24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

HIGH RESOLUTION BEAMFORMING USING COMPRESSIVE SENSING

Zhongming Xu, Shaoyu Song, Yansong He and Zhifei Zhang

State Key Laboratory of Mechanical Transmission, School of Automotive Engineering, Chongqing University, Shazhengjie 174, Chongqing, China
email: xuzm@cqu.edu.cn

Conventional beamforming(CB) has been widely used to identify sound sources. However, it suffers from grating and side lobes and results in poor spatial resolution of the beamforming map. In this paper, we propose an approach that can dramatically improve the resolution. This method, called high resolution beamforming(HRB), takes the advantages of the compressive sensing which exploits sparsity to reconstructed underlying signal. The linearized Bregman iteration, which adds residue back during each iteration, is used in the proposed method. High computational efficiency is achieved by the HRB because it involves only scalar shrinkages and matrix-vector multiplication. The proposed method is verified by synthetic and experimental data of a microphones array. Compared with the CB and the beamforming utilizing Tikhonov direct regularization method the HRB delivers superior resolution of sources map in wideband frequencies.
Keywords: beamforming, compressive sensing, high resolution

1. Introduction

Beamforming is an important sound sources identification method for manufacturing industry and transportation as it is robustness as well as convenient to implement [1, 2]. The basic idea of conventional beamforming (CB) [3] is that the sampling signal of microphones array are delayed with different time and summed coherently to enhance the strength in the focus position. By plotting the sound pressure in the scan plane, the sources positions can be detected.

Although CB achieves the desired function, this method is recognized as an approach with low resolution. Meanwhile, the grating lobes may be generated and side lobes are emerged with the beamforming outputs [4]. They together lead to so-called ghost image in the presence of sources. To improve the resolution, several techniques have been proposed. Including some deconvolution methods, for instance DAMAS [5] by Brooks and Humphreys, DAMAS2 [6] by Dougherty, DAMAS-C [7] by Brooks, and LORE algorithm [8] by Ravetta, to name just a few. These algorithms increased the resolution, whereas they remained degraded performance in other aspects. For instance, the reference solution of DAMAS and DAMAS2 is an uncorrelated monopole that is unsatisfied in many cases. DAMAS-C requires impractical computational cost, and LORE is limited by the size of scan region. The drawback of deconvolution methods is the increased amount of calculation time due to additional variables introduced in the iteration.

In the advance of signal processing, the emerging compressive sensing(CS) [9, 10] exploits the sparsity of data to recovery an underlying signal with measurements fewer than the Nyquist sampling rate. The sparsity of sound sources distribution holds in many practical cases. Consequently, a wide range of studies have emerged in the field of direction of arrive (DOA) estimation for passive sonar application. CS has been applied to single and multiple snapshots beamforming [11, 12] in which the feasibility was validated.

In this investigation, we propose a high resolution beamforming(HRB) algorithm, which takes the advantage of CS, to improve the resolution of beamforming sources map. It is achieved by eigenvalue decomposing and introducing of linearized Bregman iteration [13]. The kernel of the proposed method involves only matrix-vector multiplication and scalar shrinkages, which enjoys low computational complexity. The proposed method is formulated and verified in the following sections.

2. Methods

2.1 Mathematical formulas

Consider an array of M microphones in the free field of N point sources. The vector of complex-valued sound pressures **p**(w) received by microphones can be expressed as

$$\mathbf{p} = \mathbf{A}\mathbf{q}, \tag{1}$$

where $\mathbf{A}=[\mathbf{a}_1(k), \ldots, \mathbf{a}_M(k)]^T$ is so-called transfer matrix, $\mathbf{a}_m(k)=[e^{-jkr_{m1}}/r_{m1}, \ldots, e^{-jkr_{mN}}/r_{mN}]^T$ is the transfer function between N sources and *m*-th microphone, r_{mn} denotes the distance between the *n*-th source and the *m*-th microphone, k=w/c is the wavenumber, and $\mathbf{q}=[q_1, \ldots, q_N]^T$ is the sources strength vector. The superscript "*T*" denotes transposition operation. The conventional beamforming can be regarded as spatial filter that multiplies microphone signals by complexed-valued weight factors. It is formulated as

$$b(\mathbf{r}_s) = \mathbf{g}(\mathbf{r}_s)^H \mathbf{p}, s = 1, \ldots, S \tag{2}$$

with $\mathbf{g}(\mathbf{r}_s)^H = [e^{-jk(rs-r1)}/|r_1-r_s|, \ldots, e^{-jk(rs-r_M)}/|r_1-r_s|]/M$ denoting the spherical wave steering vector in which $\mathbf{r}_s$ and $\mathbf{r}_m$ is the location of focal point and microphone, respectively. The superscript "*H*" denotes conjugate operation. The $b(\mathbf{r}_s)$ in Eq. (2) is the output of CB that relies on the focus position of the grid. Taking all the $b(\mathbf{r}_s)$, s=1, 2, …, S together, Eq. (2) can be rewritten as

$$\mathbf{B} = \mathbf{G}^H \mathbf{p}. \tag{3}$$

Then, utilizing the cross spectral matrix(CSM) of the CB output, it can be expressed as

$$\mathbf{Q} = \mathbf{G}^H \overline{\mathbf{p}\mathbf{p}^H} \mathbf{G}. \tag{4}$$

The term $\overline{\mathbf{p}\mathbf{p}^H}$ is referred as the CSM of received pressures. Since it is a complex normal matrix, this term can be eigenvalue decomposed as

$$\overline{\mathbf{p}\mathbf{p}^H} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \tag{5}$$

where $\mathbf{U}=[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M]$ is a unitary matrix consisting of the eigenvectors, and $\mathbf{\Lambda}=\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_M)$ is a real-valued diagonal matrix. The λ_i in $\mathbf{\Lambda}$ corresponds to the eigenvector $\mathbf{u}_i$ in **U**. By defining "eigenmode", which is the product of eigenvector and its eigenvalue magnitude, orthogonal subspaces method [14] can be introduced. The eigenmode takes the form

$$\mathbf{v}_i = \sqrt{\lambda_i}\mathbf{u}_i, \quad i = 1, 2, \ldots, M. \tag{6}$$

In a specific "target domain", each eigenmode is produced by the corresponding sources distribution. The eigenmode $\mathbf{v}_i$ and corresponding sources distribution $\mathbf{a}_i$ satisfy

$$\mathbf{v}_i = \mathbf{A}\mathbf{a}_i, \quad i = 1, 2, \ldots, M, \tag{7}$$

where **A** is the transfer matrix from focal points to microphone positions. Solving all the $\mathbf{a}_i$ (i=1, 2, …, M) in Eq. (7) and summing them together, the overall noise sources distribution is collected, i.e.,

$$\bar{\mathbf{q}} = \sum_i |\mathbf{a}_i|^2, \quad i = 1, 2, \ldots, M. \tag{8}$$

The size of matrix **A** is M × S, in which M and S is the number of microphones and focal points. Each eigenmode is a linear combination of sources distribution. In the practical cases, the problem is usually underdetermined because M is generally much less than S. Meanwhile, the sources signal is sufficiently sparse, i.e., the number of nonzero elements in $\mathbf{a}_i$ K satisfies K<<S. Through optimization, we can use the sparsity of $\mathbf{a}_i$ to recover it from the sampling of Eq. (7). The solution can be obtained by solving the following basis pursuit problem

$$\overline{\mathbf{a}}_i = \min \|\mathbf{a}_i\|_1, \quad s.t. \quad \mathbf{v}_i = \mathbf{A}\mathbf{a}_i. \tag{9}$$

Many approaches have been developed to solve this convex optimization problem, including ISTA [15], FISTA [16], SISTA [17], and TWIST [18], to name a few. In this paper, we use linearized Bregman method, which is effective for large-scale L_1 minimization problem, to solve Eq. (9).

The linearized Bregman iteration (LBI) method applied to Eq. (9) takes the form of two-line algorithm

$$\mathbf{t}^{k+1} = \mathbf{t}^k - \mathbf{A}^T(\mathbf{A}\mathbf{u}^k - \mathbf{v}_i) \tag{10}$$

$$\mathbf{u}^{k+1} = \delta \cdot \text{shrink}(\mathbf{t}^k, \mu) \tag{11}$$

where $\mathbf{t}^k$ denotes an auxiliary variable and

$$\text{shrink}(x, \mu) = \begin{cases} x - \mu, x > \mu \\ 0, -\mu \le x \le \mu \\ x + \mu, x < -\mu \end{cases}. \tag{12}$$

The initial values of $\mathbf{u}^k$ and $\mathbf{t}^k$ are 0. The δ is a fixed step size. It was proved by Osher et al in [13] that the iteration of Eq. (10) and Eq. (11) converges for $0<2\delta\|AA^T\|<1$ and sufficient large μ. The soft-thresholding function Eq. (12) is illustrated in Fig. 1.

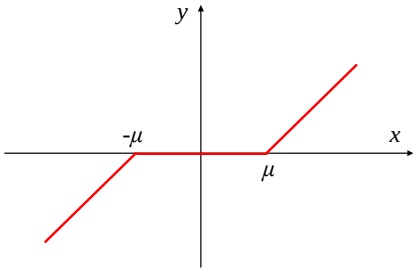

Figure 1: The function of shrink(x, μ) for each component of vector

The modulus of each component in $\mathbf{t}^k$ is individually treated as the input data of the soft-thresholding function. And the phase angle remains unchanged. In the literature of optimization, Eq. (12) is known as proximal forward-backward iterative scheme under the framework of splitting methods [15, 16]. Denoising of response can be achieved by the zeroing during the iteration. Meanwhile, the sparsity of output is realized.

3. Results

In this section, the validity of the HRB is tested. We compared the proposed method with CB and beamforming applying Tikhonov regularization to Eq. (7) in terms of source identification accuracy. The third method is referred as TRB in the following section. The dynamic range of all the display were set to be 20dB.

3.1 Numerical experiments

In the most situation of source localization application, the sources are spatially extended. The practical sound sources can be regarded as the combination of sparsely distributions of simple point sources in the space. For this reason, monopole source and two coherent sources were considered in the simulation.

In the first case, a monopole source placed at coordinate (0, 0, 1 m) was measured by a rectangular array of 81 elements at the plane of z=0. The grid spacing was 10cm. The three beamformers were performed using synthetic data at 2kHz and 5kHz. The signal-to-noise ratio(SNR) was set to 20dB to take consideration of the data disturbance. The performance of CB TRB and HRB is shown in Fig. 2. At the both frequences of 2kHz and 5kHz, the outputs of HRB had the narrowest mainlobe among three beamformers. Meanwhile, for relatively lower frequence (2kHz), all the three beamformers

delivered bigger main lobe than these of high frequency. That means relatively poor spatial resultion in the scan plane. It indicates that similar to CB and TRB, HRB is more suitable for high frequences. The secend extremum in Fig. 2, which corresponded to the sidelobes, disappeared in the results of HRB(c,f). Beasides, by contrast (a,b) and (d,e), TRB generated slightly bigger side lobes. To interpret the result of HRB, which can obtain a narrow main lobe, we have to refer to Eq. (10) and Eq. (11). During iteration Eq. (10) and Eq. (11), the linearized noise in signal is first added back to $\mathbf{t}^k$. Then, the sparse solution is obtained, and the noise is removed via applying the thresholding operator in the next step. It was proved [13] that the consequence $\mathbf{u}^k$ converges to the solution of basis pursuit problem Eq. (9) for a large μ. While regularization method of TRB, which combines side constraint with residual norm, suffers from grating lobes and side lobes as well as relatively poor resolution. Additionally, TRB applied to Eq. (7) relies heavily on regularization parameter.

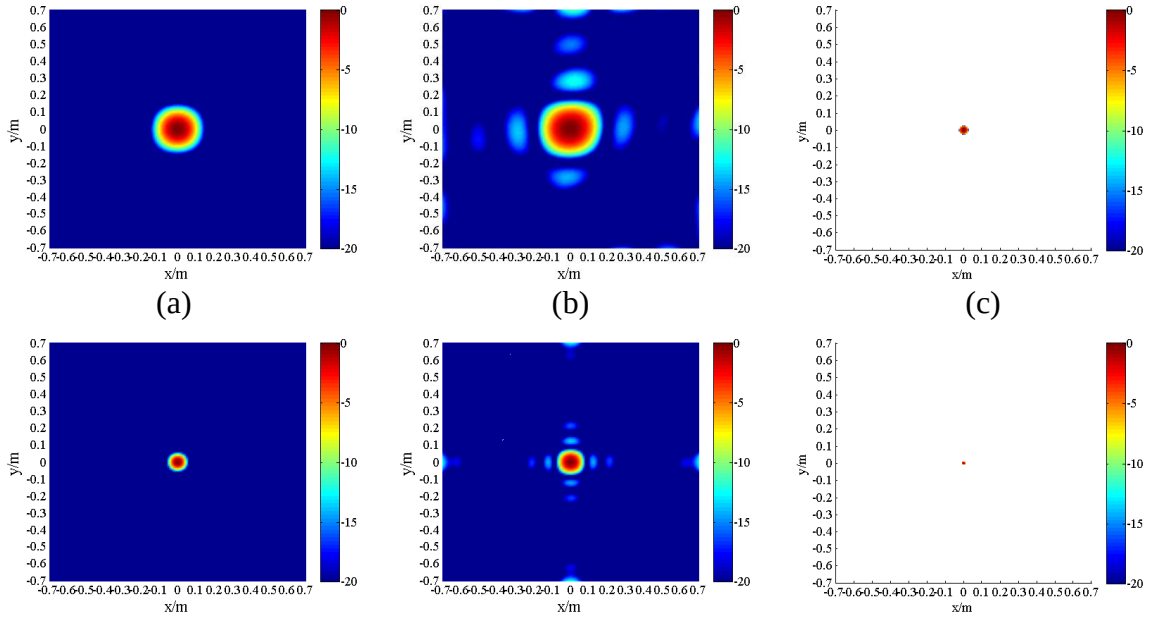

Figure 2: Beamforming outputs of the hypothetical monopole source at (0, 0, 0) with (a) CB, 2kHz; (b) TRB, 2kHz;(c) HRB, 2kHz;(d) CB, 5kHz; (e) TRB, 5kHz;(f) HRB, 5kHz.

Figure 3 illustrates the x-axial performance (normalized sound pressure level) of source identification. The source was correctly detected for all three beamformers reached a maximum at the source position. When the scan point deviated from the source position, the output of HRB decreased sharply to zero everywhere other than in the main lobes region (without the logarithmic process of the display). At frequency of 2kHz, -10 dB mainlobes width of HRB was roughly one fifth of that of CB and TRB. Compared with HRB, the other two methods produced broader main lobes. The sparsity of solution in HRB results in the short blue line in this figure. The proposed method yielded superior quality of sources identification. The zero response in Fig. 2 and Fig. 3 can be explained by shrinkage function defined by Eq. (12). As shown in Fig. 1, the output of HRB for each component whose modulus between $-\mu$ and μ is compulsively set to zero. It can be found in many literatures of signal and image processing to minimizes L_1 norm by the shrinkage [15,16,17]. As contrast, CB is governed by Gabor limit. The filter response in the scan region cannot achieve zero. Indeed, CB and TRB cannot achieve the sparsity of solution with limited number of microphones. HRB delivers zero response mathematically.

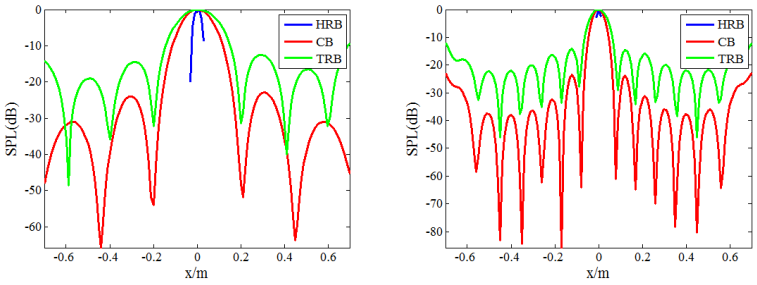

Figure 3: Performance comparison in the x-axis, where (a) 2kHz; (b) 5kHz.

In the second case, two coherent sources were defined at coordinates (-0.4 m, 0, 1 m) and (0.4 m, 0, 1 m). The simulation parameters, including the number of grid points, scan area, SNR, sources distance, and the microphones array were the same as those of the first simulation. The results are shown in Fig. 4. By comparing Fig. 4(d) and (e), it is seen that the radius of main lobe of CB is slightly narrower than that of TRB at 2kHz. Meanwhile, similar to the results of monopole source localization, HRB significantly reduced the radius of main lobe and thus improved the resolution at both frequencies of 2kHz and 5kHz. At higher frequency, HRB obtained sufficiently spare outputs in the scan plane.

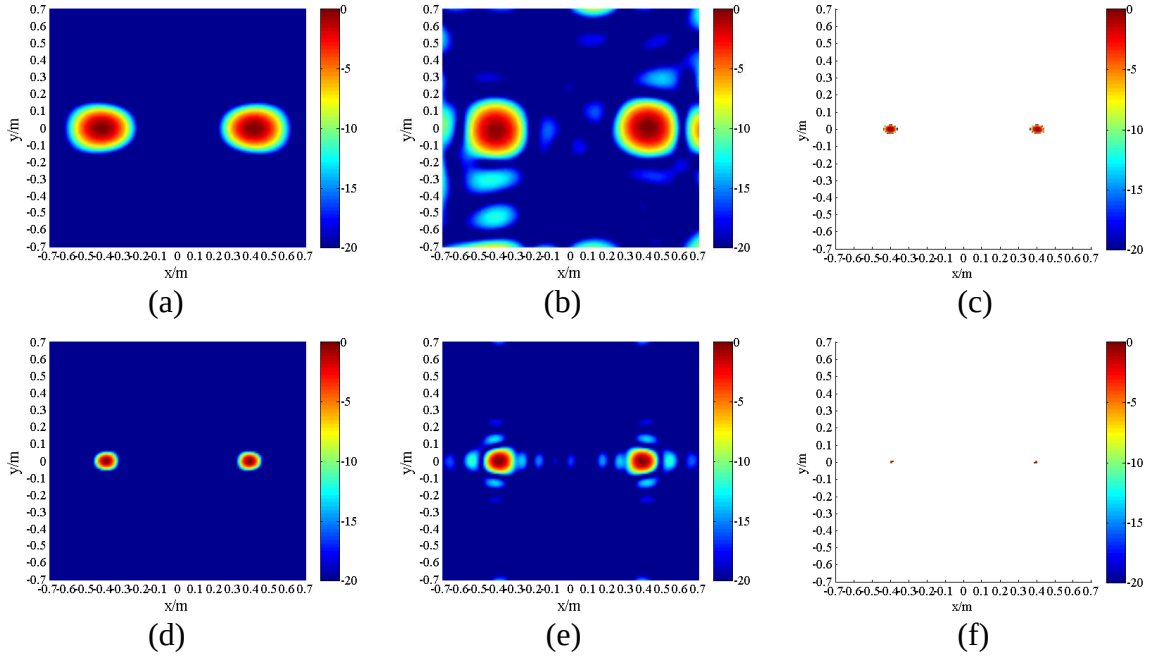

Figure 4: Beamforming outputs of the hypothetical coherent sources at (0, -0.4, 1) and (0, 0.4, 1) with (a) CB, 2kHz;(b) TRB, 2kHz;(c) HRB, 2kHz;(d) CB, 5kHz;(e) TRB, 5kHz;(f) HRB, 5kHz.

The result is more apparent in Fig. 5. At 2kHz, CB gave a fake lobe (local maximum) between the two sources, and TRB obtained stronger side lobe compared withCB. CB delivered better resolution than TRB at high frequency. In the cases of monopole and coherent sources, HRB was able to localise the sources at highest resolution among three methods.

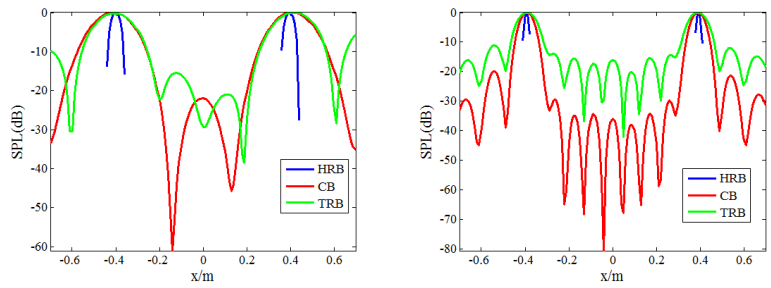

Figure 5: Performance comparison in the x-axis, where (a) 2kHz;(b) 5kHz.

To evaluated the performance of the proposed method in terms of resolution for a wide frequency band, we defined main lobe width B, which is the beam area within -12dB down from the peak value of main lobe. The simulation condition was the same as the aforementioned simulation. The resolution of the three methods are shown in Fig. 6 (the beam width is normalized by the wavelength). It also shows that HRB delivers superior resolution among three methods. Meanwhile, HRB is robust for tested frequency band for both monopole and coherent sources.

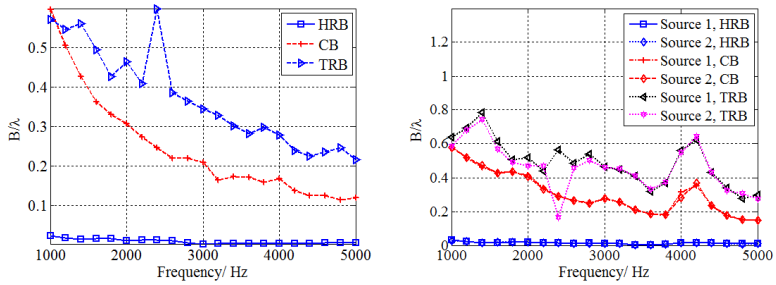

Figure 6: -12dB beam width varying radiation frequency. (a)One monopole. (b)Two coherent sources.

As for the computational cost of the proposed method, it depends on the number of iterations. In our study, the iterations are limited to 300 steps. At one simulation of our PC (AMD FX-4130), the computational time of HRB, CB and TRB is respectively 8.1s, 3.8s and 4.2s. The computational time is acceptable. Additionally, the HRB does not rely on the response of CB like deconvolution methods, which iteratively remove the side lobes of CB outputs.

3.2 Real measurements

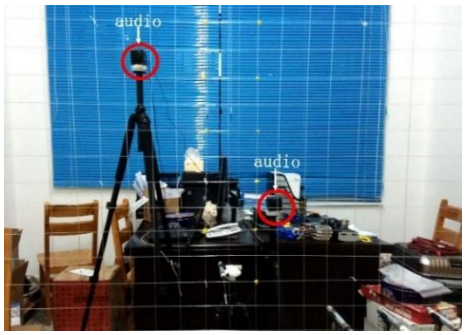

Figure 7: Setup for real measurement in a normal room.

To verify the proposed method in this paper, real measurement was performed in a normal room does not strictly satisfying a free sound field, which was more practicable than anechoic chamber. Two audios type Nogo T10 were set up 1m away from the array. The sources were coherent with single tone at 2kHz and 5kHz. The sound pressure was measured by microphones type B&K 8192-A with sampling rate 16384 Hz. The “Scan&Paint” technology [19], which can describe time stationary

fields by mixing signal and tracking information, was used to reduce the number of microphones. A virtual array of 81 elements spacing 10cm was realized by 4 microphones. Another 27 elements virtual array was previously used to perform 3D beamforming by the authors [20]. Figure 7 shows the experimental layout. The coordinates of audios were about (-0.3m, 0.3m, 1m) and (0.3m, -0.3m, 1m). The array was located at plane z=0.

Figure 8 shows the experimental results at frequencies of 2kHz and 5kHz. Both CB and HRB were able to correctly locate the audios sources. The main lobes were convolved with huge amount of side lobes produced by TRB in Fig. 8 (b) and (e). Meanwhile, it was observed that TRB failed to distinguish the two sources at 2kHz. Compared with the other two methods, result of TRB is more easily contaminated by the complex background noise. Similar to the results of simulation, the HRB method exhibits good performance in spatial denoising and delivering high resolution sources maps.

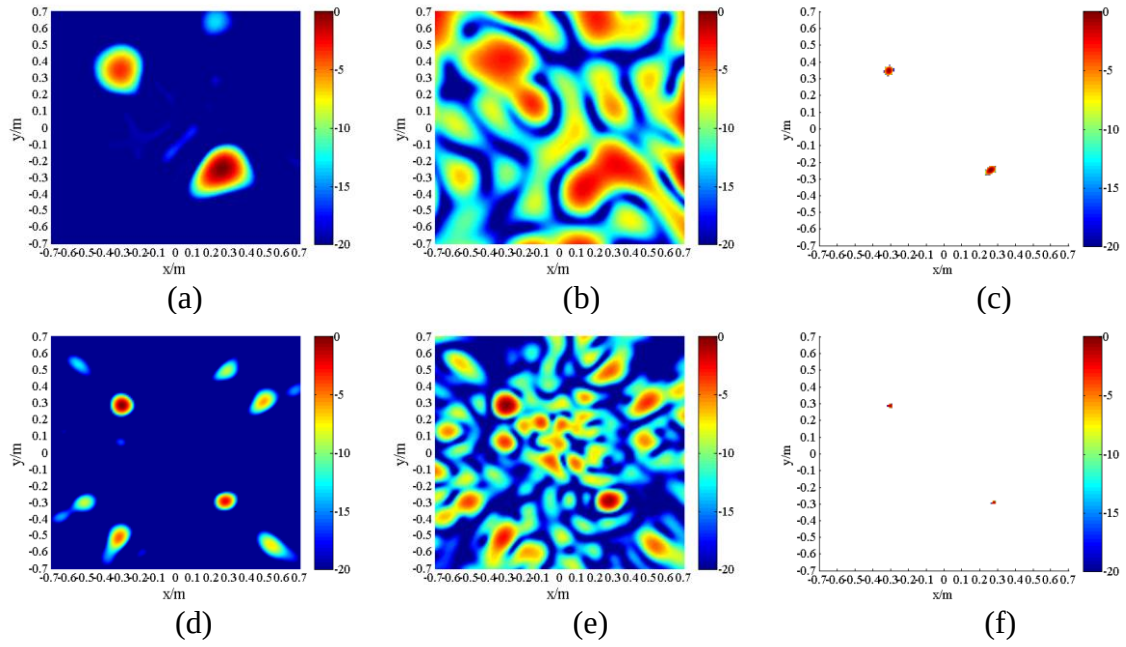

Figure 8: Beamforming outputs of the real measurements with (a) CB, 2kHz;

(b) TRB, 2kHz;(c) HRB, 2kHz; (d) CB, 5kHz; (e) TRB, 5kHz;(f) HRB, 5kHz.

4. Conclusions

In summary, a novel method called high resolution beamforming for sound sources localization based on acoustic array signal processing is proposed in this paper. It is an iterative signal recovery method based on compressive sensing. In comparison with CB and regularization method applied to Eq. (7), which are used to sound source localization, an originality of the proposed method is the using of the sparsity of sources distribution to fulfil the requirements of improving resolution. The linearized Bregman iteration, which involves only vector-matrix multiplication, is used to solve this problem. To validate the proposed method, comparison of beamforming outputs for single and coherent sources were performed. The numerical experiments and real measurements showed that the recognition accuracy is improved for a wide band frequency. The proposed method gives a feasible approach to detect sound sources in the frame of sparsity.

REFERENCES

1 Johnson, D. H., Dudgeon, D. E., *Array Signal Processing*, Prentice Hall, Englewood Cliffs, NJ (1993).

2 Bai, M. R, Lee J., Industrial Noise Source Identification by Using an Acoustic Beamforming System, *Journal of Vibration and Acoustics*, **120** (2), 426-433, (1998).

3 Porteous R, Prime Z, Doolan C. J., Three-dimensional beamforming of dipolar aeroacoustic sources, *Journal of Sound and Vibration*, **355**, 117–134, (2015).

4 Zhao, F. F., Numeric Simulation and Experimental Research of Sound Source Identification with Beamforming Method, Master of Science Thesis, Graduate Project in Mechanical Design and Theory, Shanghai Jiaotong University, (2007).

5 Brooks, T. F., Humphreys, W. M., A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays, *Journal of Sound and Vibration*, **294**, 856–879, (2006).

6 Dougherty, R. P., Extension of DAMAS and benefits and limitations of deconvolution in beamforming, *11th AIAA/CEAS Aeroacoustics Conference*, Monterey, California, 23–25 May, (2005).

7 Brooks, T. F., Humphreys, W. M., Extension of DAMAS phased array processing for spatial coherence determination (DAMAS-C), *12th AIAA/CEAS Aeroacoustics Conference*, Cambridge, Massachusetts, 8–10 May, (2006).

8 Ravetta, P. A., Burdisso, R. A., Wing, F. N., Noise source localization and optimization of phased-array results, *AIAA Journal*, **4** (11), 2520–2533, (2009).

9 Donoho, D. L., Compressed sensing, *IEEE Transactions on Information Theory*, **52** (4), 1289–1306, (2006).

10 Candès, E. J., Romberg, J., Tao, T., Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Transactions on Information Theory*, **52** (2), 489–509, (2004).

11 Li, X., Ma, X. C., Yan, S. F., Hou, C. H., Single snapshot DOA estimation by compressive sampling, *Applied Acoustics*, **74** (7), 926–930, (2016).

12 Gerstoft, P., Xenaki, A., Multiple snapshot compressive beamforming, *2015 49th Asilomar Conference on Signals*, Pacific Grove, California, 8–11 November, (2015).

13 Osher, S., Mao, Y., Dong, B., Yin, W. T., Fast linearized Bregman iteration for compressive sensing and sparse denoising, *Mathematics of Computation*, **8** (1), 93–111, (2010).

14 Long, D., Acoustic source location in wind tunnel tests via subspace beamforming, *41th Aerospace Sciences Meeting and Exhibit*, Rena, Nevada, 6–9 January, (2003).

15 Combettes, P. L., Vajs, V., Signal Recovery by Proximal Forward-Backward Splitting, *Siam Journal on Multiscale Modeling & Simulation*, **4** (4), 1168-1200, (2005).

16 Beck, A., Teboulle, M., A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *Siam Journal on Imaging Sciences*, **2** (1), 183–202, (2009).

17 Bayram I., Selesnick, A subband adaptive iterative shrinkage/thresholding algorithm, *IEEE Transactions on Signal Processing*, **58** (3),1131–1143, (2010).

18 Bioucasdias, J. M., Figueiredo, M. A. T., A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration, *IEEE Transactions on Image Processing*, **16** (12), 2992–3004, (2007).

19 Comesana, D. F., Holland, K., Wind J., Bree, H. E. D., Adapting beamforming techniques for virtual sensor arrays, *4th Berlin Beamforming Conference*, Berlin, (2012).

20 Song, S. Y., Xu, Z. M., Li S., Chen S., Zhang, Z. F., He, Y. S., Improved algorithm of three-dimensional beamforming based on spatial cross-array, *Journal of Vibroengineering*, **18** (8), 5619–5631, (2016).