

A new blind separation method for under-determined speech signals based on single source interval pre-extraction

Xiangning Hao*, Guixin Zhang, and Liqiong Deng

Air Force Communications NCO Academy, Dalian, China

Abstract. A new method is proposed to realize the blind separation of speech signals under underdetermined conditions. Before estimating the mixed parameters, the single-source interval pre-extraction operation first "filters" out a part which is obviously not a single source. The time-frequency interval of the source analysis domain category. The simulation results show the good performance of the algorithm.

1 Introduction

How to identify the sound source of interest from a noisy environment and obtain a clear voice is an urgent problem to be solved. For a speech signal, the strength of the signal and the change of intonation determine its non-stationarity, and the speech signal involved in this article is a typical non-stationary signal. Different times, occasions, and people will change this non-stationary law and make it more difficult to solve the problem.

2 Blind source separation of speech signal

In the speech processing system, blind source separation is mostly used for speech recognition, that is, to separate and reconstruct the original speech signal from a large number of mixed speech signals^[1].

The block diagram of the system model is shown in Figure 1.

In the figure, $s_1(t), s_2(t), \dots, s_N(t)$ are the N unknown source signal, $x_1(t), x_2(t), \dots, x_M(t)$ are the M observation signal output by the receiving array element, and satisfies $N > M$. The source signal and mixed signal are expressed in vector form, as follows^[2-3]:

$$\mathbf{s}(t) \triangleq [s_1(t), s_2(t), \dots, s_N(t)]^T \tag{1}$$

$$\mathbf{x}(t) \triangleq [x_1(t), x_2(t), \dots, x_M(t)]^T \tag{2}$$

* Corresponding author: hao_xiangning@163.com

The mixing process can generally be seen as a mapping from $\mathbf{s}(t)$ to $\mathbf{x}(t)$:

$$\mathbf{s}(t) \mapsto \mathbf{x}(t) \quad (3)$$

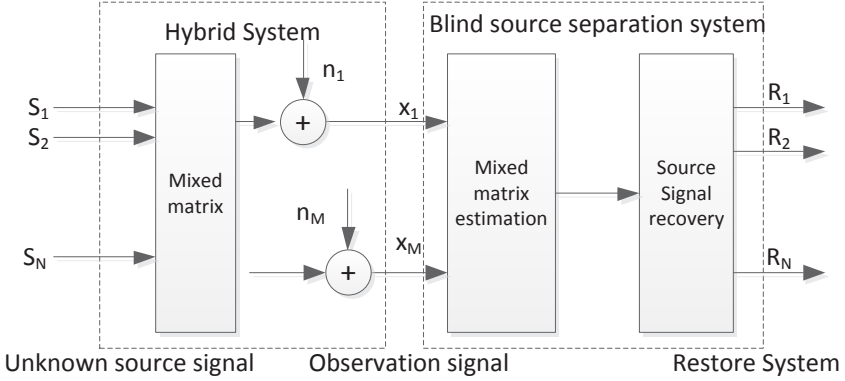


Fig. 1. Underdetermined blind source separation system model.

3 Single source interval pre-extraction

Through research of TIFROM algorithm and TIFCORR algorithm based on the two single-source interval detection algorithms^[4], it can be found that when it is used to deal with the under-determined blind source separation model, as the number of source signals increases, the two algorithms will occur errors in the search of intervals, and it is not guaranteed that the search results of a single source interval are optimal each time, resulting in poor algorithm performance^[5-6]. In order to effectively identify the single-source interval, this paper introduces the single-source interval pre-extraction algorithm. Before estimating the mixed parameters, the single-source interval pre-extraction operation is used to "filter" a part of the time-frequency that obviously does not belong to the single-source analysis domain. Interval.

$$\eta_i(\tau, \omega) = \frac{\alpha_{i1}e^{-j\varphi_{i1}}S_1(\tau, \omega) + \dots + \alpha_{iN}e^{-j\varphi_{iN}}S_N(\tau, \omega)}{\alpha_{11}e^{-j\varphi_{11}}S_1(\tau, \omega) + \dots + \alpha_{1N}e^{-j\varphi_{1N}}S_N(\tau, \omega)} \quad (4)$$

If the time-frequency interval $(T, \omega) (T = 1, 2, \dots, P)$ is the single-source interval $S_k(\tau, \omega)$ of the source signal, there are:

$$\eta_i(T, \omega) = \frac{\alpha_{ik}e^{-j\varphi_{ik}}S_k(T, \omega)}{\alpha_{11}e^{-j\varphi_{11}}S_k(T, \omega)} = b_{ik}e^{-j\mu_{ik}} \quad (5)$$

Then for any time-frequency point $(\tau_m, \omega), (\tau_n, \omega) \in ([\tau_1, \tau_2, \dots, \tau_p], \omega)$, there must be:

$$b_i(\tau_m, \omega) - b_i(\tau_n, \omega) = 0 \quad (6)$$

$$\mu_i(\tau_m, \omega) - \mu_i(\tau_n, \omega) = 0 \quad (7)$$

In practical applications, the above standards are too harsh and can be relaxed as follows:

$$\begin{cases} |b_i(\tau_m, \omega) - b_i(\tau_n, \omega)| \leq \xi \\ |\mu_i(\tau_m, \omega) - \mu_i(\tau_n, \omega)| \leq \zeta \end{cases} \quad \forall (\tau_m, \omega), (\tau_n, \omega) \in ([\tau_1, \tau_2, \dots, \tau_P], \omega) \quad (8)$$

Among them, ξ and ζ are the single-source interval decision threshold.

In practical applications, the process $\forall (\tau_m, \omega), (\tau_n, \omega) \in ([\tau_1, \tau_2, \dots, \tau_P], \omega)$ of discriminating any two points in the interval that satisfies equation (8) requires a large amount of calculation. Therefore, this paper takes the absolute value of the maximum difference in the interval, then the time-frequency interval that satisfies the formula (9) must satisfy the formula (8).

$$\begin{cases} |\max(b_i(\tau, \omega)) - \min(b_i(\tau, \omega))| \leq \xi \\ |\max(\mu_i(\tau, \omega)) - \min(\mu_i(\tau, \omega))| \leq \zeta \end{cases} \quad \forall (\tau, \omega) \in ([\tau_1, \tau_2, \dots, \tau_P], \omega) \quad (9)$$

Equation (9) is the single-source interval extraction criterion, and any time-frequency interval in the real-time frequency domain that satisfies equation (9) will be extracted as a single-source interval (the value of P is generally $3 \sim 5$).

4 Simulation verification

Next, take the under-determined hybrid system with four source signals and two receiving sensors as an example, and verify the article by comparing the time-frequency ratio and phase difference histograms of the two observation signals before and after the single source interval pre-extraction. The validity of the single-source interval extraction criterion.

It can be seen from Figures 2 to 5 that before the single source interval extraction, the distribution of the modulus ratio and phase difference of the sum of the two observation signals is very dense. However, after the single-source interval extraction, the statistical distribution characteristics of the modulus ratio and the phase difference are very obvious, and most of the useless data is filtered out at the same time, which reduces the computational complexity of estimating the parameters of the mixed matrix. Figures 6 and 7 are the scatter plots of the modulus ratio and phase difference before and after the single source interval extraction.

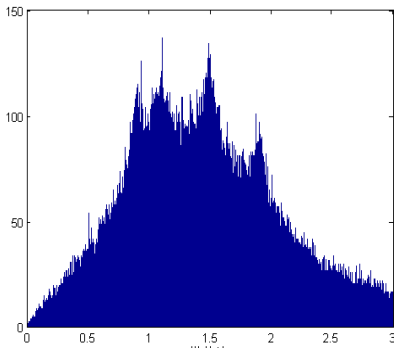


Fig. 2. Histogram of the modulus ratio of X1 and X2 before extraction.

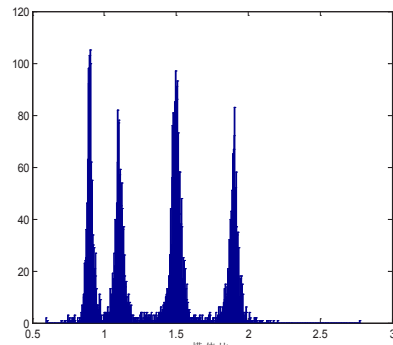


Fig. 3. Histogram of the modulus ratio of X1 and X2 after extraction.

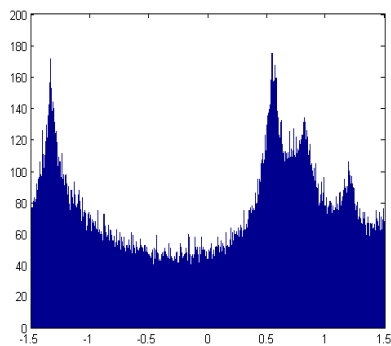


Fig. 4. Histogram of the phase difference between X1 and X2 before extraction

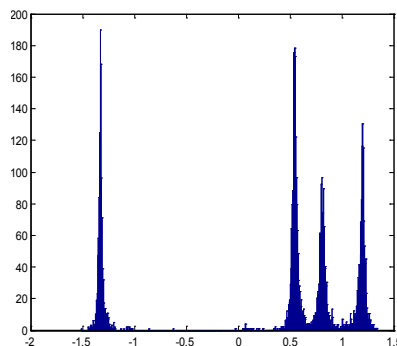


Fig. 5. Histogram of the phase difference between X1 and X2 after extraction.

5 Result analysis

It can be seen from Figure 6 that the scatter plot distribution of the modulus ratio and phase difference of the original data is irregular. The data after single-source interval detection shows good aggregation, which can prove that the single-source interval pre-processing algorithm has a good preprocessing effect. The time-frequency interval that is obviously not in the single-source analysis domain is eliminated, the efficiency of speech source signal extraction is improved, and a good foundation is laid for the correct extraction of speech signal in the next step.

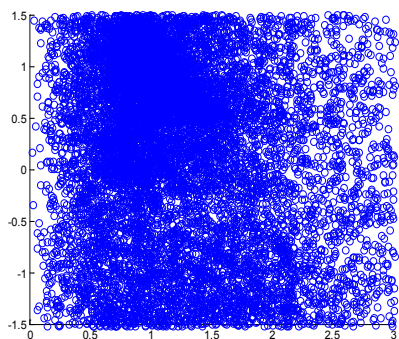


Fig. 6. Scatter plot of the Modulus ratio and phase difference before extraction.

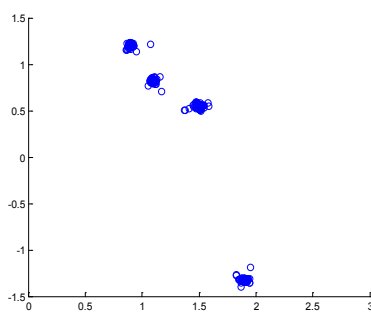


Fig. 7. Scatter plot of the Modulus ratio and phase difference after extraction.

References

1. A JBell, T JSejnowski. An information-maximization approach to blind separation and blind separation and blindde convolution[J].NeuralComputation,1995(7):1129-1159.
2. Aapo Hyvarinen, ErkkiO.ja. Independent ComponentA-nalysis:Algorithmsand Applications[J].NeuralNet-works,2000,13:411-430.
3. Xining Zhang, Yuhe Liao, Guangrui Wen, A new blind source separation method and its application, Sichuan Armory Engineering Journal, January 2010, Volume 31, Issue 1
4. Gang Wang,Bin Sun. Blind signal separation technology and algorithm research [J]. Aerospace Electronic Warfare, 2015, 31(4): 53-56. doi: 10.16328/j.htdz8511.2015.04.015.

5. Bai Lin, Chen Hao, A new method of blind separation of speech signals under underdetermined conditions, *Computer Application Research (Application Research off Computers)*, July 2010, Volume 27, Issue 7 (Vol.27No.7, Jul.2010)
6. Laljahan L, Iqbal S, Sarwar F, et al. The relationship of generalized fractional Hilbert transform with fractional Mellin and fractional Laplace transforms [J]. *American Scientific Research Journal for Engineering, Technology, and Sciences(ASRJETS)*, 2017, 38 (2):90-97.