The Digital Chaos Cover Transport and Blind Extraction of Speech Signal

Xinwu Chen, Yaqin Xie, and Erfu Wang(✉)

Key Laboratory of Electronics Engineering College, Heilongjiang University, Xuefu Road 74, Harbin 150080, China
Cxw808@qq.com, 648427372@qq.com, efwang_612@163.com

Abstract. With its nonsense, non-detection and robustness, chaotic security technology is more widely used than cryptography in the field of secure communication. In this paper, under the background of digital era, wavelet transform is used to analyze the time-frequency energy concentration of Henon chaotic signal and speech signal, and with the Henon chaotic signal as carrier, the speech signal is hidden, which has important theoretical and practical significance to improve the self-security of the chaotic secure communication system. The speech signal, which chaos is hidden, is transmitted confidentially and it is effectively made to extract blindly at the receiving end. Then similarity coefficient is compared and analyzed under different SNR, which to verify the validity of the algorithm.

Keywords: Henon chaotic · Speech signal · Wavelet transform
Blind extraction · Masking

1 Introduction

Under the background of increasingly complex communication environment, information has become one of the most important strategic resources in today's society. In order to make the eavesdropper can not intercept the real and effective information, it is necessary to take the secure transmission of the speech signal. Meanwhile, the blind source separation technique could be used to extract the speech signal in the case of uncertain channel situation. The key point to secure the speech signal is the carrier signal which we take. In recent years, the emergence of digital chaos has provided an effective means for speech information hiding.

Compared with the traditional analog chaotic system, digital chaotic system [1–3] retaining the excellent characteristics of analog chaotic system, based on this, it strengthen the security and reliability of the system, improve the ability of anti-channel interference of chaotic system and the distortion of channel, and then improve the confidentiality and robustness of chaotic communication system. Under this background, based on the chaos signal processing, it has become a hot spot in the field of chaotic secure communication, and it has presented a more and more obvious intersection and fusion [4–6]. The application of the chaotic characteristics in the field of secure communication has become an emerging direction of research in recent years.

X. Gu et al. (Eds.): MLICOM 2017, Part II, LNICST 227, pp. 612–621, 2018.
https://doi.org/10.1007/978-3-319-73447-7_66

Scholars between home and abroad have proposed many projects such as chaotic masking method, chaos switching method, chaotic modulation and other programs.

There are many scholars at home and abroad studied on the blind separation [7–11]. Some researchers took advantage of ICA algorithm to separate two different sounds from a non-accompaniment chorus recording, and realized blind separation in the background of multiple linearly mixed chaotic signals. Then some scholars used the geometric property of chaotic attractor and realized the separation of weak signal and chaotic interference by means of the concept of differential manifold tangent space. However, the above researches are based on analog signals. Neither do they take any advantage of digital technology, nor do they take the hidden and secure transmission of speech signals into account in the chaotic context.

2 Time Frequency Characteristics of Speech and Chaos

Chaos signal is easy to generate, and it possesses some characteristics, such as randomness and wide spectrum. So chaos signal is hard to decipher. Therefore, in this paper, with chaotic system as carrier, speech signal is hidden in chaos signal, which help information transmit confidentially. The Henon chaotic system is the most used to generate pseudo-random number sequence, and the theoretical basis is the sensitivity of chaotic dynamical systems to initial values and parameters.

The dynamic equation of Henon map is

$$\begin{cases} x_{n+1} = 1 + by_n - ax_n^2 \\ y_{n+1} = x_n \end{cases}. \tag{1}$$

Henon chaotic system exhibits different states with different values of parameter x, y. In this paper, assuming initial value of the system is $x_0 = 0.4, y_0 = 0.4$, When $a = 1.4, b = 0.3$, the time domain waveform of Henon chaotic system could be gotten, which is shown in Fig. 1.

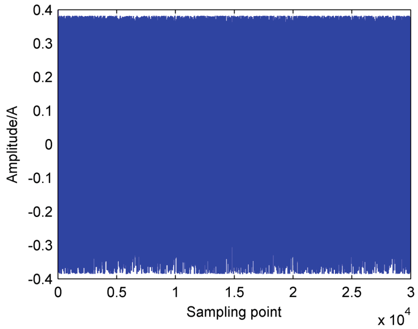

Fig. 1. Time domain waveform of Henon chaotic

In order to hide speech signal in chaos signal successfully, wavelet transform should be used to analyze the time-frequency characteristic of the chaotic signal. In Fig. 2, the energy distribution of the Henon chaotic system is relatively uniform. In the frequency range of 0–4000 Hz, the energy intensity is mainly about 0.4 J, but the speech signal is a small signal, and its energy and amplitude are relatively low, so the Henon chaotic system can be used as carrier to realize the concealment of the speech signal.

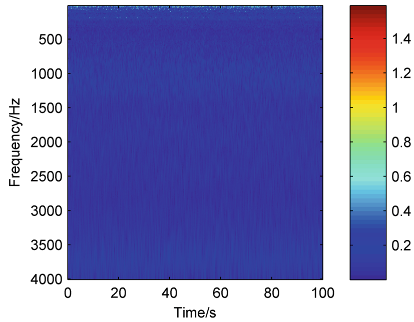

Fig. 2. Wavelet analysis of Henon

The speech signal could be gotten in our daily life. But in order to eliminate the interference of noise and ensure the validity and comparability of this algorithm, the voice bank, which embodies vowel phoneme, is selected. This paper takes the SA2 in the TIMIT voice bank as a hidden signal. As shown, the Fig. 3 is the waveform of SA2, while the Fig. 4 represents the time-frequency distribution of wavelet transform of SA2.

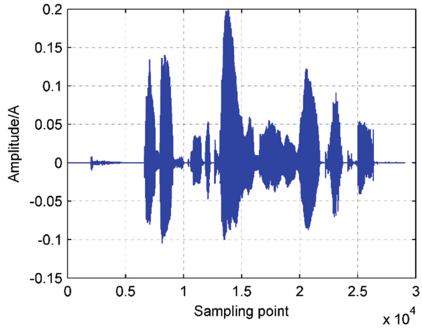

Fig. 3. SA2 speech signal waveform

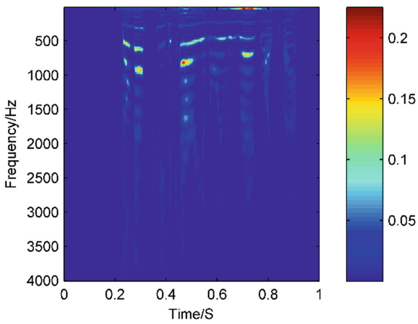

Fig. 4. Wavelet analysis of SA2 speech signal

From the wavelet analysis of the speech signal in Fig. 4, it could be seen that the power distribution of the speech signal is relatively wide, however, it is mainly concentrated in the frequency range of 500–1000 Hz.

3 System Model Establishment and Algorithm Evaluation Standard

3.1 The Establishment of the Model

In this paper, based on the digital processing, considering the speech signal which hide the chaotic signal which was blindly extract by the channel positive definite system. The system model can be abstracted as shown in Fig. 5:

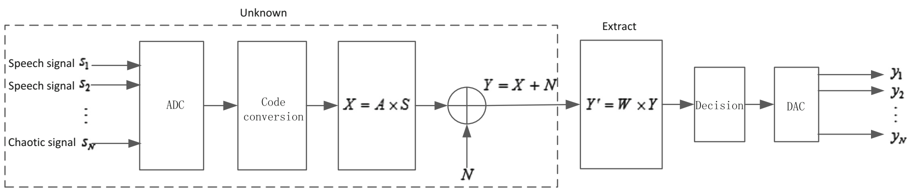

Fig. 5. The system model diagram

The mathematical expression of the model is

$$Y = A \times S + N. \tag{2}$$

Where A is an unknown channel of mixed matrix of $N \times N$, and $S = [s_1(t), s_2(t), \cdots, s_N(t)]^T$ represents N unknown source signal vectors. In order to achieve the purpose of signal secrecy transmission, one of the vectors selected as chaotic signal, while the others are the voice signals. After A/D conversion, the signals are mixed in the noisy channel, and N represents the additive white Gaussian noise in the channel. The mathematical expression of the separation model is

$$Y' = W \times Y = W \times A \times S + W \times N. \tag{3}$$

Where W is the separation matrix obtained by using the blind extraction algorithm at the receiver, and Y represents the observation signal vector at the receiving end.

In this paper, the speech signal as a hidden and extracted the desired signal, which needs digital coding. Then, the chaotic system is used as the hidden carrier, and the speech signal which is encoded hide in the binarized chaotic signal, with the help of the uncertainty of the chaotic signal to make the eavesdropper can not intercept the real information, so as to realize the purpose of confidential communication. The digital coding method of speech signal can be divided into two categories: waveform coding and parameter coding. Among them, Pulse Code Modulation (PCM) is one of the simplest and earliest methods of speech waveform coding. In this paper, the encoding of A law 13 segments is used to digitize the speech signals First, he voice signal is sampled, and set the sampling frequency $f_s = 8\,\text{KHz}$, then the sampled values are quantized. Finally, a set of binary code pulse sequences is used to represent the quantized sampling values. The speech signal and the chaotic signal are mixed after the transmitter sends the A/D conversion to obtain the corresponding symbol sequence, meanwhile adds the additive white Gaussian noise N.

The receiver applies the classic fast ICA algorithm of the blind source separation theory to separate the mixed signal. The separation matrix and the normalized separation matrix expressions are

$$\begin{aligned} W^* &= E\{Xg(W^T X)\} - E\left\{g'(W^T X)\right\}W \\ W &= W^*/\|W^*\|. \end{aligned} \tag{4}$$

Mathematical expectations in (4) must be replaced by their statistical values.

3.2 Evaluation Criteria and Simulation Process of the Algorithm

In this paper, the bit error rate (BER) and the similarity coefficient are used to evaluate the separation performance of the algorithm. The bit error rate is a measure of the accurate indicators which used to measure the data transmission over a specified period of time. The reason why the bit error emerged is due to the fact that in the signal transmission, the decay changes the voltage of the signal, thus the signal is destroyed during the process of transmission. While pulse, transmission equipment failure and other factors caused by noise, alternating current or lightning also will generate errors. The bit error rate is expressed as

$$P_e = \frac{1}{2}\left[1 - \text{erf}\left(\frac{A}{\sqrt{2}\sigma_n}\right)\right]. \tag{5}$$

Where A is the peak of the signal, σ_n is the noise rms value. In the same probability of transmission and under the optimal threshold level, the total bit error rate of the bipolar base band system only depends on the ratio of A and σ_n, regardless of the type of signal being used, at the same probability of transmission and at the optimum threshold level. The correlation coefficient is used to measure the degree of similarity between the separation signal and the source signal. Assuming that S_j denotes the source signal and y_i' represents the separation signal. The mathematical expression of the similarity coefficient is

$$\xi_{ij} = \xi(S_j, y_i') = \frac{\left|\sum_{i=1}^{n} y_i'(t)S_j(t)\right|}{\sqrt{\sum_{i=1}^{n}(y_i')^2(t)\sum_{i=1}^{n}S_j^2(t)}}. \quad (6)$$

Where ξ_{ij} is regarded as a standard for verifying the performance of the separation algorithm. When $\xi_{ij} = 1$ means the two signals are fully correlated. When $\xi_{ij} = 0$ the two signals are completely independent. When ξ_{ij} become closer to 1 which indicates the higher the similarity between the source signal and the separation signal. The better the separation performance two signals are digitized and then mixed separation, after the chaotic system and the voice signal are selected, and then the D/A conversion of the separated symbol sequence is required to obtain the desired signal.

4 Simulation Experiment and Result Analysis

In this paper, the positive definite mixed model is considered, which is under the additive white Gaussian noise. At first, assuming that the signal-to-noise ratio of noise is $\mathrm{SNR} = 18\,\mathrm{dB}$, supposing $N = 2$, and then randomly generating a full-rank $(rank(A) = N)$ channel mixing matrix, which is

$$A = \begin{bmatrix} 0.6706 & 0.1977 \\ 0.5845 & 0.7163 \end{bmatrix}$$

In the processing of speech signal, at first, the simulative speech signal is processed by the encoding of A law 13 segments, which makes it develop into digital 0/1 sequence, and its sampling frequency is $f_s = 8\,\mathrm{KHz}$. Then 8-bit quantification is adopted, and its quantized binary amplitude value is presented by a set of pulse sequences. Therefore the corresponding symbol sequence could be obtained and the output waveform could be finally decoded. In Fig. 6, it is the decoded waveform of SA2 speech signal, which compared with the waveform of the source signal in Fig. 1, and it is shown that the encoding method is successful in encoding the SA2 signal.

Also, chaos signal is digitized. Given the initial value of the Henon chaotic system, and the corresponding real valued sequence will be generated by its iterative equation. Since the mean of the real-valued sequence is μ=0.4980, therefore, 0.5 is selected as

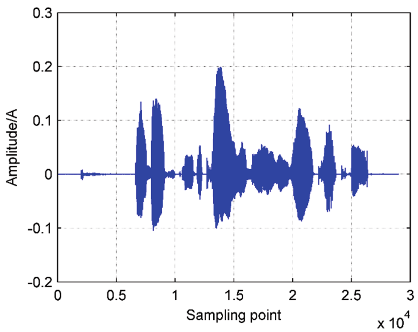

Fig. 6. SA2 speech signal decoding waveform

the criterion for quantization of the sequence, and the real-valued sequence is quantized into a 0/1 sequence.

In order to verify the successful concealment and confidential transmission of the speech signal, wavelet transform is applied to the mixed signal. The Figs. 7 and 8 show the wavelet analysis of mixed first signal and wavelet analysis of mixed second signal respectively. By comparing the wavelet analysis diagram of the chaotic signal and the speech signal (Figs. 2 and 4), the time-frequency distribution characteristic of the

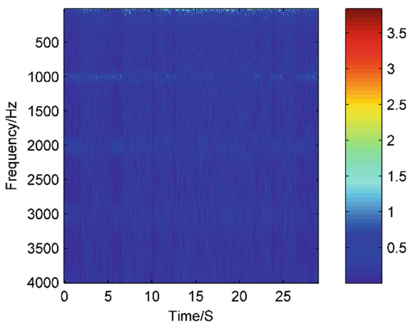

Fig. 7. Wavelet analysis of mixed first signal

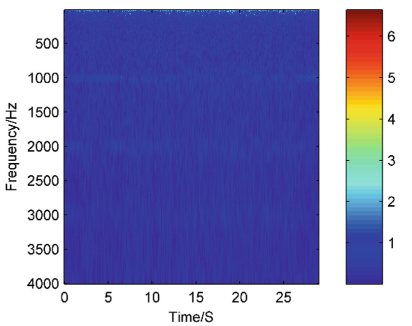

Fig. 8. Wavelet analysis of mixed second signal

speech signal is not displayed on the time-frequency analysis chart of the mixed signal, so the speech signal is successfully hidden.

The next step is to adopt the FastICA algorithm at the receiver to blindly separate the mixed signal. Because of the influence of channel and noise, the sequence after separation is a real-valued sequence, so it is necessary to convert the real-valued sequence into 0/1 sequence, and this simulation selects 0 as the decision threshold. For the probability and polarity of the opposite signal [12], the optimal decision criteria:

$$z(T) \underset{H_2}{\overset{H_1}{\gtrless}} = 0. \tag{7}$$

After the separated signal is converted into a 0/1 sequence by a threshold decision, it is necessary to verify the separation performance of the algorithm. The bit error rate and similarity coefficient are two criteria for evaluating the separation performance. When SNR $= 18\,\text{dB}$, the BER of speech signal and chaotic signal are 6.5894×10^{-5} and 8.9978×10^{-4} respectively; Similarity coefficient matrix is:

$$C = \begin{bmatrix} 0.9783 & 0.0501 \\ 0.0217 & 0.9499 \end{bmatrix}$$

The matrix C can be seen that the first column has a maximum of 0.9783, while the second column is 0.9499, and the remaining numbers are close to zero. So it can be explained that when SNR $= 18\,\text{dB}$, the algorithm realizes the blind separation of the speech signal hidden in the chaotic signal.

Finally, the separated speech coding sequence is converted to parallel output by serializer. And then its PCM decoding is reduced to an analog signal and the wavwrite function is used to save the speech signal into an audio file, which is affected by noise. Although there are some murmurs in the process, but the the content of talking of source language signal would be restored clearly. Figure 9 shows the Waveform of the SA2 speech signal, which has been decoded.

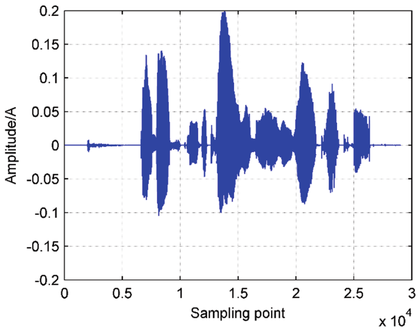

Fig. 9. Waveform of speech signal after mixed separation

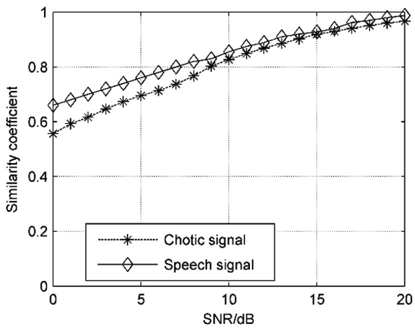

Fig. 10. Similarity coefficient of two kinds of signals with different SNR

Figure 10 is the curves of similarity coefficients of two signals at different SNR. It can be seen that with the increasing SNR, and the similarity coefficient shows a trend of growth. So it can verify the effectiveness of this separation algorithm.

5 Conclusion

Chaotic confidential transmission is widely used in a variety of security communications, and the chaotic hidden transmission of speech signals is one of the most basic problems. In this paper, based on the purpose of confidential transmission, the chaotic system is used to encrypt the speech signal on the basis of digitization. And compared with the traditional analog hybrid method, it is greatly improved from the aspects of confidentiality, robustness and reliability. The receiver adopts the FastICA, the classical algorithm in the blind source separation algorithm, to separate the mixed signal. And the separation performance of the proposed algorithm is verified by analyzing the BER and similarity coefficients, and comparing the auditory effects of the source and the separated signals. The follow-up work will study the hiding and separation of multi-channel speech signals in varied chaotic backgrounds.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (grant 61571181), Postdoctoral Research Foundation of Heilongjiang Province (grant LBH-Q14136), and Graduate Student Innovation Research Project Foundation of Heilongjiang University (grant YJSCX2017-148HLJU).

References

1. Muthukumar, P., Balasubramaniam, P.: Feedback synchronization of the fractional order reverse butterfly-shaped chaotic system and its application to digital cryptography. Nonlinear Dyn. **74**(4), 1169–1181 (2013)
2. Stavrinides, S.G.: Digital chaotic synchronized communication system. J. Eng. Sci. Technol. Rev. **2**(1), 82–86 (2009)

3. Zhou, W.J.: Chaotic digital communication system based on field programmable gate array technology—design and implementation. Acta Physica Sinica **58**(1), 113–119 (2009)
4. Yang, T., Chua, L.O.: Impulsive stabilization for control and synchronization of chaotic systems: theory and application to secure communication. IEEE Trans. Circuits Syst. I Fundam. Theory App. **44**(10), 976–988 (1997)
5. Lin, T.C., Huang, F.Y., Du, Z.: Synchronization of fuzzy modeling chaotic time delay memristor-based Chua's circuits with application to secure communication. Int. J. Fuzzy Syst. **17**(2), 206–214 (2015)
6. Wang, B., Zhong, S.M., Dong, X.C.: On the novel chaotic secure communication scheme design. Commun. Nonlinear Sci. Numer. Simul. **39**, 108–117 (2016)
7. Brown, B.T., Zebrowski, P.M., Spencer, J.P.: Blind separation of instantaneous mixture of sources via the Gaussian mutual information criterion. Signal Process. **81**(4), 855–870 (2015)
8. Douglas, S.C.: Blind Signal Separation and Blind Deconvolution. CRC Press, New York (2002)
9. Mansour, A., Jutten, C., Loubaton, P.: Adaptive subspace algorithm for blind separation of independent sources in convolutive mixture. IEEE Trans. Signal Process. **48**(2), 583–586 (2013)
10. Yeredor, A.: Performance analysis of the strong uncorrelating transformation in blind separation of complex-valued sources. IEEE Trans. Signal Process. **60**(1), 478–483 (2012)
11. Bao, G., Ye, Z., Zhou, Y.: A compressed sensing approach to blind separation of speech mixture based on a two-layer sparsity model. IEEE Trans. Audio Speech Lang. Process. **21** (5), 899–906 (2013)
12. Bernard, S.: Digital Communications Fundamentals and Applications, 2nd edn. Publishing House of Electronics Industry, Beijing (2010)