LEAST SQUARES FREQUENCY-INVARIANT BEAMFORMING

Lucas C. Parra

Biomedical Engineering Department
City College of New York
New York, NY, 10031
parra@ccny.cuny.edu

ABSTRACT

Frequency invariant beamforming aims to parametrize array filter coefficients such that the spectral and spatial response profiles of the array can be adjusted independently. Solutions to this problem have been presented for particular array geometries and rely often on inversion formulas for Fourier or spherical harmonics. These decompositions are analytically appealing but require a larger number of sensors and/or a regular microphone spacing. However, in practical applications the number and location of sensors are often restricted. This paper proposes to use a linear basis that optimally reproduces a desired spatial response pattern for each frequency. This numerical least-squares inversion is applicable to arbitrary sensor configurations for which typically no exact analytical inverses are available. This basis can be combined further with spherical harmonics resulting in a readily steerable and low dimensional parametrization. This solution to frequency invariant beamforming effectively decouples the array geometry from the steering geometry. Here the method is demonstrated for the optimal design of the far-field response of an irregular linear array with as few as 3 microphones combined with Legendre polynomials to control the azimuth orientation of the frequency-invariant beam.

1. INTRODUCTION

An array of spatially distributed sensors can be made selective in space and frequency by filtering and summing the output of multiple sensors. The spectro-spatial response profile is determined by the filter coefficients. Changing a given coefficient will typically affect both the frequency as well as the spatial response profile. This spectro-spatial coupling complicates filter design as well as adaptive algorithms. The goal of this work is to find a parametrization of the filter coefficients that decouples the spatial selectivity from the frequency selectivity for arbitrary array configurations.

2. FREQUENCY INVARIANT BEAMFORMING

Denote the filter coefficients of the n-th sensor in the frequency domain as $c_n(k)$. Frequency here will be given in terms of the wave number, $k = 2\pi/\lambda = \omega/c$. Consider a plane wave impinging on the array with orientation $\Omega = (\vartheta, \varphi)$, where ϑ and φ are elevation and azimuth.[1] We will refer to Ω also as the look direction. The response of the filter array with N sensors is given by

$$f(k, \Omega) = \sum_{n=1}^{N} c_n(k) g_n(k, \Omega), \tag{1}$$

where $g_n(k, \Omega)$ is the response of the n-th sensor to the plane wave with orientation Ω at frequency k. For sensors with omnidirectional response located at $\mathbf{r}'_n = (r'_n, \Omega'_n)$ this is given by [1]

$$g_n(k, \Omega) = \sum_{l=0}^{\infty} i^l j_l(kr'_n) \sum_{m=-l}^{l} Y_l^m(\Omega) Y_l^{m*}(\Omega'_n), \tag{2}$$

where $j_l(kr)$ are spherical Bessel functions of the first kind, $Y_l^m(\Omega)$ are spherical harmonics, and $i^2 = -1$. The response of a spherical array is given by (2) with $r'_n = r'$. For a circular array with $\vartheta'_n = \pi/2$ and $\vartheta = \pi/2$ Equation (2) becomes [2]

$$g_n(k, \Omega) = \sum_{l=-\infty}^{\infty} i^l J_l(kr') e^{il(\varphi - \varphi'_n)}, \tag{3}$$

where $J_l(kr)$ are Bessel functions of the first kind. For a linear array with $\varphi'_n = 0$ it is [3]

$$g_n(k, \Omega) = e^{-ikr'_n \cos\vartheta}. \tag{4}$$

Equation (1) can be seen as parametrization of the filter-array response for each frequency k with coefficients $c_n(k)$, and basis functions $g_n(k, \Omega)$. Modifying coefficients $c_n(k)$ will affect the frequency and spatial response simultaneously because $g_n(k, \Omega)$ depends on both the frequency k and look direction Ω. The goal of frequency-invariant beamforming is to find a new parametrization for $c_n(k)$

$$c_n(k) = \sum_{m=1}^{M} b_{nm}(k) \tilde{c}_m(k) \tag{5}$$

such that the basis transform $b_{nm}(k)$ converts the array response into a frequency-invariant array response

$$\sum_{n=1}^{N} g_n(k, \Omega) b_{nm}(k) = \tilde{g}_m(\Omega). \tag{6}$$

The basis transform replaces the N sensors indexed with n by a new set of M *virtual* sensors indexed with m. These virtual sensors are now frequency-invariant. This basis transform, in turn, factorizes the filter-array response, which is seen by combining Equations (1), (5), and (6)

$$f(k, \Omega) = \sum_{m=1}^{M} \tilde{c}_m(k) \tilde{g}_m(\Omega). \tag{7}$$

[1] The notation here will consider only the far-field response. But it is possible to generalize the argument to the near-field by replacing plane waves of orientation Ω with point sources located at positions $\mathbf{r}$.

Comparing this equation with (1) it becomes clear that now modifying the new parameters $\tilde{c}_m(k)$ will affect the spatial response profile only at frequency k. In fact, the spatial response profile of the filter array is fully determined by coefficients $\tilde{c}_m(k)$. In particular, a *frequency-invariant beamformer* is obtained by choosing the same coefficients for all frequencies, $\tilde{c}_m(k) = \tilde{c}_m$.

3. ANALYTIC INVERSION APPROACHES

Obviously the challenge lies in finding a basis transform $b_{nm}(k)$ that satisfies (6) – even if only approximately. Elegant approximations based on analytic inversion formulas have been presented for spherical [1], circular [4, 2], linear [3], and rectangular [5] arrays. Equations (2)-(4) may already give an indication of possible approaches. To give a rough idea of these methods consider the case of a circular array. For a spherical array one may choose

$$b_{nm}(k) = N^{-1} i^{-m} J_m^{-1}(kr)\, e^{-im\varphi_n}. \tag{8}$$

Inserting this in (6) gives indeed a frequency-invariant function of space, $\tilde{g}_m(\varphi) = e^{-m\varphi}$, assuming that the following orthogonality condition holds for any l, m

$$\frac{1}{N}\sum_{n=1}^{N} e^{i(l-m)\varphi_n} = \delta_{lm}. \tag{9}$$

Unfortunately this is only correct for $|l| \leq N$, and only if the N sensors are placed on an equidistant lattice along the circular array. As a result, equation (6) is only approximately correct. For circular and similarly for spherical arrays the approximation can only be improved with a larger number of sensors and the lattice must be carefully arranged to match the analytic inversion formulas [1, 4]. The same is true for linear and rectangular arrays. In addition, for those configurations the angular response profile at any given frequency only partially determines the required Fourier basis coefficients [3]. Arbitrary Fourier coefficients are chosen outside the determined range, further compromising the accuracy of the approximations. Finally, spherical and circular arrays may not be feasible in some practical applications.

4. LEAST SQUARES SOLUTION

In this work we aim to directly minimize the approximation error resulting from the restricted number of sensors and overcome the restrictions on sensor locations imposed by the analytic inversions. We suggest to simply invert Equation (6) numerically. To this end discretize the look directions with L angles Ω_j and write the equation (6) in matrix notation

$$\mathbf{G}(k)\mathbf{B}(k) = \tilde{\mathbf{G}}, \tag{10}$$

where $[\mathbf{G}(k)]_{jn} = g_n(k, \Omega_j)$, $[\mathbf{B}(k)]_{nm} = b_{nm}(k)$, and $[\tilde{\mathbf{G}}]_{jm} = \tilde{g}_m(\Omega_j)$. The new desired spatial basis vectors can be defined in the columns of $\tilde{\mathbf{G}}(k)$ while $\mathbf{G}(k)$ is determined by the array configuration. For each frequency k Equation (10) specifies ML conditions with MN unknowns. Typically we have only a few sensors and would like to parametrize the response for many different look directions. With $L > N$ the problem is therefore over-determined. The *Least Squares* solutions to this problem, i.e. the $\mathbf{B}$ that will reproduce $\mathbf{G}$ with the smallest square error, is computed with the pseudo-inverse, $\mathbf{G}^\dagger = (\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H$,

$$\mathbf{B}(k) = \mathbf{G}^\dagger(k)\tilde{\mathbf{G}}. \tag{11}$$

Note that the discretization of the look directions is only used to compute the new basis $b_{in}(k)$. When applying the basis the look direction can take on any arbitrary values. For a given array geometry the basis is computed once for each frequency k and remains unaltered by subsequent beam design.

The accuracy of the proposed method depends on how well $\mathbf{G}$ can be inverted, i.e. how well can $\tilde{\mathbf{G}}$ be represented by $\mathbf{b}$: $\varepsilon = \|\mathbf{G}\mathbf{B} - \tilde{\mathbf{G}}\|^2 = \|\mathbf{G}\mathbf{G}^\dagger\tilde{\mathbf{G}} - \tilde{\mathbf{G}}\|^2$. This is the minimum attainable square error in satisfying the factorization condition (6). Any other basis transform is suboptimal (in the least squares sense).

5. BEAM STEERING

Ideally the new spatial basis $\tilde{g}(\Omega)$ should be easily steerable. With steering we mean that the overall orientation of a beam can be rotated without changing its spatial profile. The convolution theorem of the spherical harmonics provides the required property: Multiplying the transformed domain coefficients is equivalent to convolving in angular space. To rotate a given beam pattern by and angle Ω_o one multiplies therefore with the coefficients of the corresponding shift operator, i.e. the delta function $\delta(\Omega - \Omega_o)$. The natural choice for $\tilde{g}_m(\Omega)$ will be therefore spherical harmonics [1] (omitting index l here for consistency)

$$\tilde{g}_m(\Omega) = Y_l^m(\Omega). \tag{12}$$

In case of a linear array one can only steer in ϑ direction (because of the radial symmetry in φ) in which case the spherical harmonics simplify to the Legendre polynomials

$$\tilde{g}_m(\Omega) = P_m(\cos\vartheta). \tag{13}$$

A basis transform $b_{nm}(k)$ that converts the array response $g_n(k, \Omega)$ to this desired array response $\tilde{g}_m(\Omega)$ effectively decouples the problem of steering the geometry from that of the array geometry. In previous approaches the ability to steer seemed inevitably linked to the choice of array architecture and the corresponding analytic inversion formulas.

6. APPLICATION TO ADAPTIVE BEAMFORMING DESIGN

The time domain output $y(t)$ of the filter array in response to sensor readings $x_n(t)$ is given by the convolution and sum

$$y(t) = \sum_{n=1}^{N} c_n(t) * x_n(t) = \sum_{m=1}^{M} \tilde{c}_m(t) * \tilde{x}_m(t). \tag{14}$$

The second equality here results from Equation (5) and the following definition of the *frequency-invariant virtual sensor* readings

$$\tilde{x}_m(t) = \sum_{n=1}^{N} b_{nm}(t) * x_n(t). \tag{15}$$

This suggests that adaptive algorithms which are driven by the sensor observations can be applied to the signals of the newly defined frequency-invariant virtual sensors. Instead of optimizing filter parameters $c_n(k)$ based on sensor readings $x_n(t)$ the adaptive algorithm now adapts parameters $\tilde{c}_m(k)$ based on virtual sensor readings $\tilde{x}_m(t)$. All known adaptive beamforming algorithms are therefore immediately applicable without further modification such as generalized sidelobe-canceling, blind sources separation, and others.

To find an optimal *frequency-invariant* response the adaptive algorithm will now optimize the (frequency-invariant) parameters $\tilde{c}_n$. If the algorithm is based on a gradient of some cost function $E(\{\tilde{c}_n(k)\})$ with a set of frequency dependent parameters $\{\tilde{c}_m(k)\}$ the frequency-invariant gradient is then simply the original gradient summed over all frequencies

$$\frac{\partial E}{\partial \tilde{c}_m} = \sum_k \frac{\partial E}{\partial \tilde{c}_m(k)}. \tag{16}$$

Note that this has the potential to significantly reduce the number of free parameters. For most adaptive algorithms this will result in a significant improvements in convergence speed as well as estimation accuracy. This comes as additional benefit to the potential advantage of a frequency-invariant response.

7. IMPLEMENTATION FOR IRREGULAR LINEAR ARRAY

The proposed method was implemented for a linear array of omnidirectional sensors and Legendre polynomials as the beamforming basis i.e. Equation (4) and (13). Notice that Equation (4) does not require equidistant sensor placement. The least-squares solution of Equation (11) is used to compute the basis transform. To this end the angular space (which now contains only elevation ϑ) was discretized with L samples $z_j = \cos(\vartheta_j)$ in the range $z \in [-1, 1]$ and equidistant spacing Δz.[2] Choosing Legendre polynomials (13) as the frequency-invariant virtual array response in Equation (11) means that we set with $[\mathbf{P}]_{mj} = P_m(z_j)$

$$\tilde{\mathbf{G}} = \mathbf{P}^T. \tag{17}$$

To demonstrate the effectiveness of the resulting basis we will now determine the optimal coefficients $\tilde{\mathbf{c}}$ that produce a desired frequency-invariant response. Assume the prescribed response is specified as a vector $\mathbf{f}$ with coefficients f_d, each of which represents the desired response for angles $\theta_d, d = 1, \ldots, D$. The response of the array at those angles can be written in matrix notation as

$$\mathbf{f} = \mathcal{G}\mathbf{B}\tilde{\mathbf{c}}, \tag{18}$$

where the coefficients of matrix $\mathcal{G}$ are given by $[\mathcal{G}]_{dn} = g_n(\theta_d)$ specifying the response of the n-th sensor for the desired angle θ_d. The goal is to find the coefficient $\tilde{\mathbf{c}}$ that reproduce the desired response $\mathbf{f}$. With N sensor one can satisfy at most N conditions on the response. A larger number of conditions can only be satisfied approximately. The coefficients that reproduce $\mathbf{f}$ with the least squares error are given by

$$\tilde{\mathbf{c}} = (\mathcal{G}\mathbf{B})^{\dagger}\,\mathbf{f}. \tag{19}$$

[2] Equidistant spacing in z rather than ϑ is not essential for equation (11). However, it is beneficial for the parametrization with the Legendre polynomials. With this discretization we can transform a function $f(z)$ to its Legendre coefficients $\tilde{c}_m$ by applying a matrix $\mathbf{P}$ to vector $[\mathbf{f}]_j = f(z_j)$: $\tilde{\mathbf{c}} = \mathbf{P}\mathbf{f}$. This transformation is then inverted by applying the transpose $\mathbf{P}^T$ to vector $[\tilde{\mathbf{c}}]_m = \tilde{c}_m$: $\mathbf{f} \approx \mathbf{P}^T\tilde{\mathbf{c}}$. This last approximation is exact if $\dim(\mathbf{f}) = \dim(\tilde{\mathbf{c}})$, while the error is well controlled for $\dim(\mathbf{f}) > \dim(\tilde{\mathbf{c}})$ [6]. Note that this error is not essential for the current method. It only matters in terms of defining a desired spatial profile with a limited number of coefficients. The true limitation and accuracy of the proposed approach lies is the error ε discussed above.

To demonstrate the advantage of using the Legendre basis we will compare this with the coefficient obtained in a *naive* frequency-invariant basis with $[\tilde{G}]_{jm} = \delta_{jm}$ in Equation (11). In that case the previous equation becomes

$$\tilde{\mathbf{c}} = \left(\mathcal{G}\mathbf{G}^{\dagger}\right)^{\dagger}\mathbf{f}. \tag{20}$$

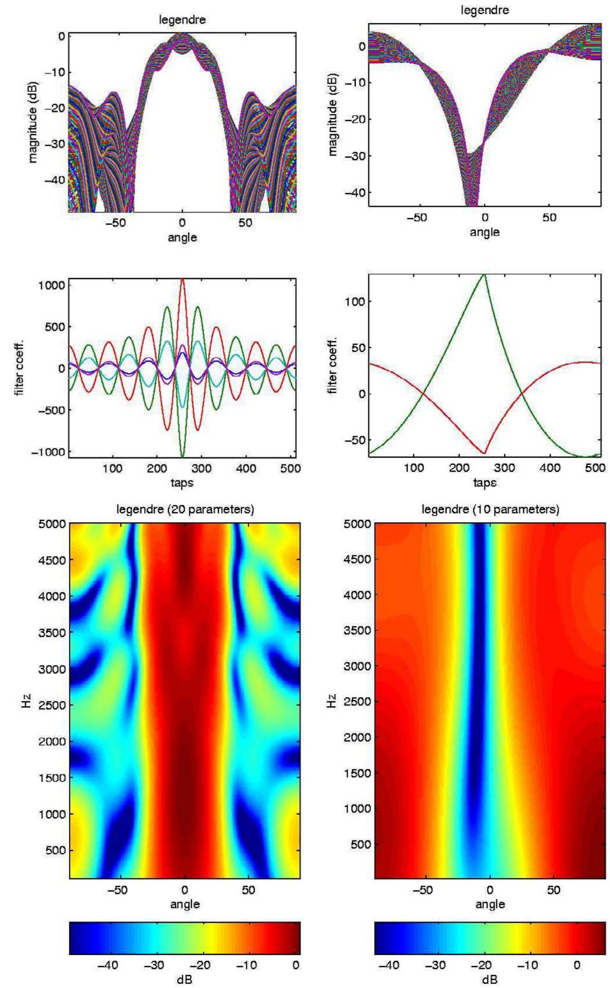

Figure 1: Magnitude response of frequency-invariant filter array as a function of azimuth and frequency (top and bottom rows). Middle row shows corresponding time domain coefficients. Left: 5 sensors positioned at $r_n = 0, 5, 7.5, 17.5, 20$ in cm. Desired response specified as $f_d = 0, 0, 1, 0, 0$ for angles $\vartheta_d = -80^o, -40^o, 0^o, 40^o, 80^o$. Right: 3 sensors positioned at $r_n = 0, 3, 6$ in cm. Desired response specified as $f_d = 1, 0, 1$ for angles $\theta_d = -80^o, -10^o, 60^o$.

8. RESULTS

In Equations (18)-(20) the dependence on frequency k was omitted for simplicity. We compute the optimal parameters with (19) or

(20) and use their average value across frequencies. Though sub-optimal we find this result not only simpler but also more robust compared to the 'optimal' solution which would require combining all frequencies prior to computing the pseudo-inverses in (19) or (20).

When computing the basis $\mathbf{B}$ with (11) we find that for the lowest frequencies $\mathbf{G}^H(k)\mathbf{G}(k)$ is ill-conditioned and therefore cannot be accurately inverted in the pseudo-inverse $\mathbf{G}^\dagger(k)$. This is to be expected since for very low frequencies (large wavelength) a limited aperture prevents effective spatial resolution. Similarly, for high frequencies the finite spacing of sensors (recall the goal of using only a small number of sensors) generates aliasing side-lobes, once again leading to a non-invertible $\mathbf{G}^H(k)\mathbf{G}(k)$. This limitation is inherent to *any* beamforming design and is typically resolved by restricting the frequency band of operation. This has to be taken into account in particular for frequency-invariant beamforming designs since otherwise non-invertible frequency bands will disrupt the estimates of coefficients which are applicable across all frequencies. For the examples described below we assume acoustic sensors with a sound propagation speed of $c = 342\text{ms}^{-1}$. For an aperture of $20cm$ we find a useful frequency range of at least 20Hz-5000Hz.

Figure 1 shows the result obtained for a linear array of omni-directional sensors (5 or 3 sensors) with irregular spacing and an aperture of 20 cm or 6 cm. The arrays were optimized to generate a maximum or a zero for a given respectively. We specified only $D = N$ conditions to guarantee that these would be satisfied exactly. Only 20 (and 10) Legendre coefficients were needed to specify the $512 \times N$ time/frequency domain filter coefficients. This represents a significant reduction in the number of free parameters. Figure 2 compares the results obtained with a Legendre basis and the 'naive' basis as described in equations (19) and (20) respectively. The Legendre basis uses 20 times fewer coefficients without compromising the desired spatial response pattern.

We close by noting that the instability in the inverse discussed above results in rather large gains which may arbitrarily magnify sensor noise. Future work will present a solution that regularizes the inverse by considering additive noise. This will also prevent the need to manually select an operating bandwidth.

9. REFERENCES

[1] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on anorthonormal decomposition of the soundfield," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2002, pp. 1781–1784.

[2] H. Teutsch and W. Kellermann, "EB-ESPRIT: 2D localization of multiple wideband acoustic sources using eigen-beams," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, 2005, pp. 89–92.

[3] T. Sekiguchi and Y. Karasawa, "Wideband beamspace adaptive array utilizing FIR fan filters for multibeam forming," *IEEE Transactions on Signal Processing*, vol. 48, no. 1, pp. 277–284, January 2000.

[4] S. Chan and C. Pun, "On the design of digital broadband beamformer for uniform circular array with frequency invariant characteristics," in *IEEE International Symposium on Circuits and Systems*, vol. 1, 2002, pp. 693–696.

[5] W. Liu and W. Stephan, "A new class of broadband arrays with frequency invariant beam patterns," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2004, pp. 185–188.

[6] R. Kennedy, T. Abhayapala, and D. Ward, "Broadband nearfield beamforming using a radial beampattern transformation," *IEEE Transactions Signal Processing*, vol. 46, no. 8, pp. 2147–2156, August 1998.

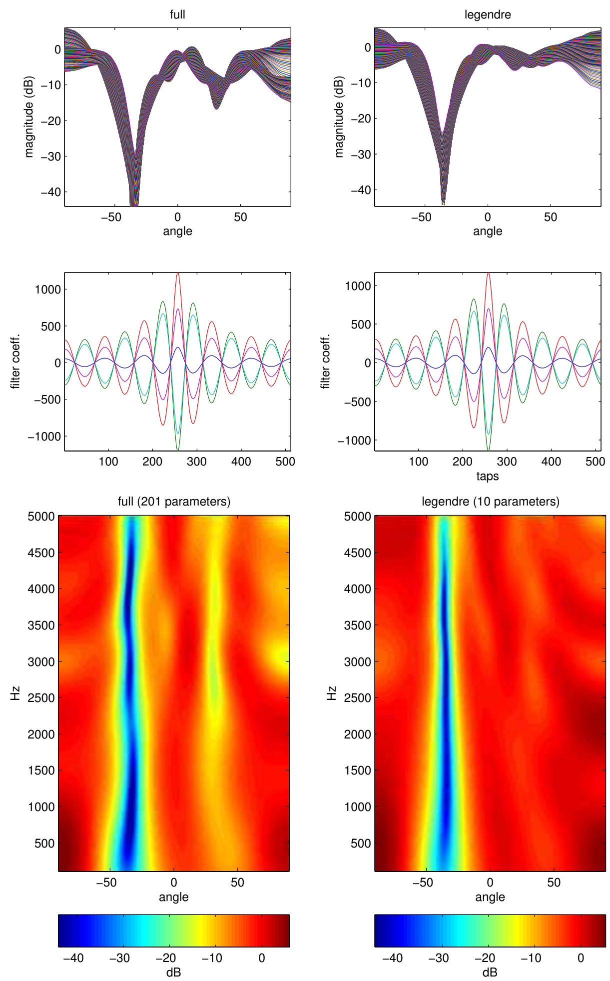

Figure 2: Magnitude response of frequency-invariant filter array. Left showing 'naive' parametrization and right the parametrization with Legendre basis. The 5 sensors are positioned at $r_n = 0, 5, 7.5, 17.5, 18.7$ in cm. Desired response specified as $f_d = 1, 0, 1, 1$ for angles $\theta_d = -81^o, -36^o, 9^o, 54^o$.