This paper appeared in Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, May 2002.

A VARIATIONAL TECHNIQUE FOR SOURCE LOCALIZATION BASED ON A SPARSE SIGNAL RECONSTRUCTION PERSPECTIVE

Müjdat Çetin, Dmitry M. Malioutov, and Alan S. Willsky

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
77 Massachusetts Ave., Cambridge, MA 02139, USA

ABSTRACT

We propose a novel non-parametric technique for source localization with passive sensor arrays. Our approach involves formulation of the problem in a variational framework where regularizing sparsity constraints are incorporated to achieve superresolution and noise suppression. Compared to various source localization schemes, our approach offers increased resolution, significantly reduced sidelobes, and improved robustness to limitations in data quality and quantity. We demonstrate the effectiveness of the method on simulated data.

1. INTRODUCTION

Source localization using sensor arrays has been an active research area, playing a fundamental role in many applications involving electromagnetic, acoustic, and seismic sensing. A basic source localization problem is that of direction-of-arrival (DOA) estimation, which we focus on in this paper. Conventional beamformers for DOA estimation have a limited resolution, and this has led to the development and successful application of more advanced techniques. Examples are Capon's minimum variance method [1], and a variety of superresolution methods based on eigenvalue decomposition, such as MUSIC [2]. Such methods exploit the presence of *a small number* of sources in order to focus the estimated signal energy towards the source DOAs to achieve superresolution.

We propose a different perspective for exploiting such structure to achieve superresolution DOA estimation. We formulate the problem in a variational framework, where we minimize a regularized objective function for finding an estimate of the signal energy as a function of angle. The key is to use appropriate non-quadratic regularizing functionals (such as ℓ_p-norms), which lead to sparsity constraints (analogous to assuming a small number of sources) and superresolution. Variational methods based on such constraints have recently found application in various domains such as image restoration [3], computed tomography [4], and radar imaging [5]. Our work extends the use of such sparse signal reconstruction methods to DOA estimation. In the literature, there exist some recent ideas in a similar direction to ours. In [6], a high-resolution Cauchy-Gaussian spectral analysis method has been proposed and its application in array signal processing has been suggested. In [7], the link between optimization theoretic sparse solution methods and the MUSIC algorithm has been investigated.

This work was supported by the Army Research Office under Grant DAAD19-00-1-0466, and the Office of Naval Research under Grant N00014-00-1-0089.

We formulate an optimization problem for this task and propose a computationally efficient numerical method for its solution. Our experimental analysis on simulated data demonstrates that the proposed technique offers increased resolution, reduced sidelobe levels, and robustness to noise and limitations in data quantity, as compared to methods such as MUSIC. Furthermore, the proposed method can readily handle coherent signal sources.

2. OBSERVATION MODEL

In this paper, we consider narrowband source signals in the far-field impinging upon an array of M omnidirectional sensors. Let $\{\theta_1, ..., \theta_{N_\theta}\}$ be a sampling grid of all directions of arrival. Then we can represent complex amplitudes of the signal field at the t-th time sample by an $N_\theta \times 1$ vector $\mathbf{s}(t)$, where the i-th element $\mathbf{s}_i(t)$ of $\mathbf{s}(t)$ is non-zero only if there is a source at direction θ_i. This leads to the following observation model:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where the $M \times 1$ vectors $\mathbf{n}(t)$ and $\mathbf{y}(t)$ denote the measurement noise and the data collected at the sensors at time t, and the $M \times N_\theta$ matrix $\mathbf{A}$ is composed of steering vectors corresponding to each potential DOA. Note that $\mathbf{A}$ differs from the steering matrix representation used in many array processing methods in the sense that it contains steering vectors for *all* potential DOAs, rather than only the (unknown) source signal DOAs. Hence, in our framework $\mathbf{A}$ is known and does not depend on the actual source locations. The reason behind using such a redundant-looking representation is our desire to formulate the problem in a sparse signal reconstruction framework. This can also be viewed as representing the observations as combination of elements from an overcomplete signal dictionary, as in adaptive signal representation methods such as basis pursuit [8].

3. SOURCE LOCALIZATION SCHEME

3.1. Overview

As in numerous non-parametric source localization techniques, our approach consists of forming an estimate of the signal energy as a function of angle, which ideally contains dominant peaks at source DOAs. We need to obtain such an estimate of the signal field $\mathbf{s}$ (hence its energy) through the sensor observations $\mathbf{y}$, which is in general an ill-posed inverse problem. The central idea in our approach is to solve this inverse problem via regularizing it by favoring sparse signal fields, where energy is concentrated around a small number of DOAs. In particular, we find $\mathbf{s}$ as the minimizer of an objective function of the following form:[1]

$$J(\mathbf{s}) = J_1(\mathbf{s}) + \alpha J_2(\mathbf{s}) \tag{2}$$

where $J_1(\mathbf{s})$ is an ℓ_2-norm-based data fidelity term, $J_2(\mathbf{s})$ reflects the regularizing sparsity constraint we would like to impose, and α is a scalar parameter. Choice of $J_2(\mathbf{s})$ is critical for attaining the objectives of superresolution and noise suppression. In sparse signal reconstruction problems such as nuclear magnetic resonance (NMR) spectroscopy [9] and astronomical imaging [10], similar objectives have previously been achieved by using maximum entropy methods. These approaches provide reconstructions with good energy concentration (i.e. most elements are small and a few are very large). It has been shown that similar behavior can be obtained using minimum ℓ_1-norm reconstruction [11]. In spectral analysis, ℓ_p-norm constraints, where $p < 2$, have been shown to result in higher resolution spectral estimates compared to the ℓ_2-norm case (which is proportional to the periodogram) [12]. Based on these observations, we choose $J_2(\mathbf{s})$ based on ℓ_p-norms where $p < 2$. Use of other non-quadratic functions is also possible.

We present two versions of an objective function $J(\mathbf{s})$ which differ in the way they combine observations temporally. The first version formulates an optimization problem at each time instant:

$$J(\mathbf{s}(t)) = \|\mathbf{y}(t) - \mathbf{A}\mathbf{s}(t)\|_2^2 + \alpha\|\mathbf{s}(t)\|_p^p \tag{3}$$

where $\|\cdot\|_p$ denotes the ℓ_p-norm. The optimization problem in (3) produces a signal field reconstruction at a single time point. Individual reconstructions at various time points can then be combined to yield a signal energy estimate as a function of DOA. There are two disadvantages of this approach. First, the computational load of multiple optimizations is high. Second, since observations at different times are processed separately, the resulting estimates may not be robust to high noise levels. The second version of $J(\mathbf{s})$ combines the temporal data prior to processing to define a single optimization problem:

$$J(\bar{\mathbf{s}}) = \frac{1}{T}\left(\sum_{t=1}^{T} \|\mathbf{y}(t) - \mathbf{A}\bar{\mathbf{s}}\|_2^2\right) + \alpha\|\bar{\mathbf{s}}\|_p^p \tag{4}$$

Here we use $\bar{\mathbf{s}}$ to emphasize that this signal field represents an aggregate estimate over time. Note that defining the problem as in (4) makes sense only if the complex envelopes of the source signals are non-zero-mean signals. This is the case, e.g. when the source signals are sinusoids with some variation around the amplitudes.

3.2. Numerical Solution

We outline the solution of the optimization problem in (4). Solution of (3) is similar as we will point out at the end of this section. In order to avoid problems due to non-differentiability of the ℓ_p-norm around the origin when $p \leq 1$, we use the following smooth approximation to the ℓ_p-norm in (4) [3]:

$$\|\mathbf{z}\|_p^p \approx \sum_{i=1}^{K} (|\mathbf{z}_i|^2 + \epsilon)^{p/2} \tag{5}$$

[1]Note that we suppress the time dependence here, which we will address later in this section.

where $\epsilon \geq 0$ is a small constant, K is the length of the complex vector $\mathbf{z}$, and $\mathbf{z}_i$ denotes its i-th element. For numerical purposes, we thus use the following slightly modified cost function:

$$J_\epsilon(\bar{\mathbf{s}}) = \frac{1}{T}\left(\sum_{t=1}^{T} \|\mathbf{y}(t) - \mathbf{A}\bar{\mathbf{s}}\|_2^2\right) + \alpha \sum_{i=1}^{N_\theta} (|\bar{\mathbf{s}}_i|^2 + \epsilon)^{p/2} \tag{6}$$

Note that $J_\epsilon(\bar{\mathbf{s}}) \to J(\bar{\mathbf{s}})$ as $\epsilon \to 0$. The minimization of $J(\bar{\mathbf{s}})$ or $J_\epsilon(\bar{\mathbf{s}})$ does not yield a closed-form solution in general, so numerical optimization techniques must be used.

For solution of this optimization problem, we use the half-quadratic regularization method of [13]. Half-quadratic regularization converts a non-quadratic optimization problem into a series of quadratic problems. We skip the derivations here and present the resulting iterative algorithm:

$$\mathbf{H}\left(\hat{\bar{\mathbf{s}}}^{(n)}\right)\, \hat{\bar{\mathbf{s}}}^{(n+1)} = \frac{1}{T}\left(\sum_{t=1}^{T} \mathbf{A}^H \mathbf{y}(t)\right) \tag{7}$$

where n denotes the iteration number, and:

$$\mathbf{H}(\mathbf{z}) \triangleq \mathbf{A}^H\mathbf{A} + \alpha\Lambda(\mathbf{z}) \tag{8}$$

$$\Lambda(\mathbf{z}) \triangleq \text{diag}\left\{\frac{p/2}{(|\mathbf{z}_i|^2 + \epsilon)^{1-p/2}}\right\}$$

where diag$\{\cdot\}$ is a diagonal matrix whose i-th diagonal element is given by the expression inside the brackets. We run the iteration (7) until $\frac{\|\hat{\bar{\mathbf{s}}}^{(n+1)} - \hat{\bar{\mathbf{s}}}^{(n)}\|_2^2}{\|\hat{\bar{\mathbf{s}}}^{(n)}\|_2^2} < \delta$, where $\delta > 0$ is a small constant. Compared to standard optimization tools, the above scheme yields an efficient method matched to the structure of our optimization problem. Convergence properties of algorithms of this type have been analyzed, and convergence from any initialization to a local minimum is guaranteed [4, 14]. Note that a solution to (3) could similarly be obtained by the following iteration:

$$\mathbf{H}\left(\hat{\mathbf{s}}^{(n)}(t)\right)\, \hat{\mathbf{s}}^{(n+1)}(t) = \mathbf{A}^H\mathbf{y}(t). \tag{9}$$

4. SIMULATION RESULTS

We consider a uniform linear array of $M = 8$ sensors separated by half a wavelength of the actual narrowband source signals. We consider two narrowband signals in the far-field impinging upon this array. The total number of snapshots is $T = 200$. We use the objective function in (4) in our experiments. We first consider uncorrelated sources, and present plots of signal power versus DOA produced by various techniques. In these experiments, we use $p = 0.1$ in our technique. We define the signal-to-noise ratio (SNR) as the ratio of the signal power to the noise power at the sensors. Figure 1 shows the case when the two sources have a wide separation. In this case all methods considered can resolve the two sources. However, the proposed method exhibits the best sidelobe suppression. Next, we consider the case when the two sources lie within a Rayleigh resolution cell. Figure 2 contains results for various SNRs. These plots demonstrate the relatively superior robustness of the proposed method to high levels of noise.

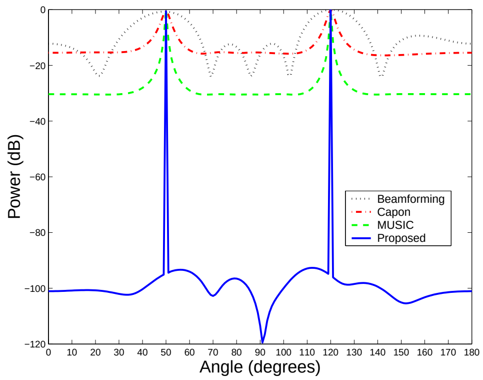


Fig. 1. Spatial spectra of two sources with DOAs of 50° and 120° (SNR = 10 dB). Broadside corresponds to 90°.

Figure 3 illustrates source localization results where the two sources are coherent. Note that Capon's method and MUSIC, which were able to resolve *uncorrelated* sources with this level of separation and SNR, fail in the coherent case, whereas the proposed method is still able to resolve the two sources.

The examples so far were based on single trials. Now we characterize the performance of the proposed method (for localizing uncorrelated sources) over 200 independent trials, as a function of SNR and the number of snapshots. We consider two performance metrics. The first one is the probability of detecting the two sources with 1° accuracy. The second one is the root-mean-squared-error (in angles) in locating the sources. The two measures convey very similar information, and the experimental results are very similar, so we present only the probability of detection results here. Figure 4(a) presents results for the case when the sources are separated by 15°. This plot shows that the proposed method has a significantly better performance than Capon's method and MUSIC at low SNR values. Next we consider the case when the separation is reduced to 10°. The plot in Figure 4(b) shows a potential weakness of the proposed approach. When the sources are too close, the proposed method may cause some bias in localization. This bias causes the relative performance of the proposed method to be worse than the other techniques at high SNRs. Such shifts in peak locations have previously been noted for ℓ_p-norm-based methods such as basis pursuit in the context of spectral estimation [8]. Gaining a better understanding and resolution of such issues is a matter of our current research. Figure 5 shows the probability of detection performance as a function of the number of snapshots used. These plots demonstrate the robustness of the proposed method to reduced amounts of data.

5. CURRENT WORK

Our current work involves the examination and extension of various aspects of our approach. An important issue in our method is the selection of the parameters p and α, which we have so far done manually based on subjective qualitative assessment. Developing techniques for automatic choice of these parameters is of interest. Another issue worth investigation is the bias in source locations that the algorithm may sometimes cause. Extensions of the technique to cases involving non-linear array configurations, near-field sources and broadband signals are subjects of our current work. We are also interested in developing extensions of our variational

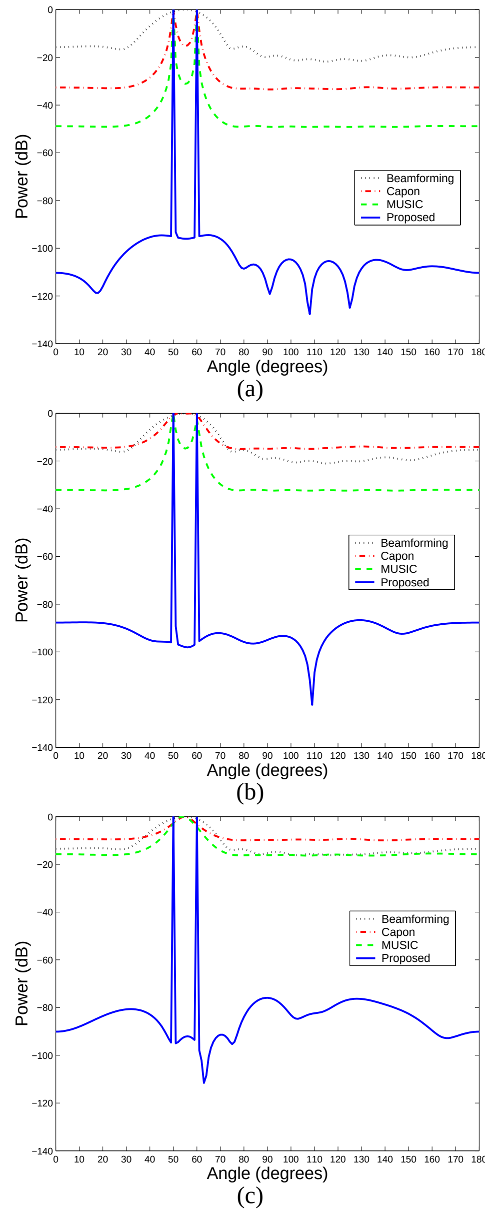


Fig. 2. Spatial spectra of two sources with DOAs of 50° and 60°. (a) SNR = 20 dB. (b) SNR = 10 dB. (c) SNR = 5 dB.

framework to take into account uncertainties in sensor locations.

6. CONCLUSIONS

We have approached the source localization problem with a sparse signal reconstruction perspective. We have developed and demonstrated the viability of a non-parametric superresolution source localization method in a variational framework, using ℓ_p-norm-based functionals as regularizing constraints. The approach effectively deals with difficulties such as sidelobes and coherency in signals, and improves upon the source localization accuracy of currently used methods, especially in low-SNR and limited-data scenarios. Our framework is generalizable to more general source localization problems than that considered in this paper.

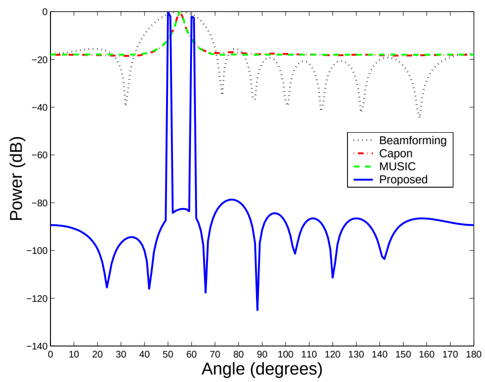
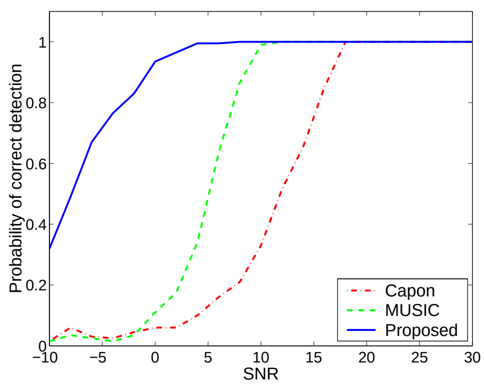
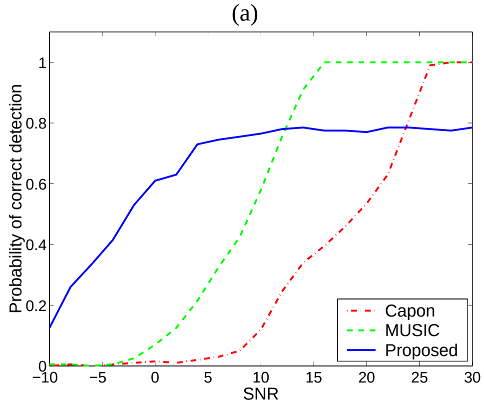


Fig. 3. Spatial spectra of two coherent sources with DOAs of 50° and 60° (SNR = 20 dB).



(a)



(b)

Fig. 4. Probability of correct detection for two sources as a function of SNR. (a) DOAs: 50° and 65°. (b) DOAs: 50° and 60°.

7. REFERENCES

[1] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[2] R. O. Schmidt, *A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation*, Ph.D. thesis, Stanford Univ., Stanford, CA, 1981.

[3] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Image Processing*, vol. 7, no. 6, pp. 813–824, June 1998.

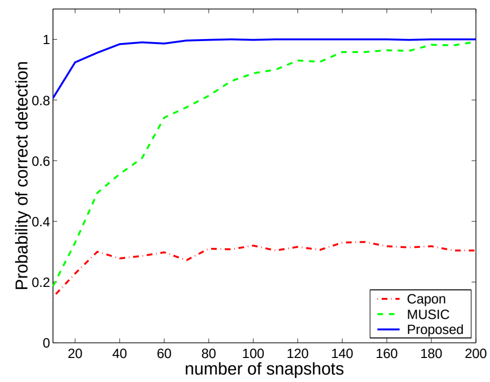


Fig. 5. Probability of correct detection for two sources with DOAs of 50° and 65° as a function of the number of snapshots (SNR = 10 dB).

[4] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 298–310, Feb. 1997.

[5] M. Çetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Trans. Image Processing*, vol. 10, no. 4, pp. 623–631, Apr. 2001.

[6] M. D. Sacchi, T. J. Ulrych, and C. J. Walker, "Interpolation and extrapolation using a high-resolution discrete Fourier transform," *IEEE Trans. Signal Processing*, vol. 46, no. 1, pp. 31–38, Jan. 1998.

[7] B. D. Jeffs, "Sparse inverse solution methods for signal and image processing applications," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1998, vol. 3, pp. 1885–1888.

[8] S. S. Chen, "Application of basis pursuit in spectrum estimation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1998, vol. 3, pp. 1865–1868.

[9] S. Sibisi, J. Skilling, R. G. Brereton, E. D. Laue, and J. Staunton, "Maximum entropy signal processing in practical NMR spectroscopy," *Nature*, vol. 311, pp. 446–447, Oct. 1984.

[10] R. Narayan and R. Nityananda, "Maximum entropy restoration in astronomy," *Ann. Rev. Astron. Astrophys.*, vol. 24, pp. 127–170, 1986.

[11] D. L. Donoho, I. M. Johnstone, J. C. Koch, and A. S. Stern, "Maximum entropy and the nearly black object," *J. R. Statist. Soc. B*, vol. 54, no. 1, pp. 41–81, 1992.

[12] P. Ciuciu, J. Idier, and J. F. Giovannelli, "Markovian high resolution spectral analysis," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Phoenix, AZ, USA, Mar. 1999, vol. 3, pp. 1601–1604.

[13] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.

[14] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography," *IEEE Trans. Image Processing*, vol. 7, no. 2, pp. 204–221, Feb. 1998.