A Microphone Array DOA Estimation Method

Mingyang Ma[a, *], Lie Chen[b], Jincheng Zhang[c], Yapeng Mao[d], Jucai Lin[e] and Jun Yin[f]

Zhejiang Dahua Technology Co., Ltd. 1199 Bin'an Road, Hangzhou 310000, China

[a]mamingyang16@163.com, [b]195517048@qq.com, [c]1456934003@qq.com, [d]1490894376@qq.com, [e]lin_jucai @dahuatech.com, [f]yin_jun@dahuatech.com

*Corresponding author

Keywords: Microphone Array, DOA Estimation, OPAST Algorithm, MUSIC Algorithm

Abstract: This paper applied OPAST algorithm to DOA estimation of microphone array, then compared with traditional MUSIC algorithm by simulation and data processing. MUSIC algorithm needs eigenvalue decomposition when estimating DOA, it's computationally expensive and complex, so it's not easy to implement in engineering. OPAST algorithm obtains the signal subspace by tracking iteration, doesn't need eigenvalue decomposition, the complexity is low and has great advantages in engineering implementation. The root mean square error of DOA estimation of MUSIC algorithm and OPAST algorithm under different SNR is verified by simulation, OPAST algorithm can achieve the same DOA estimation accuracy as MUSIC algorithm when SNR is high, but the complexity is greatly reduced.

1. Introduction

DOA estimation is an important topic in speech signal processing, it is widely used in multimedia systems, smart home systems, and video conferencing systems and so on, and DOA estimation is a technique for determining the location of sound sources by collection sound signals through microphones and then processing them.

The purpose of DOA estimation is to extract the signals received by the array and their characteristic information, and to suppress the interference noise or uninteresting information. The speech signal received by microphone array is wideband signal and non-stationary signal. The DOA estimation theory of wideband signals is developed on the basis of DOA estimation of narrowband signals, but for wideband signals, different signal subspaces due to different array manifold in different frequency domains, which makes the narrowband high resolution DOA estimation method can't be directly applied to DOA estimation of wideband signals. There are two main types of wideband high resolution DOA estimation algorithms: ISM and CSM [1]. ISM algorithm decomposes wideband signals into several narrowband signals, and then processes each narrowband signal separately. Finally, the processing results of each narrowband signal are weighted and synthesized to obtain the DOA estimation results of wideband signals.

The MUSIC and OPAST algorithms studied in this paper are both subspace-based DOA estimation methods based on ISM algorithm, the wideband speech signal is divided into several sub-bands, which are approximately processed as narrowband signals. MUSIC algorithm needs eigenvalue decomposition of the covariance matrix of the received data before estimating the spatial spectrum. Then the signal subspace and noise subspace can be obtained, which requires a large amount of computation. OPAST algorithm obtains the minimum point of the objective function by tracking iteration to approximate the signal subspace, this method can reduce the computational complexity and achieve the same DOA estimation effect as MUSIC algorithm.

2. DOA estimation method for wideband signals

2.1 MUSIC algorithm

MUSIC algorithm is based on narrowband far-filed signals, it is assumed that there are D narrowband far-filed signals incident on uniform linear arrays in space, the number of elements is M , and the space of elements is d .

The received signal of the array can be written in matrix form as follow

$$\mathbf{X}(t) = \mathbf{AS}(t) + \mathbf{N}(t) \tag{1}$$

In the upper form, $\mathbf{X}(t)$ represents M dimension received data vector, $\mathbf{S}(t)$ represents D dimension signal vector, $\mathbf{N}(t)$ represents M dimension noise vector, and $\mathbf{A}$ is $M \times D$ dimension array manifold matrix.

The covariance matrix of the received data can be expressed in the following form

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{XX}^{\mathrm{H}} \tag{2}$$

By eigenvalue decomposition of $\hat{\mathbf{R}}$, signal subspace $\hat{\mathbf{U}}_{\mathrm{S}}$ and noise subspace $\hat{\mathbf{U}}_{\mathrm{N}}$ are obtained.

From the theory of spatial spectral estimation, the spectral estimation formula of MUSIC algorithm can be expressed as [2]

$$P_{MUSIC}(\theta) = \frac{1}{\mathbf{a}^{\mathrm{H}}(\theta)\hat{\mathbf{U}}_{\mathrm{N}}\hat{\mathbf{U}}_{\mathrm{N}}^{\mathrm{H}}\mathbf{a}(\theta)} \tag{3}$$

The calculation steps of MUSIC algorithm are given below

1) Computing covariance matrix of array received data $\hat{\mathbf{R}}$;

2) Eigenvalue decomposition of $\hat{\mathbf{R}}$;

3) obtain signal subspace $\hat{\mathbf{U}}_{\mathrm{S}}$ and noise subspace $\hat{\mathbf{U}}_{\mathrm{N}}$;

4) Searching for the peak of the spectrum, the angle at which the spectrum reaches its peak is the angle at which the signal arrivals.

For narrowband signals, the frequency is relatively constant, so $\mathbf{A}$ depends on the azimuth of the source θ , when the source is a broadband signal, $\mathbf{A}$ depends on frequency f and angle θ . Therefore, it is necessary to construct multiple narrowband models in frequency domain, and then process them with narrowband DOA estimation method.

For wideband signals, the output of the M -th element at sampling time t is

$$\mathbf{X}_m(t) = \sum_{i=1}^{D}\mathbf{s}_i(t-\tau_{mi}) + \mathbf{n}_m(t) \tag{4}$$

Where, τ_{mi} represents the delay of the element m to the signal i relative to the reference element.

DFT transform to frequency domain

$$\mathbf{X}_m(f)=\sum_{i=1}^{D}\mathbf{S}_i(f)e^{-j2\pi f\tau_{mi}}+\mathbf{N}_m(f) \tag{5}$$

The frequency domain representation of the array received data is as follows

$$\begin{bmatrix}\mathbf{X}_1(f)\\ \mathbf{X}_2(f)\\ \vdots\\ \mathbf{X}_M(f)\end{bmatrix}=\begin{bmatrix}e^{-j2\pi f\tau_{11}} & \cdots & e^{-j2\pi f\tau_{1D}}\\ \vdots & & \vdots\\ e^{-j2\pi f\tau_{M1}} & \cdots & e^{-j2\pi f\tau_{MD}}\end{bmatrix}\begin{bmatrix}\mathbf{S}_1(f)\\ \mathbf{S}_2(f)\\ \vdots\\ \mathbf{S}_D(f)\end{bmatrix}+\begin{bmatrix}\mathbf{N}_1(f)\\ \mathbf{N}_2(f)\\ \vdots\\ \mathbf{N}_M(f)\end{bmatrix} \tag{6}$$

Can be expressed by matrix as follows

$$\mathbf{X}(f)=\mathbf{A}(f,\theta)\mathbf{S}(f)+\mathbf{N}(f) \tag{7}$$

Where, $\mathbf{X}(f),\mathbf{S}(f),\mathbf{N}(f)$ represent frequency domain data of array receiving data, signal and noise after DFT.

$$\mathbf{A}(f,\theta)=[\mathbf{a}_1(f,\theta),\mathbf{a}_2(f,\theta),...,\mathbf{a}_D(f,\theta)]$$

$$\mathbf{a}_i(f,\theta)=\left[e^{-j2\pi f\tau_{1i}},e^{-j2\pi f\tau_{2i}},...,e^{-j2\pi f\tau_{Mi}}\right]^T$$

The wideband signal can be divided into several sub-bands in frequency domain, and DOA estimation can be achieved by using narrowband model, this paper mainly uses ISM algorithm to divide the wideband signal into several sub-bands.

2.2 ISM algorithm

Firstly, the signals collected in observation time T_0 are divided into several segments [3], then DFT is performed on each segment of data to obtain J groups uncorrelated narrowband frequency domain components, K is frequency domain snapshot, thus K snapshots can be obtained, writing as $\mathbf{X}_k(f_j)\ (k=1,2,...,K;\ j=1,2,...,J)$. The purpose of ISM algorithm is to estimate the azimuth of multiple targets from these K frequency domain snapshots.

The cross-correlation spectral density in the frequency domain is

$$\boldsymbol{R}_X(f_i)=\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{X}_k(f_i)\boldsymbol{X}_k^H(f_j),\ 1\le j\le J \tag{8}$$

Eigenvalue decomposition of $\boldsymbol{R}_X(f_i)$, we have

$$\boldsymbol{R}_X(f_i)=\sum_{i=1}^{M}\lambda_i\mu_i\mu_i^H \tag{9}$$

Where, the eigenvectors corresponding to eigenvalue $\lambda_i>\sigma^2(i=1,2,...,P)$ constitute the signal subspace $\boldsymbol{U}_s$, the eigenvectors corresponding to eigenvalue $\lambda_i\approx\sigma^2(i=P+1,...,M)$ constitute noise subspace $\boldsymbol{U}_n(f)$, thus, MUSIC spatial spectrum writing as

$$P(\theta)=\frac{1}{\frac{1}{J}\sum_{i=1}^{J}\left\|\boldsymbol{a}^{H}(f_i,\theta)\boldsymbol{U}_n(f_i)\right\|^2} \tag{10}$$

Main steps of ISM algorithm [4]

Step1: Divided the array data into several segments, each segment is sampled N times;

Step2: DFT in every segment, divided the wideband signal into J sub-bands, the central frequency of each sub-band is: $f_1, f_2, ..., f_j$;

Step3: For each frequency point, the frequency domain matrix $\boldsymbol{X}(f_j)$ of the received data of the array can be obtained, then the covariance matrix $\boldsymbol{R}(f_j), j=1,2,...,J$ is calculated;

Step4: To get spatial spectrum by using narrowband DOA estimation for each sub-band, statistical average to get estimation angle.

2.3 OPAST algorithm

Let $\mathbf{x}(t)$ be a sequence of $n\times 1$ random vectors with covariance matrix $\mathbf{C}=E\left\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\right\}$.Consider the problem of estimating the principal subspace spanned by the sequence of dimension $r<n$, assumed to be the span of the r principal eigenvectors of the covariance matrix. For that, consider the scalar function

$$\begin{aligned} J(\mathbf{W}) &= \mathrm{E}\left\{\left\|\mathbf{x}-\mathbf{W}\mathbf{W}^{\mathrm{H}}\mathbf{x}\right\|^2\right\} \\ &= \mathrm{tr}(\mathbf{C})-2\,\mathrm{tr}(\mathbf{W}^{\mathrm{H}}\mathbf{C}\mathbf{W})+\mathrm{tr}(\mathbf{W}^{\mathrm{H}}\mathbf{C}\mathbf{W}\mathbf{W}^{\mathrm{H}}\mathbf{W}) \end{aligned} \tag{11}$$

Where $\mathbf{W}$ is a $n\times r$ matrix, rank is r. Then we consider the minimization problem of $J(\mathbf{W})$.

1) Is there a global minimum $\mathbf{W}$ of scalar function? $J(\mathbf{W})$

2) What is the relationship between the minimum $\mathbf{W}$ and the signal subspace of matrix? $\mathbf{C}$

3) Is there any other local minimum of? $J(\mathbf{W})$

Yang proved the following theorems and gave the answers to the above questions. It is shown in [6] that

Theorem1: $\mathbf{W}$ is a stationary point of $J(\mathbf{W})$ it and only if $\mathbf{W}=\mathbf{U}_r\mathbf{Q}$, where $\mathbf{U}_s$ is a $n\times r$ matrix containing any r distinct eigenvectors of $\mathbf{C}$, and $\mathbf{Q}$ is any $r\times r$ unitary matrix.

Theorem2: All stationary points of $J(\mathbf{W})$ are saddle points, except when $\mathbf{U}_r$ contains the r dominant eigenvectors of $\mathbf{C}$. In this case, $J(\mathbf{W})$ attains the global minimum.

OPAST algorithm is proposed on the basis of PAST algorithm [5]. Minimizing (11) iteratively leads to the following abstract form [6] of the PAST method [5]

$$\mathbf{W}(i)=\mathbf{C}\mathbf{W}(i-1)(\mathbf{W}^{H}(i-1)\mathbf{C}\mathbf{W}(i-1))^{-1} \tag{12}$$

where $\mathbf{C}\mathbf{W}(i-1)$ is the power term and the matrix inverse serves as a normalizer: In tracking applications, one can simply replace the covariance matrix $\mathbf{C}$ with its recursive version $\mathbf{C}(i)=\alpha\mathbf{C}(i-1)+r(i-1)r^{H}(i)$ at the ith iteration, where $\mathbf{C}(i-1)$ is the sample covariance matrix using the data available up to time $i-1$, and α is a forgetting factor chosen between (0,1].

In [6], a fast implementayion is proposed based on the projection approximation $\mathbf{C}(i)\mathbf{W}(i)\approx\mathbf{C}(i)\mathbf{W}(i-1)$, which is clearly valid if the weight matrix product $\mathbf{C}(i)\mathbf{W}(i-1)$, as well as the matrix inverse $(\mathbf{W}^H(i-1)\mathbf{C}(i)\mathbf{W}(i-1))^{-1}$ can be computed. More precisely, if we define $\mathbf{Z}(i)=(\mathbf{W}^H(i-1)\mathbf{C}(i)\mathbf{W}(i-1))^{-1}$, the PAST algorithm can be obtained.

The OPAST algorithm consists of the above algorithm plus an orthonormalization step of the weight matrix at each iteration (i.e., using informal notation) [5]

$$\mathbf{W}(i)=\mathbf{W}(i)(\mathbf{W}(i)^{\mathrm{H}}\mathbf{W}(i))^{-1/2} \tag{13}$$

Where $(\mathbf{W}(i)^{\mathrm{H}}\mathbf{W}(i))^{-1/2}$ denotes an inverse square root of $\mathbf{W}(i)^{\mathrm{H}}\mathbf{W}(i)=\mathbf{I}$. To compute the latter, we use the updating equation of $\mathbf{W}(i)$. Keeping in mind that $\mathbf{W}(i-1)$ is now an orthonormal matrix, we have

$\mathbf{W}^{\mathrm{H}}(i)\mathbf{W}(i)=\mathbf{I}+\|\mathbf{p}(i)\|^2\mathbf{q}(i)\mathbf{q}^{\mathrm{H}}(i)=\mathbf{I}+\mathbf{x}\mathbf{x}^{\mathrm{H}}$

Where we have used the fact that $\mathbf{W}^{\mathrm{H}}(i-1)\mathbf{p}(i)=0$, $\mathbf{I}$ is the identity matrix, and $\mathbf{x}=\|\mathbf{p}(i)\|\mathbf{q}(i)$. Thus

$$\begin{aligned}(\mathbf{W}(i)\mathbf{W}(i))^{-1/2}&=\mathbf{I}+\frac{1}{\|\mathbf{x}\|^2}\left(\frac{1}{\sqrt{1+\|\mathbf{x}\|^2}}-1\right)\mathbf{x}\mathbf{x}^{\mathrm{H}}\\&=\mathbf{I}+\boldsymbol{\tau}(i)\mathbf{q}(i)\mathbf{q}^{\mathrm{H}}(i)\end{aligned} \tag{14}$$

Where

$$\boldsymbol{\tau}(i)=\frac{1}{\|\mathbf{q}(i)\|^2}\left(\frac{1}{\sqrt{1+\|\mathbf{p}(i)\|^2\|\mathbf{q}(i)\|^2}}-1\right)$$

Using (13), (14), and the updating equation of $\mathbf{W}(i)$, we obtain

$$\begin{aligned}\mathbf{W}(i)&=(\mathbf{W}(i\text{-}1)+\mathbf{p}(i)\mathbf{q}^{\mathrm{H}}(i))(\mathbf{I}+\boldsymbol{\tau}(i)\mathbf{q}(i)\mathbf{q}^{\mathrm{H}}(i))\\&=\mathbf{W}(i\text{-}1)+\mathbf{p}'(i)\mathbf{q}^{\mathrm{H}}(i)\end{aligned} \tag{15}$$

Where $\mathbf{p}'(i)=\boldsymbol{\tau}(i)\mathbf{W}(i-1)\mathbf{q}(i)+(1+\boldsymbol{\tau}(i)\mathbf{q}(i)\mathbf{q}^{\mathrm{H}}(i))\mathbf{p}(i)$. Thus, the OPAST algorithm can be written as Table 1.

Table 1. OPAST Algorithm

$$\mathbf{q}(i)=\frac{1}{\alpha}\mathbf{Z}(i-1)\mathbf{y}(i)$$

$$\mathbf{y}(i)=\mathbf{W}^{\mathrm{H}}(i-1)\mathbf{r}(i)$$

$$\gamma(i)=\frac{1}{1+\mathbf{y}^{\mathrm{H}}(i)\mathbf{q}(i)}$$

$$\mathbf{p}(i)=\gamma(i)(\mathbf{r}(i)-\mathbf{W}(i-1)\mathbf{y}(i))$$

$$\mathbf{Z}(i)=\frac{1}{\alpha}\mathbf{Z}(i-1)-\gamma(i)\mathbf{q}(i)\mathbf{q}^{\mathrm{H}}(i)$$

$$\boldsymbol{\tau}(i)=\frac{1}{\|\mathbf{q}(i)\|^2}\left(\frac{1}{\sqrt{1+\|\mathbf{p}(i)\|^2\|\mathbf{q}(i)\|^2}}-1\right)$$
$$\mathbf{p}'(i)=\boldsymbol{\tau}(i)\mathbf{W}(i-1)\mathbf{q}(i)+(1+\boldsymbol{\tau}(i)\mathbf{q}(i)\mathbf{q}^{\mathrm{H}}(i))\,\mathrm{p}(\mathrm{i})$$
$$\mathbf{W}(i)=\mathbf{W}(i-1)+\mathbf{p}'(i)\mathbf{q}(i)$$

3. Verification of simulation and measured data

3.1 Simulation by MATLAB

(1) DOA estimation performance of MUSIC and OPAST

Simulation conditions: The incoming signal is two chirp signals, the angle is -30°and 60°, working frequency band is $[f_1, f_2]$, where $f_1=200\text{Hz}$, $f_2=2000\text{Hz}$, the center frequency $f_0=1100\text{Hz}$. There is a 8 elements uniform linear array, the spacing of array elements is half of the corresponding wavelength of frequency f_0, SNR=30dB, sampling frequency $f_s=16\text{kHz}$, estimating DOA using MUSIC algorithm and OPAST algorithm respectively, the results are shown in Fig 1 and Fig 2.

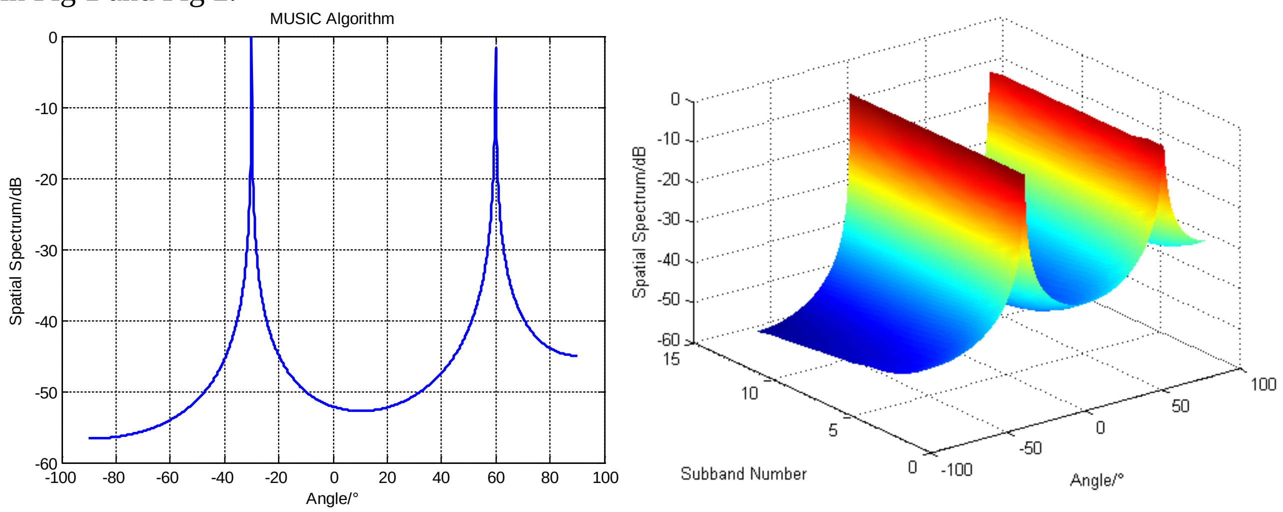

Figure 1. Result of MUSIC algorithm DOA estimation

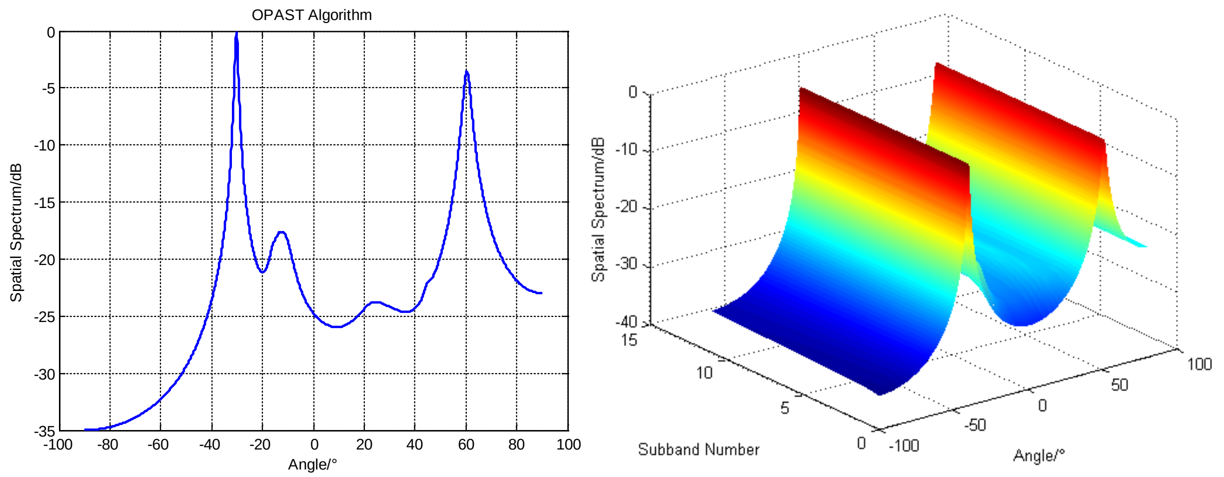

Figure 2. Result of OPAST algorithm DOA estimation

Simulation experiments show that both MUSIC and OPAST can estimate the angle of incoming signal well. It can be seen from the three dimensional images, each sub-band can estimate the signal angle very well, the performance of the two algorithms is similar.

(2) Root mean square error of two algorithms

The root mean square error is the square root of the ratio of the square of the deviation between the predicted value and the real value and the number of observations. The root mean square error is defined as [1]

$$\sqrt{\frac{1}{D}\sum_{i=1}^{D}\left(\hat{\theta}_i-\theta_i\right)^2} \tag{16}$$

Where D represents the number of signals, $\hat{\theta}_i$ is the estimated angle, and θ_i is the real angle. The curve of root mean square error with SNR is shown in Fig 3.

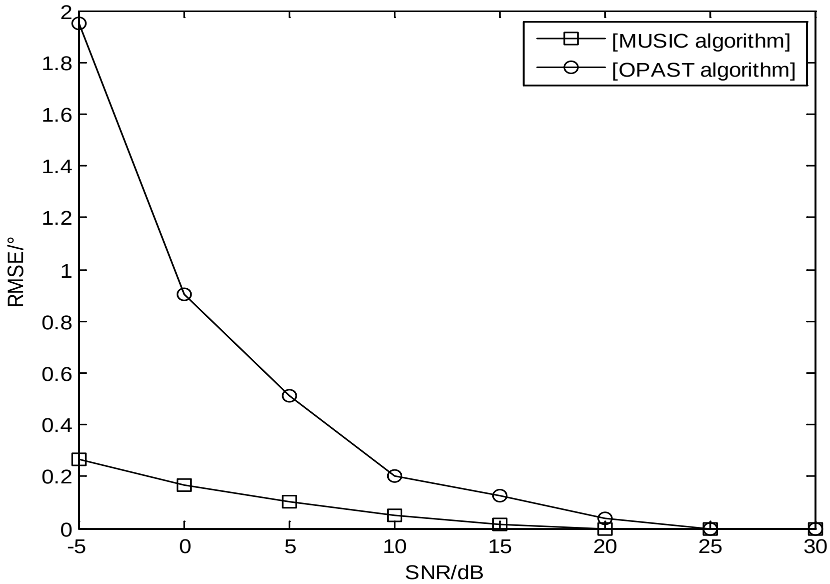

Figure 3. RMSE under different SNR

The MUSIC algorithm has a small DOA estimation error at low SNR, when SNR is high, reaching 20dB, two algorithms have the same DOA estimation performance.

3.2 Measured data processing

Experimental conditions: 4 microphones are arranged in a linear array with a distance of 4 centimeters. One person stands about 1 meter away from the microphone and reads aloud, the indoor environment is quiet, DOA estimation using MUSIC algorithm and OPAST algorithm respectively, the results are shown in the figures below.

First experiment, the person stands at the microphone array 0°position and reads aloud, choosing one of the voices, the waveforms and spectrogram as follows

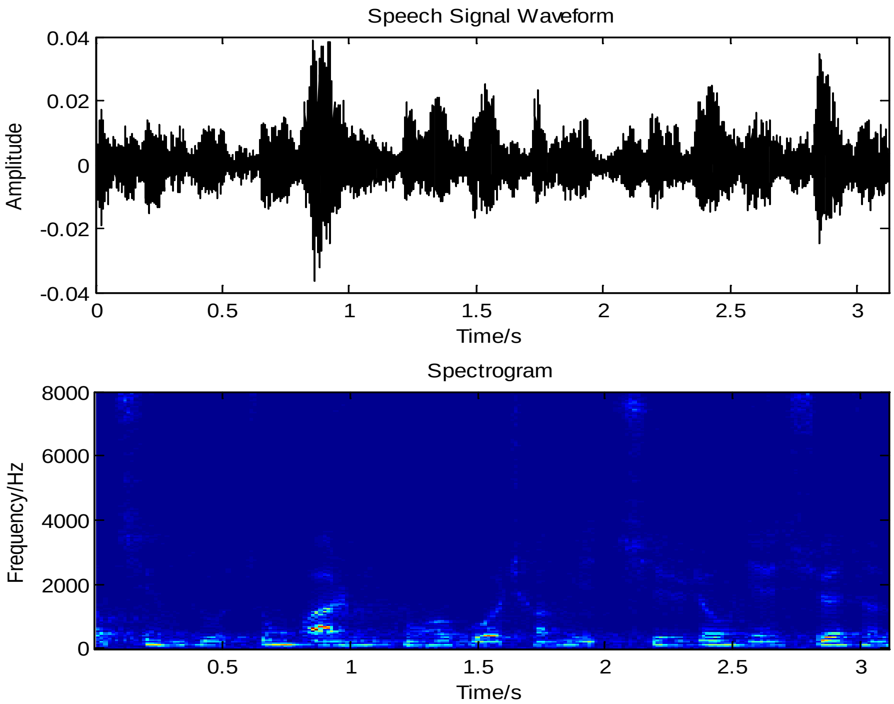

Figure 4. Wave form and spectrographic of speech signal

From the spectrogram, the frequency components of this speech signal are mostly concentrated in the frequency range of 400 to 3000 Hz, so when dividing sub-bands, the frequency components of 400 to 3000 Hz are chosen. The DOA estimation results obtained by two algorithms for the first speech signal are shown in Fig 5.

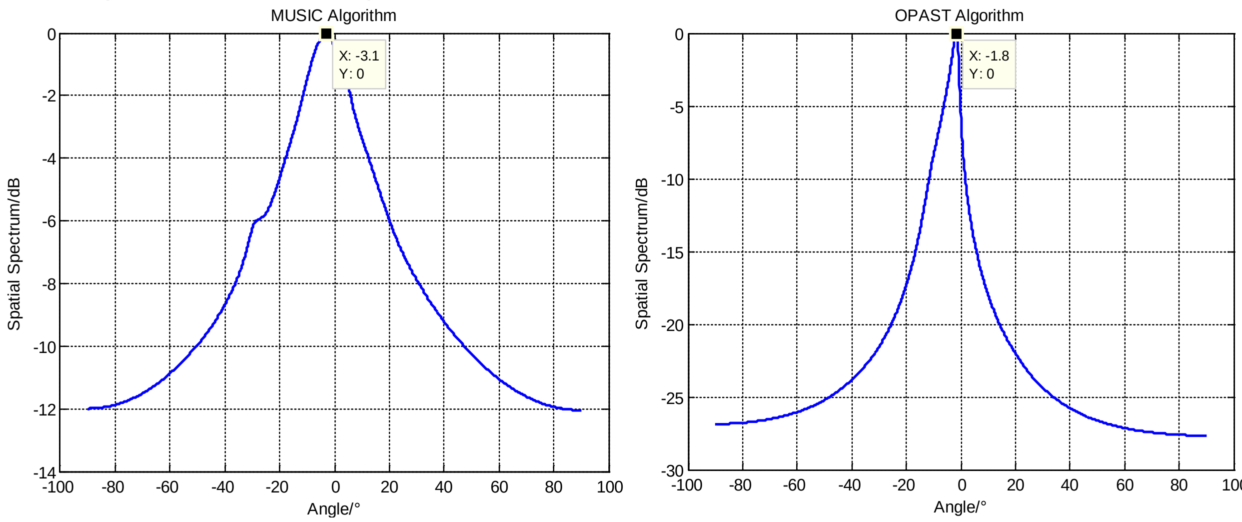

Figure 5. MUSIC algorithm and OPAST algorithm DOA estimation results

From Fig 5 we can know that the angel of incoming signal is 0, MUSIC algorithm and OPAST algorithm have similar performance.

Second experiment, the person stands at the microphone array 45°position and reads aloud, choosing one of the voices, the waveforms and spectrogram as follows

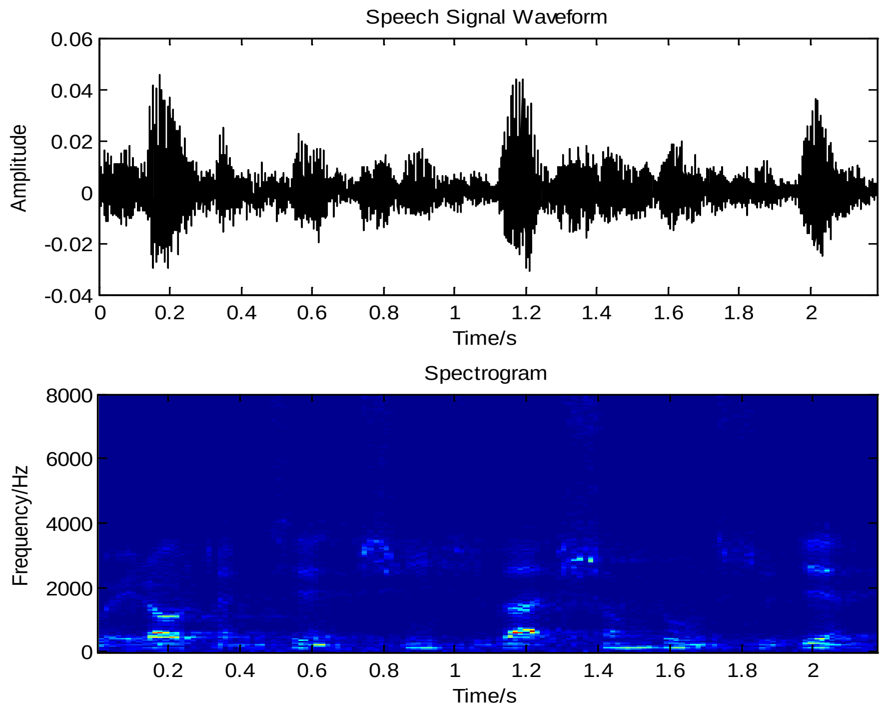

Figure 6. Wave form and spectrographic of speech signal

From the spectrogram, the frequency components of this speech signal are mostly concentrated in the frequency range of 400 to 3500 Hz, so when dividing sub-bands, the frequency components of 400 to 3500 Hz are chosen. The DOA estimation results obtained by two algorithms for the first speech signal are shown in Fig 7.

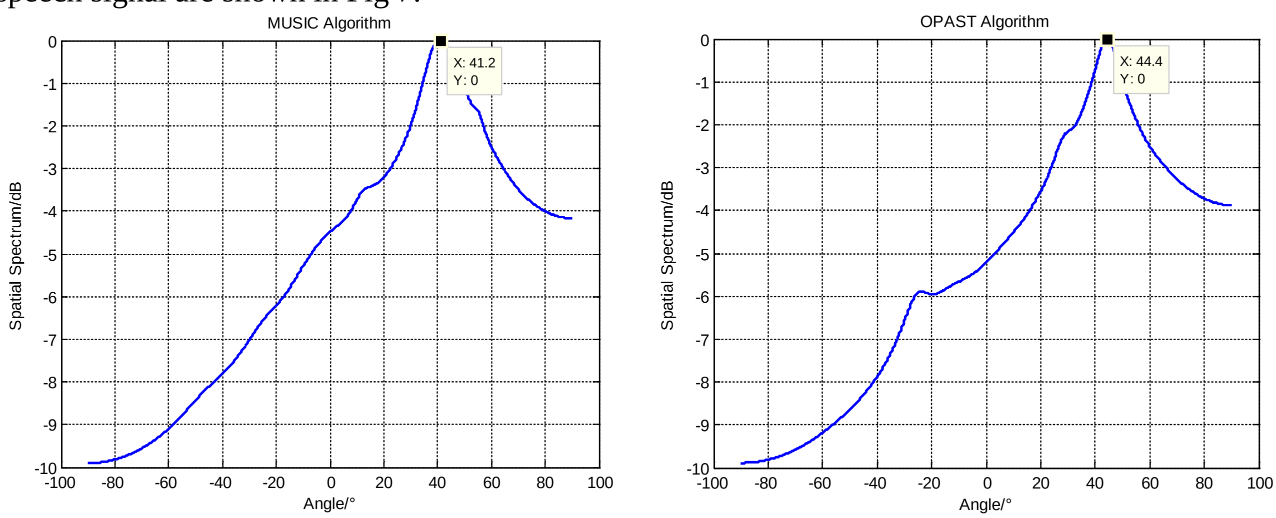

Figure 7. MUSIC algorithm and OPAST algorithm DOA estimation results

From Fig 7 we can know that the angel of incoming signal is 45, MUSIC algorithm and OPAST algorithm have similar performance.

4. Conclusion

OPAST algorithm can achieve the same DOA estimation result as MUSIC algorithm when estimating DOA of speech signals, but when OPAST algorithm solves the signal subspace and the noise subspace, no eigenvalue decomposition is required, which greatly reduces the computational complexity and is easy to implement in engineering. Next step, we will verify the performance of two algorithms when there are multiple sources arrive the microphone array.

References

[1] Xiaofei Zhang. Array Signal Processing and Realization with MATLAB [M]. Publishing House of Electronics Industry, pp. 216 - 242, 2015.

[2] Fuqin Chen. The Method of DOA Estimation for Wideband Signals Based on MUSIC [J]. Scientific Paper Online, 2013.

[3] Xia Zou. Microphone Array Signal Processing [M]. National Defense Industry Press, pp. 162-163, 2016.

[4] Yongjun Zhao. Wideband Array Signal Direction of Arrive Estimation Theory and Methods [M]. National Defense Industry Press, pp. 65-207, 2013.

[5] K. Abed-Meraim. Fast Orthonormal PAST Algorithm [J]. IEEE Signal Processing, Vol. 7, No. 3, pp. 60-62, March, 2000.

[6] B. Yang. Projection Approximation Subspace Tracking [J]. IEEE Trans. Signal Processing, Vol.44, pp. 95-107, Jan. 1995.