

Article

Direction-of-Arrival Estimation Based on Joint Sparsity

Junhua Wang *, Zhitao Huang and Yiyu Zhou

School of Electronic Science and Engineering, NUDT, Changsha 410073, China;

E-Mails: zhitao.huang@nudt.edu.cn (Z.H.); zhouyiyu@nudt.edu.cn (Y.Z.)

* Author to whom correspondence should be addressed; E-Mail: skysword_wjh@126.com;
Tel.: +86-1363-7417-940; Fax: +86-0731-572518.

Received: 16 August 2011; in revised form: 17 September 2011 / Accepted: 18 September 2011/

Published: 21 September 2011

Abstract: We present a DOA estimation algorithm, called Joint-Sparse DOA to address the problem of Direction-of-Arrival (DOA) estimation using sensor arrays. Firstly, DOA estimation is cast as the joint-sparse recovery problem. Then, norm is approximated by an arctan function to represent joint sparsity and DOA estimation can be obtained by minimizing the approximate norm. Finally, the minimization problem is solved by a quasi-Newton method to estimate DOA. Simulation results show that our algorithm has some advantages over most existing methods: it needs a small number of snapshots to estimate DOA, while the number of sources need not be known *a priori*. Besides, it improves the resolution, and it can also handle the coherent sources well.

Keywords: joint-sparse; compressed sensing; Direction-of-Arrival; quasi-Newton methods; multiple measure vectors

1. Introduction

Direction-of-Arrival (DOA) estimation using sensor arrays has been an active research area, playing a fundamental role in many applications involving electromagnetic, acoustic, and communication systems [1]. Many classical algorithms are available, and the popular methods include beamforming [2], MUSIC [3], ESPRIT [4] and the maximum likelihood method [5], *etc.* The beamforming method has low angle resolution and suffers from the Rayleigh resolution limit. MUSIC, ESPRIT and the maximum likelihood method all rely on the statistical properties of the data, and thus, require a

sufficiently large number of samples for accurate estimation. Besides, MUSIC and ESPRIT cannot handle strongly coherent sources, while the maximum likelihood method has high computation costs.

The problem of sparse recovery has evolved rapidly recently [6,7] and it has been applied in DOA estimation with array processing. Gorodnitsky *et al.* [8] used a weighted least-squares algorithm named FOCUSS for DOA estimation, but this algorithm can only be used for single snapshots. Cotter [9] combined multiple measurement vectors (MMV) and matching pursuit (MP) to solve the joint-sparse recovery problem in DOA estimation, but it has low angle resolution. JLZA-DOA is proposed in [10]; it minimizes a mixed $L_{2,0}$ norm to deal with the joint-sparse recovery problem, and a fixed point method is used for DOA estimation. This algorithm doesn't satisfy numerical stability, as matrix inversion is inevitable in every iteration. Stoica *et al.* [11] presented a novel SParse Iterative Covariance-based Estimation approach, abbreviated SPICE. However, this algorithm needs more snapshots to estimate DOA. Wide-band covariance matrix sparse representation (W-CMSR) is proposed in [12] for DOA estimation of wideband signals. So far, the most successful joint-sparse recovery algorithm for DOA estimation is L1-SVD [13,14]. It combines the SVD step of the subspace algorithms with a sparse recovery method based on $l_{2,1}$ -norm minimization. However, the number of sources needs be known *a priori*.

In this paper, we present Joint-Sparse DOA estimation, abbreviated as JSDOA, for sensor array DOA estimation. First, DOA estimation is cast as a joint-sparse recovery problem. Then, $L_{2,0}$ norm is approximated by the arctan function to represent spatial sparsity and DOA estimation can be obtained by minimizing the approximate $L_{2,0}$ norm. Finally, the minimization problem is solved by a quasi-Newton method to estimate DOA. The proposed algorithm has some advantages over most existing methods: it needs a small number of snapshots to estimate DOA, and the number of sources need not be known *a priori*. Besides, it improves the probability of resolution, and it can also handle coherent sources well.

The outline of the paper is as follow. In Section 2, the DOA estimation problem is formulated. The new algorithm, called JSDOA, is proposed in Section 3. In section 4, the validity of the proposed algorithm is proved by a number of simulations. Finally, conclusions are presented in Section 5.

2. Problem Formulation

2.1. DOA Estimation Problem

Consider a linear array consisting of M identical sensors and receiving signals from K narrowband signals $s_1(t), s_2(t), \dots, s_K(t)$, which arrive at the array from directions $\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_K$ with respect to the line of array. The received signal $y_m(t)$ at the m^{th} sensor can be written as:

$$y_m(t) = \sum_{k=1}^K a(\bar{\theta}_k) s_k(t) + n_m(t) \quad (1)$$

where $\{a(\bar{\theta}_k)\}_{k=1}^K$ denote steering vectors, $n_m(t) (m=1, 2, \dots, M)$ stand for the additive noise.

Let $y(t) = [y_1(t), y_2(t), \dots, y_M(t)]^T$, $n(t) = [n_1(t), \dots, n_M(t)]$, Equation (1) can be written as:

$$y(t) = A(\bar{\theta})s(t) + n(t) \quad (2)$$

where the manifold matrix $A(\bar{\theta})$ consists of the steering vectors $\{a(\bar{\theta}_k)\}_{k=1}^K$:

$$A(\theta) = [a(\bar{\theta}_1), a(\bar{\theta}_2), \dots, a(\bar{\theta}_K)]$$

Therefore, DOA estimation is to find K and $\bar{\theta}_k$ from T snapshots $\{y(t)\}_{t=t_1}^{t_T}$.

2.2. Joint-Sparse Recovery for DOA Estimation Problem

Because sources are sparse in space, DOA estimation with sensor arrays can be expressed as a joint-sparse recovery problem. Let Ω denote the set of possible locations, $\{\theta_n\}_{n=1}^N$ denotes a grid that covers Ω . We assume that the grid is fine enough such that the true location parameters of the existing sources lie on the grid. Let:

$$x(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T$$

Then the received signal $y_m(t)$ at the m^{th} sensor can be written as:

$$y_m(t) = \sum_{n=1}^N a(\theta_n) x_n(t) + n(t) \quad (3)$$

where the n^{th} element $x_n(t)$ of $x(t)$ is nonzero only if $\theta_n = \bar{\theta}_k$, $k \in [1, 2, \dots, K]$ and in that case $x_n(t) = s_k(t)$:

Let $\Phi = [a(\theta_1), a(\theta_2), \dots, a(\theta_N)]$, (3) can be expressed as:

$$y(t) = \Phi x(t) + n(t) \quad (4)$$

when the number of snapshots is denoted as T , Equation (4) can be written as:

$$Y = \Phi X + N \quad (5)$$

where $Y = [y(t_1), y(t_2), \dots, y(t_T)]$, $X = [x(t_1), x(t_2), \dots, x(t_T)]$. As we know, X is row-sparse and only K rows have nonzero elements, so DOA estimation can be obtained by a joint-sparse recovery problem, which is also called a multiple measurement vectors (MMV) problem. Using $l_{p,q}$ norm to express joint sparsity, the MMV problem can be converted to $l_{p,q}$ norm minimization:

$$\begin{cases} \min & \|X\|_{p,q} \\ \text{s.t.} & Y = \Phi X + N \end{cases} \quad (6)$$

where p and q are non-negative, and $\|X\|_{p,q}$ is defined as:

$$\|X\|_{p,q} = \sum_{i=1}^M \left(\|X(i,:)\|_p \right)^q \quad (7)$$

Concretely, Eldar and Mishali [15] use $p = 2$, $q = 1$. Chen and Huo [16] study for $p = 1$, $q = 1$. The above algorithms do not perform very well for $\|X\|_{2,1}$ and $\|X\|_{1,1}$ can't sufficiently reflect joint sparsity. Considering $\|X\|_{2,0}$ can reflect joint sparsity sufficiently, we minimize $\|X\|_{2,0}$ norm to solve the MMV problem. However, $\|X\|_{2,0}$ norm minimization can hardly be solved directly. Therefore, in this paper we approximate $\|X\|_{2,0}$ norm by an arctan function and estimate DOA by solving an approximate $\|X\|_{2,0}$ norm minimization problem.

3. Joint-Sparse DOA Estimation Algorithm

In this section, the DOA estimation problem is converted into an approximate $\|X\|_{2,0}$ norm minimization problem. Firstly, the $L_{2,0}$ norm is approximated by an arctan function to construct an

approximate $\|X\|_{2,0}$ norm minimization problem. Then this problem is solved by a quasi-Newton method to estimate DOA.

3.1. Basic Idea of the Proposed Method

Let $\xi_i = \|X(i,:)\|_2$, $i = 1, 2, \dots, M$, we have:

$$\|X\|_{2,0} = \sum_{i=1}^M (\|X(i,:)\|_2)^0 = \|\xi\|_0 \quad (8)$$

where $\xi = (\xi_1, \dots, \xi_M)$. Considering the following arctan function:

$$f_\delta(s) = \frac{2}{\pi} \arctan\left(\frac{s^2}{2\delta^2}\right) \quad (9)$$

where δ is a positive parameter and s is a variable parameter. Then $f_\delta(s)$ has the following property:

$$\lim_{\delta \rightarrow 0} f_\delta(s) = \begin{cases} 1 & s \neq 0 \\ 0 & s = 0 \end{cases} \quad (10)$$

Let

$$F_\delta(\xi) = \sum_{i=1}^N f_\delta(\xi_i)$$

From Equation (10), we have:

$$\lim_{\delta \rightarrow 0} F_\delta(\xi) = \|\xi\|_0 \quad (11)$$

From Equations (8) and (10), then:

$$\lim_{\delta \rightarrow 0} F_\delta(X) = \|X\|_{2,0}$$

So DOA can be obtained by solving the approximate $\|X\|_{2,0}$ norm minimization:

$$\begin{cases} \min & F_\delta(X) \\ \text{s.t.} & Y = \Phi X + N \end{cases} \quad (12)$$

where δ is a small positive constant. Because noise is unknown, we synchronously hope to minimize $F_\delta(X)$ and $\|\Phi X - Y\|_F^2$. Then Equation (12) is converted into a multiple objective optimization:

$$\begin{cases} \min_X & F_\delta(X) \\ \min_X & \|\Phi X - Y\|_F^2 \end{cases} \quad (13)$$

Using the linear weighting method, Equation (13) can be written as:

$$\min_x L_{\delta,\lambda}(X) = F_\delta(X) + \lambda \|\Phi X - Y\|_F^2 \quad (14)$$

where $\|\cdot\|_F$ denotes the Frobenious norm and the parameter λ will be discussed in Section 3.2. When the parameter δ is very small, the objective function $F_\delta(X)$ is highly unsmoothed and contains a lot of local minimization, so its global minimization is not easy. On the other hand, if the parameter δ is larger, the objective function $F_\delta(X)$ is smoother and contains less local minimization. In order to obtain global minimization, we select a decreasing sequence for δ , denoted as:

$$\delta = [\delta_1, \delta_2, \dots, \delta_J], \quad \delta_{j+1} < \delta_j$$

where δ_1 is a relatively large value, and δ_J is a small value. For $\delta = \delta_{J-1}$, the solution of Equation (14) is denoted as $x_{\delta_{j-1}}$, and $x_{\delta_{j-1}}$ is used as the initial value for $\delta = \delta_j$, thus we hope the proposed

algorithm can escape from getting trapped into a local minimum and reach the global minimization for a small value $\delta = \delta_J$.

For some fixed value $\delta = \delta_j$, minimization problem Equation (14) is solved by the quasi-Newton method in this paper. One of the most successful quasi-Newton methods is the BFGS algorithm, which is second order convergent and has good numerical stability. Therefore, we use the BFGS algorithm to solve Equation (14) for some fixed value $\delta = \delta_j$.

The conjugate gradient for a matrix variable is defined as:

$$\frac{\partial L_{\delta,\lambda}(X)}{\partial X^*} = \frac{1}{2} \left(\frac{\partial L_{\delta,\lambda}(X)}{\partial X_R} + i \frac{\partial L_{\delta,\lambda}(X)}{\partial X_I} \right)$$

where X_R and X_I denote the real part and the imaginary part respectively, X^* denotes the conjugate of X . Then the conjugate gradient of $L_{\delta,\lambda}(X)$ (for its derivation see Appendix I) can be expressed as:

$$\frac{\partial L_{\delta,\lambda}(X)}{\partial X^*} = \frac{1}{2\delta^2} \Lambda X - \lambda \Phi^*(Y - \Phi X) \quad (15)$$

where

$$\Lambda = \text{diag} \left\{ \delta^2 / (\xi_1^4 + 4\delta^4), \delta^2 / (\xi_2^4 + 4\delta^4), \dots, \delta^2 / (\xi_n^4 + 4\delta^4) \right\}$$

Solving by the BFGS algorithm, the main iterative steps are as follows:

(1) Search step length t_k , satisfy: $t_k = \min_t L_{\delta_j,\lambda}(x - t H^k \nabla L_{\delta_j,\lambda}(X))$

(2) Iteration: $X^{k+1} = X^k - t_k H^k \frac{\partial L_{\delta,\lambda}(X^k)}{\partial X^*}$

(3) BFGS adjustment:

$$H^{k+1} = H^k + \left(1 + \frac{(g^k)^T H^k g^k}{(g^k)^T s^k} \right) \frac{s^k (s^k)^T}{(g^k)^T s^k} - \frac{s^k (g^k)^T H^k - H^k g^k (s^k)^T}{(g^k)^T s^k} \quad (16)$$

where $s^k = X^{k+1} - X^k$, $g^k = \frac{\partial L_{\delta,\lambda}(X^{k+1})}{\partial X^*} - \frac{\partial L_{\delta,\lambda}(X^k)}{\partial X^*}$

3.2. Algorithm Description

Based on the above idea, JSDOA can be described as shown in Table 1.

Table 1. The main steps of JSDOA.

Algorithm 1: joint-sparse DOA estimation
Input: A, y
Initialization:
(1) Set $X^0 = A^T (AA^T)^{-1} Y$
(2) Select a decreasing sequence $\delta = [\delta_1 \ \delta_2 \ \dots \ \delta_J]$, set ε and parameter λ
Iteration:
(1) for $j = 1, 2, \dots, J$

Table 1. Cont.

(2) solving (14) by BFGS algorithm
(2-1) $x = x^{j-1}$, $H = I$ (I is unit matrix);
(2-2) while $\text{norm}(\partial L_{\delta_j, \lambda}(X)/\partial X) > \delta_j \varepsilon$;
(2-3) Search step length $\tilde{t}$: $\tilde{t} = \min_t L_{\delta_j, \lambda}(X - t H \partial L_{\delta_j, \lambda}(X)/\partial X)$
(2-4) Let $X = X - \tilde{t} H \partial L_{\delta_j, \lambda}(X)/\partial X$
(2-5) Update the matrix by (16)
(2-6) end
(3) Let $X^j = X$
(4) end
Output: $\hat{X} = X^J$, spatial spectrum $P(\bar{\theta}_i) = 10 \log_{10}(\ \hat{X}(i, :)\ _2)$

Remark:

(1) Select parameter λ : in this paper, we select the parameter λ by the α - method [17]. Set $f_1(X) = F_\delta(X)$, $f_2(X) = \|AX - y\|_F^2$, and minimizing $f_1(X)$ and $f_2(X)$, we have;

$$f_i(X^{(i)}) = \min_x f_i(X) \quad i=1, 2$$

It is easy to know that $X^{(i)} (i=1, 2)$ is 0 and $A^T(AA^T)^{-1}Y$ respectively. Then we have:

$$f_{ij} = f_i(X^{(j)}) \quad i, j=1, 2$$

Set $\lambda = \lambda_2/\lambda_1$ and introduce auxiliary parameter α , we can have:

$$\begin{cases} \sum_{i=1}^2 f_{ij} \lambda_i = \alpha \\ \sum_{i=1}^2 \lambda_i = 1 \end{cases}$$

Setting the coefficient matrix $f(i, j) = f_{ij}$, from the above equations, the solution is:

$$\begin{cases} [\lambda_1, \lambda_2] = \frac{e^T f^{-1}}{e^T f^{-1} e} \\ \alpha = \frac{1}{e^T f^{-1} e} \end{cases} \quad (17)$$

where $e = [1, 1]^T$, and $\lambda = \lambda_2/\lambda_1$.

(2) Select parameter δ : In this paper, we set $\delta_j = \gamma \delta_{j-1}$, $j=2, \dots, J$, $\gamma \in (0.5, 1)$. Let $\tilde{x} = \max_i |x_i^0|$, we hope parameter δ_1 satisfies:

$$f_\delta(\tilde{x}) = \frac{2}{\pi} \arctan\left(\frac{\tilde{x}^2}{2\delta_1^2}\right) \leq \frac{1}{2} \Rightarrow \delta_1 \geq \frac{\tilde{x}}{\sqrt{2}}$$

In order to save computation cost, we set $\delta_1 = \max_i |x_i^0|/\sqrt{2}$. When $\delta_j \rightarrow 0$, $F_{\delta_j}(x) \rightarrow \|x\|_0$. However, if δ_j is too small, $F_{\delta_j}(x)$ will be sensitive to noise, so δ_j shouldn't be too small.

(3) For coherent sources: Many of popular methods, such as MUSIC and ESPRIT require the assumption that sources are uncorrelated, because the source covariance matrix remains nonsingular so long as none of these sources are coherent. However, the algorithm proposed in this paper doesn't need to compute the source covariance matrix, and joint sparsity is used to estimate DOA, so the proposed algorithm can handle highly coherent sources as well.

(4) Limitation: The algorithm proposed in this paper suffers from two problems. First, there exist no clear-cut guidelines for the selection of δ_J , especially when knowledge of noise is unknown. Second, the proposed algorithm can't perform well when the SNR is relatively low.

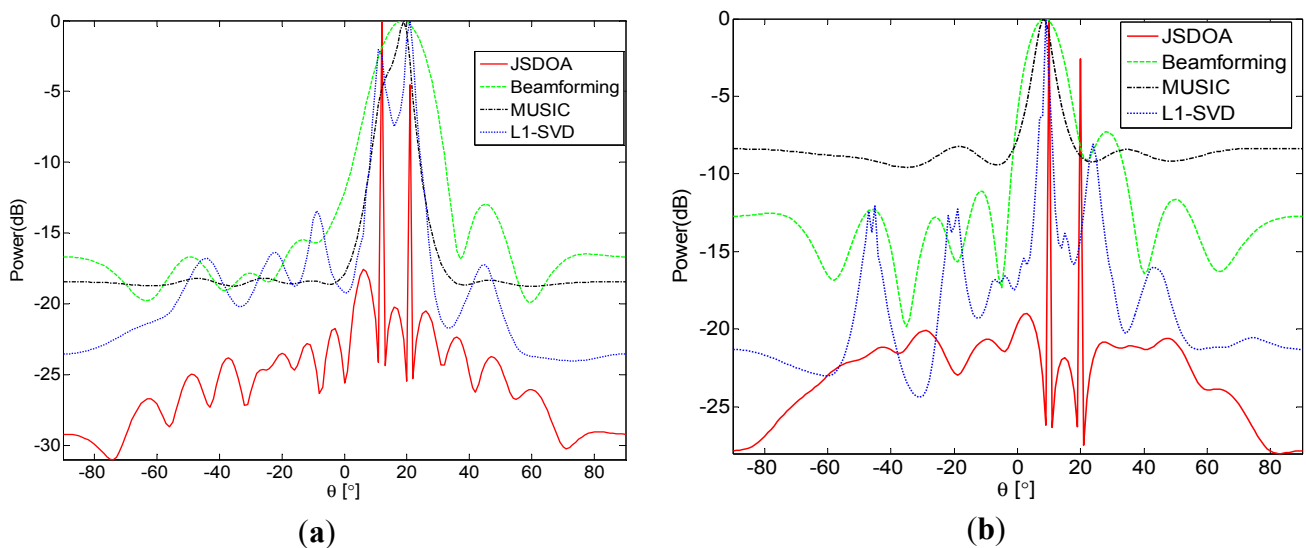
4. Simulation Results

In this section, we present several numerical simulation results to illustrate the performance of the proposed algorithm. First, we compare the spatial spectrum of JSDOA to those of beamforming [2], MUSIC [3], and L1-SVD [13] under various snapshot scenarios. Next, we compare the spatial spectrum with more than two sources. Then, the probability of resolution is compared. Finally, we compare the spatial spectrum for coherent sources. In the following numerical simulations, we consider a uniform linear array consisting of $M = 8$ identical sensors and receiving signals are narrowband signals. The sensors are uniformly placed with a spacing of half a wavelength. The interval for the DOA is $\Omega = [-90, 90]$, we use a uniform grid $\{\theta_n\}_{n=1}^N$ to cover Ω with a step of 1° , which means $N = 181$.

4.1. Spatial Spectrum Comparison under Various Snapshots

In this simulation, we compare spatial spectrum of different algorithms under various snapshots. We consider two uncorrelated sources located at 13° and 20° with $SNR = 10$. For JSDOA, we set $\delta_1 = \max_i |x_i^0|/\sqrt{2}$, $\delta_j = \gamma\delta_{j-1}$, $\gamma = 0.5$, $J = 7$. In Figures 1(a,b), we set $T = 10$ and $T = 5$, respectively. It can be seen from Figure 1 that JSDOA and L1-SVD can resolve the two sources as $T = 10$. However, only JSDOA can resolve the two sources as $T = 5$.

Figure 1. Spatial spectrum comparison under various snapshots for two uncorrelated sources: (a) $T = 10$; and (b) $T = 5$.

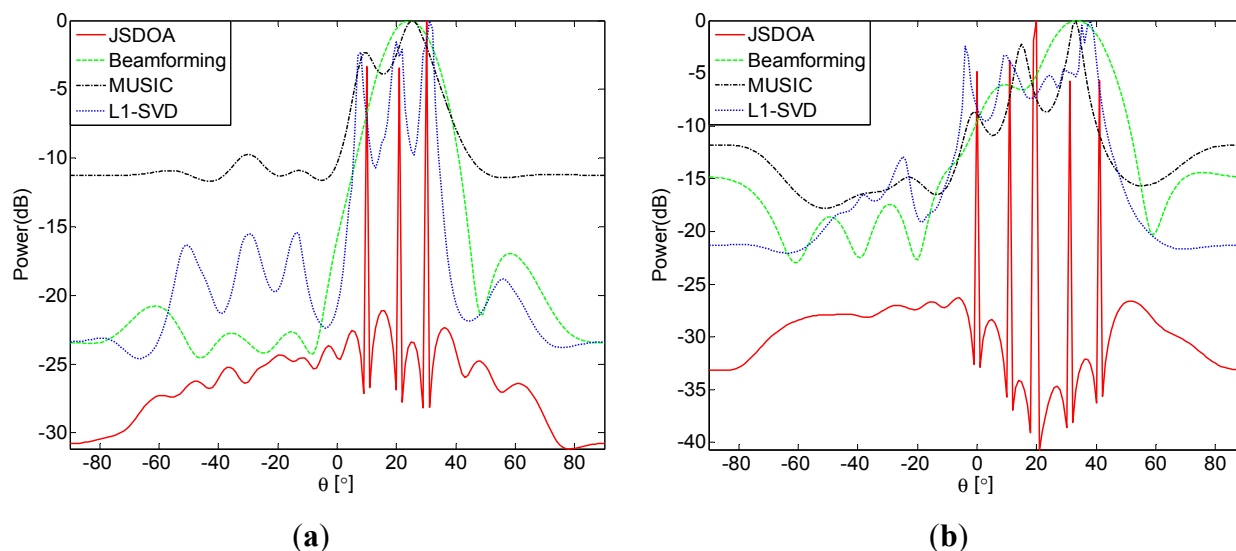


4.2. Spatial Spectrum Comparison with More Than Two Sources

In this simulation, we compare the spatial spectrum of different algorithms with more than two uncorrelated sources. Set $SNR = 10$, $T = 20$, $\delta_1 = \max_i |x_i^0|/\sqrt{2}$, $\delta_j = \gamma\delta_{j-1}$, $\gamma = 0.5$, $J = 7$. In

Figure 2(a), there are three sources located at 10, 20 and 30. In Figure 2(b), there are five sources located at 0, 10, 20, 30 and 40. It is shown from Figure 1 that JSDOA and L1-SVD can resolve the three sources, but only JSDOA can resolve the five sources.

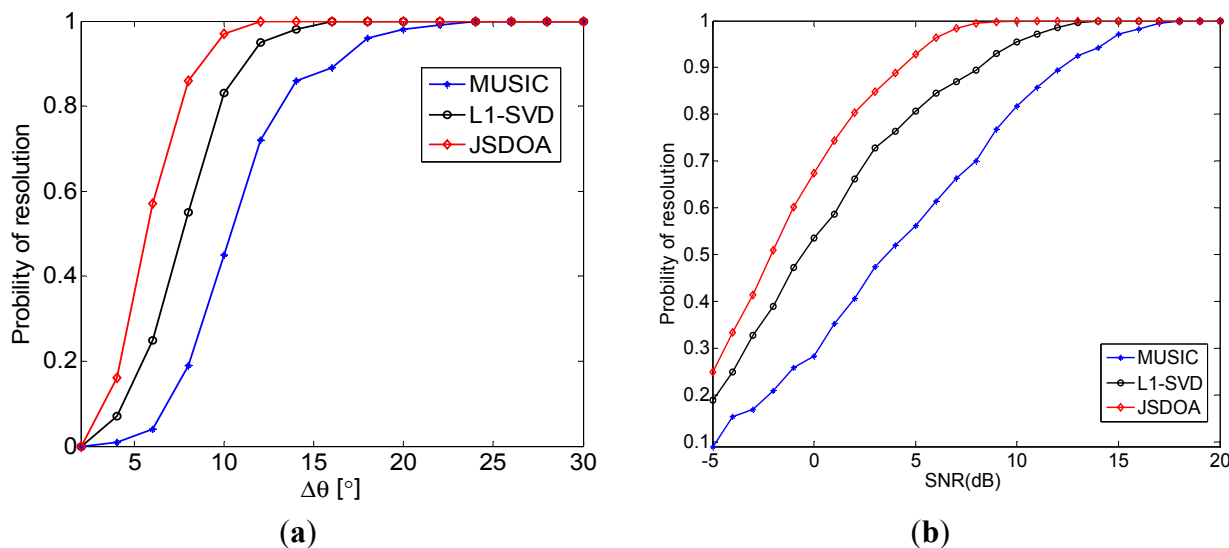
Figure 2. Spatial spectrum comparison for more than two uncorrelated sources: (a) Three sources; and (b) Five sources.



4.3. Probability of Resolution Comparison under Various Conditions

In this simulation, we compare probability of resolution for two uncorrelated sources. JSDOA, MUSIC and L1-SVD are used for comparison and we set $T = 5$, $\delta_j = \gamma \delta_{j-1}$, $\gamma = 0.5$, $J = 7$. For each $\Delta\theta$ or SNR, 200 independent simulation are carried out. In Figure 3(a), one source is fixed at 10° , the other source is located at $10^\circ + \Delta\theta$, vary from 0 to 30° . In Figure 3(b), two sources are located at 10° and 20° . SNR vary from -5 to 20 . It is shown from Figure 3 that JSDOA has the higher probability of resolution than MUSIC and L1-SVD.

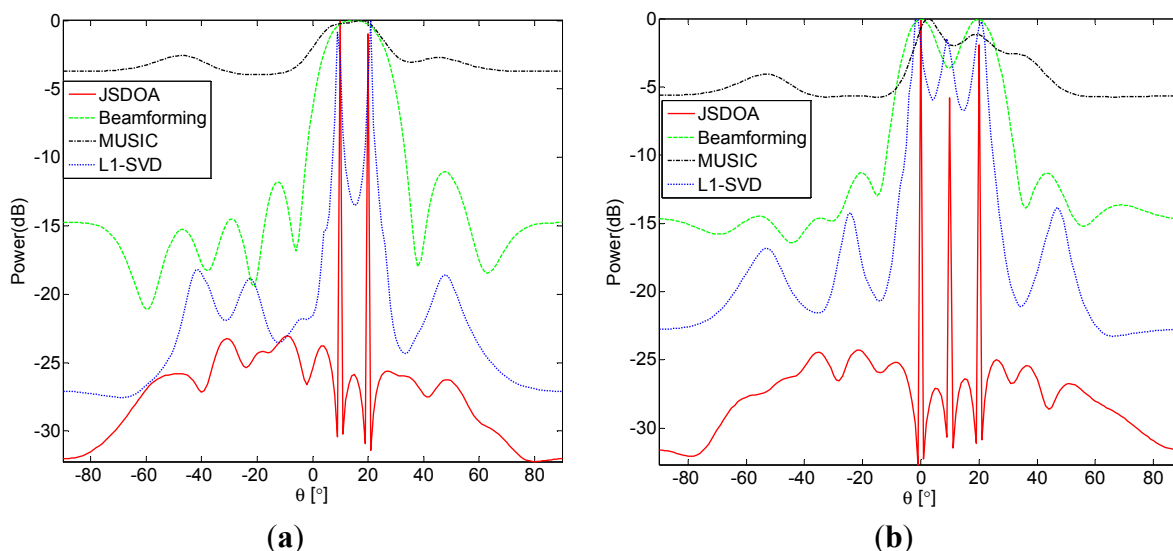
Figure 3. Probability of resolution comparison under various conditions. (a) Against $\Delta\theta$; and (b) Against input SNR.



4.4. Spatial Spectrum Comparison for Coherent Sources

In this simulation, we compare spatial spectrum for completely coherent sources. Set $SNR = 10$, $T = 20$, $\delta_1 = \max_i |x_i^0|/\sqrt{2}$, $\delta_j = \gamma\delta_{j-1}$, $\gamma = 0.5$, $J = 7$. In Figure 4(a) two coherent sources, located at 10° and 20° , are considered.

Figure 4. Spatial spectrum comparison for coherent sources. (a) Two coherent sources; and (b) Three coherent sources



In Figure 4(b) three coherent sources, located at 10° and 20° , are considered. It is shown from Figure 4 that JSDOA and L1-SVD can be used to estimate DOA for coherent sources and that JSDOA has higher resolution than L1-SVD.

5. Conclusions

In this paper, a joint-sparse recovery algorithm, called JSDOA, is proposed to estimate DOA with sensor arrays. We pose the DOA estimation problem as a joint-sparse recovery problem, which can be solved by minimizing the approximate $L_{2,0}$ norm. In particular, the $L_{2,0}$ norm is approximated by an arctan function to represent spatial sparsity and the approximate $L_{2,0}$ norm minimization problem is solved by a quasi-Newton method. Finally, the proposed algorithm is examined by simulations. Several advantages over existing DOA estimation algorithms were identified. It can perform well with a limited number of snapshots, while the number of sources need not be known *a priori*. Besides, it improves the resolution, and it can also handle the coherent sources well.

Acknowledgment

This work was supported by the National Natural Science Foundation (No. 61072120), the program for New Century Excellent Talents in University (NCET).

References

1. Krim, H.; Viberg, M. Two decades of array signal processing research: The parametric approach. *IEEE Signal Process. Mag.* **1996**, *13*, 67-94.
2. Johnson, D.H.; Dudgeon, D.E. *Array Signal Processing—Concepts and Techniques*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1993.
3. Schmidt, R. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **1986**, *34*, 276-280.
4. Roy, R.; Kailath, T. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 984-995.
5. Stoica, P.; Sharman, K.C. Maximum likelihood methods for direction-of-arrival estimation. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1132-1143.
6. Hyder, M.; Mahata, K. An improved smoothed l_0 approximation algorithm for sparse representation. *IEEE Trans. Signal Process.* **2010**, *58*, 2194-2205.
7. Mourad, N.; Reilly, J. P. Minimizing Nonconvex Functions for Sparse Vector Reconstruction. *IEEE Trans. Signal Process.* **2010**, *58*, 3485-3496.
8. Gorodnitsky, I.F.; Rao, B.D. Sparse signal reconstructions from limited data using FOCUSS: A re-weighted minimum norm algorithm. *IEEE Trans. Signal Process.* **1997**, *45*, 600-616.
9. Cotter, S. Multiple Snapshot Matching Pursuit for Direction of Arrival (DOA) Estimation. In *Proceedings of the European Signal Processing Conference*, Poznan, Poland, 3–7 September 2007; pp. 247-251.
10. Hyder, M.M.; Mahata, K. Direction-of-Arrival estimation using a mixed $L_{2,0}$ norm approximation. *IEEE Trans. Signal Process.* **2010**, *58*, 4646-4655.
11. Stoica, P.; Babu, P.; Li, J. SPICE: A sparse covariance-based estimation method for array processing. *IEEE Trans. Signal Process.* **2011**, *59*, 629-638.
12. Liu, Z.M.; Huang, Z.T.; Zhou, Y.Y. Direction-of-Arrival estimation of wideband signals via covariance matrix sparse representation. *IEEE Trans. Signal Process.* **2011**, *59*, 4256-4270.
13. Malioutov, D.; Cetin, M.; Willsky, A. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.* **2005**, *53*, 3010-3022.
14. Sun, K.; Lin, Y.; Meng, H. Sparse representation for source location with gain/phase errors. *Sensors* **2011**, *11*, 4780-4793.
15. Chen, J.; Huo, X. Theoretical results on sparse representations of multiple-measurement vectors. *IEEE Trans. Signal Process.* **2006**, *54*, 4634-4643.
16. Berg, E.; Friedlander, M.P. Theoretical and empirical results for recovery from multiple measurements. *IEEE Trans. Inf. Theory* **2010**, *56*, 2516-2527.
17. Chankong, V.; Haimes, Y.Y. *Multiobjective Decision Making Theory and Methodology*; Elsevier Science Publishing: New York, NY, USA, 1983.

Appendix—The Conjugate Gradient of $L_{\delta,\lambda}(X)$

From expression of $F_{\delta}(X)$, we have:

$$\frac{\partial F_{\delta}(X)}{\partial X_R} = \frac{1}{\delta^2} \Lambda X_R \quad (18)$$

$$\frac{\partial F_{\delta}(X)}{\partial X_I} = \frac{1}{\delta^2} \Lambda X_I \quad (19)$$

where

$$\Lambda = \text{diag} \left\{ \delta^2 / (\xi_1^4 + 4\delta^4), \delta^2 / (\xi_2^4 + 4\delta^4), \dots, \delta^2 / (\xi_n^4 + 4\delta^4) \right\}$$

According the definition of conjugate gradient, we get:

$$\frac{\partial F_{\delta}(X)}{\partial X^*} = \frac{1}{2\delta^2} \Lambda X \quad (20)$$

Set

$$G(X) = \|\Phi X - Y\|_F^2 = \text{tr} \left[(\Phi X - Y)' (\Phi X - Y) \right]$$

where $\|\bullet\|_F$ denotes Frobenious norm and $\text{tr}[\bullet]$ denotes the sum of diagonal entries. We have:

$$\frac{\partial G(X)}{\partial X_R} = -2 \text{Re} [\Phi' (Y - \Phi X)] \quad (21)$$

$$\frac{\partial G(X)}{\partial X_I} = -2 \text{Im} [\Phi' (Y - \Phi X)] \quad (22)$$

Then the conjugate gradient of $G(X)$ can be written as:

$$\frac{\partial G(X)}{\partial X^*} = \Phi' (Y - \Phi X) \quad (23)$$

From (20) and (23), we have:

$$\frac{\partial L_{\delta, \lambda}(X)}{\partial X^*} = \frac{\partial F_{\delta}(X)}{\partial X^*} + \frac{\partial G(X)}{\partial X^*} = \frac{1}{2\delta^2} \Lambda X + \lambda \Phi' (Y - \Phi X)$$