IMPROVED DOA ESTIMATION WITH ACOUSTIC VECTOR SENSOR ARRAYS USING SPATIAL SPARSITY AND SUBARRAY MANIFOLD

Bo LI and Yue Xian ZOU

Advanced Digital Signal Processing Laboratory, Peking University Shenzhen Graduate School, Shenzhen 518055, China
Email: boli@pku.edu.cn, zouyx@szpku.edu.cn

ABSTRACT

The performance of DOA estimation with scalar sensor arrays using spatial sparse signal reconstruction (SSR) technique is affected by the grid spacing. In this paper, we formulate the DOA estimation with the acoustic vector sensor (AVS) arrays under SSR framework. A coarse-to-fine DOA estimation algorithm has been developed. The source spatial sparsity and the inter-relations among the manifold matrices of the AVS subarrays are jointly utilized to eliminate the grid effect in the SSR technique and the improvement of the overall DOA estimation performance is achieved at low complexity. Simulation results show that the proposed method effectively mitigates the DOA estimation bias caused by off-grid sources. Interestingly, our method gives good DOA estimation accuracy when sources are closely located.

Index Terms— direction of arrival estimation, acoustic vector sensor, sparse signal reconstruction, manifold vector, signal subspace.

1. INTRODUCTION

Direction of Arrival (DOA) estimation using the acoustic vector sensor (AVS) arrays firstly has been investigated in the field of acoustic direction-finding applications [1]. Generally, each AVS unit consists of an omnidirectional sensor spatially collocated with two to three orthogonally directional sensors, where particle velocity sensors or differential microphones are often used as the directional sensors according to its applications. The AVS arrays are able to provide more information than the commonly used scalar sensor arrays. Many DOA estimation algorithms with the scalar sensor arrays can be extended for AVS arrays, with which the better performance of DOA estimation is expected [1]-[4]. For example, in [3], Wong *et al.* proposed a novel non-spatial realization of ESPRIT using AVS array, where the interrelations among the sensors in one AVS unit has been explored. Recently, a new DOA estimation framework based on sparse signal reconstruction (SSR-DOA) with scalar sensor arrays [5] quickly gains intensive attention. Using the same framework as that in [5], Malioutov *et al*. developed a ℓ_1-SVD method by imposing ℓ_1-norm penalties to enforce sparsity and using the singular value decomposition (SVD) of the multi-snapshot data matrix to achieve joint processing with less computation [6]. Research outcomes show that the SSR-DOA estimation methods are able to achieve super-resolution and robustness to noise. However, the performance of the SSR-DOA estimation methods is affected by the grids sampled in the space, which is termed as *grid effect*. Candidate grids need to be dense for high DOA estimation accuracy. The iterative grid refinement strategy [6] is a compromise between the estimation accuracy and the computation cost. From the view of model error, sparse total least square (STLS) [7] is proposed to resolve the off-grid DOAs. Unfortunately, STLS is non-convex. Off-grid error linearization and constraint relaxation are used in [8] to effect a convex problem.

The basic idea of this study lies on eliminating the grid effect by exploring the inter-relations between the subarrays of an AVS array under SSR framework. We proposed a coarse-to-fine multisource DOA estimation algorithm. In the coarse stage, according to the computational requirement, a relative low accuracy (coarse) source elevation and azimuth angles are estimated by SSR-DOA estimation methods using the omnidirectional subarray. In the fine stage, based on the initial estimates, we exploit the inter-relations of subarray manifold between the omnidirectional subarray and each directional one and a closed form of the DOA estimation is derived. Simulations show that the proposed algorithm is able to effectively reduce the DOA estimation bias caused by grid effect with low complexity. In addition, it gives good DOA estimation accuracy when sources are closely located.

2. DATA MODEL FOR AN AVS ARRAY

An AVS array of M identically oriented AVS units is considered. Each AVS unit is composed of J constituent sensors, where each sensor is termed as an *AVS-component*. Without loss of generality, J is chosen as *four* in this study. Supposing there are K ($K<M$) narrowband acoustic signals

This work was supported by the Chinese National 863 Program (2007AA11Z224), Shenzhen Science & Technology Program (SZKJ-200716) and the special foundation of president of PKUSZ (2010009).

$s_k(t)$ (k=1,…,K) impinging upon the AVS array. An AVS unit with J AVS-components has the following J×1 manifold vector for the kth spatial source coming from (θ_k, ϕ_k) [3]:

$$\mathbf{a}^{(J)}(\theta_k,\phi_k) \equiv [u_k, v_k, w_k, 1]^T, \quad (1)$$

where $(.)^T$ denotes the matrix transposition, $\theta_k \in [0°, 180°)$ is the elevation angle, $\phi_k \in [0°, 360°)$ is the azimuth angle, and $u_k = \sin\theta_k \cos\phi_k$, $v_k = \sin\theta_k \sin\phi_k$ and $w_k = \cos\theta_k$ is the x, y, z direction-cosine. Hence, a JM×1 array *manifold vector* of the AVS array for the kth spatial source coming from (θ_k, ϕ_k) is given as

$$\mathbf{a}(\theta_k,\phi_k) \equiv \mathbf{a}^{(J)}(\theta_k,\phi_k) \otimes \mathbf{q}_k, \quad (2)$$

where $\otimes$ denotes the Kronecker-product operator, and $\mathbf{q}_k$ is the M×1 steering vector for a spatial source located at (θ_k, ϕ_k) [9]. Therefore, for K spatial sources, the output data of the AVS array at time t can be modeled as:

$$\mathbf{x}(t) = \sum_{k=1}^{K} \mathbf{a}(\theta_k,\phi_k)s_k(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (3)$$

$$\mathbf{A} \equiv [\mathbf{a}(\theta_1,\phi_1),...,\mathbf{a}(\theta_K,\phi_K)], \quad (4)$$

$$\mathbf{s}(t) \equiv [s_1(t)...s_K(t)]^T, \ \ \mathbf{n}(t) \equiv [n_1(t)...n_{JM}(t)]^T, \quad (5)$$

where $\mathbf{x}(t)$ is the JM×1 measurement vector, $\mathbf{n}(t)$ is the JM×1 additive zero-mean white Gaussian noise vector, and $\mathbf{A}$ is the AVS array manifold matrix for all K sources. When L snapshots (t=t_1,…,t_L) are taken, the output signal of the AVS array can be written into a matrix form

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N}, \quad (6)$$

where $\mathbf{X}$=[$\mathbf{x}(t_1)$...$\mathbf{x}(t_L)$], $\mathbf{S}$=[$\mathbf{s}(t_1)$…$\mathbf{s}(t_L)$] and $\mathbf{N}$=[$\mathbf{n}(t_1)$…$\mathbf{n}(t_L)$]. As a result, the DOA estimation problem using an AVS array is formulated to determine (θ_k, ϕ_k), (k=1,…,K) from the JM×L data matrix $\mathbf{X}$.

From another point of view, an AVS array can be viewed as J co-located subarrays. Each subarray is composed of M AVS components of the same type. We denote them as u-, v-, w-, o-subarray, respectively. For instance, the o-subarray is formed by M omnidirectional components in the AVS array. Following the formulation from (1) to (6) and with some simple derivations, the output data from each subarray can be expressed as follows

$$\begin{aligned} \mathbf{X}_u &= \mathbf{A}_u\mathbf{S} + \mathbf{N}_u, \mathbf{X}_v = \mathbf{A}_v\mathbf{S} + \mathbf{N}_v, \\ \mathbf{X}_w &= \mathbf{A}_w\mathbf{S} + \mathbf{N}_w, \mathbf{X}_o = \mathbf{A}_o\mathbf{S} + \mathbf{N}_o, \end{aligned} \quad (7)$$

where block matrices $[\mathbf{X}_u^T \mathbf{X}_v^T \mathbf{X}_w^T \mathbf{X}_o^T]^T$, $[\mathbf{N}_u^T \mathbf{N}_v^T \mathbf{N}_w^T \mathbf{N}_o^T]^T$, and $[\mathbf{A}_u^T \mathbf{A}_v^T \mathbf{A}_w^T \mathbf{A}_o^T]^T$ are row partitions of $\mathbf{X}$, $\mathbf{N}$ and $\mathbf{A}$, respectively. Specifically, as an example, $\mathbf{X}_o$, $\mathbf{N}_o$ and $\mathbf{A}_o$ represents the sub-matrices of the output data, the additive noise and the subarray manifold associated with o-subarray, respectively. It is clear that the o-subarray is identical to a scalar sensor array in [6]. Inserting (2) into (4) and comparing to the row partition of $\mathbf{A}$, the following relations can be derived

$$\mathbf{A}_u = \mathbf{A}_o\mathbf{\Lambda}^u, \quad \mathbf{A}_v = \mathbf{A}_o\mathbf{\Lambda}^v, \quad \mathbf{A}_w = \mathbf{A}_o\mathbf{\Lambda}^w, \quad (8)$$

where $\mathbf{\Lambda}^u$=$diag(u_1,…,u_K)$, $\mathbf{\Lambda}^v$= $diag(v_1,…,v_K)$ and $\mathbf{\Lambda}^w$= $diag(w_1,…,w_K)$ are diagonal matrices and $\mathbf{A}_o$=[$\mathbf{q}_1$,…,$\mathbf{q}_K$]. It is interesting to see that the equation (8) shows the subarray manifold inter-relations between the o-subarray and other directional subarrays. Substitute (8) into (7), we get

$$\begin{aligned} \mathbf{X}_u &= \mathbf{A}_o\mathbf{\Lambda}^u\mathbf{S} + \mathbf{N}_u, \mathbf{X}_v = \mathbf{A}_o\mathbf{\Lambda}^v\mathbf{S} + \mathbf{N}_v, \\ \mathbf{X}_w &= \mathbf{A}_o\mathbf{\Lambda}^w\mathbf{S} + \mathbf{N}_w, \mathbf{X}_o = \mathbf{A}_o\mathbf{S} + \mathbf{N}_o. \end{aligned} \quad (9)$$

3. PROPOSED METHOD

In this section, we will develop a coarse-to-fine DOA estimation algorithm under SSR framework.

3.1. The coarse stage with o-subarray based on SSR

Since the o-subarray can be viewed as a scalar sensor array, DOA estimation with the o-subarray can be implemented under the SSR framework using the methods proposed in [6]. First of all, the whole spatial space is sampled by N (N>>M) grids and an angle set Θ=$\{(\bar{\theta}_1\ \bar{\phi}_1),...,(\bar{\theta}_N\ \bar{\phi}_N)\}$ is formed accordingly. Then a predefined overcomplete manifold matrix of the o-subarray can be constructed according to the predefined angle set Θ:

$$\mathbf{\Psi} \equiv [\mathbf{a}(\bar{\theta}_1\ \bar{\phi}_1),...,\mathbf{a}(\bar{\theta}_N\ \bar{\phi}_N)], \quad (10)$$

where $\mathbf{\Psi}$ is of JM×N. Obviously, if bigger N is used, then the smaller grid spacing is formed and it is more probably to match the true spatial source DOA with angle set Θ. Therefore, with the assumption of sufficient small grid spacing, using $\mathbf{\Psi}$ instead of $\mathbf{A}_o$, the last equation in (9) can be re-formulated as

$$\mathbf{X}_o = \mathbf{\Psi}\mathbf{Z} + \mathbf{N}_o, \quad (11)$$

where $\mathbf{Z}$ is of N×L. It can be seen that $\mathbf{Z}$ should have only K nonzero rows which correspond to the K spatial sources. Therefore, DOA estimation can be achieved by locating the indices of nonzero rows in $\mathbf{Z}$. This is the core concept of SSR-DOA estimation methods [6]. In this formulation, $\mathbf{\Psi}$ is predefined, $\mathbf{X}_o$ is measured from o-subarray, and $\mathbf{Z}$ needs to be estimated. Research shows that DOA estimation using (11) is an inverse problem regularized by the row sparsity and it can be solved by several techniques [5][6]. In our research, we use the ℓ_1 -SVD technique [6] due to its computational efficiency and better robustness to noise. Supposing that the K dominant rows in recovered $\mathbf{Z}$ are with indices $I_k \in \{1,…,N\}$ (k=1,…,K), the DOA estimation results can be expressed as

$$(\hat{\theta}_k, \hat{\phi}_k) = (\bar{\theta}_{I_k}\ \bar{\phi}_{I_k}), \quad k = 1,...,K. \quad (12)$$

According to (12), an estimate of the manifold $\mathbf{A}_o$ can be expressed as

$$\hat{\mathbf{A}}_o = \mathbf{\Psi}(:, I), \quad I = \{I_1,...,I_K\}. \quad (13)$$

It is straightforward that the DOA estimation formulation for an AVS array under SSR framework can also be formulated at higher computation cost. However, we find that there are no considerable performance improvements. Analyzing (12), the smaller grid spacing. i.e., larger N, is

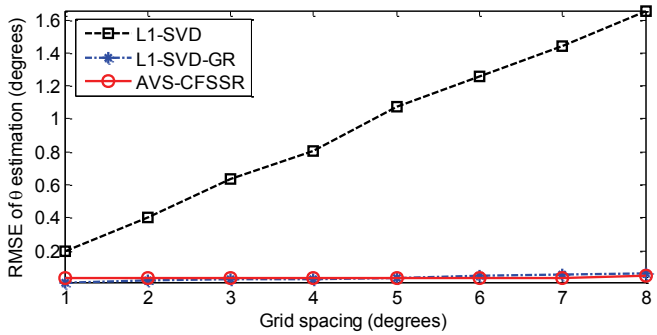

Fig. 1. RMSE of the estimation versus grid spacing.

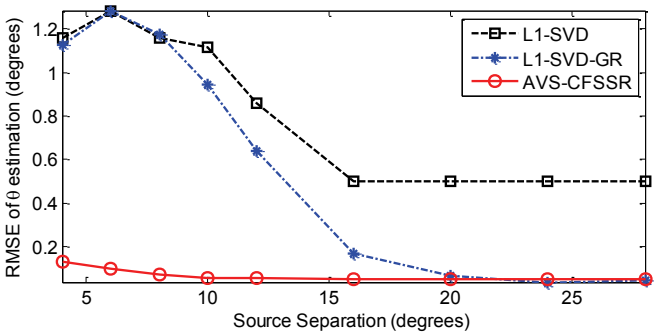

Fig. 2. RMSE of the estimation versus source separation.

required for higher DOA estimation accuracy, which leads to the higher computation cost. If we reduce N, the probability of the spatial source off-grid increases, and the resultant SSR-DOA estimation accuracy decreases. To compromise between the DOA estimation accuracy and the computation cost, the grid refinement is one of the solutions [6]. However, the grid refinement algorithm still requires relative small grid spacing to avoid numerical unstable problem [6]. Moreover, the refining process needs to solve the sparse regularized inverse problem for each iterative level. In short, the grid refinement is still computation demanding. In this study, we choose a small N in the coarse stage and the estimated DOAs are used as an initial estimation for the fine stage. In the next section, subspace technique is used to explore the inter-relations among the AVS subarrays.

3.2. The fine stage based on subspace technique

If there is no noise present, substituting the estimate of $\mathbf{A}_o$ in (13) into (9), we could obtain the diagonal matrices $\mathbf{\Lambda}^u$, $\mathbf{\Lambda}^v$ and $\mathbf{\Lambda}^w$. The DOA estimation can be derived immediately. To reduce the adverse impact of the noise on the performance of the DOA estimation, we employ the subspace technique in the fine stage. Let's define the correlation matrix of the output data $\mathbf{X}$ as $\mathbf{R}_{\mathbf{XX}}=\mathbf{X}\mathbf{X}^T$, which is of $JM\times JM$. The eigen decomposition of $\mathbf{R}_{\mathbf{XX}}$ gives the signal subspace matrix $\mathbf{E}_s$ in size of $JM\times K$, which is composed of K signal eigenvectors. Matrix $\mathbf{E}_s$ holds the following relation with $\mathbf{A}$ [3]:

$$\mathbf{E}_s \approx \mathbf{A}\mathbf{T}, \tag{14}$$

where $\mathbf{T}$ is a nonsingular $K\times K$ transform matrix. Noted that, for the noiseless case, $\mathbf{E}_s=\mathbf{A}\mathbf{T}$. Substituting the row partition of $\mathbf{A}$ into (14), $\mathbf{E}_s$ can also be partitioned as

$$[\mathbf{E}_{su}^T, \mathbf{E}_{sv}^T, \mathbf{E}_{sw}^T, \mathbf{E}_{so}^T]^T \approx [\mathbf{A}_u^T, \mathbf{A}_v^T, \mathbf{A}_w^T, \mathbf{A}_o^T]^T\mathbf{T}. \tag{15}$$

From (8) and (15), the signal-subspace matrices satisfy

$$\mathbf{E}_{su} \approx \mathbf{A}_u\mathbf{T} = \mathbf{A}_o\mathbf{\Lambda}^u\mathbf{T}, \tag{16}$$

$$\mathbf{E}_{so} \approx \mathbf{A}_o\mathbf{T}. \tag{17}$$

With the estimation of $\mathbf{A}_o$ from (13), $\mathbf{T}$ can be computed from (17), which can be substituted into (16) to yield

$$\hat{\mathbf{\Lambda}}^u = \mathbf{\Psi}(:,I)^{\perp}\mathbf{E}_{su}\mathbf{E}_{so}^{\perp}\mathbf{\Psi}(:,I), \tag{18}$$

where $(.)^{\perp}$ denotes the pseudo-inverse operator. Similarly analyzing the relations between $\mathbf{E}_{sv}$, $\mathbf{E}_{sw}$ and $\mathbf{E}_{so}$ separately, we can obtain the estimates of $\mathbf{\Lambda}^v$ and $\mathbf{\Lambda}^w$ as follows

$$\begin{aligned}\hat{\mathbf{\Lambda}}^v &= \mathbf{\Psi}(:,I)^{\perp}\mathbf{E}_{sv}\mathbf{E}_{so}^{\perp}\mathbf{\Psi}(:,I),\\ \hat{\mathbf{\Lambda}}^w &= \mathbf{\Psi}(:,I)^{\perp}\mathbf{E}_{sw}\mathbf{E}_{so}^{\perp}\mathbf{\Psi}(:,I).\end{aligned} \tag{19}$$

From the definitions of $\mathbf{\Lambda}^u$, $\mathbf{\Lambda}^v$ and $\mathbf{\Lambda}^w$, then we have the estimation for u_k, v_k and w_k (k=1,…,K) as follows:

$$\hat{u}_k = [\hat{\mathbf{\Lambda}}^u]_{kk}, \quad \hat{v}_k = [\hat{\mathbf{\Lambda}}^v]_{kk}, \quad \hat{w}_k = [\hat{\mathbf{\Lambda}}^w]_{kk}, \tag{20}$$

where $[.]_{kk}$ denotes the kth diagonal element of a matrix. Following the definitions of u_k, v_k, w_k, the closed-form of the DOA estimation can be expressed as follows

$$\hat{\theta}_k = \cos^{-1}\hat{w}_k, \quad \hat{\phi}_k = \tan^{-1}\hat{v}_k/\hat{u}_k. \tag{21}$$

It is noted that the above fine stage is an ESPRIT-like subspace method utilizing the subarray manifold inter-relations based on the estimated manifold $\hat{\mathbf{A}}_o$. Essentially, (18) is the least square solution of $\mathbf{\Lambda}^u$ under the constraints (16)-(17). The constraint (16) contains DOA information in $\mathbf{\Lambda}^u$ and thus the least square solution improves the DOA estimates over the initial ones in $\hat{\mathbf{A}}_o$, which has been verified through several experiments in section 4. The major computation cost is the eigen-decomposition of $\mathbf{R}_{\mathbf{XX}}$, which can be solved by many fast algorithms [9].

4. SIMULATION STUDY

In this section, the performance of the proposed coarse-to-fine DOA estimation algorithm, named by AVS-CFSSR, is evaluated and compared with the ℓ_1-SVD and ℓ_1-SVD-GR algorithms [6] for far-field narrowband source DOA estimation under white Gaussian noise. For AVS-CFSSR, 8 AVS units (M=8) are used to form an ULA, each AVS unit is along the z-axis, and the distance between the adjacent AVS unit is half of the source wavelength. For ℓ_1-SVD, and ℓ_1-SVD-GR, the simulation parameters are selected according to those used in [6]: 8-element ULA is formed, shrinking rate is 3 (γ=3), and 3 refinement levels are used. For comparison fairness, for these three algorithms, the azimuth angle ϕ is set to be zero, and only the estimation of elevation angle θ is considered. The number of snapshots is set to be 100. Root mean square error (RMSE) of the estimates over 200 independent trials is used as the metric.

1) Impact of grid spacing on DOA estimation accuracy

One spatial source is considered. The source elevation angle is randomly generated between 20° to 160° for each trial. To evaluate the DOA estimation accuracy of the algorithms under different grid spacing, the elevation angular space is sampled uniformly with different gird spacing varying from 1° to 8°. In such experiment setup, we can infer that the source may be off-grid with high probability. The SNR is set as 30dB. The RMSE of the elevation estimates versus the grid spacing are plotted in Fig 1. It is clear to see that the ℓ_1-SVD performs worse than ℓ_1-SVD-GR and AVS-CFSSR for all gird spacing. With the increase of the grid spacing, RMSE of ℓ_1-SVD linearly increases, which directly indicates the off-grid effect for ℓ_1-SVD algorithm. The RMSE of ℓ_1-SVD-GR is comparable to the AVS-CFSSR for all grid spacing. From this result, it can be concluded that AVS-CFSSR has similar ability to eliminate the grid effect. In addition, there exist numerical unstable cases for ℓ_1-SVD-GR during experiment trials, especially when the grid spacing is larger than 3°.

2) Performance under different source separations

Two uncorrelated narrow-band sources with different spatial separations are considered. The elevation angle of the first source is fixed to 42.5°, while the elevation angle of the second source varies from 46.5° to 70.5°. The whole elevation space is uniformly sampled with the grid spacing of 2°. The SNR equals to 30dB. The RMSE of the estimated elevation angles versus source separation are plotted in Fig. 2. It is noted that AVS-CFSSR and ℓ_1-SVD-GR perform better than ℓ_1-SVD for all source separations. When two sources separate larger than 16°, the bias of ℓ_1-SVD remains about 0.6° because sources are off-grid; meanwhile, ℓ_1-SVD-GR has much smaller RMSE, which is comparable to that of AVS-CFSSR. But when two sources closely located (the separation is less than 16°), there is large estimation bias for both ℓ_1-SVD and ℓ_1-SVD-GR; while the DOA estimation accuracy of AVS-CFSSR is excellent even when two sources are separated by only 4°. The simulation results reveal that AVS-CFSSR is able to greatly improve the DOA estimation accuracy for closely located sources where ℓ_1-SVD-GR fails.

3) Performance under different noise levels

Two uncorrelated sources are set to be located at (63.3°, 0°) and (73.3°, 0°), respectively. The angle space is uniformly sampled with the grid spacing of 2°. The SNR varies from 0dB to 40dB. The RMSE results shown in Fig. 3 are obtained by averaging over 200 independent trials for each SNR. We can clearly see that AVS-CFSSR performs best when SNR is larger than 5dB. AVS-CFSSR performs comparable to ℓ_1-SVD-GR, but better than ℓ_1-SVD when SNR is from 0dB to 5dB.

5. CONCLUSION

In this paper, a coarse-to-fine SSR-based DOA estimation algorithm (AVS-CFSSR) has been derived by employing the spatial sparse processing and subspace signal technique for an AVS array. The joint utilization of the spatial sparse information and the inter-relations among AVS subarray manifold matrices give the favorable characteristics of AVS-CFSSR. Experiments showed that AVS-CFSSR is a super-high resolution DOA estimation method with less grid effect and smaller estimation bias for the small source separation (about 4°) compared with other two algorithms with same simulation conditions. The performance of AVS-CFSSR may be limited by strong additive noise, which will be further studied.

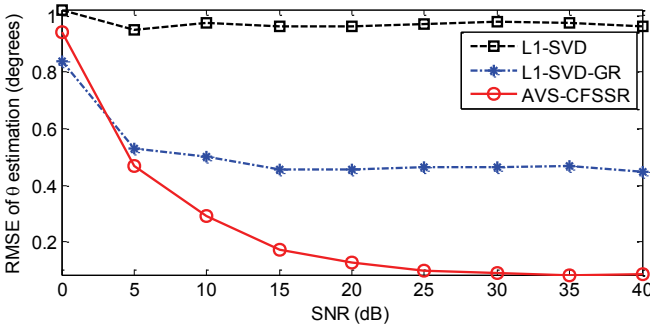

Fig. 3. Performance under different noise levels.

6. REFERENCES

[1] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Trans. Signal Process.*, pp. 2481–2491, Sept. 1994.

[2] M. Hawkes and A. Nehorai, "Acoustic vector-sensor beamforming and Capon direction estimation," *IEEE Trans. Signal Process.*, pp. 2291–2304, Sept. 1998.

[3] K. T. Wong and M. D. Zoltowski, "Closed-form underwater acoustic direction-finding with arbitrarily spaced vector-hydrophones at unknown locations," *IEEE J. Oceanic Engineering*, vol. 22, no. 3, pp. 566-575, July 1997.

[4] D. Levin, S. Gannot and E. Habets, "Direction-of-arrival estimation using acoustic vector sensors in the presence of noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp.105-108.

[5] I. Gorodnitsky and B. Rao, "Sparse signal reconstruction from limited data using focus: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, pp. 600–616, 1997.

[6] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[7] H. Zhu, G. Leus, G.B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol.59, no.5, pp.2002-2016, 2011.

[8] J. Zheng and M. Kaveh, "Direction-of-arrival estimation using a sparse spatial spectrum model with uncertainty," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp.2848-2851.

[9] H.Krim and M. Viberg, "Two decades of array signal processing research. The parametric approach," *IEEE Signal Process. Mag.*, vol.13, pp.67-94 Jul. 1996.