Mixing Matrix Estimation in Blind Source Separation Based on Generalized Gaussian Mixture Modal

Yongqiang Chen, Jun Liu

Electronic Experiment Center, Chengdu University of Information

Technology, Chengdu 610225, China

Email:chenyongq@ cuit.edu.cn

Abstract. The accurate estimation of mixing matrix is critical for blind separation, for solving the problems of traditional methods such as bad robustness and low accuracy, a method based on statistical modal is proposed. The generalized Gaussian mixture modal is used to fit the distribution of single-source-points, a new objective function for clustering is obtained from the view of maximum likelihood estimation. Constrained particle swarm optimization is used to optimize the objective function, by which the mixing matrix is estimated. This method is applicable to determined and underdetermined blind source separation. The simulation shows that the proposed method has higher estimation accuracy and is more robust than traditional methods.

Introduction

There are many blind source separation methods, such as the use of statistical independence between the signal [1,2], sparsity [3-6] and other features to separate the signals. In recent years, the use of sparsity to separate the signal has become a hotspot, such method is called sparse component analysis (SCA), it does not require the number of sensors must be greater than or equal to the number of source signals, so closer to the actual situation. In order to further improve the separation performance, speech signal is transformed by short-time Fourier transform (STFT) [4-6], because the speech signal in the time-frequency domain is sparser, so better separation is obtained. These literatures mostly use two-step methods to separate signals, that is, mixing matrix is firstly estimated, and then the source signals are estimated. Mixing matrix estimation accuracy will affect the estimation of source signals, so how to estimate the mixing matrix accurately is an important issue to be resolved. Literature [7-8] assume that there is only one source signal at each time-frequency point, and thus to estimate the mixing matrix, but the requirement is too harsh for voice and other weak sparse signals. Literature [9] estimates mixing matrix in a more relaxed conditions, and single-source-points (i.e. the time-frequency points where there is only one dominant source) detection method is proposed. However, the method in [9] is still susceptible to the initial value and the estimation accuracy is not enough. In this paper, we propose a new method for mixing matrix estimation. Firstly, the observed vectors at single-source-points are transformed into one-dimensional data, and the generalized Gaussian distribution (which is different from the Gaussian mixture model [10]) is used to fit the data distribution. From the viewpoint of maximum likelihood estimation, the objective function is derived. By constraint particle swarm optimization [11, 12] to maximize the objective function, we can get the estimated mixing matrix. The proposed method is not sensitive to the initial value, so the average mean square error of the mixing matrix estimation is smaller than traditional methods.

Instantaneous Mixture Model

Assume the number of source signals is N, and the number of sensors is M. At a certain time-frequency point (t, f), the instantaneous mixing model of the source signal is shown as below

$$X(t,f) = AS(t,f) + N(t,f).$$ (1)

where $\boldsymbol{x}(t,f)=[x_1(t,f),x_2(t,f),\cdots,x_M(t,f)]^{\mathrm{T}}$ is M-dimensional observed data vector; $\boldsymbol{s}(t,f)=[s_1(t,f),s_2(t,f),\cdots,s_N(t,f)]^{\mathrm{T}}$ is N-dimensional source signal data vector; $\boldsymbol{A}$ is the invariable mixing matrix of $M\times N$ dimension; $\boldsymbol{n}(t,f)=[n_1(t,f),n_2(t,f),\cdots,n_M(t,f)]^{\mathrm{T}}$ is M-dimensional Gaussian White Noise vector.

Mixing Matrix Estimation Based on Generalized Gaussian Mixture Modal

Literature [9] put forward a method to detect the single-source-points, in other words, if a time-frequency point satisfies the formula (2), it is considered as the single-source-point

$$\left\| \mathrm{Im}\left(\frac{X(t,f)}{x_1(t,f)}\right) \right\| < \varepsilon. \tag{2}$$

Here, single-source-points are denoted by $Y(t,f)$. Firstly, let the first component of $Y(t,f)$ no less than 0, and get $Z(t,f)$. Through the following formula，convert $Z(t,f)$ into one-dimensional data

$$z(i) = \cos^{-1}\left[Z^{(2)}(t,f) \right] \quad i=1,2,\cdots,I. \tag{3}$$

where $Z^{(2)}$ is the second element of the vector Z; $z(i)$ refers to the radian, represents the direction of the vector $Z(t,f)$; obviously, $z(i)$ is the random number between $[0,\pi]$, whose distribution mainly concentrates on several numerical values, they are several clustering centers of the data. These clustering centers correspond to the different directions of the source signals. The paper proposes to adopt the generalized Gaussian mixture distribution to fit the actual distribution

$$P(z(i)|\boldsymbol{\beta},\boldsymbol{\mu}) = \frac{1}{N}\sum_{n=1}^{N} \mathrm{v}(\beta_n)\exp\left\{ -\mathrm{c}(\beta_n)\left| z(i)-\mu_n \right|^{\frac{2}{1+\beta_n}} \right\}. \tag{4}$$

where $\boldsymbol{\beta}=[\beta_1,\beta_2,\cdots,\beta_N]$, $\boldsymbol{\mu}=[\mu_1,\mu_2,\cdots,\mu_N]$ are all the parameters of the generalized Gaussian mixture distribution. μ_n determines the position of each peak (in other words, each clustering center); β_n determines the shape of the peak, and the range of numerical value is $(-1,+\infty)$. When $\beta_n=1$, it is Laplace distribution; when $\beta_n=0$, it is Gaussian distribution; when $\beta_n>0$, it is Super-Gaussian distribution; and when $\beta_n<0$, it is Sub-Gaussian distribution. $\mathrm{v}(\beta_n)$, $\mathrm{c}(\beta_n)$ are defined as below

$$\mathrm{c}(\beta_n) = \left\{ \frac{\Gamma(1.5(1+\beta_n))}{\Gamma(0.5(1+\beta_n))} \right\}^{\frac{1}{1+\beta_n}}. \tag{5}$$

$$\mathrm{v}(\beta_n) = \frac{\left(\Gamma(1.5(1+\beta_n))\right)^{0.5}}{(1+\beta_n)\left(\Gamma(0.5(1+\beta_n))\right)^{1.5}}. \tag{6}$$

$$\Gamma(\tau) = \int_0^\infty u^{\tau-1}e^{-\varsigma}d\varsigma. \tag{7}$$

Obviously, if each μ_n can be estimated, the mixing matrix can be further estimated. Here, it should be noted that "the generalized" Gaussian mixture model is different from Gaussian mixture model[10]. Next, we will discuss the parameters estimation of the generalized Gaussian mixture distribution.

In this paper, the parameters are estimated from the point view of maximum likelihood, that is, the estimated values of the parameters $\{\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\mu}}\}$ should make the formulae (8) maximum, namely

$$\{\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\mu}}\} = \arg\max_{\boldsymbol{\beta},\boldsymbol{\mu}} \prod_{i=1}^{I} P(z(i)|\boldsymbol{\beta},\boldsymbol{u}) = \arg\max_{\boldsymbol{\beta},\boldsymbol{\mu}} \sum_{i=1}^{I} \log[P(z(i)|\boldsymbol{\beta},\boldsymbol{u})]$$

$$= \arg\max_{\boldsymbol{\beta},\boldsymbol{\mu}} \sum_{i=1}^{I} \log\left\{\sum_{n=1}^{N} v(\beta_n)\exp\left[-c(\beta_n)\left|y(i)-u_n\right|^{\frac{2}{1+\beta_n}}\right]\right\}. \tag{8}$$

Hence, all parameters can be estimated by optimizing the objective function (formula (8)).

The Objective Function Optimization

In this paper, objective function is optimized by constrained PSO algorithm [11, 12]. Firstly, J particles are initialized in the target search space of K ($K = 2N$) dimension. The flight speed of the j th particle is $V^j = [v_1^j, v_2^j, \cdots v_K^j]$ and its coordinate is $D^j = [d_1^j, d_2^j, \cdots, d_K^j]$ which is latent solution of the parameters vector $[\beta_1, \cdots \beta_N, \mu_1, \cdots, \mu_N]$. The value of fitness function is calculated according to the particle coordinate, by which the quality of each particle is evaluated. Through iteration of formulas (9) and (10), flight speed and position of particles can be renewed. At the same time, the local optimal solution of each particle $P^j = [p_1^j, p_2^j, \cdots, p_K^j]$ and the global optimal solution of the whole particle swarm $P_g = [p_{g1}, p_{g2}, \cdots, p_{gK}]$ can be renewed iteratively. After several iterations, P_g is the optimal solution of the target function, namely $\{\hat{\beta}, \hat{\mu}\} = P_g$.

$$v_k^j(t+1) = wv_k^j(t) + c_1 r_1\left[p_k^j(t) - d_k^j(t)\right] + c_2 r_2\left[p_{gk}(t) - d_k^j(t)\right]. \tag{9}$$

$$p_k^j(t+1) = p_k^j(t) + v_k^j(t). \tag{10}$$

Where w is inertia weight; c_1 and c_2 denote acceleration coefficients; r_1 and r_2 are uniformly distributed in (0, 1).

As cluster centers differ from each other to a certain extent, the solution will be meaningless if there are too little differences between the last N components of D^j. Hence, equation (8) is optimized by constrained particle swarm algorithm. Practical procedure is as follows: calculate corresponding fitness function value of the particle by equation (8) when the distance between any two of the last N components of D^j is lager than a certain constant, that is, $\left|d_p^j - d_q^j\right| > \lambda$, or else the value of fitness function is given a very small constant, and the particle will be initialized again.

Finally, the proposed algorithm is summarized as follows:

1) Find all time-frequency points where the length of observation vectors is lager than ρ and satisfy the formula (2) , the real parts and imaginary parts of the observation vectors at these points constitute the set $\{Y(t,f)\}$;

2) The vectors in set $\{Y(t,f)\}$ are transformed into one-dimensional data $z(i)$ using formula (3);

3) Generate J particles randomly in the target space, and optimize the objective function using formulae (9) and (10) iteratively. After several iterations, PSO eventually converge, then the global optimal solution P_g is the maximum likelihood estimation of the generalized Gaussian mixture modal parameters $[\hat{\beta}_1, \cdots \hat{\beta}_N, \hat{\mu}_1, \cdots, \hat{\mu}_N]$;

4) Calculate $\begin{bmatrix} \sin(\hat{\mu}_1) & \sin(\hat{\mu}_2) & \cdots & \sin(\hat{\mu}_N) \\ \cos(\hat{\mu}_1) & \cos(\hat{\mu}_2) & \cdots & \cos(\hat{\mu}_N) \end{bmatrix}$, the estimation $\hat{A}$ of mixing matrix can be obtained.

Computer Simulation

Speech signals last for 9 seconds and are sampled at 16 KHz .We used STFT basis to decompose the observations, using a hanning window and 50% overlap. In the simulation, ε in formulae (2) is 0.01.

Transform all observed vectors at single-source-points into one-dimensional data $z(i)$ according to formula (3). Its distribution is shown in Fig.1 (a). As shown in the figure, it is the distribution of $z(i)$ when there are two sensors and four source signals .The true direction of each source is 147.67°,106.27°,11.30°,62.69°, respectively. Fig.1 (b) is the generalized Gaussian mixture density function curve. It can be seen that the location of each peak is consistent with actual clustering directions very well. The estimated value of various parameters are respectively $\beta = \begin{bmatrix} 6.55 & 9.07 & 8.29 & 7.73 \end{bmatrix}$, $\mu = [147.88° \quad 105.91° \quad 11.37° \quad 62.85°]$. It is thus clear that estimated directions are very close to actual ones.

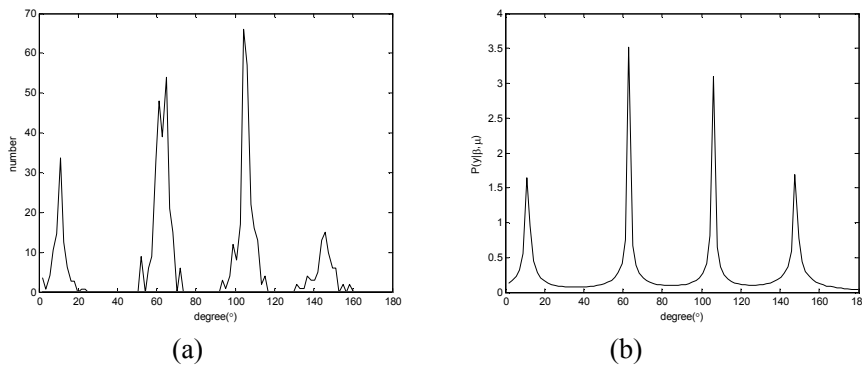(a)   (b)

Fig.1. Distribution of single-source-points and fitting figure

Fig.2 shows the curve of the relation between signal-noise-ratio (*SNR*) and estimation error of mixing matrix calculated by different algorithms when there are four and six source signals. The estimation error is calculated as follows

$$MSE = \frac{1}{N} \sum_{m=1}^{M} \sum_{n=1}^{N} \left[\widehat{A}(m,n) - A(m,n) \right]^2 . \tag{11}$$

Herein a comparison is made between the proposed algorithm and the method reported in [9] which is called SSD-Kmeans. It can be seen that the estimation error becomes higher with the decrease of *SNR* no matter which algorithm is used. Nevertheless, the algorithm put forward in this paper is obviously better than SSD-Kmeans.
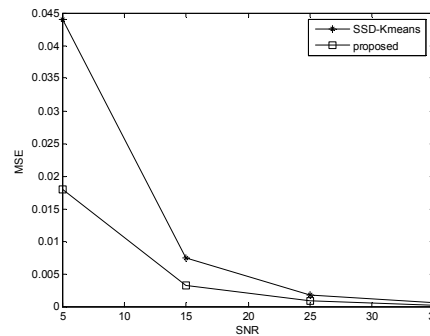

Fig.2. The curve of *MSE-SNR*

Summary

In this paper, the generalized Gaussian mixture modal is used to fit the distribution of single-source-points, and a new objective function for mixing matrix estimation is derived from the viewpoint of maximum likelihood. After optimizing the objective function by constrained particle swarm optimization, the mixing matrix is estimated. Because particle swarm optimization has strong global searching ability, the mixing matrix estimation can obtain higher precision.

References

[1] S Choi, A Cichocki, H. M. Park, et al. Blind Source Separation and Independent Component Analysis: A Review [J]. Neural Information Processing, 2005, 6(1):1-57.

[2] Chen Y Q, Tian Y F. Blind signal separation based on ppso-grads algorithm[J].Journal of Electronic& Information Technology,2007,29(8):1817-1820.

[3]A Belouchrani, M.G. Amin. Blind source separation based on time–frequency signal representations [J].IEEE Trans. on Signal Processing , 1998,46 (11):2888–2897.

[4] S Rickard, O Yilmaz. On the approximate W-disjoint orthogonality of speech[C]// Proceedings of ICASSP, 2002: 529–532.

[5] O Yilmaz, S Rickard. Blind separation of speech mixtures via time–frequency masking [J]. IEEE Trans. on Signal Processing, 2004, 52 (7):1830–1847.

[6] M Pedersen, T Lehn-Schioler, J Larsen. Blues from music: Blind underdetermined extraction of Sources from music[C]// Proceedings of Independent Component Analysis and Blind Signal Separation, 2006:392–399.

[7] N Linh-Trung, A Belouchrani, Abed-Meraimet K, et al. Separating more sources than sensors using time–frequency distributions[C]//Proceedings of the International Symposium on Signal Processing and its Applications, 2001: 583–586.

[8] A Aissa-El-Bey, N Linh-Trung, Abed-Meraimet K, et al. Underdetermined blind separation of nondisjoint sources in the time- frequency domain [J]. IEEE Trans. on Signal Processing, 2007, 55(3): 897–907.

[9] S. G. Kim, C. D. Yoo. Underdetermined Blind Source Separation Based on Subspace Representation [J]. IEEE Trans. on Signal Processing, 2009, 57(7):2604-2614.

[10]S Arberet, A Ozerov, R Gribonval, et al. Blind spectral-GMM estimation for underdetermined instantaneous audio source separation[C]// Proceedings of 8th International Conference on Independent Component Analysis and Signal Separation, 2009:751-758.

[11]G. T. Pulido, C. A. C. Coello. A Constraint-Handling Mechanism for Particle Swarm Optimization[C]// Proceedings of Evolutionary Computation, 2004, 2: 1396 – 1403.

[12]S Zhang, S.X. Gong. A modified PSO for low sidelobe concentric ring arrays synthesis with multiple constraints [J]. Journal of Electromagnetic Waves and Applications, 2009, 23:1535-1544.