Study of the Estimation of Sound Source Signal Direction Based on MUSIC Algorithm

Bao-Hai YANG[1,a,*], Ze-Liang LIU[1,b] and Dong CHEN[1,c]

[1]College of Electronic Engineering, Jiujiang Univercity,Jiujiang,332005,China

[a]dzxy2010@163.com, [b]dz_liuzelang@jju.edu.cn, [c]dz_chendong@jju.edu.cn

Keywords: Equally space dlinear array, MUSIC algorithm, Sound source signal, Direction estimation.

Abstract. The multiple signal classification, MUSIC algorithm, introduces linear space to direction estimation, which realizes the breakthrough of the sound source direction resolution and lays an important foundation for the method and theory of spacial spectrum estimation. In order to better use the algorithm to estimate the direction of the sound source signal, this paper illustrates equally spaced linear array based on the MUSIC algorithm model and carries out the comparative simulation study on the narrow-band and the broad-band sound source signal direction estimation. Computer simulation shows that MUSIC algorithm can achieve super resolution direction estimation of the narrowband signals, and when MUSIC algorithm is used to estimate the direction of broadband signals, its performance is greatly degraded. And finally the conclusion is that in the narrowband sound source signal direction estimation, MUSIC algorithm functions remarkably, and when handling voice and other broadband source signal direction estimation, it needs to be further improved.

Introduction

Sound source direction estimation is one of the hot areas in signal processing study in recent decades. It first collects voice signals, then uses digital signal processing technology to analyze and process them and achieves the signal spatial spectrum, and finally estimates the sound source direction. The sound source direction estimation mainly researches super-resolution estimation, multiple signal classification (MUSIC) algorithm proposed by Schmidt and others -- one of the algorithms to achieve super-resolution direction estimation[1-3]. MUSIC algorithm relates to not only the direction vector and array structure, but the signal frequency as well [4-5].

Characteristics and principles of MUSIC algorithm

MUSIC algorithm is characterized by high spatial spectrum resolution and asymptotical unbiased estimation for the number of sources, DOA, background noise level and signal strength. It can achieve multi-channel DOA estimation. If signal subspace is smaller than noise subspace, that is the source number is less than half of array elements, MUSIC algorithm has better estimation performance[6,7].

The basic principle of MUSIC algorithm is as follows. When signal and background noise satisfy certain conditions, and the source number is less than that of the array elements, eigenvalue decomposition can be used to decompose the correlation matrix data output from the matrix elements and they are decomposed into signal subspace and noise subspace, which are orthogonal[8]. When the signal to noise ratio is sufficient, the signal subspace's eigenvalue will be larger than the noise subspace eigenvalue, thus it is easy to distinguish the signal subspace and noise subspace, the subspace rank is less than that of the autocorrelation matrix. Using signal or noise subspace and the vector direction to form a spatial spectrum, and scan, when scanning DOA, the spacial spectrum will achieve larger value and the signal direction estimation will be acquired[9].

MUSIC algorithm

Currently, the theory and method of using array to estimate the direction has layed theoretical foundation for estimation of the sound source signal direction. MUSIC algorithm based on traditional model of equally spaced linear array (Figure 1) introduces the concept of the linear space into direction estimation. This will realize the breakthrough of the resolution in sound source direction.

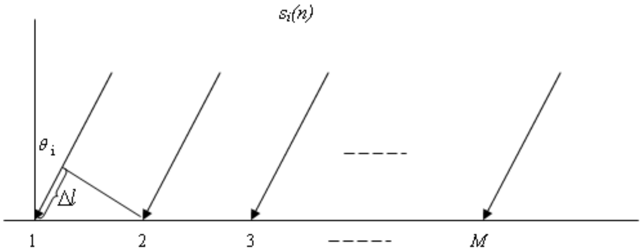


Fig.1 Traditional model of equally spaced linear array

One sample is called one snapshot. Suppose N snapshots of the received signals were observed in each array element. Using these observed values to construct correlation matrix, the desired signal s(n) 's DOA -- θ -- can be achieved.Suppose

$$\hat{R}_{XX} = \frac{1}{N}\sum_{n=1}^{N} x(n)x^H(n) \qquad (1)$$

Equation (1) is the sample autocorrelation matrix of the observed signal vector. The ideal autocorrelation matrix is impossible to get, so more snapshots should be used to estimate the autocorrelation matrix, that is the sample autocorrelation matrix. When the snapshot number tends to infinity, the sample autocorrelation matrix will converge to an ideal auto-correlation matrix with the probability 1, because the signal is stationary,For the array signal observation model described by equation $x(n) = \sum_{i=1}^{p} \alpha(\omega_i)s_i(n) + e(n) = A(\omega) + e(n)$, the following assumptions can be made.

The equally spaced linear array, for different ω_i, vectors α (ω_i) are unrelated. That is the direction matrix is a full column rank matrix. Each element of additive noise vector is zero mean of Gaussian white noise, uncorrelated, and has the same variance, satisfying (2); In which σ^2 is the additive noise variance.

$$E\{e(n)\} = 0, E\{e(n)e^H e(n)\} = \sigma^2 I \qquad (2)$$

Singal sources are independent of each other, that is to say the signal source derived from the correlation matrix is the non-singular matrix, but full matrix, satisfying (3);The signal (3)source and noise are independent of each other and satisfy equation (4).Under the assumed conditions, from equation (1), the following (5) is achieved.

$$rank\left(s(n)s^H(n)\right) = p \qquad (3)$$

$$E\left\{s(n)e^H n\right\} = 0, E\left\{e(n)s^H(n)\right\} = 0 \qquad (4)$$

$$Rxx = E\left\{x(n)x^H n\right\} = E\left\{[A(\omega)s(n) + e(n)][A(\omega)s(n) + e(n)]^H\right\} = APA^H + \sigma^2 I \qquad (5)$$

Among which, $A = A$ (ω) $P = E\left\{s(n)s^H(n)\right\}$. So Rxx is the symmetric matrix of a full matrix. When decompose its eigenvalue, the following equation can be acquired.

$$Rxx = U\Sigma U^H, \Sigma = diag(\sigma_1{}^2, \cdots, \sigma_M{}^2) \qquad (6)$$

The total number of Rxx 's eigenvalues is M. (5) multiplies U^H left and U right, equation (7) can be acquired, (6) multiplies U^H left and U right, equation (8) can be acquired.

$$U^H RxxU = U^H(APA^H + \sigma^2 I)U \qquad (7)$$

$$U^H RxxU = U^H U \Sigma U^H U = \Sigma \qquad (8)$$

With equations (7) and (8), (9) can be acquired. The vector matrix is a full matrix, so

$$\Sigma = U^H APA^H U + U^H \sigma^2 IU = U^H E\left\{As(n)s^H(n)A^H\right\}U + U^H \sigma^2 IU = diag(\sigma_1{}^2, \cdots \sigma_M{}^2) \qquad (9)$$

$$U^{H}APA^{H}U = U^{H}E\{As(n)s^{H}(n)A^{H}\}U = \text{diag}(a_1^2,\cdots a_p^2,0,\cdots 0) \quad (10)$$

In which $\alpha_1^2,\cdots\alpha_P^2$ are autocorrelation matrix APA^H's eigenvalues of the observed signals $As(n)$, when there is no additive noise. So with equations (9) and (10), the eigenvalues of the observed signals $x(n)$'s autocorrelation matrix will be achieved.

$$\lambda_i = \sigma_i = \begin{cases} \alpha i^2 + \sigma^2, i = 1,\cdots p \\ \sigma^2, i = p+1,\cdots M \end{cases} \quad (11)$$

Equation (11) shows that, when there is additive observed white noise in the background, the autocorrelation matrix's eigenvalues of the observed vectors $x(n)$ are composed by two parts. The former p's eigenvalues are the total values of α_i^2 and σ^2 -- the variance of the additive white noise, and the latter M-p's are the variance of the additive white noise's variance σ^2.

When the signal to the noise ratio is sufficient, $\alpha_i^2 > \sigma^2$, it is easy to distinguish Rxx's larger p eigenvalues $\alpha_i^2 + \sigma^2$ and M-P smaller ones σ^2. These main eigenvalues P are the signal eignevalues, and M-p are the noise eigenvalues. With these signal eigenvalues and noise ones, characteristic matrix U's column vectors can be divided two parts. They are

$$S = [s_1,\cdots s_p] = [u_1,\cdots u_p] \quad (12)$$

$$G = [g_1,\cdots g_{M-p}] = [u_{p+1},\cdots u_M] \quad (13)$$

$$U = [SG] \quad (14)$$

In which, u is U's column vector, S is the signal eigenvector and G is the noise eigenvector. The space $span\{s_1,\cdots,s_p\}$ composed by S is called the signal subspace and $span\{g_1,\cdots g_{M-p}\}$ by G is noise subspace. With (6), (15) can be achieved.

$$UU^H = U^H U = I \quad (15)$$

So U is a unitary matrix, according the feathure of the unitary matrix, the signal subspace and the noise subspace are orthogonal. That is $span\{s_1,\cdots s_p\} \perp span\{g_1,\cdots g_{M-p}\}$

With equations (12), (13) and (15), equation (16) can be achieved.

$$span\{u_1,\cdots u_p\} \perp span\{u_{p+1}, u_M\} \quad (16)$$

If G multiplies Rxx right, with equations (2.5), (2.6), (14) and (16), the following two equations can be acquired.Thus (18)

$$RxxG = APA^H G + \sigma^2 GAPA^H G + \sigma^2 G = [SG]\Sigma\begin{bmatrix} S^H \\ G^H \end{bmatrix} G = [SG]\Sigma\begin{bmatrix} 0 \\ I \end{bmatrix} = \sigma^2 G \quad (17)$$

$$APA^H G = 0, G^H APA^H G = 0 \quad (18)$$

P is full rank matrix, the determinant is not 0, so $A^H G = 0$

When equation (11) is substituted into $A^H G = 0$, (19) is acquired.

$$\alpha^H(\omega)G = 0, \omega = \omega_1 \cdots \omega_p \quad (19)$$

When $\omega \neq \omega_1,\cdots,\omega_p, \alpha^H(\omega)G \neq 0$.Putting (19) into the form of scalar, spacial spectrum is (20)

$$P(\omega) = \frac{1}{\alpha^H GG^H \alpha(\omega)} \quad (20)$$

From equation (20), the p max $\omega_1,\cdots, \omega_p$ can be achieved and the p DOA are $\theta_1,\cdots, \theta_p$.

And finally, MUSIC algorithm is concluded as follows.

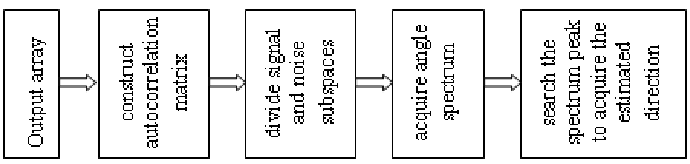


Fig.2 MUSIC algorithm processes

Computer simulation

The simulation conditions are as follows. 8 array elements equally spaced linear array, angle searching ranges are $-90^0 \sim 90^0$, searching step width is 1; incidence angle is 45^0 ,the background noise is the Gaussian white noise, the signal to noise ratio is 20dB; signal center frequency is 1KHz, space separation between array elements is half of the center frequency to wave length, that is 0.171m.

The modulating signal(21),The sound carrier(22),FM signal(23)

$$u_\Omega(t) = 2\cos(2\pi 10 t) \qquad (21)$$

$$u_c(t) = 14\cos(2\pi 10^3 t) \qquad (22)$$

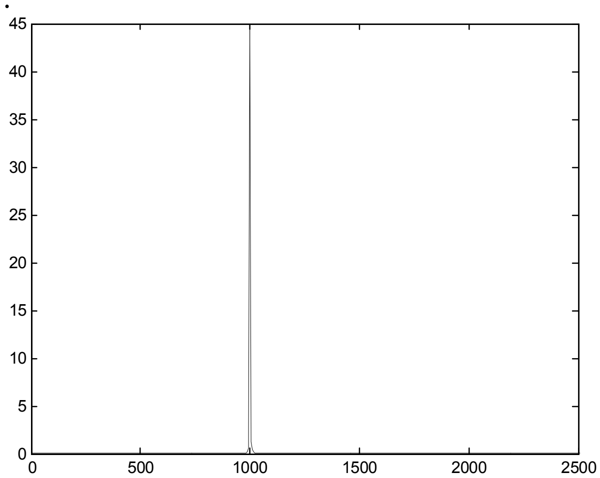
$$u_{FM}(t) = 14\cos(2\pi 10^3 t + m_f \sin 2\pi 10 t) \qquad (23)$$

In the equations m_f is the FM index. When m_f is far smaller than 1, it is called narrowband FM signal; and when much larger than 1, the broadband one.

(1) Estimation of narrowband signal direction

When $m_f = 0.01$, the narrowband signal power spectrum is as fig 3（Narrow band signal power spectrum）, and the direction estimation simulation is as fig 4(Direction estimation using MUSICAlgorithm for narrow-band signal based on linear array)

.



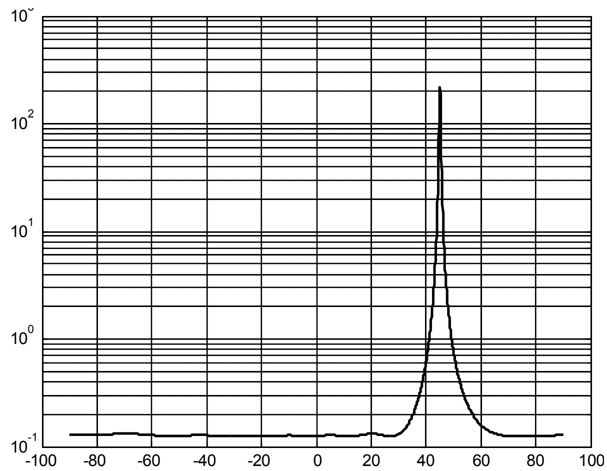
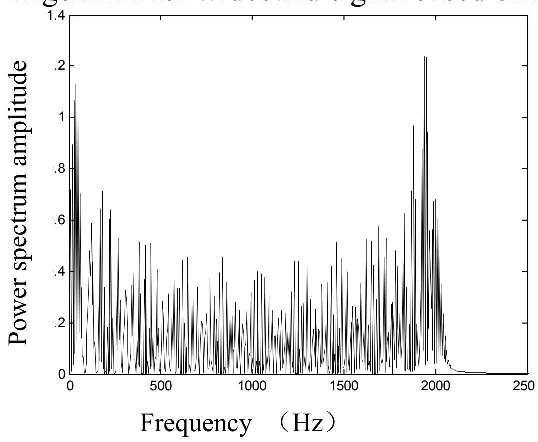


Fig.3 Narrow band signal power spectrum Fig.4 Direction estimation using MUSIC algorithm

(2) Estimation of broadband signal direction

When $m_f = 100$, the broadband signal power spectrum is as Fig 5（Wideband signal power spectrum）, and the direction estimation simulation is as Fig 6(Direction estimation using MUSIC Algorithm for wideband signal based on linear array).



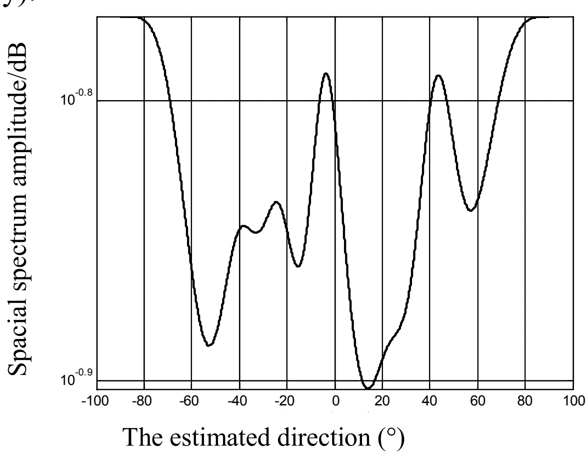


Fig.5 Wideband signal power spectrum Fig. 6 Direction estimation using MUSIC Algorithm

The causes for the incorrect estimation are as follows. The direction vector chooses the center frequency, while the broadband signal frequency covers broadly, so that the maximum signal frequency and the lowest number of frequency will greatly deviated from center frequency. For the high-frequency components, it results in the large array element spacing, which leads to false peaks. For the low-frequency components, it leads to much smaller array element spacing, which results in peak broadening. This results in MUSIC algorithm's significant performance degradation when estimating broadband signal direction.

Conclusion

Analysis shows that the MUSIC algorithm can achieve super-resolution signal direction estimation of the narrowband. If the signal bandwidth increases, MUSIC algorithm estimation performance will be greatly degraded. In practice, there are lots of broadband signals, such as estimation of the voice signal direction. It can improve MUSIC algorithm with the combination of the characters of the voice signal processing. In order to improve voice signal and other broadband signal direction estimation performance, the algorithm can be further studied and improved to adapt to source signal direction estimation of the broadband array.

Acknowledgment

This work is supported by Science Foundation of Jiangxi Provincial Department of Education （GJJ14739、GJJ14721）, National Science Foundation of Jiangxi Province (20142BAB207006，20151BAB207013),Science Foundation of Jiujiang Univercity (2013KJ02、2013KJ01).

References

[1]Carter C.The generalized correlation method for estimation of time delay[J].IEEE Transactions on Aeoust Peeeroeessn,1976,24(04):24--27.

[2]Gerhard D.Beamforming with Optimized Interpolated Microphone Arrays[J].Institute of Communications and Radio-Frequency Engineering,2008,1:33-36.

[3]Ju T L,Peng Q C,Shao H Z.Speech source localization sub-space algorithm research based on microphone array in near field[C].ICEMI NJTU,Beijing,2005:45-48.

[4]Ephraim Y,Van Trees H.A signal subspace approach for speech enhancement[J].IEEE Transactions on Speech and Audio Processing,1995,3(4):251-266.

[5]Rubsamen M,Gershman A B. Direction-of-arrivalestimation for nonuniform sensor arrays:from mani-fold separation to Fourier domain music methods[J] .IEEE Trans Signal Process, 2009, 57 (2) :588-599 .

[6]Unnikrishna S,Kwon B H.Forward/Backward Spatial Smoothing Techniques for Coherent Signal Identification[J].IEEE Transactions on ASSP,1989,37(1):8-15.

[7]Kundu D.Modified MUSIC Algorithm for Estimating DOA of Signals[J]. Signal Processing, 1996,48(1):85-90.

[8]Flanagan J L,Johnston J D,Zahn R,et al.Computer steered microphone arrays for sound transduction in large rooms[J].JAcoust Soc Amer,1985,78(5):1508-1518.

[9]Asano F,Asoh H,Matsui T.Sound source localization and separation in near field[J].IEEE Transactions on Fundamentals of Electronics, 2000, E83-A(11):2286-2294.