BEAMFORMING WITH OPTIMIZED INTERPOLATED MICROPHONE ARRAYS

Gerhard Doblinger

Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology
Gusshausstr. 25/389, A-1040 Vienna, Austria
gerhard.doblinger@tuwien.ac.at, www.nt.tuwien.ac.at/staff/gerhard-doblinger/

ABSTRACT

We present an optimization procedure for wideband beamforming with interpolated arrays. We intend to design a beamformer with a compact size. In addition, we want to reduce the number of sensors while maintaining a good beamforming performance. Our beamformers are implemented using FFT filterbanks. Performance is tested under far-field conditions and under sound propagation with simulated room impulse responses. In addition, we study the influence of sensor noise on the beamforming behavior.

Index Terms— Array signal processing, array interpolation, microphone arrays.

1. INTRODUCTION

The primary goal of an array beamformer design is to achieve a beam pattern with high directivity. The challenge thereby is to reach this goal with a minimum amount of sensors, with a small array size, and with a robustness against sensor position errors and sensor noise. A standard beamformer design with a reduced sensor count is based on superdirective arrays (e.g. [1]). However, with high superdirectivity a large amplification of sensor noise and sensor position errors must be taken into account. Our design technique involves array interpolation and extrapolation. The basic idea is to map an array of N sensors to an array of $M > N$ sensors by introducing virtual sensors. Originally, array interpolation has been used to cope with sensor failure [2]. Interpolated arrays are also used to improve the performance of direction-of-arrival (DOA) estimation [3, 5, 6, 7]. Nonuniform arrays can be mapped to uniform arrays to enable the use of the fast root-MUSIC estimation algorithm [4]. In this paper, we show that array interpolation can improve the beamforming performance and reduces the influence of sensor noise. The main advantage of interpolated arrays is the possibility to improve a given array by software and not by changing the array layout or increasing the number of microphones.

In Section 2, we present a new interpolation/extrapolation algorithm for wideband beamformers like microphone arrays. We discuss the influence of sensor noise on the performance of interpolated arrays. An efficient implementation using overlap-add FFT filterbanks and representative examples of one-dimensional interpolated beamformers operating under far-field and under simulated room conditions are given in Section 3.

2. ARRAY INTERPOLATION

Array interpolation is carried out by preprocessing of N input sensor signals to create $M \geq N$ sensor signals. As shown in Fig. 1, virtual sensors are introduced by multi-channel mapping system $T\{\cdot\}$. Array processing like beamforming is then applied to the M output sensor signals $y_i(t)$ instead of the N input sensor signals $x_i(t)$. As an example, sensor mapping

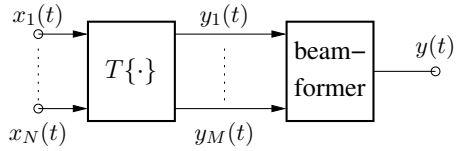

Fig. 1. Mapping of sensor signals $x_i(t)$ to virtual sensor signals $y_i(t) = T\{x_i(t)\}$ linked with beamforming.

can convert a nonuniform array to a uniform array. In this case, we normally use $N = M$ and perform a signal interpolation with $T\{\cdot\}$. Another example is extrapolation and/or interpolation where we augment a uniform array of N sensors by $M - N$ sensors to create an M sensor uniform array. We intend to improve the beamforming behavior with the M sensor array. Mapping $T\{\cdot\}$ depends on the array geometries and on the signal model describing the wavefield. We use a far-field signal model where $x_i(t)$ are time-shifted versions $s(t-\tau_i)$ of a single source signal $s(t)$. The broadband source signal with spectrum $S(j\omega)$ is emitted from a direction given by azimuth $\phi \in [0, \pi]$ (assuming a one-dimensional array). Therefore, the Fourier transform of the input signal vector $\mathbf{x}(t) = [x_1(t), \ldots, x_N(t)]^T$ is

$$\mathbf{X}(j\omega, \phi) = S(j\omega)\mathbf{d}(j\omega, \phi), \tag{1}$$

with steering vector $\mathbf{d}(j\omega,\phi) = [e^{-j\omega\tau_1(\phi)},\ldots,e^{-j\omega\tau_N(\phi)}]^T$. We implement a linear, time-invariant sensor mapping in the frequency domain using an $N \times M$ matrix $\mathbf{T}(j\omega,\phi)$:

$$\mathbf{Y}(j\omega,\phi) = S(j\omega)\mathbf{T}^H(j\omega,\phi)\mathbf{d}(j\omega,\phi) = S(j\omega)\mathbf{d}_V(j\omega,\phi), \tag{2}$$

where $\mathbf{d}_V(j\omega,\phi) = [e^{-j\omega\tau'_1(\phi)},\ldots,e^{-j\omega\tau'_M(\phi)}]^T$ is the steering vector of the desired virtual array (^{H} denotes conjugate transpose). This steering vector is then used in a minimum-variance-distortionless-response (MVDR) design of the beamformer. Usually, matrix $\mathbf{T}(j\omega)$ may be obtained by minimizing a least-squares (LS) error criterion for a set of discrete ϕ. However, we found that the standard LS method yields $\mathbf{T}$-matrices with large norms giving rise to a large amplification of sensor noise. To investigate the noise behavior, we consider uncorrelated, zero-mean, white-noise sources $u_i(t)$, $i = 1,\ldots,N$ added to the input sensor signals. These noise sources are modified by the multi-channel mapping system to output noise contributions $v_i(t)$, $i = 1,\ldots,M$. Using the cross-power spectral density $S_{u_iu_k}(j\omega) = \sigma_u^2\delta[i-k]$ of sensor noises $u_i(t)$, $u_k(t)$ as elements of the power spectral density (PSD) matrix $\mathbf{S_{uu}}(j\omega) = \sigma_u^2\mathbf{I}$ (with $N \times N$ identity matrix $\mathbf{I}$), the respective PSD matrix at the mapping system outputs is given by

$$\mathbf{S_{vv}}(j\omega) = \sigma_u^2\mathbf{T}^H(j\omega)\mathbf{T}(j\omega). \tag{3}$$

Therefore, the diagonal elements of $\mathbf{S_{vv}}(j\omega)$ in (3) are

$$S_{v_iv_i}(j\omega) = \sigma_u^2\left\|\mathbf{t}_i(j\omega)\right\|_2^2, \quad i = 1,\ldots,M \tag{4}$$

($\mathbf{t}_i(j\omega)$ is the i-th column vector of $\mathbf{T}(j\omega)$). In addition, we can show that the beamformer output noise PSD may be expressed by

$$S_{yy}(j\omega) \le \sigma_u^2\left\|\mathbf{w}(j\omega)\right\|_2^2\sum_{i=1}^{M}\left\|\mathbf{t}_i(j\omega)\right\|_2^2, \tag{5}$$

with beamformer weight vector $\mathbf{w}(j\omega)$. According to (4), (5), noise spectra are modified by the column vector norms. As a consequence, we must limit these norms to avoid an unwanted noise amplification by the mapping system. For this reason, we use the following norm-constrained LS algorithm. Suppose we would like to find $\mathbf{T}$ of the best mapping of steering vector $\mathbf{d}(j\omega,\phi)$ to $\mathbf{d}_V(j\omega,\phi)$ at L discrete azimuths $\phi_l \in [\phi_{\min},\phi_{\max}]$. We collect the steering vectors in the $N \times L$ matrix $\mathbf{D}(j\omega) = [\mathbf{d}(j\omega,\phi_1),\ldots,\mathbf{d}(j\omega,\phi_L)]$, and $M \times L$ matrix $\mathbf{D}_V(j\omega) = [\mathbf{d}_V(j\omega,\phi_1),\ldots,\mathbf{d}_V(j\omega,\phi_L)]$, respectively. Our norm-constrained LS-approach can then be formulated as

$$\begin{aligned}\mathbf{t}_i(j\omega) = \arg\min_{\mathbf{t}_i}\left\|\mathbf{D}^H(j\omega)\mathbf{t}_i(j\omega) - \bar{\mathbf{d}}_i(j\omega)\right\|_2^2 \\ \text{with } \left\|\mathbf{t}_i(j\omega)\right\|_2 \le \alpha, \quad i = 1,\ldots,M,\end{aligned} \tag{6}$$

where $\bar{\mathbf{d}}_i$ is the i-th column vector of $\mathbf{D}_V^H$. The desired norm constraint in (6) is given by parameter α. This constrained LS-problem can be solved by the following algorithm [8]:

1. compute singular value decomposition $\mathbf{D}^H = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ with singular values $\mathbf{\Sigma} = \text{diag}([\sigma_1,\ldots,\sigma_N])$
2. determine rank r of $\mathbf{D}^H$ by finding smallest $\sigma_r \ge \varepsilon$
3. if $\sum_{k=1}^{r}\frac{|b_k|^2}{\sigma_k^2} \le \alpha^2$, with $\mathbf{b} = [b_1,\ldots,b_N]^T = \mathbf{U}^H\bar{\mathbf{d}}_i$

3a. then $\mathbf{t}_i = \sum_{k=1}^{r}\frac{b_k}{\sigma_k}\mathbf{v}_k$, with $\mathbf{V} = [\mathbf{v}_1,\ldots,\mathbf{v}_N]$

3b. else find $\lambda > 0$ that satisfies $\sum_{k=1}^{r}\frac{\sigma_k^2}{(\sigma_k^2+\lambda)^2}|b_k|^2 = \alpha^2$
set $\mathbf{t}_i = \sum_{k=1}^{r}\frac{\sigma_k b_k}{\sigma_k^2+\lambda}\mathbf{v}_k$

All steps of this algorithm are repeated for all column vectors of $\mathbf{D}_V$ to obtain all column vectors $\mathbf{t}_i$, $i = 1,\ldots,M$ of $\mathbf{T}$. In addition, the whole procedure must be carried out for all desired frequency points $\omega_k \in [\omega_{\min},\omega_{\max}]$ (ω_k is omitted for readability in the outline of the algorithm). Parameter λ is computed by numerical zero-finding of the function in step 3b. We can interpret the role of λ as automatic diagonal loading (regularization) of matrix $\mathbf{\Sigma}$. It should be noted that these steps are computed only once for given array layouts.

Storage demand and computational effort can be significantly reduced if sensor positions are symmetric around the center of the array and azimuth limits $\phi_{\min},\phi_{\max}$ are symmetric to $\pi/2$. Furthermore, a special case of extrapolation is given by extending a uniform array of N sensors at both ends by K virtual sensors. For such an array only $N \times K$ matrices $\mathbf{T}_l, \mathbf{T}_r$ are needed since (2) yields

$$\mathbf{d}_V = \mathbf{T}^H\mathbf{d} = \begin{pmatrix}\mathbf{T}_l^T\\ \mathbf{I}\\ \mathbf{T}_r^T\end{pmatrix}\mathbf{d} \tag{7}$$

($\mathbf{I}$ is the $N \times N$ identity matrix). With the above mentioned symmetries the elements of $\mathbf{T}_l$ and $\mathbf{T}_r$ are real-valued and related by

$$\begin{aligned}T_r(i,k) = T_l(N-i+1,K-k+1),\\ i = 1,\ldots,N,\ k = 1,\ldots,K.\end{aligned} \tag{8}$$

Note that the N input sensor signals are not modified. Only the additional $2K$ virtual sensor signals are computed by the mapping system. A MATLAB program of the proposed algorithm is provided at the author's homepage.

3. IMPLEMENTATION AND EXPERIMENTAL RESULTS

We implement the wideband beamformer with an overlap-add FFT filterbank for each sensor signal as shown in Fig. 2. The advantage as compared with beamformers using FIR filters is the direct application of interpolation matrices $\mathbf{T}[k]$, and

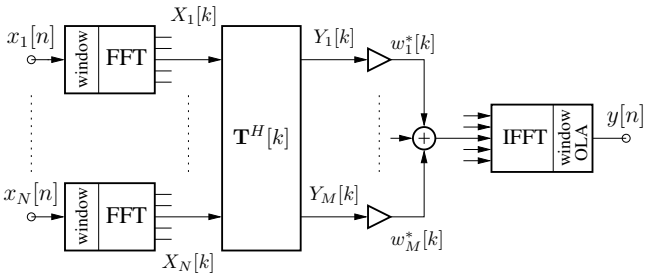

Fig. 2. Interpolated wideband beamformer as overlap-add FFT filterbank (with processing at FFT index k shown).

frequency-domain weight vectors $\mathbf{w}[k]$. MVDR weight vectors are computed at frequency points between 0.3 kHz and 5.5 kHz with a sampling frequency of 16 kHz and an FFT length of 512. Sensor spacing of uniform arrays is $d = 2.5$ cm to avoid spatial aliasing up to a frequency of $f_{\max} = \frac{c}{2d} = 6.8$ kHz (sound velocity $c = 340$ m/s). All simulation results are based on a standard MVDR design (signal model (1), diffuse noise field, regularization of the spatio-spectral correlation matrix) [1]. However, we include additional constraints in the MVDR design to obtain desired spatial nulls.

The interpolation matrices $\mathbf{T}$ are computed at the same frequency points and are applied to the FFT bins of the filterbanks. We typically use a norm constraint parameter $\alpha = 5$ and $L = 100$ discrete azimuth values $\phi_l \in [0, \pi]$ needed by the optimization algorithm in Section 2. In array interpolation for DOA estimation smaller azimuth sectors are used in order to keep the interpolation error small. In beamforming applications, however, we must span the whole azimuth range and must accept greater interpolation errors.

We begin with a representative example of a uniform array with $N = 8$ input sensors as shown in Fig. 3. The

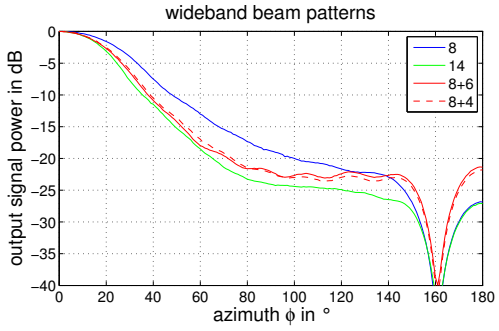

Fig. 3. Measured far-field wideband beam patterns of uniform arrays with $N = 8$, $N = 14$ input sensors, and interpolation of $N = 8$ array with 4, and 6 virtual sensors.

beamformer has an end-fire look direction ($\phi_d = 0°$) and a spatial null at $\phi_s = 160°$. Adding virtual sensors at both ends of the $N = 8$ array creates arrays with $M = 12$, and 14 sensors, respectively. Beam pattern measurements are carried out using a pseudo-random source signal with white spectrum band-limited to 0.3 kHz ... 5.5 kHz. These wideband beam patterns are obtained under far-field conditions and match the theoretical patterns. Using (1) and $S(j\omega) = 1,\ \omega \in [\omega_{\min}, \omega_{\max}]$, the theoretical wideband beam pattern of the array is given by

$$P(\phi) = \frac{1}{\pi} \int_{\omega_{\min}}^{\omega_{\max}} \left| \mathbf{w}(j\omega)^H \mathbf{d}_V(j\omega, \phi) \right|^2 d\omega, \qquad (9)$$

with beamformer weight vector $\mathbf{w}(j\omega)$, and steering vector $\mathbf{d}_V(j\omega, \phi)$. The sensors are assumed to be omnidirectional and noise-free. Fig. 3 indicates that more than 6 virtual sensors would not improve the beam pattern any more. The -6 dB-mainlobe-width is reduced from 72° (8 sensors) to 58° (8+6 sensors). With a full array of 14 sensors we get 54°. We found that a 14° - 18° reduction of the -6 dB-mainlobe-width is typical for look directions $\phi_d \in [0°, 70°]$. For broad-side directions ($\phi_d = 90°$) less improvements can be achieved. The smaller mainlobe of the interpolated array is due to a better beamforming behavior in the frequency range from 0.3 kHz to $\approx$ 2 kHz. This is caused by a strong spatial correlation of the sensor signals which in turn results in small interpolation errors.

In another experiment, we study the influence of sensor noise on the beam patterns. In general, superdirective MVDR beamformers are sensitive to uncorrelated noise at the input sensors. As a common remedy, diagonal loading (regularization) is applied to the spatio-spectral correlation matrix involved in the MVDR design [1]. Fig. 4 shows a representative example in case of an input signal-to-noise-ratio of 25 dB. As

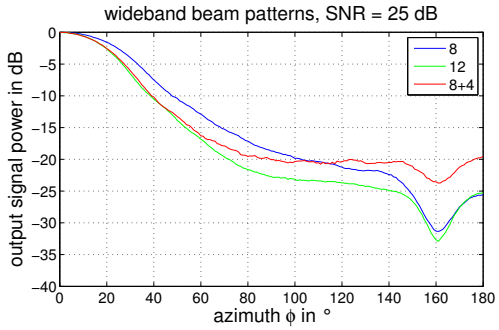

Fig. 4. Measured far-field wideband beam patterns of uniform arrays with $N = 8$, $N = 12$ input sensors, and interpolation of $N = 8$ array with 4 virtual sensors in case of noisy input sensors (SNR = 25 dB).

indicated in Fig. 4, the mainlobe of the interpolated array is barely influenced by sensor noise. In case of array interpolation, sidelobe levels are increased due to the noise by ≤ 2.5 dB at $\phi \in [80°, 180°]$. In addition, the sharpness of spatial nulls is reduced.

In order to test the beamformer design in a room environment, we investigate the influence of simulated room acous-

tics on the beam patterns. Rooms are simulated using the image method [9] in combination with fractional delay interpolation to obtain an accurate signal delay to each sensor [10]. In order to speed-up simulations, we apply delay interpolation to the direct paths only. Room reflections exhibit longer delays and thus integer sample delays are sufficient.

A typical simulation result using a room with dimensions $L \times W \times H = 6$ m $\times 5$ m $\times 3$ m, and 0.13 sec. reverberation time T_{60} is shown in Fig. 5. Moderate room reverberations increase sidelobe levels of the interpolated array by ≈ 8 dB, and remove the spatial null. However, we observe only a minor influence on the mainlobe width.

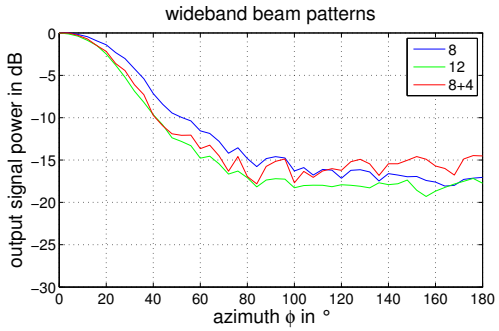

Fig. 5. Simulated room wideband beam patterns of uniform arrays with $N = 8$, $N = 12$ input sensors, and $M = 8 + 4$ interpolated array (reverberation time $T_{60} = 0.13$ sec.).

In the final example shown in Fig. 6, we compare the noise behavior of an $M = 8 + 4$ interpolated array with that of an $N = 8$ array. In order to achieve the same mainlobe width, we increase the superdirectivity (less diagonal loading in the MVDR design) of the $N = 8$ beamformer. As a result, a higher noise sensitivity of the $N = 8$ array can be observed.

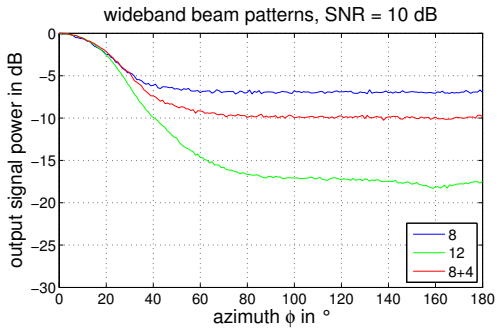

Fig. 6. Comparison of $N = 12$ uniform array, $N = 8$ array, and $M = 8 + 4$ interpolated array (noisy input sensors, SNR $= 10$ dB, $N = 8$ array with increased superdirectivity).

4. CONCLUSIONS

We have presented a wideband beamformer design technique based on array interpolation. This method exhibits a smaller number of sensors, a smaller array size, and an improved beam pattern as compared to conventional beamformer designs. However, care must be taken on the influence of input sensor noise. Therefore, we use a norm-constrained LS-algorithm to obtain the interpolation matrices involved in the beamformer design. We have tested our designs with an FFT filterbank beamformer implementation operating under far-field conditions and in simulated rooms. As a result, the beamforming performance of microphone arrays can be improved with array interpolation, provided that we can assume moderate room reverberations.

5. REFERENCES

[1] J. Bitzer, K. U. Simmer, "Superdirective microphone arrays," in *Microphone arrays*, M. Brandstein, and D. Ward (Eds.), Springer-Verlag, Berlin Heidelberg New York, 2001, ch. 2, pp. 19–38.

[2] D. N. Swingler, R. S. Walker, "Line-array beamforming using linear prediction for aperture interpolation and extrapolation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 16–30, Jan. 1989.

[3] T. P. Bronez, "Sector interpolation of nonuniform arrays for efficient high resolution bearing estimation," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 5, pp. 2285–2888, Apr. 1988.

[4] B. Friedlander, "The root-MUSIC algorithm for direction finding with interpolated arrays," *Signal Processing*, vol. 30, pp. 15-29, Jan. 1993.

[5] A. B. Gershman, J. F. Böhme, "A note on most favorable array geometries for DOA estimation and array interpolation," *IEEE Signal Processing Letters*, vol. 4, pp. 232–235, Aug. 1997.

[6] P. Hyberg, M. Jansson, B. Ottersten, "Array interpolation and bias reduction," *IEEE Trans. Signal Processing*, vol. 52, pp. 2711–2720, Oct. 2004.

[7] P. Hyberg, M. Jansson, B. Ottersten, "Array interpolation and DOA MSE reduction," *IEEE Trans. Signal Processing*, vol. 53, pp. 4464–4471, Dec. 2005.

[8] G. H. Golub, C. F. van Loan, *Matrix computations*, The Johns Hopkins University Press, Baltimore, 1985, ch. 12, pp. 405–408.

[9] J. B. Allen, D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am., vol. 64, pp. 943–950, Apr. 1979.

[10] P. M. Peterson, "Simulating the response of multiple microphones to a single acoustic source in a reverberant room," J. Acoust. Soc. Am., vol. 80, pp. 1527–1529, Nov. 1986.