A Modified Algorithm for the Logistic Sequence Based on PCA

CHUNYUAN LIU\textsuperscript{1,2} AND QUN DING\textsuperscript{1}
\textsuperscript{1}Electronic Engineering College, Heilongjiang University, Harbin 150080, China
\textsuperscript{2}Computer and Information Engineering College, Heilongjiang University of Science and Technology, Harbin 150027, China

Corresponding author: Qun Ding (qunding@aliyun.com)

This work was supported only by the National Natural Science Foundation of China under Grant 61471158.

\begin{abstract}
In order to deal with the dynamical degradation of logistic chaotic sequence caused by the structures and finite precision computing, we propose a modified algorithm based on phase space reconstruction and PCA. In the process of phase space reconstruction, it uses Cao theory after determining the delay time by the mutual information function and takes the first principal component as the result. Results of experiments demonstrate that: the algorithm erases the relevance among data points and enhances the randomness and complexity of the sequence.
\end{abstract}

\begin{keywords}
Cao theory, mutual information function, phase space reconstruction, principal component analysis, logistic map.
\end{keywords}

I. INTRODUCTION

The massive exchange of information through networks, especially the internet and mobile-phone networks, has created a great demand for real-time secure transmission. Consequently, many encryption systems have been proposed and the chaos-based approach is a promising direction and has shown some exceptionally good properties in many aspects such as security, complexity, speed and computing power [1]–[4].

In general, chaotic systems possess several properties which make them essential components in constructing cryptosystems:

1) Randomness: Chaotic systems generate a random like chaotic sequence in a deterministic way.
2) Sensitivity: They are unpredictable and strongly nonlinear, highly sensitive to initial conditions and very slight change in the starting points can lead to a radically different behavior.
3) Ergodicity: Each state variable uniformly takes all states in phase space.

These properties allow checking typical Shannon requirements of confusion and diffusion, which offer many opportunities that can be used in constructing effective pseudo-random number generator and various cryptosystems. Matthews [5] is the first who proposed a chaotic-based encryption algorithm. In [6] and [7], algorithms combining AES and chaos encryption are studied and many chaotic theory-based encryption systems for communication have been proposed in [8] and [9]. The control and synchronization of chaos are described in [10] and [11]; since then, various images encryption schemes based on chaotic map have been proposed in [12]–[15].

However, due to the computation precision of software and hardware, the chaotic complexity in real-numbered chaos is often affected by dynamic degradation (collapsing effect) and is quite different from theoretical expectations, especially in low-dimensional space, such as logistic and tent maps. Some encryption systems based on the chaos theory have been attacked and destroyed, especially in image encryption systems [16]–[18]. To be precise, some low-dimensional chaotic maps have been extensively analyzed and even confirmed [19]–[21]. At the same time, the encryption systems based on low-dimensional chaotic maps have some security defects, such as small key spaces, short periods and insensitive initial values [22]–[25]. Finally, some security problems will appear because the orbits of the discrete chaotic system will exhibit degenerate behavior near a certain cycle, especially the encryption system based on logistic and tent maps [26]–[31].

Aiming at the degradation of digital chaotic property caused by finite precision, researchers have provided some countermeasures. The error compensation method proposed by Hu et al. [45] is also an effective way to deal with the degradation of digital chaotic systems since it can drive most
properties of the compensated digital chaotic map to be approximate to the original chaotic ones. However, it seems difficult to extend this method to higher dimensional cases. The perturbation-based schemes can greatly improve the dynamical degradation of digital chaotic systems and have been proved to perform better than other methods [44], [47]. However, the properties of perturbed system are dominated by the perturbing source (like linear shift register). The other is a hybrid model based on the complementary of continuous chaos and digital chaos [46]. However due to the uncertain parameters, it is difficult to maintain long-term stable synchronization, which results in the limitation of dynamical properties. Generally, these methods cannot drive the properties of systems to achieve some levels, which may not be desirable for some applications. Therefore, one should resort to other solutions to ensure the expected properties, such as uniform distribution, higher complexity and ergodicity.

In view of the above questions, we introduce a modified logistic chaotic scheme based on phase space reconstruction theory and principal component analysis theory. The sequences generated by the algorithm have better randomness, high complexity and limited predictable long periodicity. The security of the sequence was evaluated using statistical test suites such as the NIST SP 800-22, as well as the approximate probability density, entropy and the autocorrelation analysis to determine unpredictability.

II. THEORY OF PHASE SPACE RECONSTRUCTION

In order to study the complexity of chaotic sequence, it is necessary to reconstruct phase space of it. The dynamic behavior of system evolution can be expressed without singularity from the evolution trajectory of the point in space. The space formed by the observation value and its delay value is called reconstructed phase space. In 1980, Packard et al. proposed two methods for reconstructing the original dynamical system in phase space using time series [32], the derivative reconstruction and the coordinate delay reconstruction. In principle, both of them can perform phase space reconstruction. When it comes to practical application, without knowing any prior information of a chaotic sequence, and from the perspective of numerical calculation, numerical differentiation is sensitive to errors, so phase space reconstruction of a chaotic time sequence generally uses the coordinate delayed method [33]. The essence of the coordinate delay method is to construct a one-dimensional time series \(x(n)\) into an \(m\)-dimensional phase space vector through the delay time \(\tau\), that is

\[
X(i) = \{x(i), x(i + \tau), \ldots, x(i + (m-1)\tau)\}. \tag{1}
\]

In 1981, Takens et al. proposed the embedding theorem: for the scalar time series \(x(n)\) with infinite, noiseless which generates by a \(d\)-dimensional chaotic attractor. As long as the dimension \(m \geq 2d + 1\), an \(m\)-dimensional embedded space can always be found on the sense of topological invariance [34]. From the theory of Takens, a phase space is topologically equivalent to the original dynamical system, which can be reconstructed from a one-dimensional chaotic time sequence, and the determination, analysis and prediction about the chaotic time sequence are all performed in it. Therefore, phase space reconstruction is the crucial step in the study of chaotic time series. In the process of the reconstruction, the first consideration is the selection of the optimal embedding dimension \(m\) and the delay time \(\tau\).

The delay time \(\tau\) is an important embedding parameter. If the value of \(\tau\) is too small, any two components \(x(i + j\tau)\) and \(x(i + (j+1)\tau)\) in phase space are very closed to each other, so they cannot be distinguished; but, if \(\tau\) is too large, the two coordinates are completely independent in statistic, and the projections of a chaotic attractor’s trajectories in two directions are irrelevant. The delay time can be determined by the autocorrelation function method, the mutual information method and the empirical method. The autocorrelation function method is only the way to describe the degree of a linear correlation among variables, so it is not suitable for the nonlinear dynamic system. In the mutual information function method, the time corresponding to the minimum point is taken as the delay time, which can reflect the general correlation among data points. In this paper, the mutual information method is adopted to obtain the delay time, which is generally accepted as the major method to accurately judge the delay time [33].

The optimal embedding dimension of reconstruction is confirmed by Cao’s method, which is an improved way of the false nearest neighbor method. In this, just one parameter of the delay time \(\tau\) is needed, and it can effectively distinguish between random signals and deterministic signals. The embedding dimension \(m\) can be obtained by a small amount of data. Therefore, this paper adopts the mutual information function method and Cao method to reconstruct phase space.

A. THE SELECTION OF THE DELAY TIME

The time delay is a crucial parameter for phase reconstruction. The function of it is to divide a one-dimensional logistic time series into multiple sub-sequences by \(\tau\). The goal is to reconstruct the spatial structure of the original sequence, while preserving the chaotic characteristics of the original sequence to the maximum extent. In the process of phase space reconstruction, if there is a time delay, multiple sub-sequences are projected to multiple coordinate axes, and the best result of decorrelation is retained, which is the first principal component, and the other parts with poor chaotic properties are discarded. Provided that there is no time delay, then the original sequence can project to a single coordinate axis, but the part with poor chaotic characteristics is also retained.

Fraser and Swinney proposed the mutual information method to judge the nonlinear correlation of a system [35]. That is, the time of the mutual information function’s first local minimum is selected as the delay time. The conjoint system is constructed by two discrete information systems, \(S = \{s_1, s_2, \ldots, s_n\}\) and \(Q = \{q_1, q_2, \ldots, q_n\}\), which is nonlinear. According to the information theory, the average
information of the two systems is called the information entropy, respectively as follows:

\[ H(S) = -\sum_{i=1}^{n} P_i(s_i) \log_2 P_i(s_i), \] (2)

\[ H(Q) = -\sum_{i=1}^{m} P_i(q_i) \log_2 P_i(q_i), \] (3)

where \( P_i(s_i) \) and \( P_i(q_i) \) are the probabilities of the event \( s_i \) and the event \( q_i \) in \( S \) and \( Q \) respectively.

Given \( S \), the information about system \( Q \) is called the mutual information of \( S \) and \( Q \). The formula is as follows

\[ I(Q, S) = H(Q) - H(Q|S), \] (4)

where

\[ H(Q|S) = -\sum_{j} [P_{sq}(s_j, q_j)/P_i(s_j)] \log[P_{sq}(s_j, q_j)/P_i(s_j)]. \] (5)

So

\[ I(Q, S) = \sum_{i} \sum_{j} P_{sq}(s_i, q_j) \log_2 \left[ \frac{P_{sq}(s_i, q_j)}{P_i(s_i) P_i(q_j)} \right], \] (6)

where \( P_{sq}(s_i, q_j) \) is the joint probability density of the event \( s_i \) and \( q_j \).

Define \([s, q] = [x(t), x(t+\tau)]\), that is, \( s \) stands for the time series \( x(t) \) and \( q \) is the time series \( x(t+\tau) \). The delay time is \( \tau \).

So it is obviously that \( I(Q, S) \) is the function about the delay time, denoting as \( I(\tau) \). The value of the known system \( S \) (i.e. \( x(t) \)) represents the deterministic value of the system \( Q \) (i.e. \( x(t+\tau) \)). When \( I(\tau) = 0 \), \( x(t+\tau) \) is completely unpredictable, that is, \( x(t) \) and \( x(t+\tau) \) are uncorrelated. Whereas, the value of \( I(\tau) \) is the local minimum, that means \( x(t) \) and \( x(t+\tau) \) are the most likely to be unrelated. Thus, the first local minimum used to be the optimal delay time.

B. THE OPTIMAL EMBEDDING DIMENSION

The improved FNN (Cao method) was proposed by Cao [36]. The principal advantages of this method as follows:

1) The delay time \( \tau \) is the single parameter needed for the calculation.

2) It can effectively distinguish random signals from deterministic signals.

3) The embedding dimension can be obtained by a small amount of data.

In \( d \)-dimensional phase space, each of the space phase vector \( X(i) = \{x(i), x(i+\tau), \ldots, x(i+(d-1)\tau)\} \) has the nearest point \( X^{NN}(i) \) and the distance is

\[ R_d(i) = \left\| X(i) - X^{NN}(i) \right\|
= \sum_{j=1}^{d} \left| x[i+(j-1)\tau] - X^{NN}[i+(j-1)\tau] \right|. \] (7)

As the dimension of phase space increases from \( d \) to \( d+1 \), and the distance between the two points will change to \( R_{d+1}(i) \), and

\[ R_{d+1}^2(i) = R_d^2(i) + \left| x(i+\tau d) - x^{NN}(i+\tau d) \right|^2. \] (8)

If \( R_{d+1}(i) \) is much larger than \( R_d(i) \), it can be considered that the two points are non-adjacent, which is in the higher-dimensional chaotic attractor becoming adjacent points when projected onto the lower-dimensional orbital, so the two adjacent points are false. Suppose that

\[ a_1(i, d) = \frac{x(i+\tau d) - x^{NN}(i+\tau d)}{R_d(i)}. \] (9)

If \( a_1(i, d) > R_\tau, X^{NN}(i) \) is the false adjacent point of \( X(i) \). The threshold \( R_\tau \) can be selected in \([10, 50]\). Take \( R_d(i) \) into (9), and the equation is as follows:

\[ a_1(i, d) = \frac{x(i+\tau d) - x^{NN}(i+\tau d)}{\|X_d(i) - X^{NN}_d(i)\|}. \] (10)

Cao rewrote the above formula as

\[ a_2(i, d) = \frac{\|X_{d+1}(i) - X^{NN}_{d+1}(i)\|}{\|X_d - X^{NN}_d(i)\|}, \] (11)

where \( X_d(i) \) is the \( i \)-th vector of the \( d \)-dimension phase space and \( X^{NN}_d(i) \) is the nearest adjacent point of it. \( X_{d+1}(i) \) and \( X^{NN}_{d+1}(i) \) are the \( i \)-th vector of the \( d+1 \)-dimension phase space and the nearest adjacent point respectively. Define

\[ E(d) = \frac{1}{N-d\tau} \sum_{i=1}^{N-d\tau} a_2(i, d), \] (12)

\[ E_1(d) = E(d+1)/E(d). \] (13)

If the time series is determined, the embedding dimension exists, that is, \( E_1(d) \) will not be changed till \( d \) is greater than a certain value \( d_0 \). If the time series is random, \( E_1(d) \) will be gradually increased. However, in practical, it is difficult to judge whether \( E_1(d) \) is gradually increasing or stable for finite length sequences, so the supplementary judgment criterion is

\[ E^*(d) = \frac{1}{N-d\tau} \sum_{i=1}^{N-d\tau} \left| x(i+d\tau) - x^{NN}(i+d\tau) \right|, \] (14)

\[ E_2(d) = E^*(d+1)/E^*(d). \] (15)

For the random sequence, there is no correlation between the data, and \( E_2(d) \) will always be 1, but for the non-random one, the correlations depend on the change of the embedding dimension \( d \). So there are always some values \( d \) that make \( E_2(d) \) not equal to 1.
data to extract and compress the information of the random data. Project data from the original \( N \)-dimensional space to \( m \)-dimensional space, in general \( m << N \). Therefore, it can reduce the dimension and remove the correlation of the data, but keep most of the internal information of the input data. This is the purpose of PCA. In other words, PCA transforms a large number of relevant data into a set of unrelated characteristic components. Therefore, it is a typical statistical analysis method in the statistical data analysis, feature extraction and data compression, which has a good application in image processing, face recognition and time series prediction [37], [38].

Because of the limitation of calculation accuracy, the correlation of elements in the sequence will be enhanced, which is generated by the low dimensional chaotic system, and the characteristics of the original chaotic will be reduced, resulting in the short cycle. PCA is suitable for removing the correlation in chaotic systems, which is caused by precision. The improved algorithm of logistic sequence is arranged as follows:

1) Phase space reconstruction: For the sequence \( X \) generated by chaotic system, the mutual information method is used to estimate the delay time \( \tau \), and Cao method is used to determine the optimal embedding dimension, and the one-dimensional data is reconstructed into \( m \)-dimensional phase space \( C \), that is

\[
C = \begin{bmatrix}
X_1 & X_1+\tau & \cdots & X_1+(m-1)\tau \\
X_2 & X_2+\tau & \cdots & X_2+(m-1)\tau \\
\vdots & \vdots & \ddots & \vdots \\
X_n & X_n+\tau & \cdots & X_n+(m-1)\tau
\end{bmatrix}.
\]

(16)

The components of \( C \) are \( X_1 = x(n) \), \( X_2 = x(n + \tau) \), \ldots, \( X_m = x(n + (m - 1)\tau) \).

2) Data standardization: Normalize the reconstructed multidimensional phase space, that is, subtract the mean by columns(centralization). The formula is as follows

\[
X_i \leftarrow X_i - \frac{1}{n} \sum_{i=1}^{n} X_i.
\]

(17)

3) Calculate the covariance matrix \( R_x \), that is

\[
R_x = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^T (X_i - \overline{X}),
\]

(18)

where

\[
\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i.
\]

(19)

4) Solve the eigen equation \( |R - \lambda X_p| = 0 \) of the covariance matrix \( R_x \) and the eigenvalues are \( \lambda_1, \lambda_2, \ldots, \lambda_p \), and \( \lambda_1 > \lambda_2 > \cdots > \lambda_p \). Let the orthonormal eigenvectors be \( u_1, u_2, \ldots, u_p \). In accordance with properties of a covariance matrix, \( R_x \) is a symmetric matrix. The eigenvalues of a symmetric matrix can be obtained. Therefore the orthogonal matrix \( U = (u_1, u_2, \ldots, u_p) \), where \( u_i = (u_{i1}, u_{i2}, \ldots, u_{ip}) \).

5) Transform the components of phase space into principal components.

\[
Y_j = U^T X_j, \quad j = 1, 2, \ldots, m,
\]

(20)

where \( Y_1 \) is the first principal component, and \( Y_2 \) is the second, and so on. In order to obtain the optimal decorrelation results, the first principal component of all principal components is taken as the final result of PCA.

**IV. THE ANALYSES AND TESTS OF THE IMPROVED SCHEME**

The logistic equation of chaotic system is given by

\[
x_{n+1} = \mu x_n (1 + x_n).
\]

(21)

Firstly, the delay time \( \tau \) and the embedding dimension \( m \) were obtained by the mutual information method and Cao method respectively. Fig. 1 shows that the relationship between the mutual information function \( I(\tau) \) and \( \tau \). It is the first local minimum when \( \tau = 15 \). Then the delay time \( \tau = 15 \). In Fig. 2, when \( m \) is greater than 8, both \( E1 \) and \( E2 \) are almost unchanged, the optimal embedding dimension \( m = 8 \). Secondly, we reconstructed phase space to 9895 \( \times \) 8. At last, PCA was carried out. The result of PCA is the first principal component \( y_1 \). It must be beyond the original range \([0, 1]\), because of project onto the axis. To keep the modified sequence in \([0, 1]\), the normalization will be done in practice.

**A. THE DISTRIBUTION OF CHAOTIC SEQUENCE AND THE LARGEST LYAPUNOV EXPONENT**

The bifurcation of the modified chaotic system (\( \mu \in [3, 4] \)), displayed in Fig. 3(a), shows that the improved algorithm not only removes the correlation between data points, but also retains the characteristic of the original chaotic system. In Fig. 3(b), an enlargement of Fig. 3(a), shows that the points of the system have been in chaotic and Lyapunov exponents are positive (\( \mu \in [3, 3.57] \)) which is showed in Fig. 4, just...
the complex is low. So the maximum Lyapunov exponents are all greater than zero in phase space in Fig. 4. After normalizing, in Fig. 5, it is clear that the sequence generated by the modified system is uniformly distributed in phase space without any concentration of points in special region implying a strong chaoticity and good ergodicity, when the initial value equals to 0.7 and $\mu = 3.95$.

**B. AUTOCORRELATION FUNCTION**

The waveform of the autocorrelation function with good performance should be a sharp needle without protruding side lobe, which is beneficial to the accurate detection and recognition of signals.

The random variables of Gaussian white noise sequence at any two different times are not only uncorrelated, but are also statistically independent. Fig. 6 shows the unbiased estimate of the Gaussian white noise’s autocorrelation function. The more it is similar to the autocorrelation waveform of the white Gaussian noise, the more it shows randomness. In this paper, the original sequence and the improved sequence are tested for the autocorrelation. Fig. 7 is the autocorrelation test result of the original sequence, and Fig. 8 is the autocorrelation test before the normalization of the sequence. In Fig. 8, the autocorrelation function of the improved sequence is superior to the original one, and it is more similar to the Gaussian white noise, indicating that this method removes the correlation that caused by the calculation accuracy.

**C. THE POWER SPECTRAL DENSITY**

Spectral analysis is an important method to study vibration and chaos. For the sample function of a random signal, the power spectral density function of $x(t)$ is defined as

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau)e^{-j\omega\tau}d\tau,$$

(22)
where \( R_x(\tau) \) denotes the autocorrelation function of \( x(t) \), that is
\[
R_x(\tau) = E\{x(t), E(t + \tau)\} = \lim_{t \to \infty} \frac{1}{t} \int_{0}^{t} \tilde{x}(t)\tilde{x}(t + \tau)dt,
\]
where \( \tau \) denotes the sample interval. For the periodic signal, the power spectrum peaks only arise at the fundamental frequency and doubling frequency. For quasi-periodic signals, the power spectrum peaks arise at several irreducible fundamental frequencies and superposed frequencies. Both chaotic motion and white noise are aperiodic, and their power spectrum are continuous. Because of its irregularity, Gaussian white noise has a uniform power spectrum in the entire frequency domain and all frequencies have the same energy, so it has a flat power spectrum, as shown in Fig. 9. In Fig. 10, \( x_0 = 0.7, \mu = 3.99, \) the length \( N = 10000 \) and after 10,000 iterations, the power spectrum of the sequence was obtained from (21). Energy of the power spectrum distribute uneven and the reverse peaks arise, so it is a quasi-periodic motion. The power spectrum of the modified sequence is displayed in Fig. 11. It is similar to the frequency spectrum of white Gaussian noise in the frequency domain, which is uniformly distributed in energy. The histogram of the approximate probability density function, displayed in Fig. 12, shows that a perfectly normalized distribution.

### D. THE NIST STATISTICAL TEST

For all 16 tests in the NIST suite, the significance level was set to 1%. If a \( P \)-value > 0.01, the binary sequence is accepted as random with a confidence of 99%; otherwise, it is considered as nonrandom [39] and [40]. To perform this battery of tests, we have generated up to \( 10^6 \) points by the modified scheme, and have converted the floating number obtained from this modified system to binary form. We have assessed successfully the sequence, proving the strong randomness. This implies robustness against many statistical attacks. The modified sequence is random with respect to all the 16 tests of NIST suite (see Table 1).
TABLE 1. The NIST tests for the modified logistic sequences.

| Test name                | P-value1 | Results | P-value2 | Results |
|--------------------------|----------|---------|----------|---------|
| Frequency                | 0.089131 | Success |          |         |
| Block frequency          | 0.431939 | Success |          |         |
| Runs                     | 0.543923 | Success |          |         |
| Long runs of ones        | 0.918779 | Success |          |         |
| Rank                     | 0.294177 | Success |          |         |
| FFT                      | 0.112387 | Success |          |         |
| Non-overlapping template | 0.539347 | Success |          |         |
| Overlapping template     | 0.765868 | Success |          |         |
| Universal                | 0.058371 | Success |          |         |
| Lempel-Ziv complexity    | 0.436011 | Success |          |         |
| Linear complexity        | 0.808638 | Success |          |         |
| Serial                   | 0.111645 | Success | 0.167381 | Success |
| Approximate entropy      | 0.244123 | Success |          |         |
| Cumulative sums forward   | 0.086353 | Success |          |         |
| Cumulative sums reverse   | 0.130366 | Success |          |         |
| Random excursions        |          |         |          |         |
| x=–4                     | 0.982311 | Success |          |         |
| x=–3                     | 0.742740 | Success |          |         |
| x=–2                     | 0.742740 | Success |          |         |
| x=–1                     | 0.534072 | Success |          |         |
| x= 1                     | 0.127602 | Success |          |         |
| x= 2                     | 0.262305 | Success |          |         |
| x= 3                     | 0.109172 | Success |          |         |
| x= 4                     | 0.023118 | Success |          |         |
| x= 9                     | 0.527238 | Success |          |         |
| Random excursions variant|          |         |          |         |
| x=–8                     | 0.681791 | Success |          |         |
| x=–7                     | 0.873988 | Success |          |         |
| x=–6                     | 0.748179 | Success |          |         |
| x=–5                     | 0.911033 | Success |          |         |
| x=–4                     | 0.923817 | Success |          |         |
| x=–3                     | 0.704553 | Success |          |         |
| x=–2                     | 0.300218 | Success |          |         |
| x=–1                     | 0.295996 | Success |          |         |
| x=1                      | 0.269501 | Success |          |         |
| x=2                      | 0.538933 | Success |          |         |
| x=3                      | 0.519752 | Success |          |         |
| x=4                      | 0.254175 | Success |          |         |
| x=5                      | 0.197660 | Success |          |         |
| x=6                      | 0.320782 | Success |          |         |
| x=7                      | 0.490799 | Success |          |         |
| x=8                      | 0.592944 | Success |          |         |

E. THE ENTROPY TEST

Entropy (such as K-S entropy, information entropy (EN), approximate entropy (APEN)) is one of the most important characteristics of randomness. Information entropy is a mathematical theory about data communication and storage. It is proposed by Shannon in 1949 [41]. To calculate the entropy of the modified sequence, we have

$$H(s) = - \sum_{i=0}^{2^M-1} P(s_i) \log_2 P(s_i),$$

(25)

where $P(s_i)$ represents the probability of the symbol $s$ [14], [42]. Information entropy of the original sequence is 7.7350, and the modified of it is 7.9734.

Approximate entropy (APEN) is proposed in Literature [43]. We performed APEN test to the same sequence. APEN value of the modified sequence is 2.5557, but the original is 0.6434. It can be seen that our improved algorithm increases the complexity of the sequence.

V. CONCLUSION

In this paper, we establish a scheme to deal with the influence of finite precision on the dynamical of digital chaotic systems. The improved scheme is based on phase space reconstruction technology composed of mutual information method and Cao method, and PCA. The simulation results show that the improved scheme is simple and easy to operate, and remove the correlation between the original sequence data, which not only increases the complexity of digital chaotic system, but also does not destroy the attractor structure of the original system. The scheme can be applied to any given chaotic system, low or high dimension, and can be applied to pseudo-random coding, secure communication and channel transmission, image encryption and other fields.

REFERENCES

[1] M. S. Baptista, “Cryptography with chaos,” Phys. Lett. A, vol. 240, pp. 50–54, Mar. 1998, doi: 10.1016/s0375-9601(98)00086-3.
[2] W. Yan and Q. Ding, “A new matrix projective synchronization and its application in secure communication,” IEEE Access, vol. 7, pp. 112977–112984, Oct. 2019, doi: 10.1109/ACCESS.2019.2935304.
[3] R. Lozi, “Chaotic pseudo random number generators via ultra weak coupling of chaotic maps and double threshold sampling sequences,” in Proc. ICSSA, Le Havre, France, 2009, pp. 20–24.
[4] E. Y. Xie, C. Li, S. Yu, and J. Lü, “On the crypantalysis of Fridrich’s chaotic image encryption scheme,” Signal Process., vol. 132, pp. 150–154, Mar. 2017, doi: 10.1016/j.sigpro.2016.10.002.
[5] R. Matthews, “On the derivation of a ‘chaotic’ encryption algorithm,” Cryptologia, vol. 13, no. 1, pp. 29–42, Jan. 1989, doi: 10.1080/0161-118991865745.
X.-Y. Wang, Y.-Q. Zhang, and Y.-Y. Zhao, “A novel image encryption algorithm and its implication based on FPGA,” in Proc. NEWCAS, Trois-Rivieres, QC, Canada, Jun. 2014, pp. 217–220.

Y. Kun, Z. Han, and L. Zhaohui, “An improved AES algorithm based on chaos,” in Proc. Int. Conf. Multimedia Inf. Netw. Secure, Wuhan, China, vol. 9, 2009, pp. 3721–3737, Sep. 2011, doi: 10.1109/cnms.2010.12.032.

S. H. Wang, J. Y. Kuang, J. H. Li, Y. L. Luo, H. P. Lu, and G. Hu, “Chaos-based secure communications in a large community,” Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top., vol. 66, no. 6, Dec. 2002, Art. no. 065202, doi: 10.1103/PhysRevE.66.065202.

A. A. Zaher and A. Abu-Rezq, “On the design of chaos-based secure communication systems,” Commun. Nonlinear Sci. Numer. Simul., vol. 16, no. 2, pp. 824–836, Mar. 2011, doi: 10.1016/j.cnsns.2010.12.010.

B. Vaseghi, M. A. Pourmina, and S. Mobayen, “Secure communication in wireless sensor networks based on chaos synchronization using adaptive sliding mode control,” Nonlinear Dyn., vol. 89, no. 3, pp. 1689–1704, May 2017, doi: 10.1007/s11071-017-3543-9.

S. Boccaletti, C. Grebogi, Y. C. Lai, H. Mancini, and D. Maza, “The control of chaos: Theory and applications,” Phys. Rep., vol. 329, no. 3, pp. 103–197, May 2000, doi: 10.1016/s0370-1573(99)00046-4.

C. K. Huang and H. H. Nien, “Multi chaotic systems based pixel shuffle for image encryption,” Opt. Commun., vol. 282, no. 11, pp. 2123–2127, Jun. 2009, doi: 10.1016/j.optcom.2009.02.044.

H. S. Kwo and W. K. S. Tang, “A fast image encryption system based on chaotic maps with finite precision representation,” Chaos, Solitons Fractals, vol. 32, no. 4, pp. 1518–1529, May 2007, doi: 10.1016/j.chaos.2005.11.090.

S. M. Seyedzadeh, B. Norouzi, M. R. Mosavi, and S. Mirzakuchaki, “A novel color image encryption algorithm based on spatial permutation and quantum chaotic map,” Nonlinear Dyn., vol. 81, no. 1–2, pp. 511–529, Mar. 2015, doi: 10.1007/s11071-015-1008-2.

S. Song, S. Dutta, R. Singh, A. Kotal, and S. Palit, “Confusion and fusion of color images with multiple chaotic maps and chaos-based pseudorandom binary number generator,” Nonlinear Dyn., vol. 80, no. 1–2, pp. 615–627, Jan. 2015, doi: 10.1007/s11071-014-1893-8.

G. Jakimoski and L. Kocarev, “Analysis of some recently proposed chaos-based encryption algorithms,” Phys. Lett. A, vol. 291, no. 6, pp. 381–384, Dec. 2001, doi: 10.1016/s0375-9601(01)00771-x.

C. Li, L. Y. Zhang, R. Ou, K.-W. Wong, and S. Shu, “Breaking a novel colour image encryption algorithm based on chaos,” Nonlinear Dyn., vol. 70, no. 4, pp. 2383–2388, Oct. 2012, doi: 10.1007/s11071-012-0626-5.

S. Lian, J. Sun, and Z. Wang, “Security analysis of a chaos-based image encryption algorithm,” Phys. A, Stat. Mech. Appl., vol. 351, nos. 2–4, Jun. 2005, doi: 10.1016/j.physa.2005.01.001.

Z.-H. Guan, F. Huang, and W. Guan, “Chaos-based image encryption algorithm,” Phys. Lett. A, vol. 346, nos. 1–3, pp. 153–157, Oct. 2005, doi: 10.1016/j.physleta.2005.08.006.

X. Tong and M. Cui, “Image encryption scheme based on 3D baker with dynamical compound chaotic sequence cipher generator,” Signal Process., vol. 89, no. 4, pp. 480–491, Apr. 2009, doi: 10.1016/j.sigpro.2008.09.011.

D. Xiao, X. Liao, and P. Wei, “Analysis and improvement of a chaos-based image encryption algorithm,” Chaos, Solitons Fractals, vol. 40, no. 5, pp. 2191–2199, Jun. 2009, doi: 10.1016/j.chaos.2007.10.009.

C. Li, Y. Liu, T. Xie, and M. Z. Q. Chen, “Breaking a novel image encryption scheme based on improved hyperchaotic sequences,” Nonlinear Dyn., vol. 73, no. 3, pp. 2083–2089, Aug. 2013, doi: 10.1007/s11071-013-1024-6.

C. Zhang and K. Sun, “Cryptanalysis and improving a novel color image encryption algorithm using RT-enhanced chaotic tent maps,” IEEE Access, vol. 6, pp. 18759–18770, 2018, doi: 10.1109/ACCESS.2018.2817660.

X.-Y. Wang, Y.-Q. Zhang, and Y.-Y. Zhao, “A novel image encryption scheme based on 2-D logistic map and DNA sequence operations,” Nonlinear Dyn., vol. 82, no. 3, pp. 1269–1280, Oct. 2015, doi: 10.1007/s11071-015-2234-7.

C. Liu, L. Ding, and Q. Ding, “Research about the characteristics of chaotic systems based on multi-scale entropy,” Entropy, vol. 21, no. 7, p. 663, Jul. 2019, doi: 10.3390/e21070663.

C. Dellago and W. G. Hoover, “Finite-precision stationary states at and away from equilibrium,” Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top., vol. 62, no. 5, pp. 6275–6281, Nov. 2000, doi: 10.1103/PhysRevE.62.6275.
CHUNYUAN LIU received the B.S. degree from the Computer and Information Engineering College, Heilongjiang University of Science and Technology, in 2003. She is currently pursuing the Ph.D. degree in information and communication engineering with Heilongjiang University, Harbin. Her research interests include nonlinear dynamics, image encryption, and image compression.

QUN DING received the Ph.D. degree in instrument science and technology from the Harbin Institute of Technology, China. She is currently a Professor with the College of Electronic Engineering, Heilongjiang University, China. She has published two books and over 100 scientific articles in refereed journals and proceedings. Her research interests include nonlinear dynamics and control, chaos pseudo-random sequence generator, and chaotic secure communication.

***