,\theta_0)}{\frac{1}{4\pi}\int\limits_0^{2\pi}\int\limits_0^{\pi}\Psi(f,\theta)\,\sin\theta\;d\theta\,d\varphi}\;\;, \qquad (2)$$

where θ_0 represents the direction of principal incidence. In case of a diffuse noise sound field the gain according to equation (2) is equal to the improvement of the SNR. By means of equation (1), an equivalent expression of the gain can be obtained as

$$G(f) \;=\; \frac{\Psi(f,\theta_0)}{\sum\limits_{n=1}^{N}\sum\limits_{m=1}^{N} A_n(f)\,A_m^*(f)\,h_{mn}(f)}\;\;, \qquad (3)$$

where the coefficients $h_{mn}(f)$ are given by:

$$\begin{aligned} h_{mn}(f) \;&=\; \frac{1}{4\pi}\int\limits_0^{2\pi}\int\limits_0^{\pi}\exp(j\beta\,d_{nm}\cos\theta)\,\sin\theta\,d\theta\,d\varphi \\ &=\; \begin{cases} \dfrac{\sin(\beta\,d_{nm})}{\beta\,d_{nm}} & \text{for } n \neq m \\ 1 & \text{for } n = m \end{cases} \qquad (4) \end{aligned}$$

The robustness of an array against random errors of the positions and the transfer functions of the microphones can be evaluated by the susceptibility $K(f)$. In [2] it has been shown that the deviation of the power directivity pattern averages to

$$E\{\Delta\Psi(f,\theta)\} \;=\; \Psi(f,\theta_0)\cdot\Delta^2(f)\cdot K(f)\;\;, \qquad (5)$$

where $\Delta^2(f)$ comprises the relative errors of the positions as well as of the microphone transfer functions. Therefore, to obtain a small deviation of the power directivity pattern, it is required that the susceptibility $K(f)$ does not exceed an upper limit, which depends on the variances of the microphones' transfer functions.

The design of a superdirective array aims at the maximization of the gain, while the susceptibility must not exceed a presupposed upper limit. As a result, the transfer functions $A_n(f)$ can be obtained by solving the system of linear equations

$$\begin{aligned} &\sum_{m=1}^{N} h_{nm}(f)\,A_m(f) + \mu\,A_n(f) \\ &\quad=\; \exp\left(-j\beta\,d_{\mathrm{mic}}\cdot\left(\tfrac{N+1}{2} - n\right)\cos\theta_0\right) \qquad (6) \end{aligned}$$

for $1 \le n \le N$ [2]. In equation (6), μ denotes an undetermined Lagrangian multiplier, which allows to control the superdirectivity as well as the susceptibility. The choice of a large multiplier $\mu \gg 1$ results in the conventional delay-and-sum-beamformer, whereas the superdirectivity and the susceptibility increase when μ tends towards zero. The impulse responses $a_n(k)$ are obtained by solving the system of equations (6) for several discrete frequencies $f_\nu = \nu\cdot f_S/M$ with $0 \le \nu \le M/2$, taking the inverse DFT of length M of the $\overline{A_n(f_\nu)}$, and multiplying the resulting time-domain sequences with a Hamming window. For the arrays considered in this contribution, which are designed for a sampling frequency of $f_S = 16\,\mathrm{kHz}$, impulse responses $a_n(k)$ of length 256 have been used.

Fig. 2 a) depicts the gain of a linear end-fire array, i.e. the direction of principal incidence is in parallel to the array's axis ($\theta_0 = 0$). The array considered in this example consists of four microphones placed with a spacing of $d_{\mathrm{mic}} = 5\,\mathrm{cm}$. To examine the gain resulting from the coupling of the microphones, an omnidirectional characteristic is assumed for each individual microphone. This is with no loss in generality because the results can also be applied to directional microphones.

As indicated by the dotted line, the delay-and-sum-beamformer results in an insufficient gain at low frequencies. It can be stated that for the array considered in this example the well-known rule of thumb – each doubling of the number of microphones leads to an improvement of 3 dB – is only valid for frequencies above 1500 Hz. Thus if delay-and-sum-beamforming is supposed, quality enhancement at lower frequencies requires that the number of the microphones or the dimensions of the array have to be enlarged.

An alternative solution consists in the superdirective design as shown in Fig. 2 a) by the solid line. The superdirective design yields a significantly improved

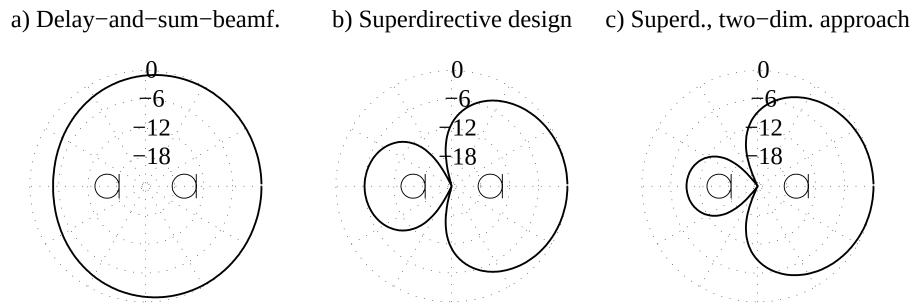

Figure 3: Power directivity pattern in dB of the two-microphone end-fire array ($d_{\text{mic}} = 5\,\text{cm}$) at 1000 Hz

gain, especially in the frequency range up to 3 kHz, which is of main importance for speech communication. The maximum gain of the superdirective array approaches $20 \log_{10} N$ dB as μ tends towards zero. For frequencies above 3.3 kHz (i.e. for small wavelengths $\lambda/2 < d_{\text{mic}}$) the superdirective array turns into the delay-and-sum-beamformer.

However, the superdirective design results in a higher susceptibility, which requires lower variances of the microphones' transfer functions. Although the solid curve in Fig. 2 a) refers to an array with a moderate superdirectivity (due to the choice of $\mu = 0.01$), the gain of the array decreases significantly if microphone transfer functions with realistic variances are considered. This is indicated by the dashed curve in Fig. 2 a), which has been obtained by averaging the gains of 100 arrays with randomly disturbed microphone transfer functions. In average, the magnitude transfer functions have been disturbed by 10 % and the phases have been affected by deviations of $0.03\,\pi$. Since the susceptibility increases as the number of microphones is enlarged, the superdirective design is not appropriate for arrays consisting of a large number of microphones. However, if only a small number of microphones is used (which is the case in most applications of digital speech communication), the susceptibility resulting from the superdirective design is small enough to cope with the variances of real microphone transfer functions.

To demonstrate this, the gains depicted in Fig. 2 b) refer to an endfire array consisting of only two microphones placed at a distance of 5 cm. According to the application of two microphones, the maximum gain of the superdirective array reaches only $20 \log_{10} 2$ dB = 6 dB. However, since the two-microphone array provides a lower susceptibility, the gain is much less reduced by the non-identical transfer functions as compared to the four-microphone array. This is indicated by the dashed curve in Fig. 2 b), which refers to the same variances of the microphone transfer functions as described before.

3 IMPROVEMENT OF THE DIRECTIVITY PATTERN

Fig. 3 depicts the power directivity pattern $\Psi(f_0, \theta)$ at $f_0 = 1000$ Hz of the delay-and-sum-beamformer (a) and of the superdirective array (b), both consisting of two omnidirectional microphones at a distance of $d_{\text{mic}} = 5$ cm. According to its insufficient gain at low frequencies, the delay-and-sum-beamformer shows an almost omnidirectional characteristic. Although the directivity of the superdirective design is much superior, there is still a distinct secondary lobe in the rear direction. The extent of the secondary lobe is caused by the fact that the superdirective design aims at the maximization of the gain. Because of the term $\sin\theta$ in the denominator of equation (2), the rear direction (i.e. $\theta = \pi$) causes a much smaller influence on the gain than any other direction. Consequently, the maximization of the gain does not force a small secondary lobe in the rear direction.

However, in many acoustic situations (e.g. if the array is applied to an electronic hearing aid) a small secondary lobe in the rear direction is of greater importance than an optimal SNR in case of an omnidirectional incidence of noise sound. For this reason, in the following a modified design for the impulse responses $a_n(k)$ is derived, which is based on a two-dimensional definition of the gain. According to this two-dimensional definition of the gain, all directions θ have the same contribution. Therefore, the influence of the rear direction is enlarged, so that the maximization of the two-dimensional gain forces a lower secondary lobe in the rear direction. The modified gain resulting from this two-dimensional approach is given by

$$G'(f) = \frac{\Psi(f, \theta_0)}{\frac{1}{2\pi} \int_0^{2\pi} \Psi(f, \theta)\, d\theta} \; . \tag{7}$$

Equation (3) also holds for the two-dimensional gain if the coefficients $h_{mn}(f)$ are replaced by

$$h'_{mn}(f) = \frac{1}{2\pi} \int_0^{2\pi} \exp(j\beta\, d_{mn} \cos\theta)\, d\theta \; . \tag{8}$$

Therefore, the resulting set transfer functions $A'_n(f)$, which maximizes the two-dimensional gain, can be obtained by solving the system of equations (6) if the coefficients $h'_{mn}(f)$ are applied. Similar to the three-dimensional approach, the impulse responses $a'_n(k)$ emerge by an inverse DFT and appropriate windowing.

Unfortunately, the integral in equation (8) can not be solved in closed form. A numerical solution can be approximated as follows:

$$\begin{aligned} h'_{mn}(f) &= \frac{1}{\pi} \int_0^{\pi} \cos(\beta\, d_{mn} \cos\theta)\, d\theta \qquad (9) \\ &= \lim_{L\to\infty} \frac{1}{L} \sum_{l=0}^{L-1} \cos\left(\beta\, d_{mn} \cos\left(\pi\, \frac{2l+1}{2L}\right)\right) \end{aligned}$$

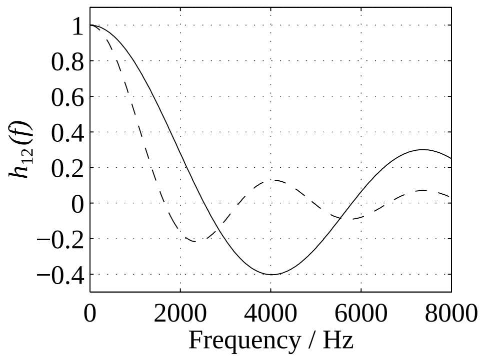

Figure 4: Coefficients $h_{12}(f)$ (– – –) and $h'_{12}(f)$ (——) for $d_{12} = 5\,\mathrm{cm}$

In Fig. 4 the functions $h'_{12}(f)$ resulting from the two-dimensional approach and $h_{12}(f)$ corresponding to equation (4) are compared. The plots refer to a two-microphone array with a spacing of $d_{12}=5\,\mathrm{cm}$. For the array dimensions and the frequency range considered here, it is possible to choose the upper limit of the sum in equation (9) as $L=20$, which has been confirmed by numerical evaluations.

Fig. 3 c) proves that the two-dimensional approach reduces the secondary lobe in the rear direction by 3 dB. Furthermore it can be observed that the improvement is at the expense of an only marginally enlarged angle of the main beam.

4 SIMULATION RESULTS

To demonstrate that the designs discussed in this contribution can successfully be applied to real microphone arrays, an array consisting of two omnidirectional microphones at a distance of $d_{\mathrm{mic}} = 5\,\mathrm{cm}$ has been assembled. The impulse responses $a_n(k)$ of length 256 are designed for a sampling frequency of 16 kHz. Using this arrangement, the microphone signals have been recorded in an anechoic chamber for different angles of incidence. To determine the narrow-band power directivity pattern depicted in Fig. 5, the microphone signal $x_1(k)$ and the output signal $\hat{s}(k)$ are applied to a band-pass filter at 1000 Hz. The directivity pattern emerges by calculating the power ratio of the narrow-band signals for different angles θ.

Fig. 5 confirms that the insufficient directivity of the delay-and-sum-beamformer at frequencies below 1000 Hz can be improved significantly by the superdirective design without any complications resulting from the variances of real microphone transfer functions. Furthermore, it can be observed that the design based on the two-dimensional approach yields a significantly reduced secondary lobe in the rear direction of the measured directivity pattern.

To evaluate the speech enhancement resulting from the optimized directivity, microphone signals have been recorded which refer to an acoustic situation where the array is aligned to a single speaker. The noise sound field is approximately diffuse. Informal listening tests confirm a significant reduction of the noise. In contrast to most adaptive noise reduction systems, the array's output signal is of high naturalness because of the absence of time-variant distortions. Since the improvement is independent of the input-SNR, the superdirective array can successfully be applied even in situations with extremely poor SNR.

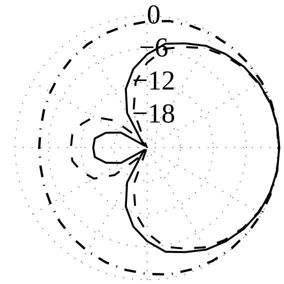

Figure 5: Power directivity pattern in dB of a real array with two microphones at 1000 Hz
· – · – · delay-and-sum-beamformer
– – – superdirective design
——— superdirective, two-dim. approach

5 CONCLUSIONS

In this contribution it has been shown that the superdirective design known from antenna arrays can successfully be applied to microphone arrays. Simulations considering idealized as well as recorded microphone signals confirm that the susceptibility resulting from the superdirective design is small enough to cope with the variances of real microphone transfer functions, if a small number of microphones is used. The directivity pattern can be further improved by a novel design, which is based on a two-dimensional definition of the gain. Listening tests confirm a significant reduction of spatially distributed noise sources. According to the application of time-invariant filtering the output signal is not affected by any time-variant distortions, which results in a very high naturalness. Therefore, the two-microphone array with optimized directivity is a powerful alternative to state-of-the-art noise suppression techniques based on adaptive filtering.

6 REFERENCES

[1] R. Martin and P. Vary, "Combined Acoustic Echo Control and Noise Reduction for Hands-Free Telephony – State of the Art and Perspectives", in *Proc. EUSIPCO*, Sept. 1996, pp. 1107–1110.

[2] E.N. Gilbert and S.P. Morgan, "Optimum Design of Directive Antenna Arrays Subject to Random Variations", *The Bell System Technical Journal*, vol. 34, pp. 637–663, May 1955.

[3] R. Zelinski, "Ein Mikrofon-Array mit superdirektiven Eigenschaften zur Geräuschreduktion bei Sprachsignalübertragung", in *Kleinheubacher Berichte*. 1995, vol. 38, pp. 507–515, Deutsche Telekom, Darmstadt.

[4] R. Zelinski, "Mikrofon-Arrays mit superdirektiven Eigenschaften zur Sprachsignalübertragung", *Frequenz*, vol. 50, no. 9-10, pp. 198–204, 1996.

[5] M. Dörbecker and Th. Moll, "Einsatz superdirektiver Mikrofonarrays in Kombination mit adaptiven Verfahren zur Reduktion akustischer Störgeräusche", in *Proc. 9. Aachener Kolloquium Signaltheorie, Bild- und Sprachsignale*, Mar. 1997, pp. 243–246.

This work has been supported by the German Research Council (DFG).