NEAR-FIELD BEAMFORMING FOR MICROPHONE ARRAYS

James G. Ryan

Institute for Microstructural Sciences
National Research Council of Canada
Ottawa, Ontario
Canada K1A 0R6
james.ryan@nrc.ca

Rafik A. Goubran

Systems and Computer Engineering
Carleton University
Ottawa, Ontario
Canada K1S 5B6
goubran@sce.carleton.ca

ABSTRACT

This paper describes the application of array optimization techniques to improving the near-field response of an arbitrary microphone array. The optimization exploits the differences in wavefront curvature between near-field and far-field sound sources and is suitable for reverberation reduction in small rooms. The optimum near-field beamformer provides increased array gain over that obtained from a uniformly weighted delay-and-sum beamformer.

1. INTRODUCTION

Conventional analysis and synthesis techniques for microphone arrays are based on the simplifying assumption that all acoustic sources are located far from the array. In this case wavefront curvature can be neglected and all waves impinging upon the array are assumed to be planar. This assumption simplifies analysis and provides a realistic framework for certain applications [1].

The assumption of plane-wave propagation results in errors, however, when speech-bandwidth arrays are deployed in small rooms. Since the size of an array is usually related to its operating wavelength, the aperture of a speech-bandwidth array can be of the same order of magnitude as the distance to the sound source. In such cases, the sound source is located in the array's near field and wavefront curvature can be detected within the array's aperture. This fact has been considered in the recent design of broadband near-field microphone arrays [2].

For a sound source in a reverberant room, the direct wavefront from a near-field source appears more curved than the wavefronts from reflections. Figure 1 shows an acoustic source located in the near field of a linear array in a reverberant room. Wall reflections are represented by the image sources behind the reflecting surfaces. Since the image sources are farther from the array than the original sound source, the corresponding wavefronts are less curved when observed at the array location.

By focusing an array on a near-field point, it is possible to provide a certain degree of range discrimination. This discrimination permits the array to reduce the effects of reflected wavefronts arriving from the same angle but a greater distance than the original sound source.

This paper describes an optimization for the near-field response of a microphone array that attempts to improve performance in reverberation. The optimization increases the array's rejection of reverberation based on differences in wavefront curvature. Using an appropriate mathematical description of the near-field performance permits existing techniques to be applied

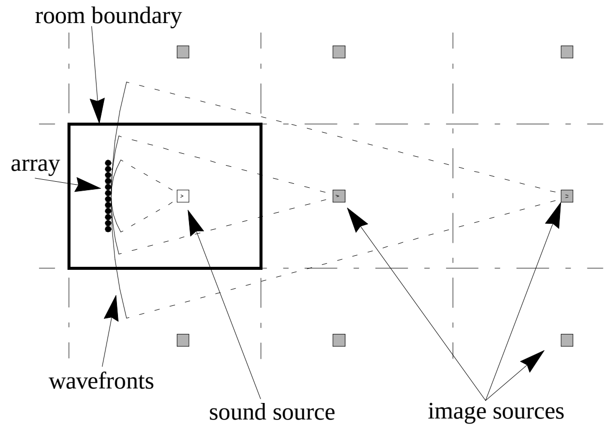

Figure 1 Acoustic source in the near field of a linear array in a reverberant room. Reflected wavefronts are less curved than the direct wavefront due to the longer acoustic path travelled.

to the array optimization.

2. ARRAY OPTIMIZATION

2.1. Gain Optimization

The gain $G(f)$ provided by an arbitrary, M-element array at a frequency f can be expressed in matrix notation as [3]

$$G(f) = \frac{\boldsymbol{w}^{\dagger}(f)\boldsymbol{R}_{ss}(f)\boldsymbol{w}(f)}{\boldsymbol{w}^{\dagger}(f)\boldsymbol{R}_{nn}(f)\boldsymbol{w}(f)}. \tag{1}$$

where $\boldsymbol{R}_{ss}(f)$ is the desired-signal correlation matrix, $\boldsymbol{R}_{nn}(f)$ is the noise correlation matrix, $\boldsymbol{w}(f)$ is the vector of complex microphone weights and the superscript $\dagger$ denotes the conjugate transpose. The matrices $\boldsymbol{R}_{ss}(f)$ and $\boldsymbol{R}_{nn}(f)$ have been normalized to the total signal and noise powers, respectively.

Equation (1) can be recognized as the ratio of two quadratic forms. Consequently, there exists a vector $\boldsymbol{w}_{\mathrm{opt}}(f)$ that maximizes the value of $G(f)$. Provided that $\boldsymbol{R}_{nn}(f)$ is non-singular, $\boldsymbol{w}_{\mathrm{opt}}(f)$ is given by

$$\boldsymbol{w}_{opt}(f) = \alpha \boldsymbol{R}_{nn}^{-1}(f)\boldsymbol{s}(f) \tag{2}$$

where $\boldsymbol{s}(f)$ is the signal propagation vector defined such that

$$\boldsymbol{R}_{ss}(f) = \boldsymbol{s}(f)\boldsymbol{s}^{\dagger}(f)$$

and α is an arbitrary complex constant. Inclusion of the complex constant α implies that the optimum weight vector is only unique to within a constant scale factor. In fact, it has been shown that the optimum weight solutions for several different processing strategies can all be expressed using equation (2) [3]. The resulting maximum array gain $G(f)_{\mathrm{opt}}$ is

$$G(f)_{\mathrm{opt}} = \boldsymbol{s}^{\dagger}(f)\boldsymbol{R}_{nn}^{-1}(f)\boldsymbol{s}(f)\,. \tag{3}$$

Specific solutions for $\boldsymbol{w}_{\mathrm{opt}}(f)$ are determined by the exact values of the signal and noise correlation matrices, $\boldsymbol{R}_{ss}(f)$ and $\boldsymbol{R}_{nn}(f)$. The specific case of interest in this paper is that for which little detailed information is available about the noise characteristics. In such cases, it is often most convenient to assume a spherically isotropic noise field. That is, assume a noise correlation matrix $\boldsymbol{Q}(f)$

$$\boldsymbol{R}_{nn}(f) = \boldsymbol{Q}(f) \tag{4}$$

where the element corresponding to the m^{th} row and l^{th} column is given by

$$q_{ml} = \frac{\sin k\left|\vec{\zeta}_l - \vec{\zeta}_m\right|}{k\left|\vec{\zeta}_l - \vec{\zeta}_m\right|}\,. \tag{5}$$

In this equation $\vec{\zeta}_m$ is a position vector representing the arbitrary location of microphone m and k is the wavenumber. Thus, $\left|\vec{\zeta}_l - \vec{\zeta}_m\right|$ is the spatial separation between microphones m and l.

Optimum array designs obtained with this technique can be classified according to the microphone spacing. Conventional array behavior is observed for microphone spacing that is large with respect to the wavelength λ. For small microphone spacing (less than $\lambda/2$) optimum array designs are super-directive producing large increases in gain at the expense of error sensitivity.

For example, when the distances between the microphones $\left|\vec{\zeta}_l - \vec{\zeta}_m\right|$ are an integer multiple of $\lambda/2$ (as in the case of a uniformly spaced linear array) then $\boldsymbol{Q}(f)$ reduces to an $M \times M$ identity matrix. The optimum microphone weights are then readily obtained

$$\boldsymbol{w}_{opt}(f) = \alpha\boldsymbol{s}(f)\,. \tag{6}$$

That is, the optimum microphone weights are equal to the scaled signal propagation vector. The optimum array gain is then

$$G(f)_{\mathrm{opt}} = \boldsymbol{s}^{\dagger}(f)\boldsymbol{s}(f) = \mathrm{trace}(\boldsymbol{R}_{ss})\,. \tag{7}$$

For an array steered to a far-field source this reduces to the familiar result for conventional, delay-and-sum beamforming

$$G(f)_{\mathrm{opt}} = M\,. \tag{8}$$

When the microphone spacing is less that $\lambda/2$, the optimum weights are not simply a scaled version of the signal propagation vector. Instead, the optimum weights include a phase and amplitude adjustment at each microphone in addition to that required for wavefront alignment. It has been shown that the theoretical limit for super-directive array gain is M^2 and is achieved for an endfire linear array as the microphone spacing approaches zero. This beamformer, however, suffers from an extreme sensitivity to errors and is of little practical use. Thus, array optimization for practical applications must also include a constraint to limit the array's sensitivity to such errors.

2.2. Constrained Optimization

To reduce the beamformer sensitivity to errors, it is necessary to constrain the weight optimization in such a way that limits the array's sensitivity to errors. This is accomplished most straight forwardly by adding a diagonal component to the noise correlation matrix $\boldsymbol{R}_{nn}(f)$. That is, define $\boldsymbol{R}_{nn}(f)$ to be

$$\boldsymbol{R}_{nn}(f) = \boldsymbol{Q}(f) + \varepsilon(f)\boldsymbol{I}\,. \tag{9}$$

where $\varepsilon(f)$ is a small positive, real constant which is unique for each frequency. The strength of the constraint is controlled by the magnitude of $\varepsilon(f)$. Setting $\varepsilon(f)$ to a large value implies that the dominant noise is uncorrelated from microphone to microphone. When uncorrelated noise is dominant, the optimum weights are those of a conventional delay-and-sum beamformer. Setting $\varepsilon(f) = 0$, of course, produces the unconstrained optimum array.

The sensitivity of an array to uncorrelated errors is indicated by a parameter known as white noise gain $G_w(f)$. It is calculated as the value of array gain assuming that the dominant noise is uncorrelated from microphone to microphone (i.e. $\boldsymbol{R}_{nn} = \boldsymbol{I}$). That is [4]

$$G_w(f) = \frac{\boldsymbol{w}^{\dagger}(f)\boldsymbol{R}_{ss}(f)\boldsymbol{w}(f)}{\boldsymbol{w}^{\dagger}(f)\boldsymbol{w}(f)} \le M\,.$$

This parameter represents the array gain *against* uncorrelated noise therefore a large value of $G_w(f)$ means that the array is less sensitive. Unfortunately, there is no simple relationship between the constraint parameter $\varepsilon(f)$ and the resulting value of white noise gain $G_w(f)$. Designing an array for a prescribed value of $G_w(f)$ requires an iterative procedure.

This constrained optimization approach has been used successfully in the design of practical super-directive microphone arrays for hearing-aid applications [6].

3. NEAR-FIELD ARRAY OPTIMIZATION

3.1. Unconstrained Optimization

Array gain optimization may also be applied to the near-field response of an array. Apply of this technique simply requires a suitable choice be made for the matrices $\boldsymbol{R}_{ss}(f)$ and $\boldsymbol{R}_{nn}(f)$.

In a reverberant environment, since the direct wavefront travels the shortest distance to the array it possesses the smallest radius of curvature. Waves which have undergone multiple reflections possess larger radii of curvature. As the distance travelled by the reflected wave increases, the reflected waves may be accurately modeled as plane waves. Of course, the reverberant field is composed of waves propagating in many different directions. Thus, a suitable model for reverberation is a spherically isotropic noise field. Insofar as the reflected wavefronts appear less curved than the direct wave, this optimization should increase the ratio of direct to reverberant energy at the output of the array.

For a point source located at $\vec{\Upsilon}$ in the near field of an M element array, the signal propagation vector is

$$\boldsymbol{s}^T = \left|\vec{\Upsilon}\right| e^{-jk\left|\vec{\Upsilon}\right|} \left[\frac{e^{jk\left|\vec{\zeta}_0 - \vec{\Upsilon}\right|}}{\left|\vec{\zeta}_0 - \vec{\Upsilon}\right|} \; \frac{e^{jk\left|\vec{\zeta}_1 - \vec{\Upsilon}\right|}}{\left|\vec{\zeta}_1 - \vec{\Upsilon}\right|} \; \cdots \; \frac{e^{jk\left|\overrightarrow{\zeta_{M-1}} - \vec{\Upsilon}\right|}}{\left|\overrightarrow{\zeta_{M-1}} - \vec{\Upsilon}\right|} \right]$$

where the factor $\left|\vec{\Upsilon}\right| e^{-jk\left|\vec{\Upsilon}\right|}$ is included to normalize the array response to that of a single omnidirectional microphone located at the coordinate origin. Using this normalization, the trace of the resulting $\boldsymbol{R}_{ss}(f)$ matrix equals

$$\mathrm{trace}(\boldsymbol{R}_{ss}) = \boldsymbol{s}^{\dagger}\boldsymbol{s} = \sum_{m=0}^{M-1} \frac{\left|\vec{\Upsilon}\right|}{\left|\vec{\zeta}_m - \vec{\Upsilon}\right|},$$

so the numerical value of the near-field response approaches the far-field response as the source distance increases.

Since the interference assumed is spherically isotropic the definition of $\boldsymbol{R}_{nn}(f)$ is identical to that in (4) and (5). Consequently, the matrix $\boldsymbol{R}_{nn}(f)$ is invertible and the optimum weight vector given by equation (2).

As before, when the distance between the microphones equals an integer multiple of a half-wavelength then the noise correlation matrix reduces to an $M \times M$ identity matrix and the optimum microphone weights are equal to the signal propagation vector. That is

$$\boldsymbol{w}_{\mathrm{opt}}^{\ T} = \left|\vec{\Upsilon}\right| e^{-jk\left|\vec{\Upsilon}\right|} \left[\frac{e^{jk\left|\vec{\zeta}_0 - \vec{\Upsilon}\right|}}{\left|\vec{\zeta}_0 - \vec{\Upsilon}\right|} \; \frac{e^{jk\left|\vec{\zeta}_1 - \vec{\Upsilon}\right|}}{\left|\vec{\zeta}_1 - \vec{\Upsilon}\right|} \; \cdots \; \frac{e^{jk\left|\overrightarrow{\zeta_{M-1}} - \vec{\Upsilon}\right|}}{\left|\overrightarrow{\zeta_{M-1}} - \vec{\Upsilon}\right|} \right]. \quad (10)$$

Note that, in contrast to the far-field case, the optimum microphone weights are not of equal magnitude. Instead, the magnitude is inversely proportional to the distance from the microphone to the focal point. Thus, to achieve the maximum near-field array gain those microphones which are closest to the focal point must be given the highest weighting. microphones which are further from the focal point are given the lowest weighting and contribute less to the array output. This is intuitively appealing since signals from the closest microphones will have the highest ratio of near/far energy prior to beamforming.

For microphone spacing much less that $\lambda/2$ the optimum weights display a super-directive behavior similar to that observed in the far-field. As in the far-field case, substantial increases in the optimization index are produced at the expense of sensitivity to wavefront perturbations and other types of uncorrelated errors. Thus, practical designs must employ constrained optimization to limit the error sensitivity.

4. OPTIMIZATION RESULTS

Although the optimization method applies to an arbitrary array, for illustrative purposes a linear array consisting of 15 uniformly spaced microphones will be used.

4.1. Half-wavelength Spacing

The optimum near-field weights and corresponding array

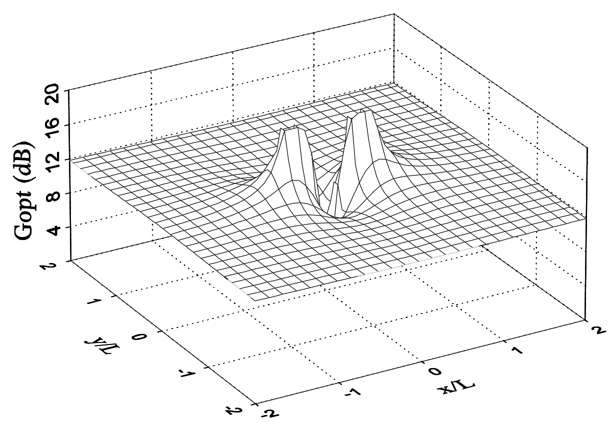

Figure 2 $G_{\mathrm{opt}}(f)$ obtained in the near-field of a 15 element linear array with microphone spacing of $\lambda/2$. The near field extends to a distance of $3.5L$.

gain were calculated for an array with a microphone spacing of $\lambda/2$. For an array of total length L, the extent of the near field region r_{nf} can be defined by [5]

$$r_{\mathrm{nf}} = L^2/2\lambda.$$

For the present example, this means that the near field extends to a distance of $3.5L$ where L is the array length.

Data were calculated for a regular grid of points over a square region surrounding the array. Position is described by the rectangular coordinates (x, y). The square region of interest is bounded by $-2L \le x \le 2L$ and $-2L \le x \le 2L$. This region is sufficient to include most of the near-field while providing sufficient detail in the plots. In all the plots to follow, the location coordinates are normalized to the array length and the array is located along the x axis from $x = -0.5$ to $x = 0.5$.

The effect of the optimization is illustrated in Figure 2 where $G_{\mathrm{opt}}(f)$ is plotted as a function of near-field position. As seen in the figure, at large distances, the near-field gain approaches the far-field gain (i.e. $10\log M = 10\log 15 = 11.8\mathrm{dB}$). For sources located at the reference point (coordinate origin) the array performance approaches that of a single microphone (i.e. 0 dB).

For positions in the immediate vicinity of the individual microphones, the gain is high due to the optimal use of the closest microphone. As the focal point approaches a microphone location, all other microphones are turned off to limit the introduction of reverberation.

To highlight the impact of the optimization, the difference between $G_{\mathrm{opt}}(f)$ and the gain obtained from a uniformly weighted delay-and-sum beamformer is illustrated in Figure 3. The difference is positive for all locations although the improvement is small (less than 1dB) for distances greater than approximately $0.2L$. The largest improvements are noted for positions close to the microphone positions further highlighting the fact that uniform weighting is not optimal for near-field sources.

4.2. Spacing Less than Half-wavelength

For microphone spacing less than $\lambda/2$, the optimization

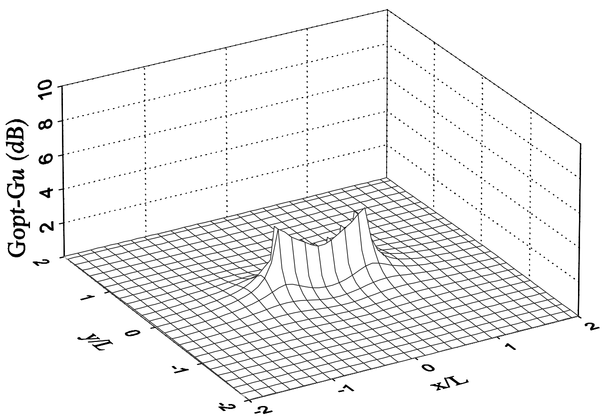

Figure 3 Difference between $G_{\text{opt}}(f)$ and the gain provided by uniform delay-and-sum beamforming $G_{\text{u}}(f)$ for a 15 element linear array with $\lambda/2$ microphone spacing. The near field extends to a distance of $3.5L$.

procedure produces a super-directive array design with unacceptably high sensitivity to errors. To avoid this, the constraint of equation (9) was used for arrays with microphone spacing less than $\lambda/2$. For the data presented here, the parameter $\varepsilon(f)$ was adjusted to provide a white noise gain of 0 dB. This should produce a beamformer whose sensitivity to errors is no greater than that of a single microphone.

The constrained optimization was applied to the design of an array with microphone spacing of $\lambda/4$. In this case, the near field extends to a distance of $1.75L$. The improvement in performance compared to a uniformly weighted delay-and-sum beamformer is illustrated in Figure 4.

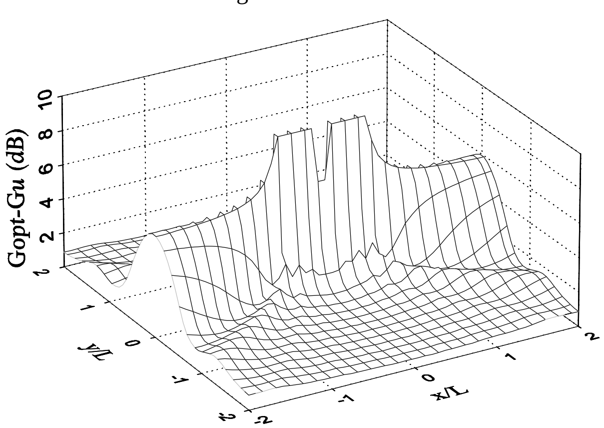

Figure 4 Difference between $G_{\text{opt}}(f)$ and the gain provided by uniform delay-and-sum beamforming $G_{\text{u}}(f)$ for a 15 element linear array with $\lambda/4$ microphone spacing. The near field extends to a distance of $1.75L$

In this case, the impact of the optimization is greater and extends over a larger near-field region than for the case of $\lambda/2$ microphone spacing. As before, the improvement is largest in the immediate vicinity of the microphones. As the distance increases, the optimum near-field gain rapidly approaches the value of the optimum far-field gain. Despite the equivalence in gains, however, the optimum near-field weights are distinct from the optimum far-field weights. That is, correcting the phase and amplitude of the optimum far-field weights for wavefront curvature will not, in general, produce the optimum near-field weights. This is illustrated in Figure 5 where the gains provided by both the near-field and far-field constrained optimum weights are shown for various distances in the broadside direction within the nearfield of a 15 element linear array. This figure shows that the optimum near field beamformer (curve NF) provides the highest gains. It also illustrates the errors made by neglecting source distance when using an optimum far-field beamformer (curve FF). Finally, compensating the amplitude and phase of the optimum far-field weights for wavefront curvature (curve FF Comp.) provides some improvement but the array gain is still less than the optimum near-field gain.

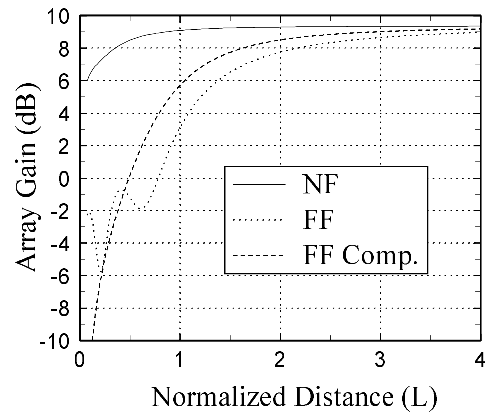

Figure 5 Gain provided by optimum weights at various distances in the broadside direction of a 15 element linear array with $\lambda/4$ microphone spacing. Gains are shown for near-field optimum (NF), far-field optimum (FF) and far-field optimum compensated for wavefront curvature (FF Comp.).

5. CONCLUSIONS

This paper describes a method for optimizing the near-field response of an arbitrary microphone array. Exisiting array optimization techniques are used to improve the ratio of near-field to far-field response. The optimum near-field weights are shown to provide increased gain for near-field sources when compared to a delay-and-sum beamformer with uniform microphone weights.

REFERENCES

1. Flanagan, J.L., Johnston, J.D., Zahn, R. and Elko, G.W., “Computer-steered microphone arrays for sound transduction in large rooms”, *J. Acoust. Soc. Am.*, **78**, 1985, pp. 1508-1518.
2. Kennedy, R.A., Abhayapala, T., Ward, D.C. and Williamson, R.C., “Nearfield Broadband Frequency Invariant Beamforming”, *Proc. IEEE ICASSP 1996*, pp. 905-908.
3. Monzingo, R.A. and Miller T.W., *Introduction to Adaptive Arrays*, John Wiley and Sons, 1980, pp. 78-105.
4. Cox, H. Zeskind, R.M. and Owen, M.M., “Robust Adaptive Beamforming”, *IEEE Trans. ASSP*, vol. ASSP-35, Oct. 1987, pp. 1365-1375.
5. Steinberg, B.D., *Principles of Aperture and Array System Design*, John Wiley and Sons, 1976, p. 12.
6. Kates, J. M., and Weiss, M.R., “A comparison of hearing-aid array processing techniques”, *J. Acoust. Soc. Am.*, vol. 99, May 1996, pp. 3138-3148.