

Estimation of Source Number for Single-Channel Received Signal

Zhen Zuo¹, Jun Peng Hu^{1,a}, Zhi Ping Huang¹ and Yi Meng Zhang¹

¹College of Mechatronics and Automation, National University of Defense Technology, Changsha, Hunan Province, China.

Abstract. A single-channel source number estimation method is presented. The single-channel received signal is firstly decomposed into several intrinsic-mode functions (IMFs) by empirical mode decomposition (EMD). Then the information theoretic criteria (ITC) based methods, MDL and AIC, are introduced to realize the source number estimation. And the diagonal loading technique is used to smooth the fluctuation of the eigenvalues. Compare with the previous works, the proposed method can effectively improve the source number detection performance.

1 Introduction

Blind source separation (BSS) has aroused a large number of researchers interest in it during the recent years because of its wide range of application. Accurate estimation of the source number is a vital issue for BSS. In the array signal process, the source number estimation has been well learned and a plethora of algorithms have been proposed [1-9].

Restricted by the cost and the limited fix space, only one receiver can be installed in some situation. The BSS problem is reduced to a single-channel blind source separation (SCBSS) problem[10]. In such cases, the conventional source separation methods can not be used directly. Many approaches have been introduced to realize the SCBSS[11, 12]. But most of them are built upon the precondition that the number of source has been known already. Without the precise estimation of the source number, the performance of the BSS or SCBSS algorithms will be severely impacted.

Various methods have been proposed to solve the source number estimation, such as the methods based on hypothesis testing [7, 9], information theoretic criteria (ITC) based approaches[1-3] and Gerschgorin's disk estimation (GDE) method[6]. The ITC based approaches are represented by Akaike Information Theory (AIC)[1] and minimum description length (MDL)[2, 3] criteria. The MDL estimator is known to be a consistent estimator[13], whereas the AIC principle is not consistent and in most cases overestimates the number of signals[14]. The ITC based methods are in essence estimating the source number by the differences of eigenvalues between the signal and the noise. They can not be used for the single-channel signal directly.

Transforming the single-channel signal to multi-channel form is a feasible way. The empirical mode decomposition (EMD) has recently gained reputation as a method for analyzing non-linear and nonstationary time

series data[15]. The EMD is a signal-analysis tool that is able to decompose any time series into a set of spectrally independent oscillatory modes, called intrinsic-mode functions (IMFs)[16]. Accordingly, the EMD tool can be used to transform the single-channel signal to multi-channel.

The purpose of this paper is to solve the source number estimation under single-channel signal condition. The EMD technique is firstly used to decompose the single-channel data into multi-channel form. And the ITC based methods, MDL and AIC, are used to estimate the source number. Then the diagonal loading technique is introduced to reduce the fluctuation of eigenvalues caused by noise.

The rest of this paper is organized as follows. Section 2 presents the single-channel received signal model. And the review of the EMD decomposition is introduced in section 3. Section 4 describes the source number estimation algorithm and the improvement realized by this paper. The experiment results show performance of the proposed algorithm in section 5, followed with the conclusions in section 6.

2. Signal Model

For the single-channel received signal, only one data can be observed at every time point. The signal model can be described as follow.

$$X(t) = AS(t) + N(t) \quad (1)$$

Where $X(t) = [x(1), x(2), \dots, x(m)]$ represents the observed data and t indicates the observation time. $A = [a_1, a_2, \dots, a_m]$ is the mixing matrix.

$S(t) = [s_1(t), s_2(t), \dots, s_m(t)]^T$ is a source signal matrix. Noise and signal are independent. The parameter estimation of single-channel signal is an underdetermined

^a Corresponding author: hujunpeng332@163.com

problem. It is difficult to process the single-channel signal directly.

3. EMD Decomposition

EMD is a signal processing tool for decomposing signal into oscillating components by empirically identifying the physical time scales intrinsic to the data[15]. The EMD decomposes the mixture signal into a collection of IMFs[17].

$$y(t) = \sum_{n=1}^N c_n(t) + r_N(t) \quad (2)$$

$c_n(t)$ is the n th IMF and N is the total number of IMFs. $r_N(t)$ is the final residue. The IMFs satisfy two fundamental conditions[15]: firstly, in the whole dataset, the number of extrema (minima and maxima) and the number of zero crossing must be same or differ at most by one. Second, the mean value of envelop defined by the local minima is always zero.

4. Source Number Estimation Method

Assuming that the noise and the source signals are mutually independent and $X(t)$ represents the multi-dimension recording data, the covariance matrix of the observed signal can be expressed as

$$R_X = E\{X(t)X^T(t)\} = AR_S A^T + \sigma^2 I \quad (3)$$

Where $R_S = E\{S(t)S^T(t)\}$ is the signal covariance matrix. $\sigma^2 I$ represents the covariance matrix contains the noise signal. Transform R_X by eigenvalue decomposition (EVD).

$$R_X = \sum_{i=1}^p \lambda_i u_i u_i^H \quad (4)$$

$\lambda_i, i = 1, \dots, p$ are the eigenvalues of R_X in descending order, $\lambda_1 > \dots > \lambda_q > \lambda_{q+1} = \dots = \lambda_p = \sigma_n^2$. $u_i, i = 1, \dots, p$ stand for the corresponding eigenvectors.

Since the noise and signal are independent, it can be deduced that $\lambda_1 = \mu_1 + \sigma^2$, $\lambda_2 = \mu_2 + \sigma^2$, ..., $\lambda_n = \mu_n + \sigma^2$, $\lambda_{n+1} = \lambda_{n+2} = \dots = \lambda_m = \sigma^2$. Under normal circumstances, the signal eigenvalues are far greater than the noise one. Setting an appropriate threshold, the source number is equal to the amount of the eigenvalues which are bigger than threshold.

In real application, due to the limited number of snapshots, eigenvalues of noise and signal are difficult to distinguish. It is impossible to directly determine the number of sources according to the size of the eigenvalues. Some criteria need to be taken to further estimate the number of sources. The ITC has been proved to be useful mean to solve the source number estimation problem. There are two parts in the ITC. The first part is the likelihood function representing the information gained from the recording data which are dependent on the assumed number of sources. The second part is the penalty term, a function of the number of snapshots[18].

AIC is a ITC method for model order selection, introduced by Akaike[1]. MDL is used to select the model based on the concept of the shortest code length, proposed by Rissanen[3, 19]. AIC and MDL estimation function can be show as follow.

$$AIC = -2 \log f(\mathbf{X}|\hat{\Theta}) + 2k \quad (5)$$

$$MDL = -\log f(\mathbf{X}|\hat{\Theta}) + \frac{1}{2}k \log N \quad (6)$$

$\hat{\Theta}$ represents the maximum likelihood estimation of the parameters Θ . k is freedom degrees of the parameter vector Θ . N is the number of samples and $f(\mathbf{X}|\hat{\Theta})$ indicates the joint probability density of X . The first term of the right end in Equation (9) and (10) denote model parameter estimation function. The second term represents the penalty function. When AIC and MDL are applied as source number estimate criteria, the model is expressed as follows[4]

$$AIC(k) = -2N(M-k) \log f(k) + 2k(2M-k) \quad (7)$$

$$MDL(k) = -N(M-k) \log f(k) + \frac{1}{2}k(2M-k) \log N \quad (8)$$

Where

$$f(k) = \left(\prod_{i=k+1}^M \lambda_i \right)^{\frac{1}{M-k}} \left(\frac{1}{M-k} \sum_{i=k+1}^M \lambda_i \right) \quad (9)$$

And the number of source can be yielded by minimizing the corresponding criterion of AIC and MDL.

$$\hat{q}_{AIC} = \arg \min_k AIC(k) \quad (10)$$

$$\hat{q}_{MDL} = \arg \min_k MDL(k) \quad (11)$$

The AIC and MDL criteria get the sources number via the eigenvalues of data covariance matrix R_X . Owing to its properties of simplicity and consistency, the MDL estimator has become the standard tool for estimating the number of sources[20]. However, on account of the impact of noise and limited snapshots of received data, the eigenvalues of R_X are fluctuant. This will affect the estimation performance. To reduce the impact of the noise, a diagonal loading technique is proposed to modify the eigenvalues of the data covariance matrix [21-23].

$$\mu_i = \lambda_i + \lambda_{DL}, i = 1, 2, \dots, p \quad (12)$$

$\lambda_i, i = 1, 2, \dots, p$ are the eigenvalues of R_X , and λ_{DL} is the loading value. μ_i are the final eigenvalues with diagonal loading. Diagonal loading produces little affect to the bigger eigenvalues corresponding to source signals. And the smaller eigenvalues corresponding to the noise will convergence near the loading value λ_{DL} . Therefore, by diagonal loading technique, the noise eigenvalues are approximately equal and the noise is smoothed.

The selection of loading value λ_{DL} makes a great impact to this method. It can be inferred that a small value of λ_{DL} play little improvement to the estimation, and an oversize λ_{DL} may bring effect to the signal eigenvalues and destroy the difference of eigenvalues between noise and signal. Remarkable improvement of the diagonal loading technique needs an appropriate selection of λ_{DL} .

The loading value λ_{DL} should satisfy the follow equations.

$$\lambda_N \ll \lambda_{DL} \ll \lambda_S \quad (13)$$

Where λ_N relates to the noise eigenvalue, and λ_S relates to the signal eigenvalue. To meet this condition, a feasible selection of λ_{DL} is described as[21]

$$\lambda_{DL} = \sqrt{\sum_{i=1}^p \lambda_i} \quad (14)$$

Then the eigenvalues of R_X are transformed to follow by diagonal loading.

$$\mu_i = \lambda_i + \sqrt{\sum_{i=1}^p \lambda_i}, i = 1, 2, \dots, p \quad (15)$$

Then λ_i in equation (9) are replaced with μ_i in equation (15).

5. Experiments

Consider the single-channel recording signals which are collected from a single sensor. The received data contains three source signals with different frequencies and initial phase. The source signals can be expressed in the form as follow.

$$\begin{cases} s_1(t) = \cos(2\pi f_1 t + \theta_1) \\ s_2(t) = \cos(2\pi f_2 t + \theta_2) \\ s_3(t) = \cos(2\pi f_1 t + \theta_1) \cos(2\pi f_2 t + \theta_2) \end{cases} \quad (16)$$

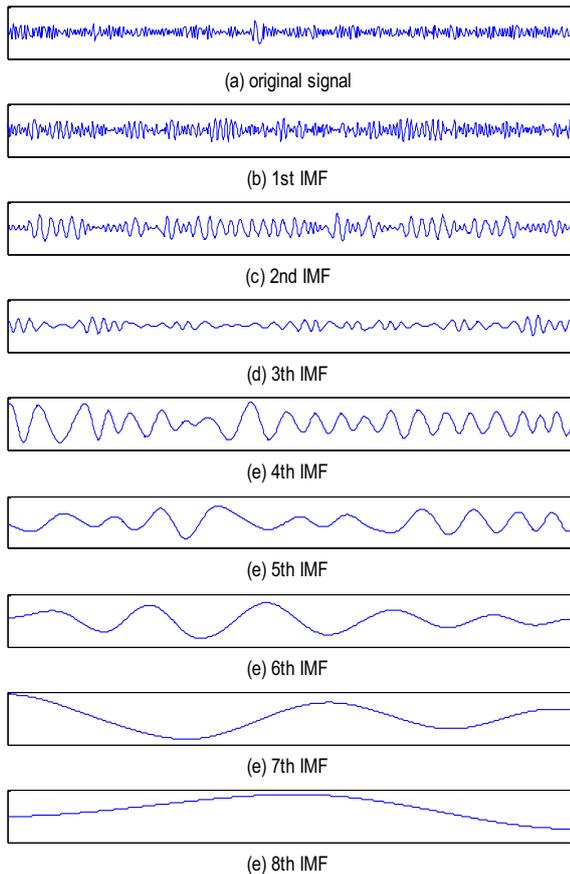


Figure 1. The single-channel signal and its IMF's component

The mixed signal is acquired by a single sensor and converted to discrete form with snapshots $L=1000$. Noise is added to the signal, making SNR increase from -10 dB to 60dB. Then the mixed signal is analyzed by EMD decomposition as discussed in section 3. The decomposed and original signals are shown in Fig 1.

It can be implied that a single-channel received signal has been transformed to multi-dimension form by EMD, that is $X = [y(n); c_1(n); c_2(n); \dots; c_N(n); r_N(n)]$. $y(n)$ is the original signal. $c_1(n), c_2(n), \dots, c_N(n)$ are the IMFs with total number of N and $r_N(n)$ is the final residue signal. Then the covariance matrix R_X can be calculated as follow.

$$R_X = XX^H \quad (17)$$

The sources number of the received signal is estimated by AIC and MDL criteria based on the eigenvalues of R_X . Fig 2 describes the estimation performance of the method with and without diagonal loading. Due to the fluctuation of the eigenvalues caused by the added noise, the method without diagonal loading can no longer estimate the source number correctly. By diagonal loading technique, the AIC and MDL estimators obtain a good estimation performance.

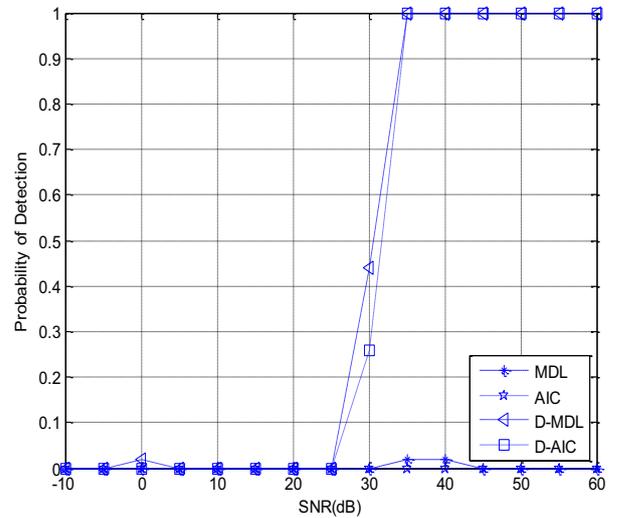


Figure 2. estimation performance comparing of MDL and AIC method and the method improved by diagonal loading technique

6. Conclusion

Accurately estimating of the sources number is essential for the blind source separation (BSS). Many algorithms have been proposed to estimate the sources number with multi-channel received signal, but not applicable for the single-channel recording data.

This paper has proposed a source number estimation method based on empirical mode decomposition (EMD), dealing with the single-channel received signal. The single-channel recording data is firstly decomposed into many intrinsic mode functions (IMFs) by EMD. The ITC based methods, AIC and MDL, are used to calculate the

source number, which obtain well performance at high SNR. The diagonal loading technique is introduced to smooth the fluctuation of the eigenvalues to reduce the effect of noise. After computing the mix matrix, this paper proposed an improved method which calculates the source number by the inner product of the eigenvectors and the mix matrix, instead of eigenvalues. The experiments results show that the improved method gets a better performance than the conventional method at low SNR.

23 Ma Ning, Goh Joo Thiam. *Acoustics Speech and Signal Processing*, **15**:341-344 (2003).

References

- 1 H. Akaike. *IEEE Trans on Automatic Control*, **19**:716-723 (1974).
- 2 G. Schwartz. *The Annals of Statistics*, **6**:461-464 (1978).
- 3 Rissanen J. *Automatica*, **14**:465-471 (1978).
- 4 Wax M., Kailath T. *IEEE TRANS Acoustic, Speech, and Signal Process*, **ASSP-33**:387-392 (1985).
- 5 Wax M, Ziskind I. *IEEE Trans on Acoustics Speech Signal Processing*, **37**:1190-1196 (1989).
- 6 Wu H T, Yang J F, Chen F K. *IEEE Trans on Acoustics Speech Signal Processing*, **4**:261-264 (1994).
- 7 R.F. Brcich, A.M. Zoubir, P. Pelin. *IEEE Trans Signal Process*, **50**:20-215 (2002).
- 8 Gu J.-F., P.Wei, Tai H.-M. *IET Signal Process*, **1**:2-8 (2007).
- 9 Xiao Manlin, Wei Ping, Tai Heng-Ming. *journal of communications and networks*, **14** (2012).
- 10 Wang Dong, Tse Peter W. *Journal of Sound and Vibration*:4956-4970 (2012).
- 11 QuanWei CAI, Ping WEI, XianCi XIAO. *SCIENCE CHINA*, **53**:147-157 (2010).
- 12 Tengtrairat N., Gao Bin, Woo W. L., Dlay S. S. *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, **24**:1722-1735 (2013).
- 13 Fishler Eran, Poor H. Vincent. *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, **53**:3543-3553 (2005).
- 14 Valaee S., Kabal P. *Ieee Transactions on Signal Processing*, **52**:1171-1178 (2004).
- 15 Gao Bin, Woo W. L., Dlay S. S. *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING*, **19**:961-976 (2011).
- 16 Mijovic Bogdan, Vos Maarten De, Gligorijevic Ivan, Taelman Joachim, Huffel Sabine Van. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, **57**:2188-2196 (2010).
- 17 Wu B.Z., Huang N.E. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc R Soc London A*2004. p. 1597-1611.
- 18 Bai X. X., He B. *Ieee Transactions on Biomedical Engineering*, **53**:1883-1892 (2006).
- 19 Cheng W., Lee S., Zhang Z. S., He Z. J. *Journal of Sound and Vibration*, **331**:5153-5167 (2012).
- 20 Xiao M. L., Wei P., Tai H. M. *Journal of Communications and Networks*, **14**:481-486 (2012).
- 21 Jiling Xie, Xicai Si. *Systems Engineering and Electronics*, **30**:46-49 (2008).
- 22 D. Calson B. *IEEE Transactions on Aerospace and Electronic Systems*, **24**:397-401 (1988).