Blind Dereverberation Wideband Multi-Source 2-D DOA Estimation Method Based on Single Channel Nonnegative Matrix Factorization

Pengju He1*, Mi Qi2, Zijiang Yu2, Qiang Fu3, Mengyang Tang2

1 Research & Development Institute of Northwestern Polytechnical University in Shenzhen, China
2 School of Automation, Northwestern Polytechnical University, Xi’an, China
3 College of Aeronautics Engineering, Air Force Engineering University, Xi’an, China

*Corresponding author’s e-mail: hepengju@nwpu.edu.cn

Abstract. There are many algorithms for DOA estimation based on blind source separation (BSS), which assumes that each sensor perceives a linear instantaneous mixed narrowband signal with known number of sources in noise-free or low noise environments. However, in practical applications, the number of mixed signal sources is unknown, and the perceived signals are often acquired in strong noise and reverberant environments. This paper proposes a novel noise reduction single-channel nonnegative matrix factorization deconvolution (NRSNMFD) wideband multi-source 2-D DOA estimation algorithm. Firstly, LMS adaptive filtering is used to remove the noise of the observation signal. Secondly, the single channel observation signal is decomposed into multi-channel signals composed of multiple intrinsic mode functions (IMFs) by empirical mode decomposition (EMD). The number of signal sources is estimated and the determined signals are reconstructed. Thirdly, the dereverberation of the signals are realized by NMFD method. The GCC-PHAT method is used to estimate the delay of the signal source. Finally, signal source location is realized based on the delay. The simulation and actual test results show that the proposed algorithm can estimate 2-D DOA with high accuracy under noise and reverberation environment, providing a new method for engineering application.

1. Introduction

DOA estimation of wideband signal sources is of great practical value in many fields, and it has application requirements in MIMO radar [1], MIMO sonar [2,3], indoor speech enhancement [4] and other fields. At present, the main methods of DOA estimation are beamforming [5], spectrum analysis [6], TDOA [7], compressed sensing (CS) [8] and BSS [9]. DOA estimation methods based on BSS has been paid attention and studied by academia because it can realize blind source location estimation when the source signal and transmission channel are unknown.

Many DOA estimation methods based on BSS have been proposed by relevant scholars. Takuya Yoshioka [10] regarded reverberation and BSS as two networks, and used WPE algorithm and FastICA algorithm to perform dereverberation and blind separation respectively, and then jointly optimized two network parameters to estimate each signal source. Yiteng [11] first decomposed the MIMO system into multiple SIMO systems, then used Bezout theorem to eliminate the influence of indoor reverberation, and finally adopted unconstrained standardized multi-channel frequency domain LMS algorithm to achieve blind separation of speech signals, but it did not involve the study of sound source location. In
[12], based on the special structure of the steering vector of multipath signals, an algorithm based on CS and ICA was proposed. The ICA algorithm was used to obtain the steering vector containing multipath component information. The CS theory was used to estimate the direct component DOA. However, this method is only applicable to blind separation of linear instantaneous mixtures, and does not consider the reverberation problem in practical application.

These DOA estimation methods based on BSS mainly focus on the theoretical simulation of multi-source linear instantaneous mixing with known number of signal sources, and the 2-D DOA estimation in noise and reverberation environment is rarely reported. In this paper, a new noise reduction single-channel nonnegative matrix factorization deconvolution (NRSNMFD) wideband multi-source 2-D DOA estimation algorithm is proposed. Firstly, the LMS adaptive filtering is used to remove the observed signal noise. Secondly, eigenvalue method is used to estimate the number of source signals and EMD is used to reconstruct multi-channel signals. Thirdly, NMFD is used to realize blind separation and dereverberation of signals. Finally, the GCC-PHAT is used to estimate the delay of each signal source, and then the position of the source signal is calculated according to the delay.

The remaining of this paper is organized as follows. The mathematical model of single-channel convolutional BSS is established in Section 2. NRSNMFD is introduced in Section 3. The time difference DOA estimation is described in Section 4. Section 5 introduces the NRSNMFD-GCC-PHAT 2-D DOA estimation algorithm proposed in this paper. The experimental validation is reported in Section 6. Conclusions are discussed in Section 7.

2. Single channel convolutional mixed model
In practical applications, due to the influence of environmental noise and reflection of indoor walls, floors and ceilings, the propagation path and time delay of sound sources are different. The signals received by each array element are convolutional mixtures of multiple signals with time delay sources. Therefore, the signal received by a single array element in reverberation environment can be represented by a single channel convolutional mixing model.

There are independent source signals, which are received by a single array element after traveling through different paths. Because of the delay of propagation time, the mixed signal $x(t)$ received by the array element is related to the source signal at the current time and the previous $T$ time. Equation (1) represents a mathematical model of single channel convolution mixing.

$$x(t) = \sum_{i=1}^{K} \sum_{\tau=1}^{T} a_i(\tau) s_i(t-\tau) + \epsilon(t)$$  \hspace{1cm} (1)

Where $x(t)$ is single channel convolution mixing signal, $a_i(\tau)$ is the FIR filter between the $i$-th source signal and the sensor, $\epsilon(t)$ is additive white Gaussian noise with mean of zero and variance of $\sigma^2$, $s_i(t-\tau)$ is the time delay component of the $i$-th source signal passing through different paths, $K$ is the number of source signals and $T$ is the order of filter.

3. Blind Dereverberation of Single Channel Nonnegative Matrix for Noise Reduction

3.1. LMS adaptive filtering for noise reduction
The algorithm uses the Minimum Mean Square Error (MMSE) criterion to minimize the mean square error between the actual output value of the filter and the expected value. The input signal at $n$ time is represented by $X(n)$, the weight coefficient of the adaptive filter at $n$ time is represented by $W(n)$, and the expected output at $n$ time is represented by $d(n)$.

Table 1. LMS adaptive filtering method

| Step | Action |
|------|--------|
| 1.   | Set the initial value of the weight coefficient: $W(0) = 0$; |
| 2.   | Calculate the estimated value of the actual output of the filter: $y(n) = W^T(n)X(n)$; |
| 3.   | Calculate estimation error: $e(n) = d(n) - y(n)$; |
4. Automatically adjust filter parameters: \( W(n+1) = W(n) + \mu(n)e(n)X(n) \);
5. Let \( W(n) = W(n+1) \), repeat steps 2-4.

The LMS adaptive filtering [13] method is shown in Table 1. where \( 0 < \mu < \lambda_{\text{max}}^{-1}, \lambda_{\text{max}} \) is the largest eigenvalue of the correlation matrix.

3.2. EMD over-determined mapping, estimation of source number and determination reconstruction
In this paper, EMD method is used to decompose the denoised single-channel mixed signal into multiple IMFs components to realize multi-channel mapping, which can be seen in [14]. On the basis of EMD method, eigenvalue method is used to estimate the number of independent components in single channel mixed signals. The correlation matrix of mixed signals is decomposed into eigenvalues by singular value decomposition, and then the number of sources is determined according to the information of eigenvalues, as shown in Table 2.

| Table 2. Estimation of the number of source signals based on EMD & eigenvalue method |
|---------------------------------|---------------------------------|
| 1. Decompose a single channel signal using EMD to obtain IMFs \( X_{\text{imf}}(t) = (c_1(t), c_2(t), \ldots, c_n(t), r_p(t)) \); |
| 2. Calculate the correlation matrix of the IMF: \( R_{\text{imf}} = E[X_{\text{imf}}(t)X_{\text{imf}}^H(t)] \); |
| 3. Perform singular value decomposition on the correlation matrix \( R_{\text{imf}} \) to obtain a set of eigenvalues \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0 \); |
| 4. Calculate the falling speed ratio of adjacent eigenvalues, \( k_i = \lambda_i / \lambda_i, \ldots, k_{m-1} = \lambda_m / \lambda_{m-1} \); |
| 5. Determine the number of source signals, if \( k_i = \max \{k_i, i=1,\ldots,m-1\} \), the number of signal sources is \( r \). |

3.3. Nonnegative matrix factorization deconvolution
NMF for any nonnegative matrix \( V = [v_1, v_2, \ldots, v_n]^T \in R^{m \times n} \), finding two nonnegative matrices \( W \in R^{m \times r} \) and \( H \in R^{r \times n} \) to satisfy equation (2). \( r \) is the rank of the matrix decomposition and \( r \leq \min(m, n) \).

\[
V = W \times H \ s.t. W, H \geq 0 \tag{2}
\]

Paris [15] proposed a NMFD algorithm on the basis of nonnegative matrix. The NMFD method converts equation (2) into the time-frequency domain so that it can use existing methods to complete deconvolution. For a nonnegative matrix \( V \in R^{m \times n} \), find a set of nonnegative matrices \( W(\tau) \in R^{m \times r} (\tau = 1, 2, \ldots, T-1) \) and \( H \in R^{r \times n} \), where \( r \) is the rank of the matrix decomposition, and \( r \leq \min(m, n) \), NMFD can be expressed as shown in equation (3).

\[
V = \sum_{\tau=0}^{T-1} W(\tau) \times H \tag{3}
\]

Where \( V \) represents a reconstructed nonnegative matrix, \( T \) is the time-expanded frame number of the spectrum, \( W(\tau) \) and \( H \) are the base matrix and the coding matrix, respectively. In the blind separation problem, the element \( W_{ij}(\tau) \) in matrix \( W(\tau) \) represents the \( \tau \)-th coefficient of the FIR filter from the \( i \)-th mixed signal to the \( j \)-th source estimated signal. \((\ast)\) is the operation of shifting the original matrix column vector to the right. The position after translation is supplemented by 0, as shown in equation (4).

\[
H = \begin{bmatrix}
h_1 & \cdots & h_{n-1} & h_n \\
0 & h_1 & \cdots & h_{n-1} \\
0 & 0 & h_1 & \cdots & h_{n-1} \\
\cdots & & & & \cdots \\
0 & 0 & \cdots & 0 & h_1 & \cdots & h_{n-1} \\
\end{bmatrix} \ast \tag{4}
\]

\[
E = \|V \odot \ln(V) - V + V\| \tag{5}
\]

In this paper, KL divergence is chosen as the objective function, then NMF can be transformed into the optimization problem shown in equation (5). \( W(\tau) \) and \( H \) are updated for each \( \tau \) during each update.
iteration. For the above optimization problem, the gradient descent method is used to iteratively update \( W(\tau) \) and \( H \). The NMF update rules are as follows.

\[
\begin{align*}
W(\tau) & \leftarrow W(\tau) - \frac{1}{\tau} \left( V H - W(\tau) H \right)
\end{align*}
\]

where \( I \) represents an \( m \times n \) matrix whose elements are all 1, \( \circ \) is a Hadamard product. Iterating according to the update rule, when \( E \) is less than a certain threshold, the algorithm is considered to converge.

4. Time difference DOA estimation

In this paper, the time difference estimation between array elements is realized based on the generalized cross correlation phase transformation (GCC-PHAT) [16]. Taking the linear array of two elements as an example, the same signal source received at both array elements can be represented by equation (7).

\[
x_j(t) = \alpha_j s(t - \tau_j) + n_j(t), \quad j = 1, 2
\]

Where \( s(t) \) is the source signal, \( \tau_j \) is the time delay from the source to the \( j \)-th microphone, \( \alpha_j \) is the attenuation amplitude from the source to the \( j \)-th microphone, and \( n_j(t) \) is additive noise.

Assuming that the time difference between the sound source and the two microphones is represented by \( \tau_{12} \), the cross-correlation function of the speech signal received by the microphone is shown in equation (8).

\[
R_{\alpha_j}(\tau) = \alpha_j \alpha_j R_{\alpha_j}(\tau - \tau_{12})
\]

Where \( R_{\alpha_j}(\tau) \) in GCC-PHAT is given by formula (9). When \( R_{\alpha_j}(\tau) \) takes the maximum value, the corresponding \( \tau_{12} \) is the target delay. Taking a single source as an example, assuming that its x-axis delay is \( \tau_{12} \) and y-axis delay is \( \tau_{13} \), its azimuth \( \theta \) and elevation \( \phi \) can be obtained by using formula (10).

5. NRSNMFD-GCC-PHAT 2-D DOA estimation algorithm

Combining the methods of Section 3 and 4, when there are \( K \) source signals, the observation signal \( x_i(t)(i = 1, 2, \cdots, M + N - 1) \) satisfies the equation (1), and the L-type array is used as the receiving array. The number of array elements is \( M + N - 1 \), and the number of array elements of the x-axis and the y-axis is \( M, N \) and the array spacing is \( d \). A novel NRSNMFD-GCC-PHAT estimation algorithm is proposed based on the idea of multi-algorithm fusion.

1) The LMS adaptive filtering method is used to filter and denoise the \( x_i(t) \), and the denoised signal is \( x'_i(t) \);

2) Decompose \( x'_i(t) \) into multiple IMF components using EMD algorithm to get \( X_{\text{IMF}}(t) = (c_1(t), c_2(t), \cdots, c_n(t), r_n(t))^T \);
3) The eigenvalue method in Section 3.2 is used to estimate the number of source signals in a single channel mixed signal, and the result is $K$.

4) According to $K$ and IMFs, the appropriate signal is reconstructed and converted into time-frequency domain by short-time Fourier transform.

5) The time-frequency signal is decomposed by NMFD algorithm, and the source signal $\phi_{ik}(t)$ is obtained by inverse short-time Fourier transform of the time-frequency spectrum of the separated signal.

6) For the same signal source separated from each array element, GCC-PHAT algorithm is used to estimate the time delay of each signal source, and the time delay of each signal source is $\tau_{mn}^k$.

7) According to the $\tau_{mn}^k$ obtained in the previous step, the two-dimensional direction of arrival is calculated by formula (10).

6. Experiments and results

6.1. Establishment of experiment
In order to verify the effectiveness and practicability of the proposed method, female signal S1 (dr1/fakso/sa1.wav) and male signal S2 (dr1/mdabo/sa2.wav) in TIMIT standard speech library are selected as wideband signal sources, and their effectiveness is verified by computer simulation and reverberation environment measurement.

6.2. Computer simulation experiment
The simulation platform of the experiment is MATLAB R2017a. S1 and S2 are used as source signals. The sampling frequency is 16KHz and the sampling points are 16000. The receiving array selects an L-shaped array with an array element spacing $d = 15cm$. The azimuth and elevation angles of signal sources S1 and S2 are set to $(22, 61)$ and $(37, 68)$. The indoor reverberation is generated by a FIR filter randomly generated by a fourth-order coefficient, and the mixing matrix is as shown in (11).

$$A = \begin{bmatrix}
\tilde{a}_{11} & \tilde{a}_{12} \\
\tilde{a}_{21} & \tilde{a}_{22} \\
\tilde{a}_{31} & \tilde{a}_{32}
\end{bmatrix}$$

$$\tilde{a}_{11} = [0.9571 \ 0.5404 \ 0.7222 \ 0.5537] ; \tilde{a}_{12} = [0.2654 \ 0.9310 \ 0.8027 \ 0.0165]$$

$$\tilde{a}_{21} = [0.4626 \ 0.4605 \ 0.3431 \ 0.2377] ; \tilde{a}_{22} = [0.0516 \ 0.2131 \ 0.3546 \ 0.3983]$$

$$\tilde{a}_{31} = [0.8147 \ 0.9058 \ 0.1270 \ 0.9134] ; \tilde{a}_{32} = [0.6323 \ 0.0975 \ 0.2784 \ 0.5469]$$

The simulation experiments are carried out in the environment of noise-free and 5dB white Gaussian noise respectively. Figure 2 shows the descent rate ratio of the eigenvalue of the covariance matrix of the observed signal at element 1, it can be seen that the number of signal sources estimated by EMD-
eigenvalue algorithm is 2. The DOA estimation results shown in Table 1 show that there is no error in the 2-D DOA estimation of two sources when there is no noise. When 5dB white Gaussian noise is added, the estimation results of S1 and S2 produce small errors. The SNR is increased to 24.23 dB after the Section 3.1 denoising, as shown on the left side of Figure 3.

Table 3. DOA Estimation of Signal Source without Noise/5dB Gauss White Noise

| Signal source | Azimuth (degree) | Elevation (degree) | Azimuth/Elevation error (degree) |
|---------------|-----------------|-------------------|----------------------------------|
| Noise-free    |                 |                   |                                  |
| S1            | 37              | 68                | 0/0                              |
| S2            | 22              | 61                | 0/0                              |
| 5dB Gaussian white noise |
| S1            | 35.7            | 70.2              | 1.3/2.2                          |
| S2            | 19.5            | 63.6              | 1.5/2.6                          |

The simulation results show that the proposed algorithm can accurately estimate the direction of arrival of the signal source in a noise-free environment. In the noise environment, the DOA estimation error of signal source can be small while the SNR can be improved.

6.3. Actual test experiment

The experiment was carried out in a laboratory of the Automation College of Northwestern Polytechnical University with a space size of 7.47 × 3.63 × 3.14 m³ and the sound velocity is 340 m/s. The experimental conditions are consistent with the computer simulation experiments. The placement positions of arrays, S1 and S2 are set to (3.0, 1.2, 0), (5.4, 3.0, 1.2), (6.0, 2.4, 1.1). PC collects data through Labview, triggers two ARM development boards placed in different positions at the same time by RS485, and runs sound playback program. DOA estimation is carried out by using 2-D ISM algorithm and NRSNMFD-GCC-PHAT algorithm, and the results of the two algorithms are compared.

Table 4. Comparison of the results of the 2-D ISM algorithm and the algorithm proposed in this paper

| 2-D ISM algorithm | NRSNMFD-GCC-PHAT algorithm |
|-------------------|-----------------------------|
| Azimuth (degree)  | Elevation (degree)  | Azimuth elevation error (degree) | Azimuth (degree)  | Elevation (degree)  | Azimuth elevation error (degree) |
| S1                | 45.6                       | 57.7                       | 8.6/10.3           | 34.8                       | 64.6                       | 2.2/3.4                          |
| S2                | 41.4                       | 46.5                       | 19.4/14.5          | 18.2                       | 53.8                       | 3.8/7.2                          |

On the right side of Figure 3 is the comparison of the observed signal before and after denoising at element 1 in the actual test environment. The SNR after denoising is increased by 27.48 dB and the denoising effect is obviously improved. In Table 2, the estimation error of 2-D ISM algorithm is large, especially the azimuth error of S2 reaches 19.4 degrees. The total estimation error of the proposed algorithm is small, which may be caused by noise, installation error and sampling rate.

The experimental results show that the proposed algorithm can also estimate the 2-D DOA of each signal source in noise and reverberation environment, and the estimation error is within allowable range of engineering application error.

7. Conclusion

This paper proposes an NRSNMFD-GCC-PHAT algorithm which can realize 2-D DOA estimation of wideband multi-signal source under reverberation environment. It has the functions of signal denoising, multi-channel mapping, signal source number estimation, dereverberation, 2-D DOA estimation. Computer simulation and experimental results show that the proposed algorithm can estimate the position of the signal source without error in a noise-free environment, and can reduce noise interference in noise environments. The algorithm proposed in this paper provides theoretical basis for DOA estimation of wideband multi-source in reverberation environment and provides a relatively complete solution for engineering application.
Acknowledgments
The research was supported by the Basic Research Projects of Shenzhen Knowledge Innovation Program (No. 20170301142145781), Subtopics of National Key R&D Program (No. 2018KF090177), Fund of MIIT(MJ-2017-F-05) and Key Industry Innovation Chain-Industrial Field Projects (No. 2018ZDCXL-G-8-7).

References
[1] Zhang, X., et al. (2010) Direction of Departure (DOD) and Direction of Arrival (DOA) Estimation in MIMO Radar with Reduced-Dimension MUSIC. IEEE Communications Letters, 14(12): 1161-1163.
[2] Huang, J., Zhang, L., Zhang, Q., Jin, Y. and Jiang, M. (2009) Performance analysis of DOA estimation for MIMO sonar based on experiments. In: 2009 IEEE/SP 15th Workshop on Statistical Signal Processing. Cardiff. pp. 269-272.
[3] Shi, W., Huang, J. and Hou, Y. (2012) Fast DOA estimation algorithm for MIMO sonar based on ant colony optimization. Journal of Systems Engineering and Electronics, 2(23): 173-178.
[4] Xenaki, A., Boldt, J.B. and Christensen, M.L. (2018) Sound source localization and speech enhancement with sparse Bayesian learning beamforming. The Journal of the Acoustical Society of America, 143(6): 3912-3921.
[5] Weiss, S., Bendoukha, S., Alzin, A., et al. (2015) MVDR broadband beamforming using polynomial matrix techniques. In: 2015 23rd European Signal Processing Conference. Nice. pp. 839-843.
[6] Jaafer, Z., Goli, S. and Elameer, A.S. (2018) Best Performance Analysis of DOA Estimation Algorithms. In: 2018 1st Annual International Conference on Information and Sciences. Fallujah. pp.235-239.
[7] Farmani, M., et al. (2015) Informed TDoA-based direction of arrival estimation for hearing aid applications. In: 2015 IEEE Global Conference on Signal and Information Processing. Orlando. pp. 953-957.
[8] Zhang, Y., et al. (2018) A new DOA estimation algorithm based on compressed sensing. Cluster Computing, pp:1-9.
[9] Dorfan, Y., et al. (2016) Multiple DOA estimation and blind source separation using estimation-maximization. In: 2016 IEEE International Conference on the Science of Electrical Engineering. Elat. pp:1-5
[10] Yoshioka, T., et al., (2011) Blind Separation and Dereverberation of Speech Mixtures by Joint Optimization. IEEE Transactions on Audio, Speech, and Language Processing. 19(1): 69-84.
[11] Huang, Y., Benesty, J. and Chen J. (2005) A blind channel identification-based two-stage approach to separation and dereverberation of speech signals in a reverberant environment. IEEE Transactions on Speech and Audio Processing, 13(5): 882-895.
[12] Zhao, L., et al. (2017) Direction-of-arrival estimation of multipath signals using independent component analysis and compressive sensing. PLOS ONE, 12(7): 1-17.
[13] Limem, M., Hamdi, M.A., Maaref, M.A. (2016) Denoising uterine EMG signals using LMS and RLS adaptive algorithms. In: 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing. Monastir. pp. 273-276.
[14] He, P. and Chen, X. (2015) A method for extracting fetal ECG based on EMD-NMF single channel blind source separation algorithm, Technol Health Care. 24 Suppl 1: S17-26.
[15] Smaragdis, P. (2004) Non-negative Matrix Factor Deconvolution: Extraction of Multiple Sound Sources from Monophonic Inputs. Proc Ica, 3195: 494-499.
[16] Knapp, C., Carter, G. (1976) The generalized correlation method for estimation of time delay. IEEE Transactions on Acoustics, Speech, and Signal Processing, 24(4): 320-327.