Journal of Beijing Institute of Technology, 2021, Vol. 30, No. 3

Speech Encryption in Linear Canonical Transform Domain Based on Chaotic Dynamic Modulation

Liyun Xu✉, Tong Zhang, Chao Wen

Abstract: In order to transmit the speech information safely in the channel, a new speech encryption algorithm in linear canonical transform (LCT) domain based on dynamic modulation of chaotic system is proposed. The algorithm first uses a chaotic system to obtain the number of sampling points of the grouped encrypted signal. Then three chaotic systems are used to modulate the corresponding parameters of the LCT, and each group of transform parameters corresponds to a group of encrypted signals. Thus, each group of signals is transformed by LCT with different parameters. Finally, chaotic encryption is performed on the LCT domain spectrum of each group of signals, to realize the overall encryption of the speech signal. The experimental results show that the proposed algorithm is extremely sensitive to the keys and has a larger key space. Compared with the original signal, the waveform and LCT domain spectrum of obtained encrypted signal are distributed more uniformly and have less correlation, which can realize the safe transmission of speech signals.

Keywords: communication security; linear canonical transform; transform domain encryption; chaotic system

1 Introduction

As a part of the nonlinear dynamic system, Chaos seems to be random and regular. Due to the features of sensitivity to the initial state and that can be mapped to the entire space through iterations. And it has been widely used in information encryption [1-3]. However, with the in-depth research of chaotic theory, it can be discovered that the pure chaos-based cryptosystems have limitations in security protection, such as being easy to be broken under selective plaintext attack. To increase the performance of security, cryptosystems on one hand, can use multiple chaotic systems to enlarge the key space, and on the other hand, can combine chaotic systems with other techniques, such as chaotic encryption in transform domain after the signal is transformed [4-8].

Signal encryption methods can be divided into three categories. The signal is encrypted in time domain, frequency domain, or time-frequency domain combination [9-11]. Time domain encryption approaches are mainly based on pure chaotic encryption. The ciphertext sometimes carries part of the original signal and is vulnerable to selected plaintext attacks. Speech recovery is sometimes unsatisfactory and has low security. The frequency domain encryption methods usually perform with Fourier transform, wavelet transform, cosine transform, etc. and then encrypt the signal in the transform domain [12-14]. The transform domain encryption methods increase the complexity of the algorithm to a

Manuscript received May 30, 2021; revised Aug. 10, 2021; accepted Aug. 30, 2021. The associate editor coordinating the review of this manuscript was Dr. Hongmei Zhu. This work was supported by the National Natural Science Foundation of China (No. 61901248) and the Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (No. 2019L0029).

Liyun Xu, Tong Zhang and Chao Wen are with the Institute of Big Data Science and Industry, School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China.

✉ Corresponding author. Email: xuliyun@sxu.edu.cn

DOI: 10.15918/j.jbit1004-0579.2021.038

certain extent. Encryption in the transform domain which is unfamiliar to the cryptanalyst can improve the security performance. The linear canonical transform (LCT) has both scale and angle stretching characteristics, and it is a broader form of many classic transforms, including Fourier transform, fractional Fourier transform, Fresnel transform, etc., which are widely used in signal and image processing [15-17]. More applications based on the LCT also have attracted much attention in recent years [18-21].

Speech communication is one of the most used information communication methods, and its confidential communication is gradually being paid attention to the individual. In this paper, for security speech communication, combining chaotic systems and LCT, a new speech encryption algorithm in LCT domain based on dynamic modulation of chaotic systems is proposed. The LCT parameters are modulated by multiple chaotic systems, and the speech signal is dynamically grouped and encrypted to increase the key space and improve the security performance. While the fast discrete LCT algorithm based on basis decomposition is used in the encryption system to improve computational efficiency.

2 Preliminaries

2.1 Linear Canonical Transform and Its Discrete Algorithm

The LCT is a linear integral transformation with three changing parameters [15]. Its parameters can be expressed as matrix $\boldsymbol{M}=(a,b;c,d)$ and satisfy $|\boldsymbol{M}|=ad-bc=1$, $a,b,c,d\in\mathbb{R}$. The continuous LCT of signal $x(t)$ under parameter $\boldsymbol{M}=(a,b;c,d)$ is expressed as

$$L_{\boldsymbol{M}}(u)=\begin{cases}\int_{-\infty}^{\infty}x(t)K_{\boldsymbol{M}}(u,t)\mathrm{d}t, & b\neq 0\\ \sqrt{d}\exp\left(\mathrm{i}\dfrac{cd}{2}u^2\right)f(du), & b=0\end{cases}\tag{1}$$

where the transform kernel function is

$$K_{\boldsymbol{M}}(u,t)=\sqrt{\frac{1}{\mathrm{i}2\pi b}}\exp\left(\mathrm{i}\frac{a}{2b}t^2-\mathrm{i}\frac{1}{b}ut+\mathrm{i}\frac{d}{2b}u^2\right)\tag{2}$$

LCT is additivity, that is, two consecutive LCTs with parameters $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are equivalent to directly performing LCT with parameters $\boldsymbol{M}_1\cdot\boldsymbol{M}_2$. Therefore, LCT is also reversible, that is, $f(t)=L_{\boldsymbol{M}^{-1}}[L_{\boldsymbol{M}}(u)](t)$.

The LCT discretization fast algorithm is the prerequisite for its wide application. A detailed analysis and comparison of the LCT discretization and fast algorithm research progress are provided in [16]. The three main types of LCT discretization methods are direct discretization, fast basis decomposition algorithms, and operator decomposition discrete algorithms. The basis decomposition-based algorithm is approximate continuous transformation and has display expressions. It also has a computational complexity $O(N\log N)$, which is equivalent to that of Fast Fourier Transform (FFT). So it is much more convenient in applications. The fast discrete LCT algorithm based on basis decomposition will be used in the encryption system to improve computational efficiency.

2.2 Chaotic System

To improve the security of the system, the logistics chaotic model will be used, which is defined by $f_j(\mu_j,x)=\mu_j x(1-x)$, $j=1,2,\cdots,5$. In the encryption algorithm, five independent chaotic systems will be used to act on different parts of the encryption link, and the specific use will be described in the following.

3 Encryption Algorithms in LCT Domain Based on Chaotic Dynamic Modulation

The overall model of the LCT domain encryption and decryption system using multiple chaotic systems for dynamic modulation is shown in Fig.1. In the encryption process, the chaotic systems are used to perform random length modulation segmentation, different LCT parameters for each segment of the signal, and the final encryption processing on the signal in the LCT domain.

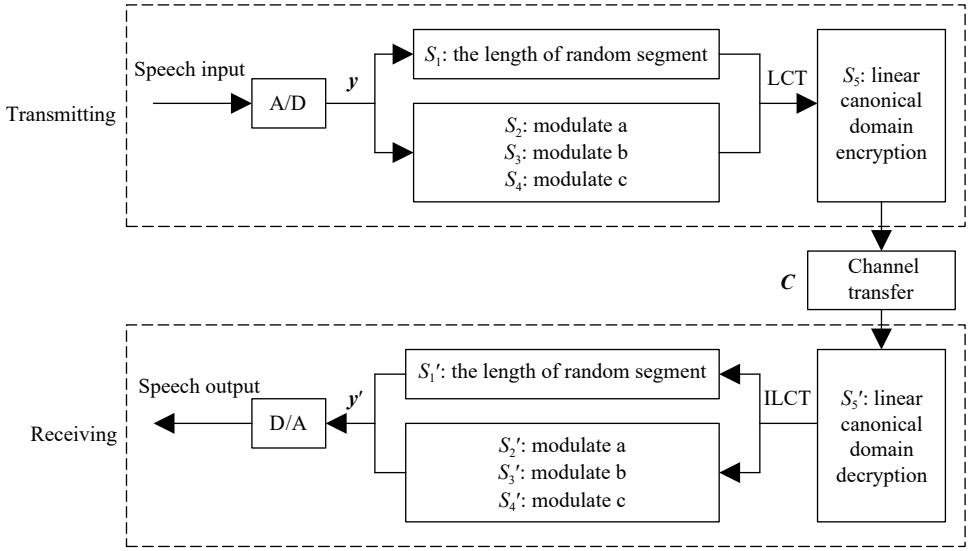

Fig. 1 The system model of speech signal encryption and decryption

The specific encryption process is as follows:

Step 1: Perform analog-to-digital (A/D) conversion on the input speech communication signal to prepare for further processing;

Step 2: Set the initial value parameters S_1 and μ_1 of the chaotic system to generate a random chaotic sequence $\boldsymbol{x_1} = \{x_{1,i}\}$, $x_{1,i} \in (0, 1)$, $i = 1, 2, 3, \cdots$, which is used to modulate the length of the random segment. The segment length is selected as 2^q. To avoid the segment length being too short or too long, the value q of the chaotic sequence modulation is set in a certain range. The q is obtained as

$$q = \lceil 10x_{1,i} \rceil \bmod (5) + 8 \tag{3}$$

And it is obvious that $q \in \{8, 9, \cdots, 13\}$.

Step 3: Set the initial value parameters of the chaotic system S_2, S_3, and S_4, generate random chaotic sequences $\boldsymbol{x}_2 = \{x_{2,i}\}, \boldsymbol{x}_3 = \{x_{3,i}\}, \boldsymbol{x}_4 = \{x_{4,i}\}$, and obtain the i-th randomly modulated LCT parameters $\{a, b, c\}_i = \{x_{2,i}, x_{3,i}, x_{4,i}\}$ respectively;

Step 4: For randomly segmented speech signals, each segment uses a different parameter array to perform LCT to obtain the transform domain spectrum of each segment of the signal;

Step 5: Set the initial value parameters of the chaotic system S_5, and use the obtained fully chaotic random sequence to encrypt the data in each segment of the LCT domain to obtain the ciphertext until all the segmented speech signals are encrypted.

After the encryption, the ciphertext is transmitted through the channel to the receiving end, and the speech information is obtained through decryption processing. The decryption process is the reverse process of encryption. When the chaotic system used in the decryption process is synchronized with all the chaotic systems in the encryption, all the information of the original speech signal can be obtained.

4 Encryption Effect Realization

4.1 Design and Realization of Numerical Experiment

In the encryption algorithm, the five chaotic systems have different modulation effects. In the experiment, the initial values and parameters of five chaotic systems need to be set first, which are shown in Tab.1.

Tab. 1 System parameters

Chaotic system	Initial value	Parameter
S_1	0.1	3.81
S_2	0.2	3.90
S_3	0.3	4.00
S_4	0.4	3.70
S_5	0.5	3.93

The chaotic system S_1 modulates the length of each segmented encrypted signal. To better present encryption and decryption effect, the experiment selects 4 segment continuous signals in the middle part of the speech signal. Based on chaotic system modulation, the numbers of sampling points of the 4 segment signals obtained by the formula (3) are 512, 256, 1024, and 1024 respectively. Fig.2 is the plaintext waveform of the selected speech signal.

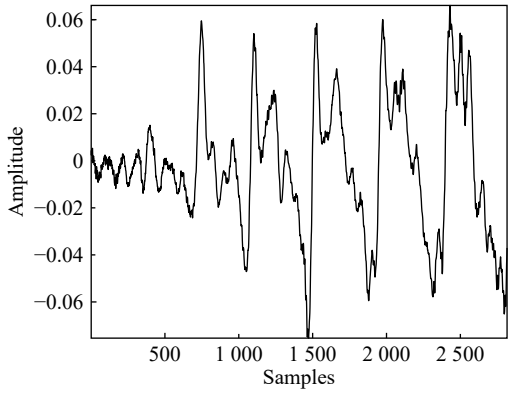

Fig. 2 The original speech signal waveform

The parameters $\{a, b, c\}$ of the LCT of 4 segments of speech signals randomly obtained by chaotic systems S_2,S_3, and S_4 are $\{0.4908, 0.9685, 0.7679\}$, $\{0.9747, 0.1220, 0.6595\}$, $\{0.0963, 0.4284, 0.8308\}$, $\{0.3393, 0.9795, 0.5200\}$, respectively. According to the parameter relationship equation $ad - bc = 1$, the value of parameter d can be obtained. The LCT of the 4 segments of speech signals are obtained under their respective corresponding parameters. The LCT domain spectrogram of the signal is shown in Fig.3.

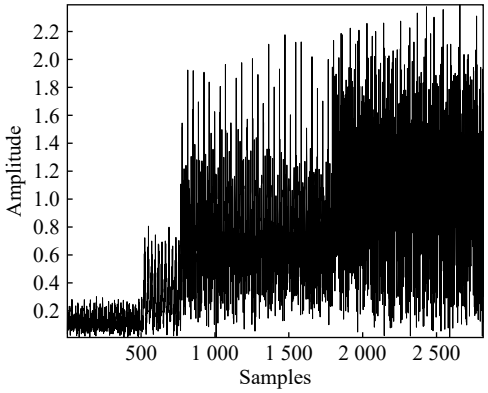

Fig. 3 LCT domain spectrum of speech signal

Under the modulation of the chaotic system S_5, the LCT domain spectrum of the speech signal is encrypted. The encrypted spectrogram is shown in Fig.4, and the corresponding encrypted speech signal waveform is shown in Fig.5. Comparing Fig.4 with Fig.3, Fig.5 and Fig.2, it can be seen that the encrypted LCT domain spectrogram is distributed more uniformly, which is completely different from the shape of the spectrum before encryption. And the encrypted time domain speech signal waveform has been distorted severely, which can hide the original information and achieve a better confidential effect.

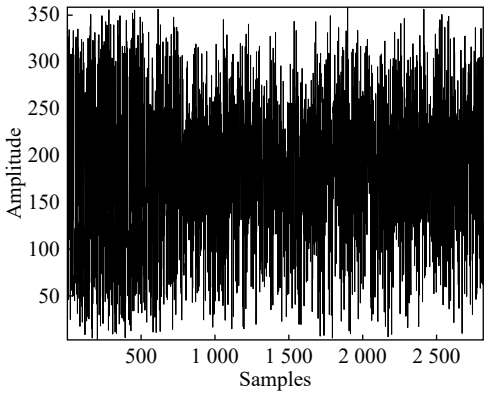

Fig. 4 Encrypted LCT domain spectrogram

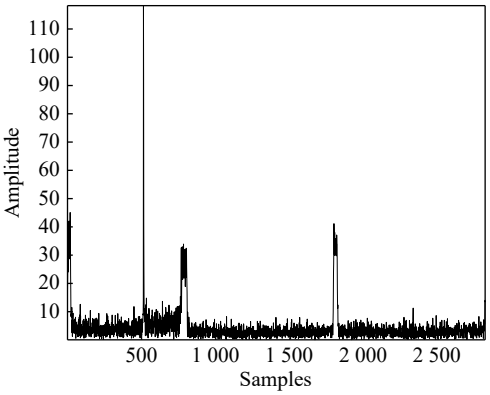

Fig. 5 Encrypted speech signal waveform

After the encryption, the encrypted LCT domain spectrum of the speech signal is transmitted as ciphertext, and it is decrypted at the receiving end. When the secret keys are correct, the correct LCT domain spectrum after decryption and the corresponding speech signal can be obtained, as shown in Fig.6 and Fig.7. When the

keys are error, the decrypted speech signal is shown in Fig.8. Comparing Fig.8 with the original signal waveform in Fig.2, the waveform of the decrypted speech signal changes greatly, and it has a serious distortion effect from an intuitive point of view.

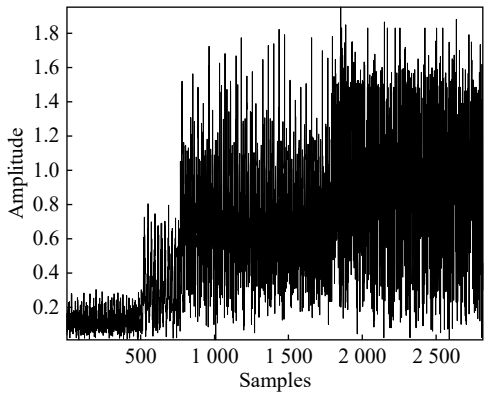

Fig. 6 Decrypted LCT domain spectrogram

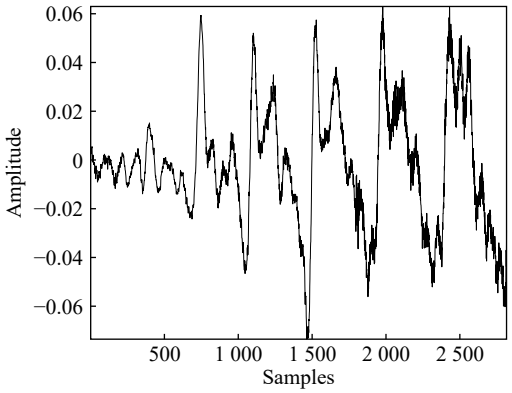

Fig. 7 Speech signal waveform after correct decryption

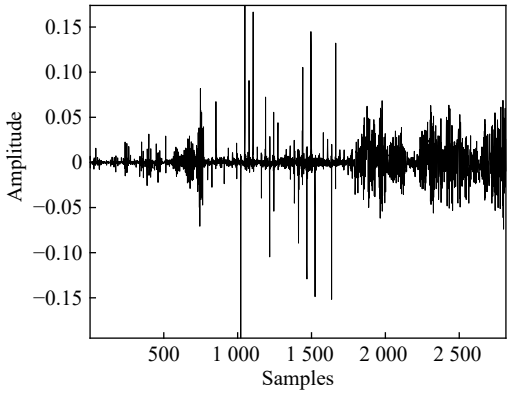

Fig. 8 Speech signal waveform after wrong decryption

4.2 Analysis of Decryption Effect

In order to evaluate the decryption effect of the algorithm objectively, the mean square error (MSE) and the normal average distortion metrics (NADM) are calculated. The calculation formulas are as follows:

$$R_{\text{MSE}} = \frac{1}{N}\sum_{n=1}^{N}|y'(n)-y(n)|^2 \tag{4}$$

$$R_{\text{NADM}} = \frac{\sum_{n=1}^{N}|y'(n)-y(n)|}{\sum_{n=1}^{N}|y(n)|} \tag{5}$$

where $y'(n)$ and $y(n)$ represent the amplitude of the decrypted signal and the original signal respectively, N represents the signal length or the total number of sampling points.

For the correct decryption:

$$R_{\text{MSE}} = 2.24\times10^{-5}, R_{\text{NADM}} = 0.1588$$

For the error decryption is:

$$R_{\text{MSE}} = 5.10\times10^{-3}, R_{\text{NADM}} = 1.2567$$

There is a difference of two orders of magnitude between the MSEs of correct decryption and wrong decryption signals. And there is also a difference of one order of magnitude between the NADMs. This further confirms the security performance of the encryption algorithm from an objective perspective.

5 Algorithm Performance Analyses

5.1 Sensitivity Analysis

As the secret keys, the parameters of the chaotic systems play a vital role in the security of the system. It is very important to analyze the influence of the parameter change on the encryption and decryption. In the experiment, one of the parameters is fluctuated and the others remain unchanged. The fluctuation change of the parameter is $\Delta\mu = 10^{-n}\,(n = 8,\ 9,\ \cdots,\ 17)$, and the MSE of the corresponding decryption signal is calculated. The change curve is shown in Fig.9. When $n \leqslant 15$, the chaotic parameter fluctuates $\Delta\mu \geqslant 10^{-15}$, the decrypted speech signal has a larger value of MSE and a distortion effect. Thus

the encryption method is extremely sensitive to the chaotic system parameters.

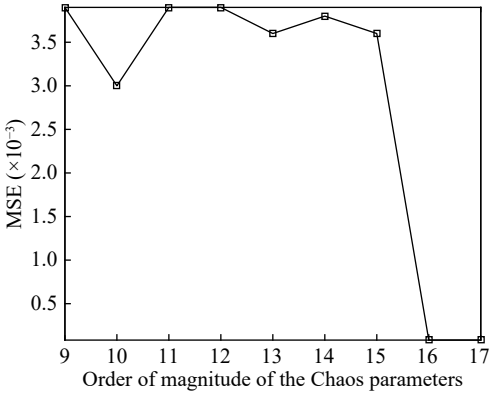

Fig. 9 The relationship between the magnitude changes of the chaotic parameter fluctuation and the MSE

When the order of the chaotic parameters is incorrect, the decryption parameters cannot correspond one-to-one and the decryption result will be distorted seriously. The signal waveform shown in Fig.10 is the decrypted result when the order of parameters μ_2 and μ_3 are swapped, that is, the LCT parameters a and b are swapped. Compared with the original signal waveform in Fig.2, the encrypted signal has been distorted completely and it cannot obtain decryption information correctly. Therefore, the encryption algorithm has certain requirements on the order of the parameters of the chaotic systems, which can further increase the difficulty of cracking the ciphertext and improve the security performance.

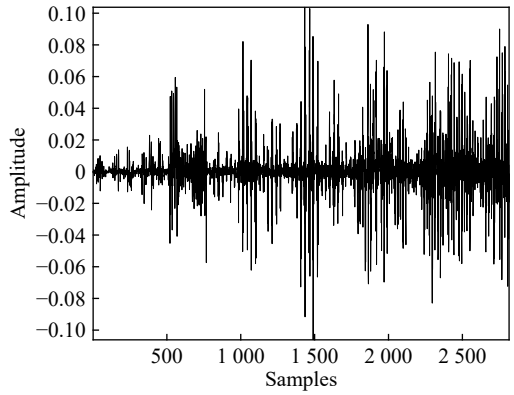

Fig. 10 Decrypted speech signal waveform with the wrong parameter sequence

In the encryption, the parameters of the LCT are obtained through the chaotic system. When the obtained parameters have fluctuations, it will affect the decryption effect of the speech signal. The deviation range of the parameter is set between $[-0.005, 0.005]$. Every time the parameter fluctuation changes by 0.001, the MSE of the decrypted signal is recorded once, and the relationship between the parameter deviation and the MSE is obtained as shown in Fig.11. It is shown that as the parameter deviation increases, the MSE between the decrypted and the original signal increases rapidly, and the degree of distortion of the decrypted speech signal increases. The encryption algorithm is extremely sensitive to LCT parameters.

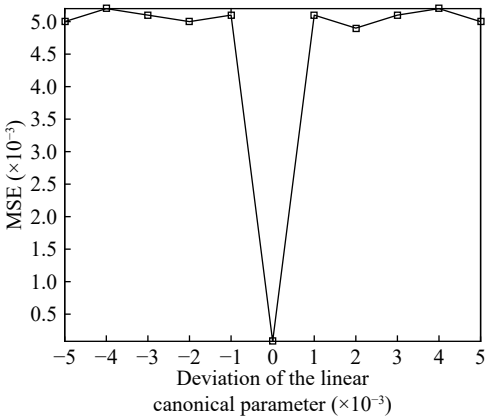

Fig. 11 Relationship between parameter deviation and the MSE

In the encryption, the signal is divided into different length of samples to encrypt based on the chaotic system and the principle of (3). The segment length is different and varies dynamically with the transmission of the signal. When the length of the grouped signal is incorrect in decryption, the recovered signal will be distorted compared with the original signal. Two segment signals with 512 and 256 samples respectively are decrypted with swapped samples after encryption. The result is compared in Fig.12. When the segment length is incorrect, the decrypted signal is totally different from the original signal. Thus the segment length is very sensitive in the encryption algorithm.

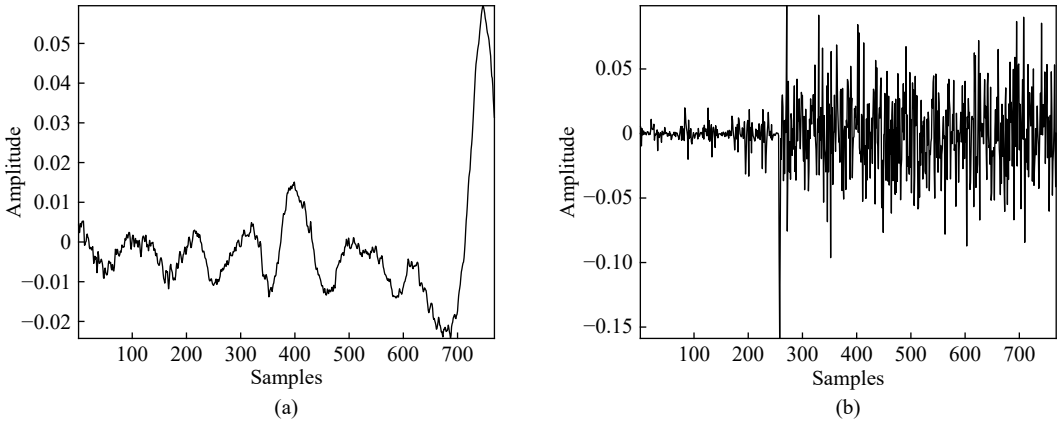

Fig. 12 Comparison of the waveforms: (a) original signal; (b) decryption signal with the incorrect length

5.2 Statistical Analysis

The time domain waveform and LCT domain spectrum corresponding to the original speech signal and the encrypted signal are displayed in histograms, which are shown in Fig. 13 and Fig. 14. By comparing the histograms of the amplitude before and after encryption, it can be seen that the amplitude of the time domain waveform of the signal before encryption is concentrated in the interval $(-0.1, 0.1)$, while the range of the time domain amplitude of the encrypted signal is significantly increased and concentrated in $(-50, 50)$. The waveform amplitude distributions before and after encryption are significantly different.

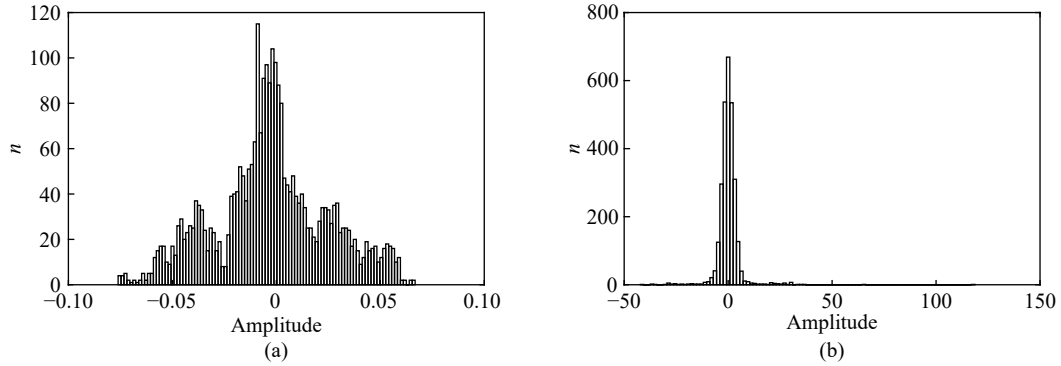

Fig. 13 Comparison of waveform histogram distributions: (a) original signal; (b) encrypted signal

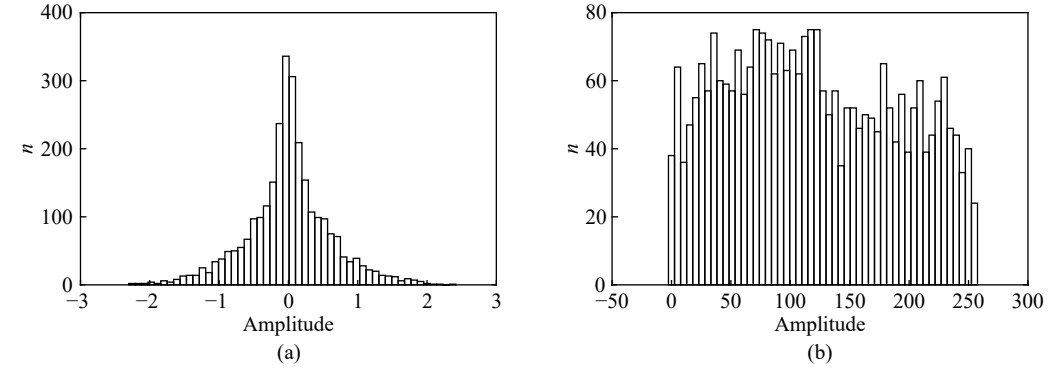

Fig. 14 Comparison of LCT domain spectrum histogram distributions: (a) original signal; (b) encrypted signal

In the encryption, the actual encryption operation is performed in the LCT domain. By comparing the LCT domain spectrum histogram of the signal before and after encryption, it is shown that the LCT domain spectrum before encryption is concentrated in $(-2, 2)$, while the

LCT domain spectrum range after encryption is significantly increased and distributed uniformly, which can hide effective spectrum information and enhance information security.

Through the comparison of histograms, the time domain waveform and LCT domain spectrum of the speech signal before and after encryption have significant differences. The distribution characteristics of the histogram after encryption show that the encryption algorithm can effectively resist statistical attacks and has better security capabilities.

Usually, speech signals have a certain correlation between adjacent amplitudes, and one of the purposes of encryption is to reduce the correlation. The correlation coefficient can be used to evaluate the correlation between the amplitudes. The correlation coefficient is calculated as

$$\gamma_{xy}=\frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{(\sum_{i=1}^{N}(x_i-\bar{x})^2)(\sum_{i=1}^{N}(y_i-\bar{y})^2)}} \tag{6}$$

where x_i and y_i represent the i-th pair of adjacent amplitudes, $\bar{x}$ and $\bar{y}$ are the mean values, N represents the signal length or the number of sampling points. Fig.15 shows the correlation diagrams of adjacent signals before and after encryption, where the abscissa represents the amplitude of the n-th sampling point of the signal, and the ordinate represents the amplitude of the $n+1$-th sampling point, $n=1,2,\cdots,N-1$. The signal before encryption has a linear correlation, which is no longer significant after encryption.

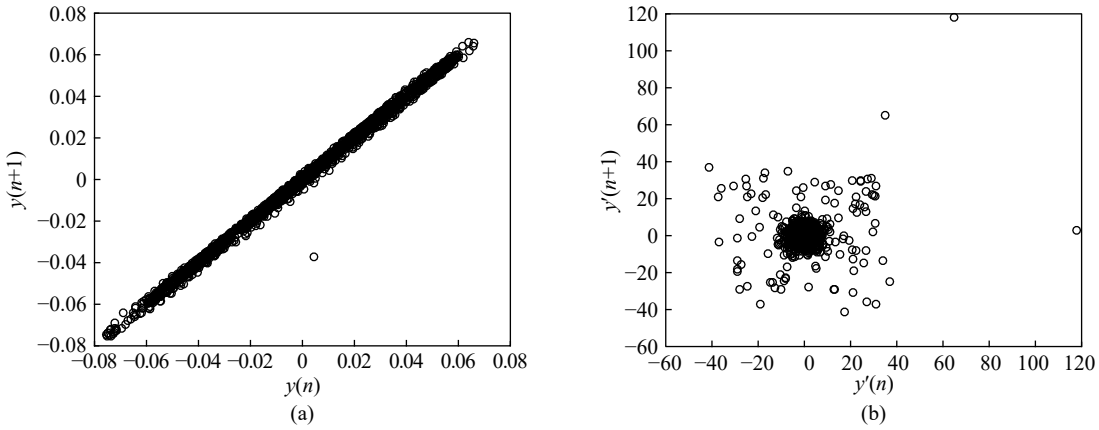

Fig. 15 Comparison of the signal correlation: (a) original signal; (b) encrypted signal

The LCT domain spectrum correlation diagrams before and after encryption are shown in Fig.16, where the horizontal and vertical coordinates correspond to the LCT domain spectrum of

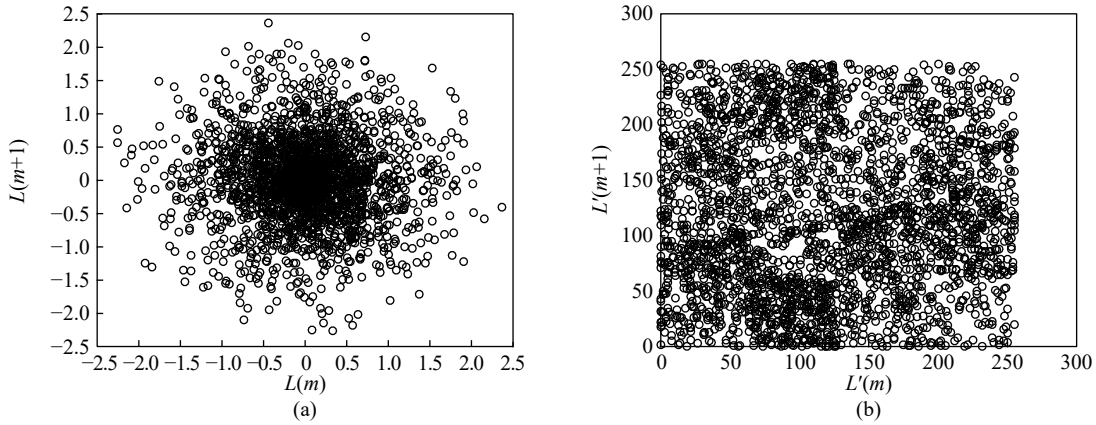

Fig. 16 Comparison of the LCT domain spectrum correlation: (a) original signal; (b) encrypted signal

adjacent points respectively. Compared with the amplitude of the signal in time domain, the LCT domain spectrum of the signal no longer has strong correlation and the characteristics of relatively concentrated distribution, while the correlation of the LCT domain spectrum after encryption shows a more significant uniform random distribution state.

The correlation analyses and comparisons show that, whether in time domain or in LCT domain, the correlation before and after encryption has significant changes. It can result in that the statistical information of the speech can be fully hidden after encryption. Thus it can resist statistical attacks more effectively.

5.3 Key Space Analysis

The keys of the encryption algorithm include the initial value x and parameter μ of 5 chaotic systems. The value range of the state value x is generally in $(-1,1)$. When double-precision floating-point data accurate to 15 digits after the decimal point is used, the key space can reach 2×10^{15}. According to the sensitivity analysis of the chaotic system parameters, it can be known that when the value of one of the parameters changes more than 10^{-16}, or the order of the chaotic system is incorrect, the encrypted signal cannot be decrypted correctly. Therefore, the entire cascaded chaotic system makes the space of the algorithm larger and can reach to $2 \times 10^{15} \times 5 \times 10^{16\times5} \times 5! = 5! \times 10^{96}$. In addition, the sensitivity of the algorithm to each parameter of the LCT is above 10^{-3}, which will further increase the size of the key space, increase the difficulty of the algorithm to decipher, and improve the security performance.

Compared with Fourier transform or fractional Fourier transform with one parameter, the LCT includes three variable parameters. Thus the proposed speech encryption method based on LCT includes much more variable parameters rather than encrypting only in frequency domain or in fractional domain [11]. As the algorithm is extremely sensitive to these parameters, the key space of the proposed algorithm is much larger obviously than that of the algorithm in [11], which has been analyzed above. The sensitiveness of the parameters and the large key space ensure the security of the encryption method.

6 Conclusion

To realize the encrypted transmission of speech signal, an encryption algorithm in LCT domain was proposed. The algorithm used multiple chaotic systems to modulate the algorithm parameters, one for the grouping of the signal samples, three for the modulation of the parameters of the LCT, and the last for the realization of the dynamic packet encryption in LCT domain. Numerical experiment results verified the effectiveness of the proposed encryption algorithm. The performance analyses showed that the algorithm was extremely sensitive to the keys, could expand the key space, and resisted statistical attacks at the same time. Thus it can achieve secure transmission of speech signal and has a good application prospect.

References:

[1] S. Yu, J. Lu, and C. Li,"Some progresses of chaotic cipher and its applications in multimedia secure communications," *Journal of Electronics & Information Technology*, vol. 38, no. 3, pp. 231-248, 2016.

[2] Z. Cheng, "Research and application of encryption technology based on Chaos theory, " M.S. dissertation, Hunan University, 2011.

[3] K. Mistry, S. Dash, and S. Tallur,"Audio encryption through synchronization of chaotic oscillator circuits: Teaching non-linear dynamics through simple electrical circuits," *American Journal of Physics*, vol. 87, no. 12, pp. 1004-1013, 2019.

[4] N. Zhou, L. Huang, and L. Gong,"Novel quantum image compression and encryption algorithm based on DQWT and 3D hyper-chaotic Henon map," *Quantum Information Processing*, vol. 19, no. 9, pp. 284, 2020.

[5] H. Ye, N. Zhou, and L. Gong,"Multi-image compression-encryption scheme based on quaternion discrete fractional Hartley transform and improved

pixel adaptive diffusion," *Signal Processing*, vol. 175, pp. 107652, 2020.

[6] Q. Wu and M. Wu,"Adaptive and blind audio watermarking algorithm based on chaotic encryption in hybrid domain," *Symmetry-Basel*, vol. 10, no. 7, pp. 284, 2018.

[7] F. Zhao and L. Yao,"Gray image encryption algorithm based on cascade chaostic," *Journal of Huaibei Normal University (Natural Science)*, vol. 38, no. 4, pp. 33-37, 2017.

[8] J. Jin, G. Ma, and J. Wei,"Encryption of fractional order Fourier domain of video using cascaded chaos," *Computer Engineering and Design*, vol. 37, no. 8, pp. 1995-2001, 2016.

[9] P. Sathiyamurthi and S. Ramakrishnan,"Speech encryption algorithm using FFT and 3D-Lorenz–logistic chaotic map," *Multimedia Tools and Applications*, vol. 79, no. 3, pp. 17817-17835, 2020.

[10] L. Zhang and D Wei,"Image watermarking based on matrix decomposition and gyrator transform in invariant integer wavelet domain," *Signal Processing*, vol. 169, pp. 107421, 2019.

[11] J. Jin, C. Chen, and M. Wei,"Audio encryption based on Chaotic modulation DFRFT rotation factor," *Computer Engineering*, vol. 38, no. 12, pp. 95-98, 2012.

[12] Y. Shen, C. Tang, and L. Zhou,"Optical single-channel cryptosystem based on the discrete wavelet transform and the chaotic standard map for multi-image," *Applied Optics*, vol. 59, pp. 9558-9567, 2020.

[13] Y. Zhang, L. Zhang, and Z. Zhong,"Hyperchaotic image encryption using phase-truncated fractional Fourier transform and DNA-level operation," *Optics and Lasers in Engineering*, vol. 143, no. 7, pp. 106626, 2021.

[14] Y. Li, D. Wei, and L. Zhang,"Double-encrypted watermarking algorithm based on cosine transform and fractional Fourier transform in invariant wavelet domain," *Information Sciences*, vol. 551, pp. 205-227, 2020.

[15] T. Xu and B. Li, *Linear canonical transform and its application.* Beijing: Science Press, 2013.

[16] Y. Sun, B. Li, and R. Tao,"Research progress on discretization of linear canonical transform," *Opto-Electronic Engineering*, vol. 45, no. 6, pp. 25-45, 2018.

[17] D. Wei and Y. Li,"Convolution and multichannel sampling for the offset linear canonical transform and their applications," *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 6009-6024, 2019.

[18] J. Wu, W. Liu, and Z. Liu,"Correlated-imaging-based chosen plaintext attack on general cryptosystems composed of linear canonical transforms and phase encodings," *Optics Communications* , vol. 338, pp. 164-167, 2015.

[19] C. Guo, I. Munirang, and J. T. Sheridan,"Phase-retrieval-based attacks on linear-canonical-transform-based DRPE systems," *Applied Optics*, vol. 55, no. 17, pp. 4720-4728, 2016.

[20] Z. Huang, S. Cheng, and L. Gong,"Nonlinear optical multi-image encryption scheme with two-dimensional linear canonical transform," *Optics and Lasers in Engineering*, vol. 124, pp. 105821, 2020.

[21] L. Zhao, J. T. Sheridan, and J. J. Healy,"Unitary algorithm for nonseparable linear canonical transforms applied to iterative phase retrieval," *IEEE Signal Processing Letters*, vol. 24, no. 6, pp. 814-817, 2017.

Liyun Xu received the Ph.D. degree in applied mathematics from Beijing Institute of Technology, China, in 2018. She is currently a lecturer with the Institute of Big Data Science and Industry, Shanxi University, China. Her research interests include signal processing, machine learning and artificial intelligence.

Tong Zhang received her B.S. degree in the School of Economics and Management from Shandong University of Science and Technology, Qingdao, China, in 2020. Now, she is currently pursuing the M.S. degree at Shanxi University. Her main research interests include deep learning, speech enhancement and speech separation.

Chao Wen received the Ph.D. degree from Xidian University, China, in 2017. He is currently a Lecturer with the Institute of Big Data Science and Industry, Shanxi University, China. His research interests include array signal processing, passive sensing and artificial intelligence.