

Received April 12, 2020, accepted April 27, 2020, date of publication May 8, 2020, date of current version May 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2993305

Blind Source Separation Algorithm for Chaotic Masking Multipath Signals Based on Spectral Peak Search Counter Permutation

SHIYU GUO^{ID}, JIAYIN YU, CHAO LI, MENGNA SHI, AND ERFU WANG

Electronic Engineering College, Heilongjiang University, Harbin 150080, China

Corresponding author: Erfu Wang (wangerfu@hlju.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61571181 and Grant 61801173, and in part by the Natural Science Foundation of Heilongjiang Province, China, under Grant LH2019F048.

ABSTRACT Since voice communication is the main method of information transmission, it is important to ensure the safety and efficiency of voice communication. In this paper, a more complex multipath channel model in a wireless environment is considered, and chaotic masking technology is used to ensure the security of voice signal transmission. Based on this, a multipath blind source separation algorithm based on cross-correlation permutation of spectral peak search counters is proposed. First, chaotic masking is performed between multiple voice signals and chaotic signal, and an impulse response filter is used to simulate the multipath transmission process. Next, the short-time Fourier transform is employed to convert the observation signal into a frequency-domain signal. Then, the joint approximate diagonalization of eigenmatrix algorithm is used to perform linear instantaneous blind source separation at each frequency. And to solve the permutation uncertainty problem in the frequency domain method, a cross-correlation permutation algorithm based on a spectral peak search counter is proposed. Finally, the permuted signal is transformed back to the time domain to get the estimated source signal by the inverse short-time Fourier transform. And simulation experiments show that the algorithm greatly reduces the calculation amount of the autocorrelation coefficient in the traditional permutation algorithm. Moreover, the permutation algorithm has stability and accuracy, and provides a solution for the secure transmission and effective separation of voice signals in multipath transmission channels.

INDEX TERMS Chaotic masking, convolutive mixture blind source separation, cross-correlation permutation, short-time Fourier transform.

I. INTRODUCTION

With the rapid development of information technology, voice transmission technology introduces an increasing amount of convenience to people's life. As the main form of human communication, a voice message is used for its strong real-time performance and easy identification [1]–[3]. Whether it is indoor or outdoor transmission, the transmission and reception of voice information will inevitably be affected by the multipath transmission effect [4], that is, after a voice message was sent, it needs to reach the receiver end via multiple transmission paths. The length (delay) and attenuation of each path are different; thus, the received signal presents a superposition of various voice signals affected by the phase,

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Zia Ur Rahman ^{ID}.

delay, and multipath [5]. On the contrary, an increasing number of fields and departments need secure and efficient voice transmission technology, and information security is facing unprecedented development opportunities [6]. Thus, in the multipath transmission channel, the safe transmission and effective extraction of voice information in and out of the laboratory environment is one of the scientific problems that is being addressed by many scholars.

In recent years, the chaos of nonlinear systems has been increasingly investigated. Chaos is an aperiodic bounded dynamic behaviour caused by a deterministic, nonlinear and dynamical system [7]. Due to its initial value sensitivity, internal randomness and unpredictability, chaos plays an important role in secure communication. Considering the secure transmission of voice information in a wireless transmission environment, chaotic encryption or masking of voice signals

can be performed at the sender. For indoor environments, a classic problem of voice communication is the cocktail parties. In a noisy indoor environment, the microphone is used to record the sound information. At this time, the receiver contains voices of multiple people speaking simultaneously, music sounds, and other sound sources. And there are also reflected sound generated by the reflection of walls and indoor objects. So a voice transmission over multipath channels can be abstracted into a convolutive mixed model in an indoor or outdoor environment. In the process of transmission, a lack of source information and channel information exists. Therefore, the separation of expected voice signals from unknown information is known as the convolutive mixed blind source separation (CMBSS) [8] problem.

Two main types of methods exist to solve the blind source separation (BSS) problem in multipath channels: time-domain method and frequency-domain method. In the time-domain method [9], [10], a separation filter in the time-domain must be calculated with an order equal to or greater than the order of the impulse response filter. However, owing to the influence of the order of the convolutive model, the complexity of the algorithm is high and the convergence speed is slow. The frequency-domain method [11], [12] transforms the time-domain signal into the frequency domain by short-time Fourier transform (STFT) [13]. Using a mature independent component analysis method [14], it is a linear instantaneous BSS at each frequency point. This process is also the key to BSS in the frequency domain. This method avoids tedious convolution operations, and has a small calculation amount and a fast convergence speed. Yet, there are uncertainties in the amplitude and permutation, which prevents accurate separation of source signal. The amplitude uncertainty can be solved by normalizing the separation matrix; however, the permutation uncertainty is always a difficult problem in the frequency-domain method.

Two methods exist to solve the uncertainty problem of frequency-domain method permutation: the first method is the geometric method based on the direction of arrival (DOA) [15]–[17], and the second method is the mutual parameter method that is based on the correlation between adjacent frequency points [18]–[21]. The DOA-based method considers the source signal and the spatial position of the sensor as prior knowledge. The advantage of this method is that it can independently solve the permutation uncertainty problem at each frequency point, and the adjacent frequency points are not related to each other. However, this method is not suitable for the accuracy of the low-frequency point and the effectiveness of the high-frequency point. Thus, achieving this method in a practical environment is difficult. The mutual parameter method, which is based on the correlation of adjacent frequency points, relies on the correlation of the same source signal or the independence of the adjacent frequency points of different source signals. The indicators of similar sequences are optimized to determine whether the separated signals are derived from the same sources. This method is more accurate. To improve the similarity of recovered signal,

X.-L. Bo and others established different influence factors. This method can control the influence of distance between the frequency points and the order of separation sequence at each frequency point, which achieved the purpose of improving the signal similarity [22]. X.-D. Ou *et al.* combined the energy-related permutation method with the DOA ordering method. And the method solved the uncertainty problem of the frequency-domain permutation algorithm, which could accurately separate the signals [23]. W. Cheng *et al.* proposed a frequency-domain convolutive BSS algorithm based on the modified conjugate gradient, which separated and arranged the source signals [24]. These algorithms all have a good effect in solving BSS in the frequency domain, but they also have a high computational complexity. Then K. Fang and others presented a low-complexity permutation alignment method based on the inter-frequency dependence of signal power ratio to resolve permutation problem [25]. This algorithm can reduce its computational complexity. Yet, we designed an effective and lower computational complexity permutation algorithm based on the idea of cross-correlation permutation.

This paper proposed an improvement on the traditional correlation permutation algorithm. We only need to calculate the cross-correlation and find the wrong frequency points in the way of spectral peak search. The autocorrelation is calculated only when the number of wrong frequency points reached the set threshold. Therefore, the cross-correlation permutation (CCP) algorithm that is based on spectrum peak search counter (SPSC) greatly reduced the computational complexity. And the algorithm effectively solved the permutation uncertainty in frequency domain of CMBSS problem. The paper employed chaotic masking technology [26]–[28] to hide the voice signal in the chaotic signal, which can realize the confidential transmission of voice information. Then, the observation signal at the receiver was transformed into the frequency domain by STFT. And the joint approximation diagonalization of eigen-matrix (JADE) [29] algorithm can realize a linear instantaneous BSS at each frequency point. Next, the proposed algorithm is used to reorder the subsignals with wrong ordering. The sequence-adjusted subsignals were transformed into the time domain by the inverse short-time Fourier transform (ISTFT), to achieve accurate recovery of voice signals. The cross-correlation permutation algorithm proposed in this paper can solve the problem of out-of-order separation of signals with less overhead and adaptive spectral peak search, which effectively reduced the complexity and calculation amount of the permutation algorithm of voice transmission over multipath channels.

The content of this paper is arranged as follows. Section II briefly introduces frequency-domain CMBSS in the multipath transmission model and fractional-order chaotic system for confidential transmission. Section III proposes a low-complexity CCP algorithm that is based on the SPSC for permutation uncertainty in the CMBSS algorithm and introduces the implementation steps of the permutation algorithm in detail. Algorithm experiment simulation is performed in

Section IV. Section V analyzes the performance of the algorithm. Section VI gives the conclusions and future research prospects based on the analysis results.

II. MATHEMATICAL MODELING AND THEORETICAL BASIS

This part will establish the signal mixing and separation model of multichannel voice signals in multipath transmission channels for the secure transmission. Further, this section introduces the fractional Chen chaotic system selected for chaos concealment in this paper.

A. MULTIPATH TRANSMISSION AND MATHEMATICAL MODEL

As the actual communication environment is more complex, the voice information will be reflected by the obstacles to form a multipath transmission. The voice information will travel to the receiver via multiple paths to get unknown observation signals. The corresponding signal separation technology can be employed for the observation signal to realize the estimation of the source signal. Fig. 1 shows an example of a three-way voice signal. The signal simulates the multipath transmission of voice information in real environment and the process of separating and extracting.

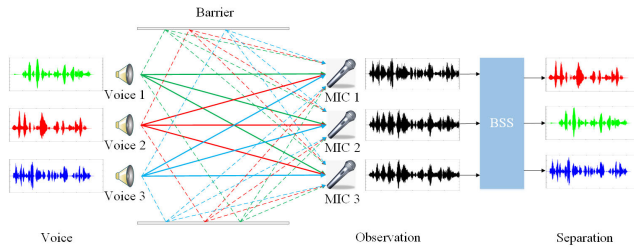


FIGURE 1. Multipath transmission blind source separation system.

The source signal is affected by multipath and reaches the receiver. The voice signal received by the microphone includes both direct wave signals and reflected wave signals. When the source signal and the channel are unknown, the estimation of the observation matrix or the source signal can only be achieved with a small amount of prior knowledge. If noise is not considered, its mathematical model is shown in Fig. 2. This model is called convolutive mixed blind source separation.

It is assumed that there are N source signals $S(t) = (S_1, S_2, \dots, S_N)^T$ and M observation signals $X(t) = (X_1, X_2, \dots, X_M)^T$ ($N \geq M$), and both are discrete signals. The mathematic model for multisource signal mixing [30] is

$$X_i(t) = \sum_{j=1}^N \sum_{p=0}^P h_{ij}(p) \cdot S_j(t-p) = h_{ij} * S_j(t), \quad (1)$$

where $h_{ij}(p)$ is the impulse response from the j th source signal to the i th microphone, $S_j(t)$ is the j th source signal, $X_i(t)$ is the i th observation signal, $*$ is the convolution operation,

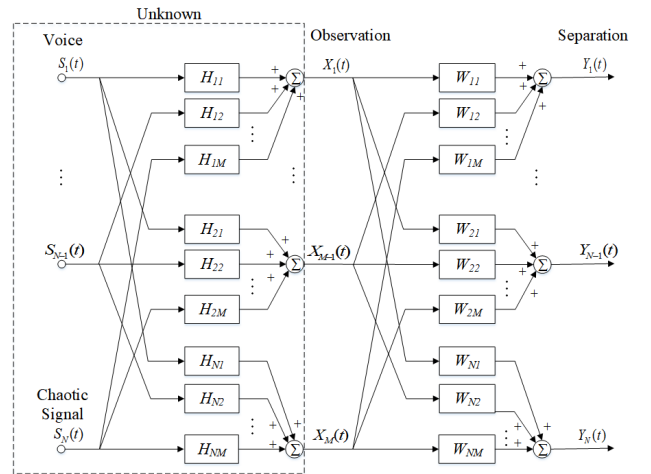


FIGURE 2. The convolutive mixture blind source separation model for multisource signals.

and p is the order of the transfer function. Each observation signal is a convolution mixture of each source signal and its corresponding impulse response. When $p = 1$, the model degenerates into a linear instantaneous mixed model. For convenience, equation (1) is expressed in vector form as

$$X(t) = \sum_{p=0}^P H(p)S(t-p), \quad (2)$$

where $H(p)$ represents the vector representation of the transfer function. In practice, a finite impulse response filter can be applied to approximate the transfer function. The observation signal is converted from the time domain to the frequency domain by STFT, and the mathematical expression of the observation signal at each frequency point [24] is

$$X(t, f) = \sum_{k=0}^K X(k+t)win(k)e^{-i2\pi ft}, \quad (3)$$

where k is the number of frequency points of STFT, $win(k)$ is the window function, which is usually the Hamming window, and the window length is L . The frequency value of each frequency point is $f = 0, \frac{1}{L}f_s, \frac{2}{L}f_s, \dots, \frac{0.5L}{L}f_s$, where f_s is the sampling frequency. When the length of the window function $win(k)$ is substantially larger than the order of the impulse response filter p , the convolutive mixture model at each frequency point can be approximated as an instantaneous mixture model

$$X(t, f) = A(f)S(t, f), \quad (4)$$

where $A(f) = [a_1(f), a_2(f), \dots, a_N(f)]$ is the mixing matrix at the frequency point f , and $a_i(f) = [a_{1i}(f), a_{2i}(f), \dots, a_{Mi}(f)]^T$. And the observation signal is $X(t, f) = [X_1(t, f), X_2(t, f), \dots, X_M(t, f)]^T$ and the source signal is $S(t, f) = [S_1(t, f), S_2(t, f), \dots, S_N(t, f)]^T$. As $M = N$, assuming that the inverse filter of the mixing filter has solution, the estimated signal is

$$Y(t, f) = W(f)X(t, f), \quad (5)$$

where $\mathbf{Y}(t, f) = [Y_1(t, f), Y_2(t, f), \dots, Y_N(t, f)]^T$ is the separation signal at the frequency point f , and $\mathbf{W}(f)$ is the separation matrix in the frequency domain. The purpose of blind signal separation is to find the corresponding observation matrix $\mathbf{W}(f)$ from the observation signal $\mathbf{X}(f)$ to achieve signal separation.

B. FRACTIONAL CHEN CHAOTIC SYSTEM

Fractional Chen chaotic systems have more complex dynamic characteristics than integer-order chaotic systems, and the security of confidential communication is higher. The chaotic system dynamics [31] equation is described as

$$\begin{cases} \frac{d^\alpha x_1}{dt^\alpha} = a(y - x) \\ \frac{d^\alpha x_2}{dt^\alpha} = (c - a)x - zx + cy \\ \frac{d^\alpha x_3}{dt^\alpha} = xy - bz, \end{cases} \quad (6)$$

where $a, b,$ and c are system variables; $x, y,$ and z are state variables. α ($0 < \alpha < 1$) is the order of the fractional-order system. Set the initial values $x_0 = -15, y_0 = -18,$ and $z_0 = 15,$ and the system variables $a = 35, b = 3,$ and $c = 28.$ When the fractional order of the chaotic system reaches $\alpha \approx 0.82,$ the fractional Chen system begins to appear chaotic. For example, the evolution of fractional Chen chaotic motion system when $\alpha = 0.88$ is shown in Fig. 3.

A fractional-order Chen chaotic system has better key space and system complexity. The chaotic signal has a large bandwidth and good security characteristics. The voice signal is hidden in the chaotic signal to form a noise-like signal. And the useful information is completely covered by the chaotic signal to realize the secure transmission of the voice signal.

III. LOW-COMPLEXITY PERMUTATION ALGORITHM IN THE FREQUENCY DOMAIN

In the frequency-domain method of convolutive mixture blind source separation, there is a problem of permutation uncertainty. There are mainly two kinds of methods to solve this problem. The first method is the geometric method based on the direction of arrival, which can solve the permutation problem at individual frequency points. However, the spatial position of the sensor is strict, and it takes a lot of calculation. The second method is a correlation coefficients method that is based on adjacent frequency points. This method optimizes the indicators of similar sequences based on the correlation of the same source signal or the independence of adjacent frequency points of different source signals. This method is more accurate, but the optimization process will inevitably increase the computational complexity of the permutation algorithm. So, this paper proposes a low-complexity permutation algorithm that is based on the cross-correlation of adjacent frequency points, while ensuring robustness. This algorithm is an improvement on the traditional method that is based on the correlation of adjacent frequency points by adding a counter to the peak value

search of the cross-correlation. This algorithm can not only effectively solve the out-of-order problem in the separation algorithm, but also reduce the computational complexity of the traditional permutation algorithm.

A. CROSS-CORRELATION PERMUTATION ALGORITHM BASED ON SPECTRAL PEAK SEARCH COUNTER

In the traditional correlation-based permutation algorithm [20], different signals have completely different signal contours at the same frequency, that is, they are completely uncorrelated. However, the signal contours between adjacent frequency points of the same signal are roughly the same, and have very extremely high correlation. Therefore, the amplitude correlation between adjacent frequencies is used as the basis for permutation. It is not difficult to infer that the energy between adjacent frequencies also has a correlation, which can also solve the permutation problem [23]. And the energy is more stable than the amplitude. First, the correlation coefficient is given as

$$cor(y_i(f), y_j(f)) = \frac{E(y_i(f)y_j(f)) - E(y_i(f))E(y_j(f))}{\sqrt{D(y_i(f))} \cdot \sqrt{D(y_j(f))}} \quad (1 \leq i, j \leq M), \quad (7)$$

where $E(\cdot)$ represents the average value of the signal, $D(\cdot)$ represents the variance of the signal, and $y_i(f)$ represents the i th signal separated at the frequency point $f.$ The amplitude of the separation subsignal $y_i(f)$ is $|y_i(f)|.$ Similarly, we define the autocorrelation and correlation energy coefficients as

$$C_{ii} = cor(|y_i(f)|^2, |y_i(f + 1)|^2), \quad (8)$$

$$C_{ij} = cor(|y_i(f)|^2, |y_j(f + 1)|^2), \quad (9)$$

where C_{ii} represents the correlation between the frequency point f and the frequency point $(f + 1)$ of the i th signal, which is called the autocorrelation energy coefficient. C_{ij} represents the correlation between the frequency point f of the i th signal and the frequency point $(f + 1)$ of the j th signal, which is called the cross-correlation energy coefficient.

In this algorithm, we need to set the comparison threshold. The purpose is to compare with the correlation which is obtained at different frequency points of the separated subsignals. Regarding the selection of the comparison threshold, the corresponding reference [30] is given. Therefore, we set the comparison threshold to $\epsilon = 0.7.$ If $\rho_{ij} > \epsilon,$ it indicates that the separated signal is misordered at the frequency points. These frequency points are called wrong frequency points. Here it is necessary to calculate the autocorrelation of the wrong frequency points, then reorder the separation signal.

Taking two-channel source signals as an example, the energy correlation coefficient is as follows:

$$\begin{cases} C_{11} = cor(|y_1(f)|^2, |y_1(f + 1)|^2) \\ C_{12} = cor(|y_1(f)|^2, |y_2(f + 1)|^2) \\ C_{21} = cor(|y_2(f)|^2, |y_1(f + 1)|^2) \\ C_{22} = cor(|y_2(f)|^2, |y_2(f + 1)|^2). \end{cases} \quad (10)$$

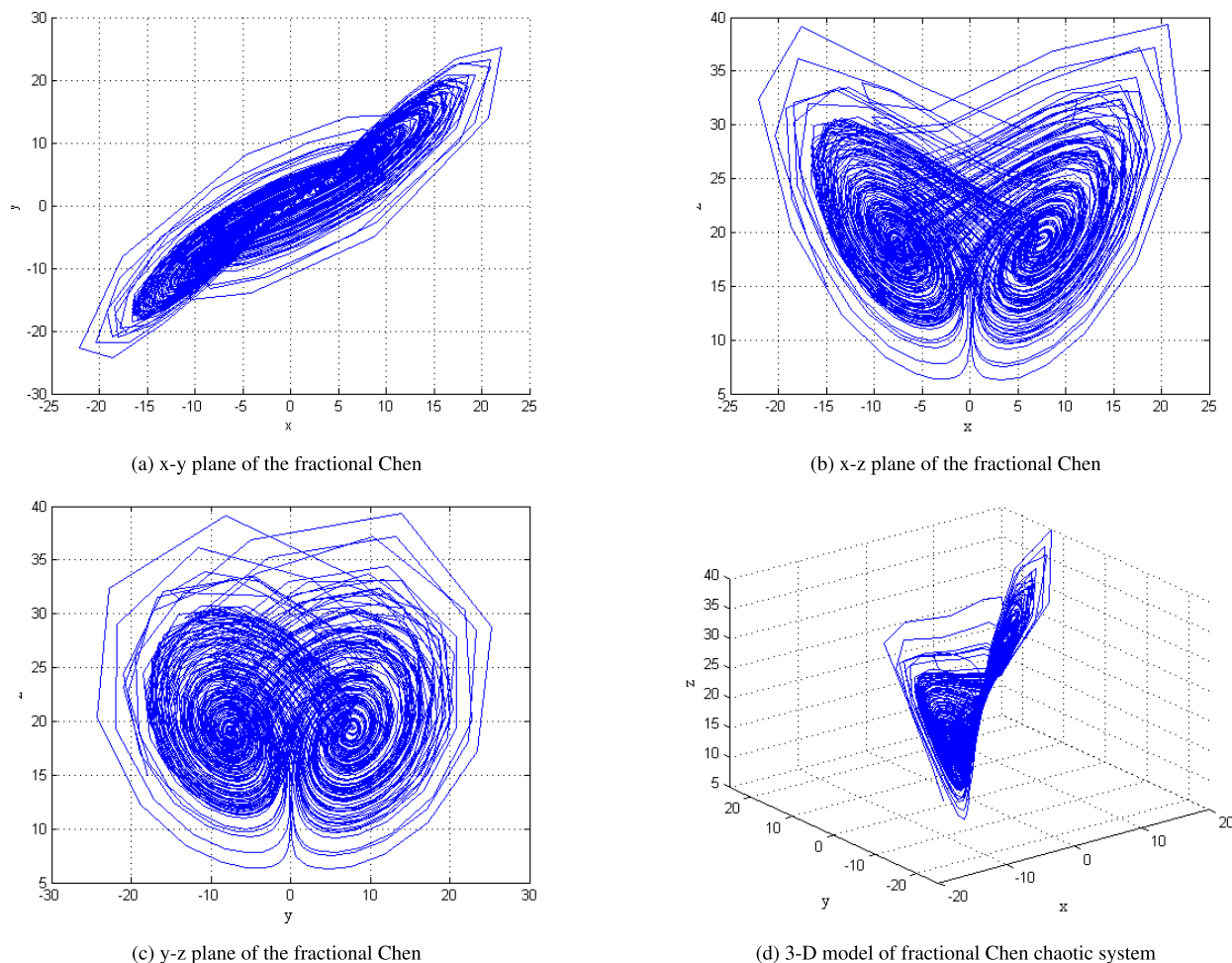


FIGURE 3. 3-D evolution diagram of fractional-order Chen chaotic system.

If $C_{12} > 0.7$ or $C_{21} > 0.7$, then calculate the autocorrelation C_{11} and C_{22} , and reverse the order of the two separated subsignals; otherwise, it remains in the original position.

It should be noted that the algorithm in this paper needs two key stages, namely the search stage and the error-correcting permutation stage. Obviously, the calculation amount of the search stage is much larger than the error-correcting permutation stage. The search process needs to find the location and number of wrong frequencies. In this stage, the calculations for autocorrelation and cross-correlation are identical. Yet, our Monte Carlo experiments show that cross-correlation is more sensitive than autocorrelation. In other words, when the subsignal of the frequency point appears out of order, the autocorrelation is relatively close to 1, while the cross-correlation is already far away from 0. Therefore, this article takes the cross-correlation as the basis of the search stage, and uses it to determine whether the frequency point is wrong. In the next error-correction permutation stage, the autocorrelation calculation is adopted.

Therefore, the basic idea of the cross-correlation permutation algorithm based on the spectrum peak search counter is as follows. It is assumed that the first frequency point is

a reference frequency point, and the signals at the frequency point are correctly ordered. Firstly, the frequency points are traversed to calculate the correlation of separated signals between adjacent frequency points. Secondly, it determines whether the frequency point is a wrong frequency point according to the comparison threshold. Then a counter is used to mark such frequency points, and record the position and quantity information. This counter is called a spectrum peak search counter. The above process is the search stage in the algorithm of this paper. Next, the error-correction permutation stage of this algorithm is designed. Here we set an upper limit for the number of wrong frequency points, and define the upper limit as the threshold of wrong frequency points (TWFP), which is N_t . When the count value is smaller than the TWFP, the position of the separated subsignals at each frequency point can be kept unchanged. And the disorder of individual frequency points will not affect the recovery effect. When the count value is greater than the TWFP, the autocorrelation of the marked frequency points needs to be calculated. Among the calculated autocorrelation and cross-correlation, we choose the largest one. The most relevant signal should be placed at the same position as the reference frequency point.

TABLE 1. Permutation algorithm in the frequency domain.

Algorithm: The CCP algorithm based on the SPSC
Input: SF $SF(i, k)$ represents the frequency-domain separated subsignal, i represents the number of separators ($1 \leq i \leq M$), and k represents the position of the frequency points.
Output: SF' $SF'(i, k)$ represents the frequency-domain separated subsignal after permutation adjustment.
Initialization: LocalL=[], $C=0$, K is the number of frequency points of STFT, M represents the number of observed signals, and N_t is the threshold of wrong frequency point.
For $k \leftarrow 1 : (K/2)$ do For $i \leftarrow 1 : M$ do For $j \leftarrow 1 : M$ do If $i \neq j$ then $W_{ij} \leftarrow W(SF(i, k), SF(j, k))$ we calculate the cross-correlation energy coefficient between each subsignals of by (10). If $W_{ij}(k) > 0.7$ then Set the cross-correlation threshold to be 0.7. $C \leftarrow C + 1$ Start counter. LocalL(C) $\leftarrow k$ End End End End If length(LocalL) $> N_t$ For $k \leftarrow 1 : \text{LocalL}$ do For $i \leftarrow 1 : M$ do For $j \leftarrow 1 : M$ do $W_{ii}(k) \leftarrow W(SF(i, k), SF(i, k))$ we calculate the autocorrelation energy coefficients of each signal by (9). $W_{imax}(k) \leftarrow \max(W_{ii}(k), W_{ij}(k))$ $SF(i, k) \leftarrow SF(\max, k)$ The signal of the path separated subsignal at the maximum value is substituted with the signal of the path. End End End End End Return SF'

Finally, a cross correlation permutation ranking algorithm based on a spectrum peak search counter is implemented. To sum up, the specific implementation process of the proposed algorithm is described in TABLE 1.

In the traditional algorithm, all wrong frequency points need to be reordered. And the autocorrelation and cross-correlation between all frequency points needs to be calculated. The algorithm in this paper only performs sequencing when the number of wrong frequency points is greater than the TWFP. At the same time, we only need to calculate the cross-correlation. The autocorrelation is only calculated when the frequency point is misordered. Therefore, the cross-correlation permutation algorithm based on the spectrum peak search counter greatly reduces the computation complexity of the algorithm. And this still can obtain a good estimated signal. The time complexity of the algorithm can be obtained by calculation and analysis, and the low complexity of the algorithm is proved theoretically.

B. CONVOLUTIVE MIXED BLIND SOURCE SEPARATION ALGORITHM BASED ON CROSS-CORRELATION PERMUTATION

Based on the spectral peak search counter and cross-correlation permutation to solve the signal separation in the

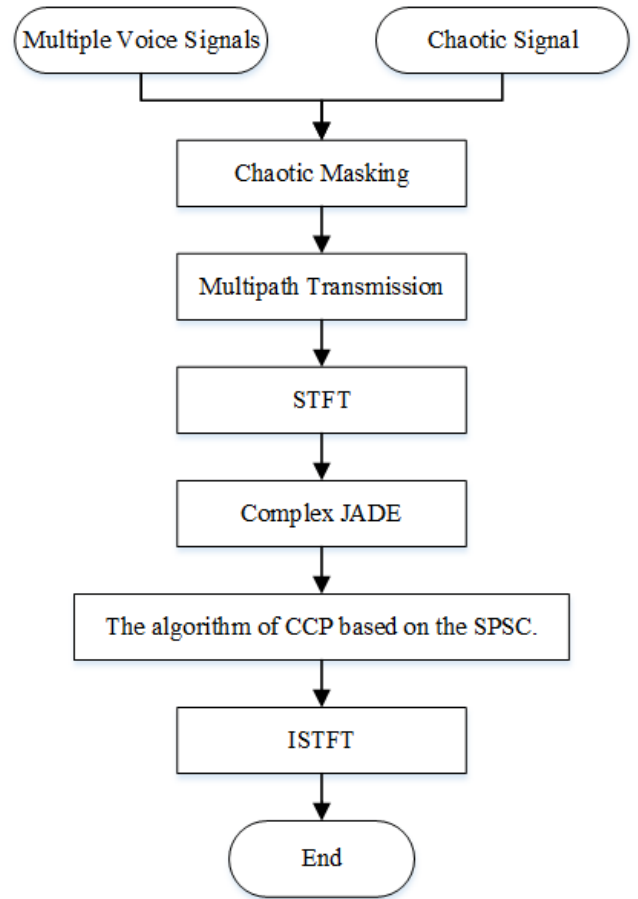


FIGURE 4. Convolutive mixture blind source separation algorithm based on cross-correlation permutation.

case of secure multipath voice transmission, the total algorithm flow description is shown in Fig.4.

At the beginning, multiple voice signals are hidden into chaotic signal to ensure the security of voice information. Considering the convolutive problem during multipath transmission, the time-domain signals are converted to the frequency domain using STFT. In the frequency domain, the complex JADE algorithm is used for instantaneous blind source separation to obtain the estimated signal. The cross-correlation permutation algorithm based on the spectral peak search counter was used to sequence the separated signals. Finally, the estimated signal is transformed into the time domain by ISTFT.

C. PERFORMANCE EVALUATION INDEX

In the process of convolutive mixture blind source separation simulation experiment, there still are some differences between the signal estimation and the source signal even in the absence of noise. The subjective and objective aspects are usually used to evaluate the estimated signal. This paper used the similarity coefficient (SC) and signal-to-interference ratio (SIR) to quantitatively analyze the performance of separated signals from an objective perspective.

SC: $s(t)$ and $y(t)$ represent the source signal and estimation of source signal respectively, and the mathematical expression is shown as follows

$$\rho(y_i, s_i) = \frac{\left| \sum_t y_i(t) s_i(t) \right|}{\sqrt{\sum_t y_i^2(t) \sum_t s_i^2(t)}}. \quad (11)$$

The similarity coefficient between the source signal $s_i(t)$ and the estimated signal $y_i(t)$ is between 0 and 1. When the two signals are completely correlated, the similarity between the source signal and estimated signal is higher. The closer the similarity coefficient is to 1, the better the separation performance of the algorithm will be. Otherwise, the smaller the similarity coefficient, the worse the separation effect of the algorithm.

SIR: $s_{t \text{ arg et}}(t)$ represents the part of the estimated signal that belongs to the source signal; $e_{interf}(t)$ indicates that the estimated signal does not meet the source signal, but belongs to the mixed signal, which is the estimation error caused by other signal sources [32], [33].

$$\begin{aligned} SIR_i &= 10 \log_{10} \frac{\|s_{t \text{ arg et}}(t)\|^2}{\|e_{interf}(t)\|^2} \\ &= 10 \log_{10} \frac{\sum_t y_{ii}^2(t)}{\sum_t \sum_{j \neq i} y_{ij}^2(t)}, \end{aligned} \quad (12)$$

where $y_{ii}(t)$ represents the energy from the source signal $s_i(t)$ in the estimated signal $y_i(t)$ of the i th source signal $s_i(t)$, and y_{ij} represents the energy from $y_i(t)$ in other sources than the source signal $s_i(t)$, which is the interference component in $y_i(t)$. The larger the value SIR, the higher the degree of crosstalk suppression, and the better performance of the separated signals.

IV. SIMULATION

The voice signal in the TIMIT database, and the x component of the fractional-order Chen chaotic system are randomly selected as the input signal. The length of the signal was 3 s, and the sampling frequency was $f_c = 16\text{KHz}$, and the filter is $p = 20$. Then STFT is performed on each observed signal, whose window length is $L = 256$, frame shift is $H = 32$, and weighting function is hamming window. Furthermore, the number of frequency points is $K = 1024$.

A. PERMUTATION ALGORITHM EXPERIMENT

The input signal is mixed through the impulse response filter and the observed signal is obtained. The voice signal is well hidden in the chaotic signal to realize safe and secret transmission. The experimental simulation is shown in Fig.5.

The simulation results show that the source signal information cannot be reflected in the observed signals. Chaotic signals have a good effect of chaotic masking on voice signals to realize the secret transmission of voice signals. A ‘‘blind’’ observation signal is simultaneously obtained at the receiver.

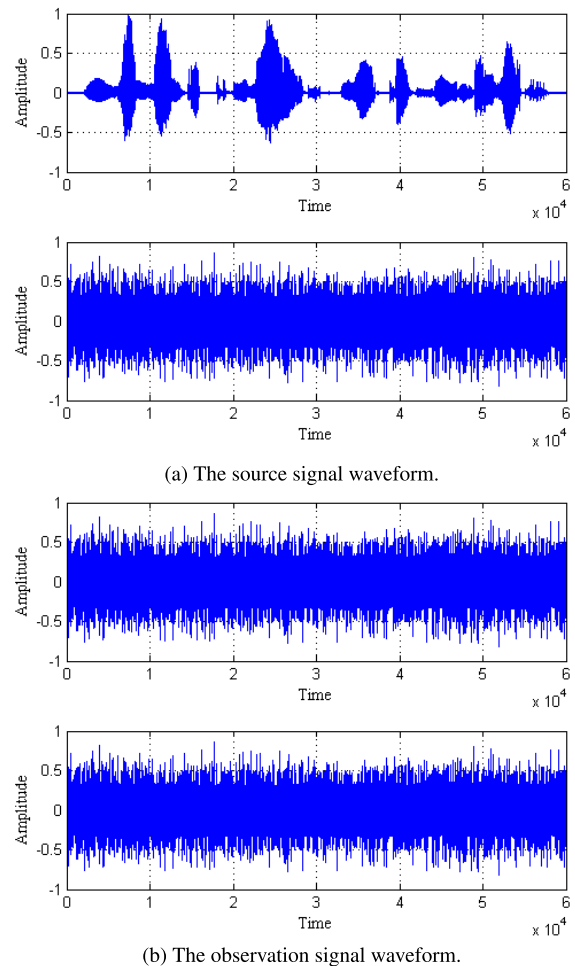


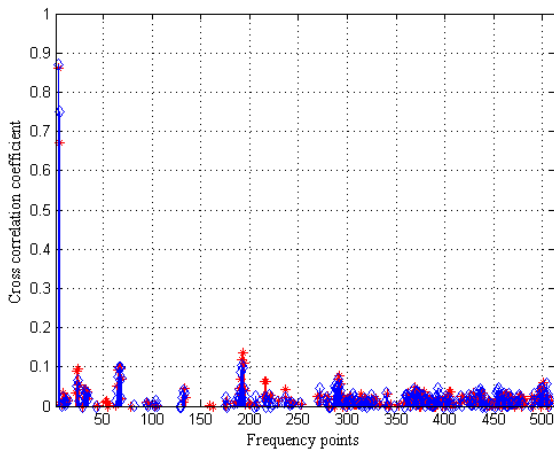
FIGURE 5. The source signal and observation signal waveform.

TABLE 2. Performance evaluation of separated signals.

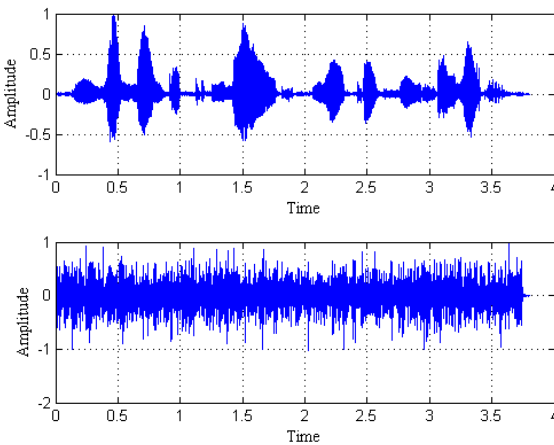
	Similarity coefficient		SIR	
	Separation 1	Separation 2	Separation 3	Separation 4
Voice	0.9977	0.0176	23.3633	-5.2915
Chaotic	0.0047	0.7159	-0.4990	2.9744

To analyze the effectiveness of the cross-correlation permutation algorithm based on the spectral peak search counter, we evaluate it with different number of wrong frequency points. When the number of wrong frequency points is small, as shown in Fig. 6, only 2 exist. At this time, the subsignal of wrong frequency points is not replaced. And the separated signal waveform remains still clearly visible. The similarity coefficient and SIR are used as the evaluation criteria for separated signals, as shown in TABLE 2.

The data in TABLE 2 reveals that the similarity coefficient between separation signal 1 and source voice signal is 0.9977 and the SIR also reaches 23.3633. The accurate estimation of the voice signal is realized. The similarity coefficient between the separation signal 2 and the chaotic signal is 0.7195, and the chaotic signal used for masking can also obtain better separation results. So, when the individual frequency points are wrong, there is no ordering of the



(a) The spectral peak search of cross-correlation at each frequency point.



(b) The separated signals.

FIGURE 6. The effect of individual wrong frequency on the result.

separated signals and its effect on the separation results is negligible. The computational complexity of the algorithm is reduced. Therefore, the individual wrong frequency points are not considered.

When the number of wrong frequency points is large, as shown in Fig. 7 (a), 54 wrong frequency points are reached. If a replacement operation is not operated for this frequency point, the recovered voice signal waveform will become so fuzzy that it is impossible to accurately recover the original signal. At this point, performing the permutation of the separation subsignal is necessary for the wrong frequency point shown in Fig. 7 (a). The spectral peak search diagram of the cross-correlation coefficients of each frequency point after permutation is shown in Fig. 7 (c). Compare the experimental results before and after replacement, and correct the wrong frequency points in turn. The separated voice signal waveform is more similar to the original signal and the similarity coefficient is increased from 0.5092 to 0.8644. This experimental analysis can verify the effectiveness of the CCP algorithm based on the SPSC. This algorithm reduces unnecessary calculation and permutation operations. And this still can accurately recover the separated signal.

B. DETERMINE THE THRESHOLD VALUE OF WRONG FREQUENCY POINTS

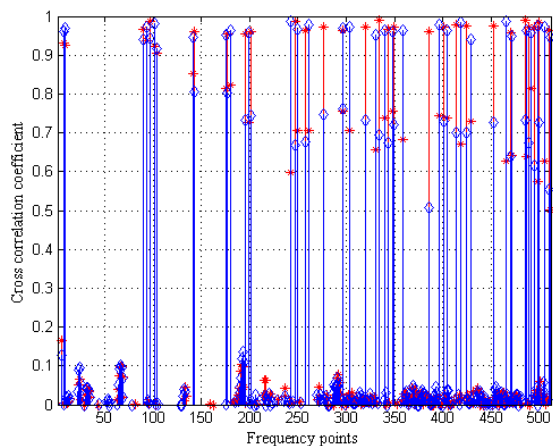
The experiments in the previous section verify the effectiveness of the CCP algorithm based on the SPSC proposed. The key point of this algorithm is how to select the threshold N_t . If N_t is too small, the influence of the wrong frequency points on the separation result will be reduced; however, it cannot reflect the advantages of this algorithm's low complexity. If N_t is too large, although the computational complexity is reduced, too many wrong frequency points will affect the recovery of the source signal. Therefore, we used the results of a large number of experiments to analyze and determine the threshold value N_t . So the 300 Monte Carlo experiments are carried out. The random part is the coefficient of the filter which is the primary parameter in the analog channel mixed process. And the other experimental conditions will not change (the experimental conditions are set as the first paragraph of section IV). Perform and record experimental data. The number of wrong frequency points in each experiment and the corresponding occurrence times in 300 experiments are shown in Fig. 8.

According to the statistical data of 300 experiments, the proportion with 0 wrong frequency point is the highest, that is, all frequency points have no disorders, which accounts for approximately 60% of the experiments. The number with wrong at other wrong frequency points is much smaller than this value, as shown in Fig. 8, which indicates that the number of required permutations is very small. Therefore, the CCP algorithm based on the SPSC needs small calculations.

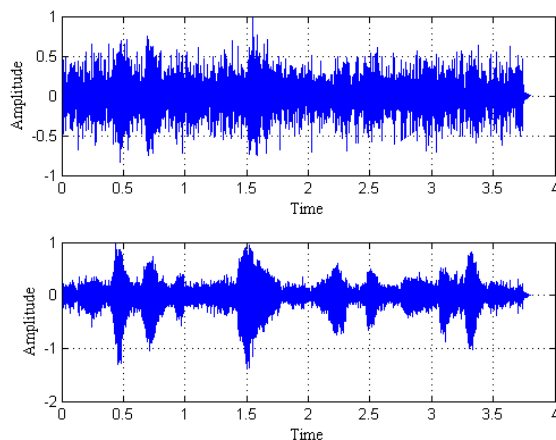
We employ the similarity coefficient as the evaluation standard and further analyze the recovered signal. The experimental results obtained are shown in Fig. 9. When the number of wrong frequency points is less than 10, the wrong frequency points are not adjusted, and the similarity coefficient also can exceed 0.7. Moreover, the difference between the recovered signal and the original signal is minimal, and the chaotic signal used for concealment can still be recovered well. This experiment again verifies that the wrong with individual frequency points has little effect on the separation results. Then the sequence of wrong frequency points is not adjusted. When the number of wrong frequency points reaches 9, the similarity coefficient starts to show a downward trend. When the number of wrong frequency points is 10, the similarity coefficient suddenly drops to approximately 0.5, and the effect of chaotic signal recovery also decreases. Therefore, when the number of wrong frequency points reaches the upper limit tolerated by the algorithm, the cross-correlation permutation for wrong frequency points should be performed. The data analysis reveals that the empirical value for judging the threshold of wrong frequency point is $N_t = 9$.

C. MULTISOURCE SIGNAL EXPERIMENT RESULTS

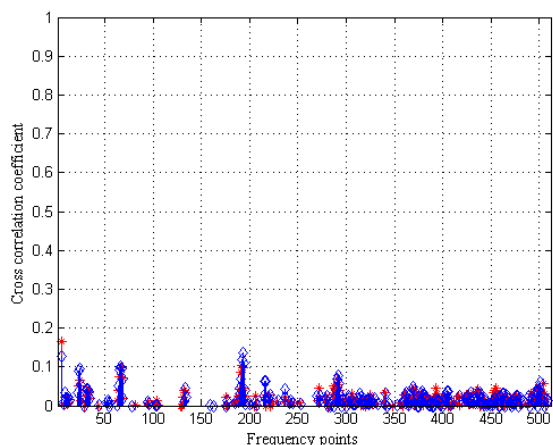
Based on the original experiment, a set of voice signals are randomly selected from TIMIT database and added to the



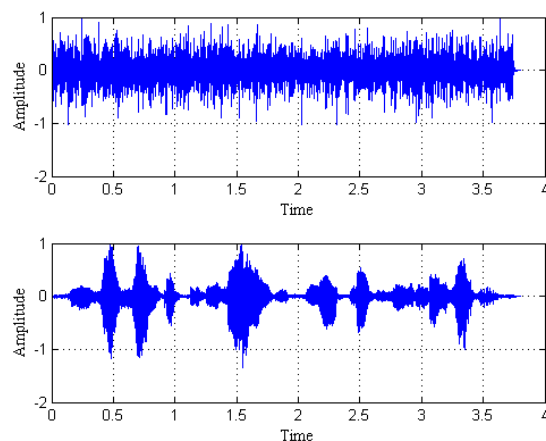
(a) The spectral peak search before sequence adjustment.



(b) The separated signals.



(c) The spectral peak search after sequence adjustment.



(d) The separated signals.

FIGURE 7. Comparison results before and after permutation.

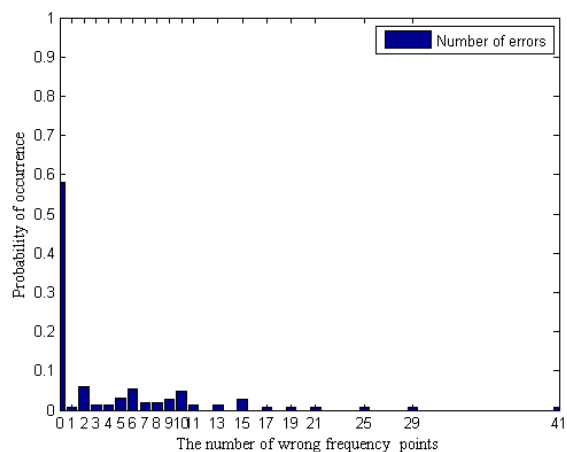


FIGURE 8. The distribution and proportion of the wrong frequency points.

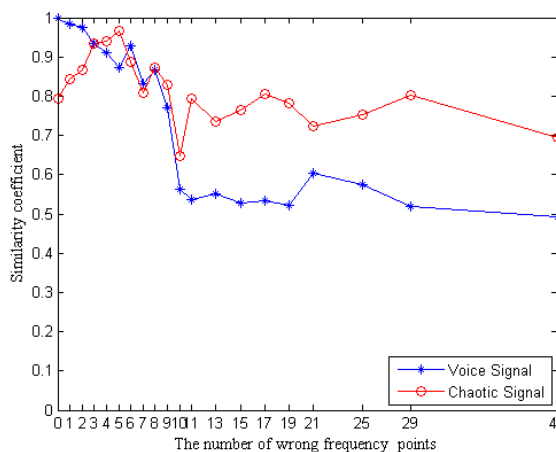


FIGURE 9. Relationship between wrong frequencies and similarity coefficient.

experiment. Three signals are mixed by impulse response filters. The simulation results are shown in Fig. 10.

When the input of the voice signal is increased, the source signal is multiplexed. The CCP algorithm based on the SPSC can adequately recover the source signal. By comparing

the waveforms of the separation signal with the source signal, this algorithm adequately solves the uncertainty problem of frequency-domain ordering and achieves the accurate estimation of the source signal. The separation effect of this algorithm is analyzed quantitatively in TABLE 3, and

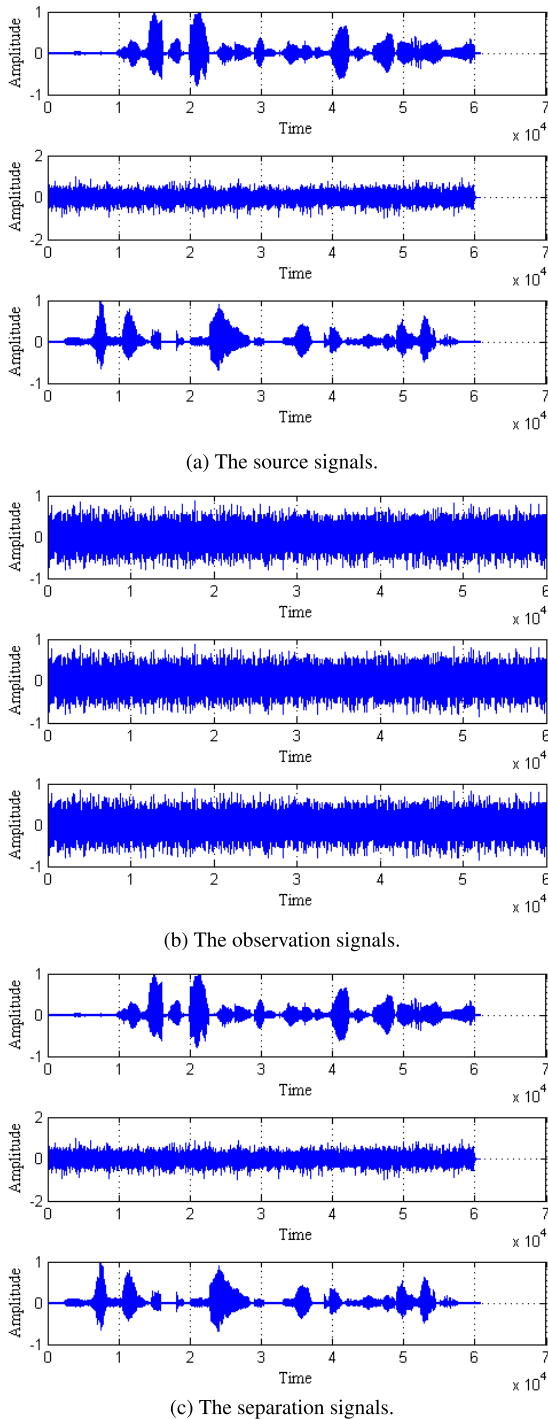


FIGURE 10. Experimental results of multiplexing signals.

Sep1, 2 and 3 represents the separation signals 1, 2 and 3, respectively.

The data in TABLE 3 indicates that the similarity coefficients of Spe.1 and Spe.3 can reach 0.9382 and 0.9366, respectively, and the signal interference ratio is 8.7817 and 8.6126, respectively. The accurate separation of the source voice signal is realized. The results of each evaluation index of Sep.2 were lower, and the similarity coefficient was 0.5252. Due to the large energy of the chaotic signal as a

TABLE 3. Performance evaluation of separated signals.

	Similarity coefficient			SIR		
	Sep.1	Sep.2	Sep.3	Sep.1	Sep.2	Sep.3
Voice 1	0.9382	0.0058	0.0044	8.7817	-6.7693	-2.0696
Voice 2	0.0072	0.0015	0.9366	-3.6444	-7.3105	8.6126
Chaotic	0.0040	0.5066	0.0189	-0.0011	0.1002	0.0011

carrier, the SIR is near 0, but it does not have an effect on the recovery of the voice signal. From the experimental data, the separation signals can be quantitatively analyzed, and the CCP algorithm based on the SPSC is determined to be applicable to multiple signals. It can obtain better separation effect and reduce the computational complexity of multichannel signals.

V. ALGORITHM PERFORMANCE ANALYSIS

The effectiveness of the cross-correlation permutation algorithm based on the spectral peak search counter is verified in Chapter IV. On the basis of these experimental conditions, a random voice signal is selected from the TIMIT standard database as the new input signal. This part will analyze the security, algorithm complexity and robustness of the voice transmission system.

A. SECURITY ANALYSIS

The fractional Chen chaotic system has good confusion and diffusion characteristics, and also has a better ability to resist statistical attacks. The voice signal is hidden in the chaotic signal to ensure the safe transmission. The histogram of the source signal and the observed signal suggests the safety of the transmission system, shown in Fig. 11.

Compared with the histogram of the source signal, as shown in Fig. 11, the histogram distribution of observed signal is uniform. It indicates that the voice signal have high security and are not easy to crack. The time-domain waveform in Fig. 10 (a) can also be proved. It can be seen that the chaotic signal have large energy and have better concealment effect, so as to ensure the confidential transmission of voice information.

B. ALGORITHM COMPLEXITY ANALYSIS

This section analyzes and calculates the time complexity of the key operation steps for the CMBSS algorithm based on the CCP Let the frequency points of STFT be K and the total frames of data be B . Assume that the number of source signals N is the same as the number of observed signals M . One division operation and one square-root operation is counted by 20 and 40 real-valued multiplication operations for a general digital signal processor, respectively. The real-valued multiplication operations required for the main operation process are shown in TABLE 4, where N_p represents the cycle index of the JADE algorithm process.

The number of multiplications required by the algorithm is

$$C_{proposed} = C_1 + C_2 + C_3. \tag{13}$$

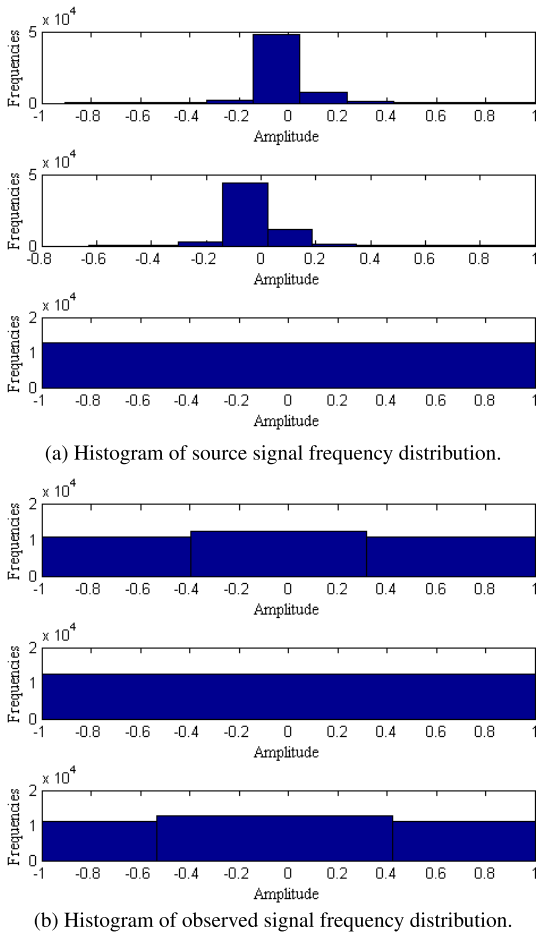


FIGURE 11. The histogram distribution.

TABLE 4. Real-valued multiplication required by the main program.

Step	Computations
1.STFT	$C_1 = 4(K/2)NB\log_2(K)$
2.JADE	$C_2 = 4N_p N^4 B$
3.Permutation	$C_3 = KN^3$

TABLE 5. Algorithm running time.

Method	JADE	Proposed	Ref.[25]
Run Time	3.4 s	9.6 s	10.3 s

To compare the computational complexity of various algorithms, the running time of the frequency-domain permutation algorithm in reference [25] is given. Select the same operating environment. The number of voice signals is $N = 4$, the length of the voice signals is 10 s, and the sampling frequency is $f_s = 8000\text{Hz}$. As shown in TABLE 5, JADE runtime is 3.4 s. Meanwhile, the running time of the algorithm proposed in literature [25] and this paper is 10.3 s and 9.6 s, respectively. Compared with JADE, the calculation of the permutation algorithm cannot be disregarded. Compared with reference [25], this algorithm has a shorter running time.

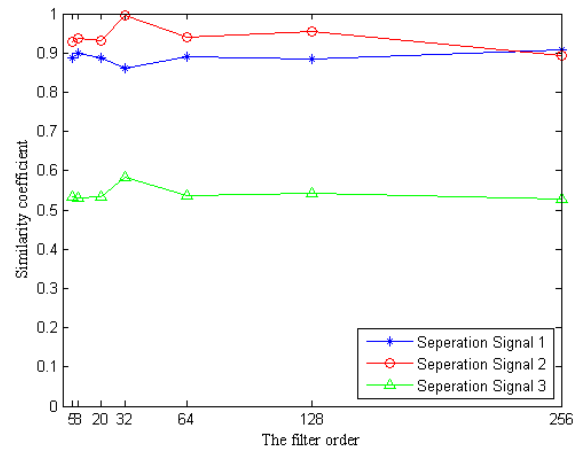


FIGURE 12. The effect of filter order on separation signals.

Based on this analysis of the computational complexity of this algorithm, while ensuring accurate separation of the source voice signal, the computational complexity is reduced. And a new low-complexity frequency-domain permutation method is provided for solving the CMBSS method.

C. ROBUSTNESS ANALYSIS

The cross-correlation permutation algorithm based on the spectral peak search counter not only shows the superiority in reducing the computational complexity but also better robustness. The robustness of this paper specifically refers to the multipath transmission environment. When the order of impulse response filter and the proportional coefficient of direct wave intensity change, the algorithm can still maintain stable performance, but not refers to robustness against noise. Because of the similarity between chaotic signal and Gaussian noise, the relationship between the signal to noise ratio (SNR) and performance cannot be described. To show the robustness in this paper, we change the order of impulse response filter and proportional coefficient of direct wave intensity in the experiment. Other experimental conditions remain constant, and the following simulation experiments are performed.

Change the order of the impulse response filter, that is, change the length of the mixed response during the convolutive mixed process. The set order was 5, 8, 20, 32, 64, 128, and 256. The experimental results were shown in Fig. 12.

The simulation results showed that the order of impulse response filter was increased from the 5th to the 256th order. The similarity coefficients of separation signal 1 and 2 both reached approximately 0.9 and remained stable. It can be seen that as the order of filter increased, the separation results remained almost constant. In other words, when the number of multipath paths is large, the CCP algorithm based on the SPSC accurately recovers the source signal with a lower computation and complexity. It has good robustness.

In multipath transmission, the intensity of direct wave is also an important channel parameter. When the intensity of direct wave is large, the attenuation of the voice signal

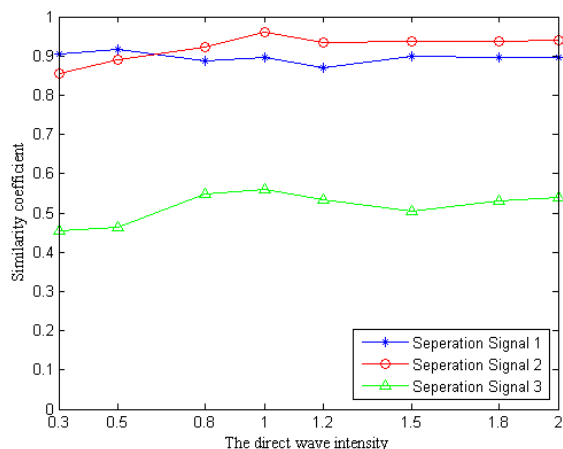


FIGURE 13. The effect of direct wave intensity on separation signals.

during transmission is weak. And the source voice signals are stronger in the observed signals. When the intensity of direct wave is small, the signals are more attenuated in the transmission process. And the intensity of observed signals are relatively weak at the receiver. The intensity of direct wave relative to other path signals in the source signal was changed, and its proportional coefficients were set to 0.3, 0.5, 0.8, 1, 1.2, 1.5, 1.8, and 2, respectively. The experimental results are shown in Fig.13.

Simulation experiments shown that when the proportional coefficient of direct wave intensity is 1, the similarity coefficient of the separation signals 1 and 2 were approximately 0.9. The experimental results did not change significantly with the proportional coefficient of weakening or strengthening the direct wave intensity. Thus, when the signal strength changed in the transmission environment, the CCP algorithm based on the SPSC can still ensure its robustness.

VI. CONCLUSION

In the complex multipath channel transmission environment, chaos signal was introduced as the information carrier to ensure the safe transmission of voice information. Then the mixed signals were separated and extracted at the receiver. The key problem of the convolutive blind source separation in the frequency domain is to solve the permutation uncertainty in the separation signal. Thus, a cross-correlation permutation algorithm based on the spectral peak search counter was proposed in this paper. The cross-correlation energy coefficient was employed as the basis of the frequency point replacement to determine whether the frequency point was wrong and counted by a counter. By established the threshold value of wrong frequency points and compared the count value with it, we could determine whether the wrong frequency points need to be reordered. As this method greatly reduced the calculation of the autocorrelation energy coefficient in the ordering, it can effectively reduce the total calculation complexity. Simulation experiments showed that this permutation algorithm could better solve the permutation problem of multipath voice transmission, and achieve

accurate estimation of voice signals. In future work, the application of this algorithm will be extended to study the noisy and underdetermined model, so as to solve the practical problem that the number of microphones is less than the number of source signals and the transmission environment contains noise.

REFERENCES

- [1] H. I. Shahadi, "Covert communication model for speech signals based on an indirect and adaptive encryption technique," *Comput. Electr. Eng.*, vol. 68, pp. 425–436, May 2018.
- [2] D. Qi, N. Longmei, and X. Jinfu, "A speech privacy protection method based on sound masking and speech corpus," *Procedia Comput. Sci.*, vol. 131, pp. 1269–1274, May 2018.
- [3] C. Ntantogian, E. Veroni, G. Karopoulos, and C. Xenakis, "A survey of voice and communication protection solutions against wiretapping," *Comput. Electr. Eng.*, vol. 77, pp. 163–178, Jul. 2019.
- [4] X.-L. Cao, W.-H. Jiang, and F. Tong, "Time reversal MFSK acoustic communication in underwater channel with large multipath spread," *Ocean Eng.*, vol. 152, pp. 203–209, Mar. 2018.
- [5] S. Leglaive, R. Badeau, and G. Richard, "Student's t source and mixing models for multichannel audio source separation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 6, pp. 1154–1168, Jun. 2018.
- [6] F. I. Abro, F. Rauf, U. R. Mobeen, B. S. Chowdhry, and M. Rajarajan, "Towards security of GSM voice communication," in *Proc. 18th Global Conf. Wire. Optic. Commun.*, Madrid, Spain, vol. 108, 2019, pp. 1933–1955.
- [7] Y.-J. Chen, H.-G. Chou, W.-J. Wang, S.-H. Tsai, K. Tanaka, H. O. Wang, and K.-C. Wang, "A polynomial-fuzzy-model-based synchronization methodology for the multi-scroll chen chaotic secure communication system," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103251.
- [8] C.-T. Leung and W.-C. Siu, "A general contrast function based blind source separation method for convolutively mixed independent sources," *Signal Process.*, vol. 87, no. 1, pp. 107–123, Jan. 2007.
- [9] K. Zhang and L.-W. Chan, "Convolutive blind source separation by efficient blind deconvolution and minimal filter distortion," *Neurocomputing*, vol. 73, nos. 13–15, pp. 2580–2588, Aug. 2010.
- [10] H. Zhang, G. Wang, P. Cai, Z. Wu, and S. Ding, "A fast blind source separation algorithm based on the temporal structure of signals," *Neurocomputing*, vol. 139, pp. 261–271, Sep. 2014.
- [11] T. Mei, A. Mertins, F. Yin, J. Xi, and J. F. Chicharo, "Blind source separation for convolutive mixtures based on the joint diagonalization of power spectral density matrices," *Signal Process.*, vol. 88, no. 8, pp. 1990–2007, Aug. 2008.
- [12] M. Pinchas, "A new efficient expression for the conditional expectation of the blind adaptive deconvolution problem valid for the entire range of Signal-to-Noise ratio," *Entropy*, vol. 21, no. 1, p. 72, 2019.
- [13] C. Mateo and J. A. Talavera, "Short-time Fourier transform with the window size fixed in the frequency domain," *Digit. Signal Process.*, vol. 77, pp. 13–21, Jun. 2018.
- [14] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, Apr. 1994.
- [15] Y. Xie, K. Xie, and S. Xie, "Underdetermined convolutive blind separation of sources integrating tensor factorization and expectation maximization," *Digit. Signal Process.*, vol. 87, pp. 145–154, Apr. 2019.
- [16] M. Z. Ikram and D. R. Morgan, "Permutation inconsistency in blind speech separation: Investigation and solutions," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 1, pp. 1–13, Jan. 2005.
- [17] Q. Fu, B. Jing, and P. He, "Blind DOA estimation in a reverberant environment based on hybrid initialized multichannel deep 2-D convolutional NMF with feedback mechanism," *IEEE Access*, vol. 7, pp. 179679–179689, 2019.
- [18] N. Murata, S. Ikeda, and A. Ziehe, "An approach to blind source separation based on temporal structure of speech signals," *Neurocomputing*, vol. 41, nos. 1–4, pp. 1–24, Oct. 2001.
- [19] L. Zhang, Y. Jiang, and L. He, "Bidirectional coupled noise sources separation over the same frequency band using convolutive blind separation method," *Appl. Acoust.*, vol. 74, no. 3, pp. 301–306, Mar. 2013.

- [20] I. Durán-Díaz, A. Sarmiento, S. Cruces, and P. Aguilera, "Solving permutations in frequency-domain for blind separation of an arbitrary number of speech sources," *J. Acoust. Soc. Amer.*, vol. 131, no. 2, pp. EL139–EL144, Feb. 2012.
- [21] L. Wang, H. Ding, and F. Yin, "A region-growing permutation alignment approach in frequency-domain blind source separation of speech mixtures," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 3, pp. 549–557, Mar. 2011.
- [22] X.-L. Bo, Y.-G. He, and B.-Q. Yin, "Blind source separation sorting algorithm in frequency domain based on influence factor," *Acta Electronica Sinica., China*, vol. 42, no. 2, pp. 360–365, Feb. 2014.
- [23] X.-D. Ou, T.-Q. Zhang, and Z.-H. Yan, "Blind source separation for speech convolution based on multi-band energy-correlated sorting," *Appl. Res. Comput., China*, vol. 295, pp. 207–211, May 2016.
- [24] W. Cheng, Z. Jia, X. Chen, and L. Gao, "Convulsive blind source separation in frequency domain with kurtosis maximization by modified conjugate gradient," *Mech. Syst. Signal Process.*, vol. 134, Dec. 2019, Art. no. 106331.
- [25] F. Kang, F. Yang, and J. Yang, "A low-complexity permutation alignment method for frequency-domain blind source separation," *Speech Commun.*, vol. 115, pp. 88–94, Dec. 2019.
- [26] Ö. Morgül and M. Feki, "A chaotic masking scheme by using synchronized chaotic systems," *Phys. Lett. A*, vol. 251, no. 3, pp. 169–176, Jan. 1999.
- [27] S. Çiçek, A. Ferikoğlu, and I. Pehlivan, "A new 3D chaotic system: Dynamical analysis, electronic circuit design, active control synchronization and chaotic masking communication application," *Optik*, vol. 127, no. 8, pp. 4024–4030, Apr. 2016.
- [28] E. E. Mahmoud and B. H. AL-Harhi, "A phenomenal form of complex synchronization and chaotic masking communication between two identical chaotic complex nonlinear structures with unknown parameters," *Results Phys.*, vol. 14, Sep. 2019, Art. no. 102452.
- [29] W.-J. Zeng, X.-L. Li, H. Zou, and X.-D. Zhang, "Near-field multiple source localization using joint diagonalization," *Signal Process.*, vol. 89, no. 2, pp. 232–238, Feb. 2009.
- [30] T. Jan, W. Wang, and D. Wang, "A multistage approach to blind separation of convulsive speech mixtures," *Speech Commun.*, vol. 53, no. 4, pp. 524–539, Apr. 2011.
- [31] M. M. Asheghan, M. T. H. Beheshti, and M. S. Tavazoei, "Robust synchronization of perturbed Chen's fractional-order chaotic systems," *Commun. Nonlinear Sci.*, vol. 16, no. 2, pp. 1044–1051, Feb. 2011.
- [32] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 4, pp. 1462–1469, Jul. 2006.
- [33] L. Wang, "Multi-band multi-centroid clustering based permutation alignment for frequency-domain blind speech separation," *Digit. Signal Process.*, vol. 31, pp. 79–92, Aug. 2014.



JIAYIN YU was born in Heilongjiang. She received the B.E. degree from the School of Electronic Engineering, Heilongjiang University, Harbin, China, in 2018, where she is currently pursuing the M.E. degree in information and communication engineering. Her current research interests include secure communication and noise analysis during information transmission.



CHAO LI was born in Anhui. He received the B.E. degree from the School of Electrical and Mechanical Engineering, Chizhou College, China, in 2018. He is currently pursuing the M.E. degree in information and communication engineering with Heilongjiang University, Harbin. His research interests include blind source separation and secret communication.



MENGA SHI received the B.E. degree from the Baoji University of Arts and Sciences, China, in 2019. She is currently pursuing the M.E. degree in information and communication engineering with Heilongjiang University. Her main research interests include chaotic secure communication and image encryption based on compressive sensing and chaos.



ERFU WANG was born in Harbin, Heilongjiang, in 1980. She received the B.E. degree in engineering from the School of Communication Engineering, Jilin University, in 2003, and the M.E. and D.E. degrees in information and communication engineering from the Harbin Institute of Technology, in 2005 and 2009, respectively.

Since 2009, she has been teaching in communication engineering with the School of Electronic Engineering, Heilongjiang University, China. In 2010, she entered the Post-doctoral Mobile Station in microelectronics and solid electronics. She is currently a Professor and a master's Tutor. She is also the Chairman of the Department of Communication Engineering. She has authored a monograph. She holds one invention patent. She has published more than 70 articles. Her research interests include chaotic secure communication, blind signal separation, and interference protection.

• • •



SHIYU GUO was born in Liaoning. She received the B.E. degree from the School of Electronic Engineering, Heilongjiang University, Harbin, China, in 2017, where she is currently pursuing the M.E. degree in information and communication engineering. Her research interests include blind source separation, chaotic marking, and voice communication.