Buenos Aires – 5 to 9 September 2016
Acoustics for the 21st Century…

PROCEEDINGS of the 22nd International Congress on Acoustics

Acoustic Array Systems: Paper ICA2016-139

Analysis and design of time-domain first-order circular differential microphone arrays

Yaakov Buchris[a], **Israel Cohen**[b], **and Jacob Benesty**[c]

[a]Technion, Israel Institute of Technology, Israel, bucris@tx.technion.ac.il
[b]Technion, Israel Institute of Technology, Israel, icohen@ee.technion.ac.il
[c]INRS-EMT, University of Quebec, Canada, benesty@emt.inrs.ca

Abstract

Circular differential microphone arrays (CDMAs) are characterized as compact superdirective beamformers whose beampatterns are almost frequency invariant. In contrast to linear differential microphone arrays (LDMAs) where the optimal steering direction is at the endfire, CDMAs provide almost perfect steering for all azimuthal directions. Herein, we present the design of a first-order CDMA in the time domain which is motivated by several aspects. First, time-domain implementation is important in some applications where minimal delay is required, such as real-time communications. Moreover, direct design in the time domain can reduce the computational efforts compared to the frequency-domain design, especially when short filters are sufficient. We present a design example for the time-domain first-order CDMA illustrating some of its fundamental properties as well as the equivalence to the frequency-domain alternative.

Keywords: Circular differential microphone arrays, time-domain broadband array processing

Analysis and design of time-domain first-order circular differential microphone arrays

1 Introduction

Differential microphone arrays (DMAs) can be integrated into several real-world beamforming applications involving speech signals, e.g., hands-free telecommunication, mobile phones, and others. DMAs, which is a family of small size-array beamformers and include the well-known superdirective beamformer [1] as a particular case, have beampatterns that are almost frequency invariant, leading to greatly intelligible signals even in heavy reverberant and noisy environments. Due to these benefits, DMAs have attracted a significant amount of interest in the field of broadband microphone array processing during the past decade [2]-[7].

Broadband array processing algorithms can be implemented both in the time and frequency domains. Design in the time domain is important for real-time applications that require small delays [9]. Furthermore, in some cases the implementation of time-domain filters is computationally more efficient than the equivalent frequency-domain filters, especially when short filters are sufficient. The advantage of the frequency-domain implementation is mainly due to the ability to implement some frequency-dependent processing algorithms, like frequency-selective null-steering.

Previous work on DMAs dealt with linear array geometry which is optimal only at the endfire direction. In some applications like teleconferencing and 3D sound recording where the signal of interest may come from any direction, it is necessary for the microphone array to have similar, if not the same response from one direction to another. In this case, circular arrays are often used. Recently, Benesty et al. [8] introduced an innovative approach for the design and implementation of CDMAs. This approach ignores the traditional differential structure of DMAs and develops the fundamental theory and algorithms for broadband frequency-domain CDMAs up to any order from a signal processing perspective. The proposed solution allows perfect steering in the directions of the array sensors. Several examples are presented, showing the equivalence between the traditional design of DMAs and the proposed design.

In this work, we present a framework for a broadband time-domain implementation of first-order CDMAs which enables perfect steering to any azimuthal direction. First, the array input signal is manipulated and represented in a separable form as a product between a desired signal dependent term and a second term which depends only on the array geometry. Then we derive a closed-form solution for time-domain first-order CDMAs. Due to the DMA assumption, the derived solution is very simple with respect to other methods usually employed in the design of general arrays where some constraints that ensure the frequency invariance should be imposed. We also establish the time-domain equivalent of widely used quality measures like the beampattern, the white noise gain (WNG), and the directivity factor (DF). Finally, we evaluate the performance of the time-domain DMAs and compare it with that of the frequency-domain implementation recently proposed by Benesty et al. [3].

The paper is organized as follows. In Section 2, we formulate the signal model. In Section 3, we define a general broadband beamformer and in Section 4, we develop a closed-form solution for the first-order DMA filters. In Section 5, we define some useful performance measures to evaluate time-domain DMAs. In Section 6, we present a design example of the first-order CDMAs along with some simulation results confirming the validity of the developed time-domain solution.

2 Signal model

We consider a broadband source signal, $s(n)$, in the far-field, where n is the discrete-time index, that propagates in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s, and impinges on a uniform circular array (UCA) of radius r, consisting of M omnidirectional microphones, where the distance between two successive sensors is

$$\delta = 2r\sin\left(\frac{\pi}{M}\right) \approx \frac{2\pi r}{M}. \tag{1}$$

The direction of the source signal to the array is parameterized by the angle θ, where $\theta = 0^\circ$ corresponds to the endfire direction. We assume that the center of the UCA coincides with the origin of the Cartesian coordinate system and serves also as the virtual reference sensor. Assuming a far-field propagation, the time delay between the mth microphone and the center of the array is

$$\tau_m(\theta) = \frac{r}{c}\cos(\theta - \psi_m), \quad m = 1, 2, ..., M, \tag{2}$$

where

$$\psi_m = \frac{2\pi(m-1)}{M} \tag{3}$$

is the angular position of the mth array element. In this scenario, the signal measured at the mth microphone is given by

$$y_m(n) = s\left[n - \Delta - f_s\tau_m(\theta)\right] + v_m(n), \tag{4}$$

where Δ is the propagation time from the position of the source $s(n)$ to the center of the array, f_s is the sampling frequency, and $v_m(n)$ is the noise picked up by the mth sensor. For the general case where $f_s\tau_m(\theta)$ is not an integer, we may apply the Shannon's sampling theorem [10], which implies that

$$\begin{aligned} y_m(n) &= \sum_{l=-\infty}^{\infty} s\left[n - \Delta - l\right]\text{sinc}\left[l - f_s\tau_m(\theta)\right] + v_m(n) \\ &\approx \sum_{l=-P}^{P+\mu L_h} s\left[n - \Delta - l\right]\text{sinc}\left[l - f_s\tau_m(\theta)\right] + v_m(n), \end{aligned} \tag{5}$$

where $P \gg f_s\tau_m(\theta)$, μ is a fraction, and L_h is the length of the FIR filter to be defined later. Hence, we can also express (4) as

$$y_m(n) = \mathbf{g}_m^T(\theta)\mathbf{s}(n - \Delta) + v_m(n), \tag{6}$$

where the superscript ^{T} is the transpose operator, the vector $\mathbf{s}(n-\Delta)$ contains $L = 2P + \mu L_h$ successive samples of the signal $s(n-\Delta)$, and $\mathbf{g}_m(\theta)$ is a vector containing the coefficients of the interpolation kernel function. By considering L_h successive time samples of the mth microphone signal, (6) becomes a vector of length L_h:

$$\mathbf{y}_m(n) = \mathbf{G}_m(\theta)\mathbf{s}(n-\Delta) + \mathbf{v}_m(n), \tag{7}$$

where $\mathbf{G}_m(\theta)$ is a Sylvester matrix of size $L_h \times L$ created from the vector $\mathbf{g}_m^T(\theta)$, and $\mathbf{v}_m(n)$ is a vector of length L_h containing the noise samples.

Now, by concatenating the observations from the M microphones, we get a vector of length ML_h:

$$\begin{aligned} \underline{\mathbf{y}}(n) &= \begin{bmatrix} \mathbf{y}_1^T(n) & \mathbf{y}_2^T(n) & \cdots & \mathbf{y}_M^T(n) \end{bmatrix}^T \\ &= \underline{\mathbf{G}}(\theta)\mathbf{s}(n-\Delta) + \underline{\mathbf{v}}(n), \end{aligned} \tag{8}$$

where

$$\underline{\mathbf{G}}(\theta) = \begin{bmatrix} \mathbf{G}_1(\theta) \\ \mathbf{G}_2(\theta) \\ \vdots \\ \mathbf{G}_M(\theta) \end{bmatrix} \tag{9}$$

is a matrix of size $ML_h \times L$ and

$$\underline{\mathbf{v}}(n) = \begin{bmatrix} \mathbf{v}_1^T(n) & \mathbf{v}_2^T(n) & \cdots & \mathbf{v}_M^T(n) \end{bmatrix}^T \tag{10}$$

is a vector of length ML_h.

Like in LDMAs, we also assume in CDMAs that δ is small relative to the wavelength. Yet, in contrast to the linear case, herein, we allow the desired signal to arrive from all azimuthal directions and not only from the endfire. We denote the desired signal's direction as θ_d, so that the observations are

$$\underline{\mathbf{y}}(t) = \underline{\mathbf{G}}(\theta_\mathrm{d})\mathbf{s}(t-\Delta) + \underline{\mathbf{v}}(t). \tag{11}$$

Then, our objective is to design all kind of broadband CDMAs, where the main lobe is at the angle $\theta = \theta_\mathrm{d}$, with a real-valued spatiotemporal filter of length ML_h:

$$\underline{\mathbf{h}} = \begin{bmatrix} \mathbf{h}_1^T & \mathbf{h}_2^T & \cdots & \mathbf{h}_M^T \end{bmatrix}^T, \tag{12}$$

where $\mathbf{h}_m,\ m = 1, \ldots, M$ are temporal filters of length L_h.

3 Broadband beamforming

By applying the filter $\underline{\mathbf{h}}$ to the observation vector $\underline{\mathbf{y}}(n)$, we obtain the output of the broadband beamformer:

$$z(n) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{y}_m(n) = \underline{\mathbf{h}}^T \underline{\mathbf{y}}(n) = x_\mathrm{fd}(n) + v_\mathrm{rn}(n), \tag{13}$$

where

$$x_{\mathrm{fd}}(n) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{G}_m(\theta_{\mathrm{d}}) \mathbf{s}_L(n-\Delta) \\ = \underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta_{\mathrm{d}}) \mathbf{s}(n-\Delta) \tag{14}$$

is the filtered desired signal and

$$v_{\mathrm{rn}}(n) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{v}_m(n) = \underline{\mathbf{h}}^T \underline{\mathbf{v}}(n) \tag{15}$$

is the residual noise. We see from (14) that the distortionless constraint is

$$\underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta_{\mathrm{d}}) = \mathbf{i}^T, \tag{16}$$

where $\mathbf{i}$ is a column vector of length L with all its elements equal to zero except for one element. The decision which element will be non-zero can be made empirically. This constraint is always required in the design of CDMAs.

4 Design of First-Order CDMAs

In order to design first-order CDMAs we need at least $M \geq 3$ (the case of $M = 2$ coincides with the linear case already discussed in [3], [4]). Thus, for first-order CDMAs, we have three constraints to fulfill; the distortionless one given in (16) and two more symmetric constraints with nulls in the directions $\theta_{\mathrm{d}} + \theta_1$ and $\theta_{\mathrm{d}} - \theta_1$ where $\theta_1 \in \left[\frac{\pi}{2}, \pi\right]$, i.e.,

$$\underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta_{\mathrm{d}} + \theta_1) = \mathbf{0}^T \tag{17}$$

and

$$\underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta_{\mathrm{d}} - \theta_1) = \mathbf{0}^T, \tag{18}$$

where $\mathbf{0}$ is a zero vector of length L. Combining all these constraints together, we get the following linear system:

$$\begin{bmatrix} \mathbf{G}_1^T(\theta_{\mathrm{d}}) & \mathbf{G}_2^T(\theta_{\mathrm{d}}) & \cdots & \mathbf{G}_M^T(\theta_{\mathrm{d}}) \\ \mathbf{G}_1^T(\theta_{\mathrm{d}}+\theta_1) & \mathbf{G}_2^T(\theta_{\mathrm{d}}+\theta_1) & \cdots & \mathbf{G}_M^T(\theta_{\mathrm{d}}+\theta_1) \\ \mathbf{G}_1^T(\theta_{\mathrm{d}}-\theta_1) & \mathbf{G}_2^T(\theta_{\mathrm{d}}-\theta_1) & \cdots & \mathbf{G}_M^T(\theta_{\mathrm{d}}-\theta_1) \end{bmatrix} \underline{\mathbf{h}} = \underline{\mathbf{i}}_1, \tag{19}$$

or, equivalently,

$$\mathbf{C}_{1,M}(\theta) \underline{\mathbf{h}} = \underline{\mathbf{i}}_1, \tag{20}$$

where $\mathbf{C}_{1,M}(\theta)$ is a matrix of size $3L \times ML_h$ and

$$\underline{\mathbf{i}}_1 \triangleq \begin{bmatrix} \mathbf{i} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{21}$$

is a vector of length $3L$. We can solve (20) using the pseudo-inverse of $\mathbf{C}_{1,M}(\theta)$:

$$\underline{\mathbf{h}} = \mathbf{P}^{\dagger}_{\mathbf{C}_{1,M}}(\theta)\underline{\mathbf{i}}_1, \tag{22}$$

where

$$\mathbf{P}^{\dagger}_{\mathbf{C}_{1,M}}(\theta) = \left[\mathbf{C}^T_{1,M}(\theta)\mathbf{C}_{1,M}(\theta) + \lambda \mathbf{I}\right]^{-1} \mathbf{C}^T_{1,M}(\theta) \tag{23}$$

is the pseudo-inverse of the matrix $\mathbf{C}_{1,M}(\theta)$ and the scalar λ is a regularization parameter. Later, in simulations we show that this simple solution yields a frequency-invariant beampattern although no specific constraints were imposed. This is due to the fact that we deal with the DMA model which inherently provides the frequency-invariance property.

5 Performance Measures

Herein, we present some useful quality measures which we use in simulations in order to assess the performance. Assuming microphone 1 to be the reference sensor, the gain in signal-to-noise ratio (SNR) is

$$\mathscr{G}(\underline{\mathbf{h}}) = \frac{\mathrm{oSNR}(\underline{\mathbf{h}})}{\mathrm{iSNR}} = \frac{\underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta_{\mathrm{d}})\underline{\mathbf{G}}^T(\theta_{\mathrm{d}})\underline{\mathbf{h}}}{\underline{\mathbf{h}}^T \Gamma_{\underline{\mathbf{v}}} \underline{\mathbf{h}}}, \tag{24}$$

where $\Gamma_{\underline{\mathbf{v}}} = \frac{\mathbf{R}_{\underline{\mathbf{v}}}}{\sigma^2_{v_1}}$ is the pseudo-correlation matrix of $\underline{\mathbf{v}}(t)$.

The WNG is obtained by taking $\Gamma_{\underline{\mathbf{v}}} = \mathbf{I}_{ML_h}$, where $\mathbf{I}_{ML_h}$ is the $ML_h \times ML_h$ identity matrix, i.e.,

$$\mathscr{W}(\underline{\mathbf{h}}) = \frac{\underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta_{\mathrm{d}})\underline{\mathbf{G}}^T(\theta_{\mathrm{d}})\underline{\mathbf{h}}}{\underline{\mathbf{h}}^H \underline{\mathbf{h}}}. \tag{25}$$

We can also define the broadband beampattern or broadband directivity pattern as

$$|\mathscr{B}(\underline{\mathbf{h}}, \theta)|^2 = \underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta)\underline{\mathbf{G}}^T(\theta)\underline{\mathbf{h}}. \tag{26}$$

Finally, we define the DF of the array which is the gain in SNR for the case of spherical diffuse noise using the direct definition of the DF (see for example [11, ch.2]):

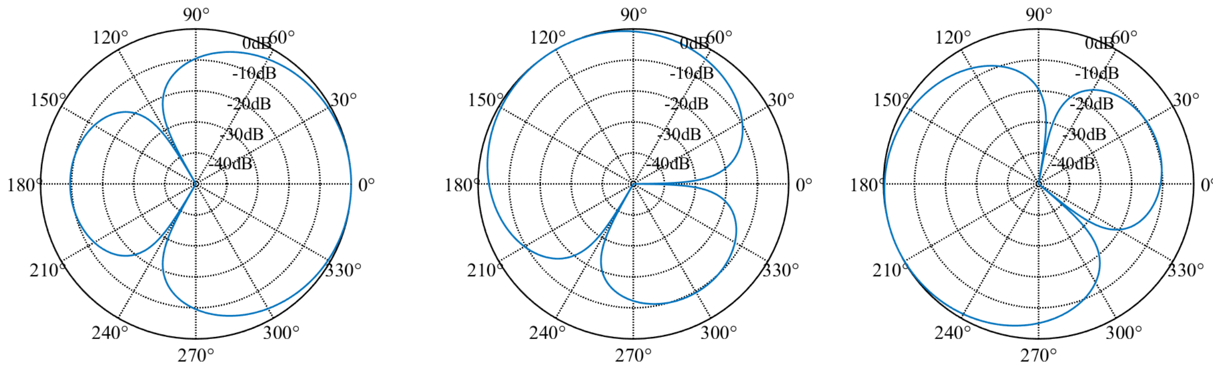
$$\mathscr{D}(\underline{\mathbf{h}}) \approx \frac{2}{\int_{\theta_{\mathrm{d}}}^{\pi+\theta_{\mathrm{d}}} |\mathscr{B}(\underline{\mathbf{h}}, \theta)|^2 \sin\theta d\theta}, \tag{27}$$

where $\mathscr{B}(\underline{\mathbf{h}}, \theta)$ is defined in (26). The last definition is a good approximation for small orders of DMAs.

6 A design example

In this section, we study the design of a first-order hypercardioid circular differential microphone array (CDMA) directivity pattern. For the case of CDMA, the hypercardioid has a distortionless response in the direction θ_d and two more null constraints: the first one is in the direction $\theta_\mathrm{d}+\theta_\mathrm{Hc}$ and the second is in the direction $\theta_\mathrm{d}-\theta_\mathrm{Hc}$, where $\theta_\mathrm{Hc}=\frac{4\pi}{3}$. We choose a sensor spacing of $\delta=$ 1cm and examine the case of $M=3$ sensors. We choose the filter length to be $L_h=12$ taps and the sampling frequency to be $f_\mathrm{s}=8000$ Hz. We choose $P=6$ taps, $\mu=0.2$, and get $L=18$ taps. The regularization parameter is set to be $\lambda=10^{-4}$.

Figure 1 shows the broadband beampattern of the time-domain implementation (26) of a first order hypercardioid for different values of the angle θ_d. These patterns are similar to those obtained with the frequency-domain implementation in [8, ch.3]. One can see that the directivity pattern is identical for each value of the presented steering direction, θ_d. Note also that for the case of $\theta_\mathrm{d}=120^\circ$, the vectors $\mathbf{h}_i, i=1,...,M$ are permutations of the same vectors for the case of $\theta_\mathrm{d}=0^\circ$. This is because both the angles $\theta_\mathrm{d}=0^\circ$ and $\theta_\mathrm{d}=120^\circ$ are the directions of two out of the three array sensors. Therefore, we can exploit this symmetry for scenarios in which only the steering angles of the M sensors' directions are required and calculate the vectors $\mathbf{h}_i, i=1,...,M$ only for the endfire direction, then, just permute between filters. This is not the case for the third case of $\theta_\mathrm{d}=200^\circ$ which is not one of the sensors' directions.



(a) (b) (c)

Figure 1: Beampatterns for the time-domain first order hypercardioid CDMA with $M=3$ sensors in different steering angles: (a) $\theta_\mathrm{d}=0^\circ$, (b) $\theta_\mathrm{d}=120^\circ$, and (c) $\theta_\mathrm{d}=200^\circ$.

We also plot in Fig. 2 the time-domain WNG and the time-domain DF as a function of the number of sensors, M, for the case of a first-order CDMA hypercardioid. One can see that the WNG is increased with the number of sensors while the DF is slightly above the value of 5 dB and does not vary at all.

7 Conclusions

We have presented a framework for time-domain implementation of first-order CDMAs, which is desirable in some applications such as real-time communications. Due to the DMA as-

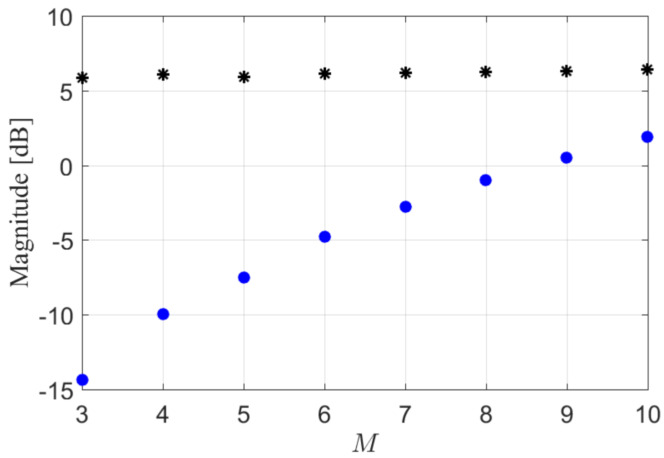


Figure 2: WNG (circles) and DF (stars) vs. M, for the case of a first-order hypercardioid.

sumption, we get a very simple solution that provides a frequency-invariant beampattern. The quality measures widely used for assessment of beamformers were also defined in the time domain. Simulation results of the proposed implementation demonstrate that it is equivalent to the frequency-domain implementation, thus providing a large amount of flexibility in the design considerations of practical systems employing CDMAs.

Acknowledgements
This Research was supported by Qualcomm Research Fund and MAFAAT-Israel Ministry of Defense.

References

[1] Elko, G. W.; Meyer, J. Microphone arrays, in Springer Handbook of Speech Processing, J. Benesty, M. M. Sondhi, and Y. Huang, Eds. Berlin, Germany: Springer-Verlag, 2008, Chapter 50, Part I, pp. 1021-1041.

[2] Benesty, J.; Souden, M.; Huang, Y. A perspective on differential microphone arrays in the context of noise reduction, IEEE Trans. Audio, Speech, Language Process., vol. 20, no. 2, pp. 699-704, Feb. 2012.

[3] Benesty, J.; Chen, J. Study and Design of Differential Microphone Arrays. Berlin, Germany: Springer-Verlag, 2012.

[4] Elko, G. W. Superdirectional microphone arrays, in Acoustic Signal Processing for Telecommunication, S. L. Gay and J. Benesty, Eds. Boston, MA: Kluwer Academic Publishers, 2000, Chapter 10, pp. 181-237.

[5] Elko, G. W.; Pong, A. T. N. A steerable and variable first-order differential microphone array, in Proc. ICASSP, 1997, pp.223-226.

[6] Chen, J.; Benesty, J. A general approach to the design and implementation of linear differential microphone arrays, in Signal and Information Processing Association Annual Summit and Conference (APSIPA), Oct, 2013, Asia-Pacific, pp. 1-7.

[7] Zhao, L.; Benesty, J.; Chen, J. Design of robust differential microphone arrays, IEEE Trans. Audio, Speech, Language Process., vol. 22, no. 10, pp. 1455-1464, Oct. 2014.

[8] Chen, J.; Benesty, J.; Cohen, I. Design of Circular Differential Microphone Arrays. Berlin, Germany: Springer-Verlag, 2015.

[9] Godara, L. C. Application of the fast Fourier transform to broadband beamforming, J. Acoust. Soc. Amer., vol. 98, no. 1, pp. 230-240, July 1995.

[10] Shannon, C. E. Communications in the presence of noise, Proc. IRE, vol. 37, pp. 10–21, 1949.

[11] Van-trees, H. L. Detection, Estimation and Modulation Theory, Part IV - Optimum Array Processing. New York: Wiley Interscience, 2002.