Research on the chaos recognition method based on differential entropy*

Zhang Shu-Qing(张淑清)†, Zhao Yu-Chun(赵玉春), Jia Jian(贾 健), Zhang Li-Guo(张立国), and Shangguan Han-Lu(上官寒露)

Institute of Electrical Engineering, Measurement Technology and Instrumentation Key Lab of Hebei Province, Yanshan University, Qinhuangdao 066004, *China*

(Received 30 September 2009; revised manuscript received 18 November 2009)

Phase space reconstruction is the first step to recognizing the chaos from observed time series. On the basis of differential entropy, this paper introduces an efficient method to estimate the embedding dimension and the time delay simultaneously. The differential entropy is used to characterize the disorder degree of the reconstructed attractor. The minimum value of the differential entropy corresponds to the optimum set of the reconstructed parameters. Simulated experiments show that the original phase space can be effectively reconstructed from time series, and the accuracy of the invariants in phase space reconstruction is greatly improved. It provides a new method for the identification of chaotic signals from time series.

Keywords: chaos, differential entropy, embedding dimension, time delay
PACC: 0545, 0547

1. Introduction

The chaotic phenomenon exists extensively in continuous and discrete systems. Chaotic time series include abundant dynamic information; the accurate identification of chaotic signals contained in a complex time series is crucial in the field of chaotic application. Many developments on chaotic feature extraction from chaotic time series have been obtained, such as: the Poincaré surface of section, the Lyapunov exponent analysis and the fractal dimension method, etc. But chaotic time series analysis and prediction are often carried out in a phase space, so phase space reconstruction is an important subject for processing chaotic time series and their application.[1−3]

According to Takens' embedding theorem, a time series $\{x_i, i = 1, 2, \ldots, n\}$ can be embedded into a reconstructed delay phase space with an embedding dimension and a time delay; a point in the reconstructed delay phase space is

$$X_i = [x_i,\ x_{i+\tau},\ \ldots,\ x_{i+(m-1)\tau}]^{\mathrm{T}}, \qquad (1)$$

where m is the embedding dimension and τ is the time delay. Takens' theorem is valid for an infinite noise-free data set; in this case, the time delay can be selected arbitrarily. However, the observed time series are finite and noisy and an arbitrary choice of the two parameters would influence the quality of the reconstructed phase space. So it is important to select the embedding dimension m and time delay τ, though it is difficult to do so.[4]

For the estimation of embedding dimension, the most commonly used techniques are false nearest neighbours (FNN), singular value analysis and Cao's method.[5] For nonlinear systems, the autocorrelation function and average mutual information (AMI)[6] are often used to determine optimal time delay τ. But all methods are from the view-point of the uncorrelation of m and τ. Some researchers have suggested that m and τ are correlated with each other. Kim *et al.*[7] have proposed an available technique called the C-C method for estimating τ by the correlation integral, and m will be determined simultaneously. Aiming at its limitation, some authors presented[8,9] some improved methods. At present, the achievements on information theory have been introduced to determine the best reconstructed parameters.[10,11] Entropy is one kind of measure to characterize the uncertainty of time series, and the differential entropy can be used to characterize the disorder degree of the attractor in different states.

In this paper, we present a method based on differential entropy. By optimizing the reconstructed

*Project supported by the National Natural Science Foundation of China (Grant Nos. 50775198 and 60102002).

†Corresponding author. E-mail: zhshqyd@yahoo.cn

© 2010 Chinese Physical Society and IOP Publishing Ltd

chaotic attractor, the inconsistency between m and τ is eliminated.

2. AMI/FNN

Average mutual information (AMI)/false nearest neighbours (FNN) is one of the popular techniques to estimate τ and m.

The AMI $I(\tau)$ is defined as

$$I(\tau) = \sum_{x_i, x_{i+\tau}} P(x_i, x_{i+\tau}) \log_2 \frac{P(x_i, x_{i+\tau})}{P(x_i)P(x_{i+\tau})}, \quad (2)$$

where $P(x_i)$ and $P(x_{i+\tau})$ denote probability density functions of x_i and $x_{i+\tau}$ respectively, $P(x_i,x_{i+\tau})$ denotes the joint probability density function of x_i and $x_{i+\tau}$. These probability density functions are calculated from histograms of each time series. When τ becomes large, x_i and $x_{i+\tau}$ become independent, and $I(\tau)$ will tend to be zero. The optimal time delay τ is the first minimum of function $I(\tau)$.

The most popular method for the estimation of m is FNN. Suppose the delay coordinate vector in the m dimension is $X_i = [x_i,\ x_{i+\tau}, \ldots,\ x_{i+(m-1)\tau}]^{\mathrm{T}}$, one vector X_i^{NN} is nearest to the vector X, $X_i^{\mathrm{NN}} = [x_i^{\mathrm{NN}},\ x_{i+\tau}^{\mathrm{NN}},\ \ldots,\ x_{i+(m-1)\tau}^{\mathrm{NN}}]^{\mathrm{T}}$. The Euclidean distance between X_i^{NN} and X_i in the m dimension is ε_i, and the Euclidean distance between X_i^{NN} and X_i in the $m+1$ dimension is ε_i', then the relative distance between X_i^{NN} and X_i in the $m+1$ dimension is

$$\varepsilon = \sqrt{\frac{\varepsilon_i'^2 - \varepsilon_i^2}{\varepsilon_i^2}}. \quad (3)$$

When ε is larger than some heuristic threshold, X_i^{NN} is a false neighbour of X_i. The optimal embedding dimension, which is sufficiently high for reconstructing the attractor, is the dimension in the case where the fraction of false neighbours is zero or at least sufficiently small.

3. Differential entropy method

Suppose the possible value of continuous random variables X is over the field of real numbers, and $f(x)$ is a probability density function of X; it is obvious that $f(x) \geq 0$, and $f(x)$ satisfies

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1. \quad (4)$$

The definition of differential entropy can be obtained by partitioning the range of X into bins of length Δx with associated sample points x_i within the bins; for X being Riemann integrable, $X \in (x_i, x_i + \Delta x)$, the approximate value of the entropy is defined as

$$\begin{aligned} H(X) &= -\sum_i f(x_i)\Delta x \log(f(x_i)\Delta x) \\ &= -\sum_i f(x_i)\Delta x[\log f(x_i) + \log \Delta x], \end{aligned} \quad (5)$$

when $\Delta x \to 0$, the second term $\log\Delta x \to \infty$, while the first term approximates the differential entropy, so the differential entropy $H(x)$ is defined as

$$H(x) = -\int_{-\infty}^{+\infty} f(x) \log f(x)\mathrm{d}x. \quad (6)$$

Accurate estimation of the entropy is of great importance for parameter estimation. Differential entropy is based on an averaging of the nearest neighbour distances between random points in the logarithm domain and the topological structure of the attractor will be changed along with the changes of the nearest neighbour distances in the m dimension, which represents the necessary number of coordinates to unfold observed orbits from self-overlaps.

The purpose is to estimate $H(x)$ from a time series $\{x_i, i = 1, 2, \ldots, n\}$ with unknown density function $f(x)$. The entropy is the average of $-\log f(x)$. If one had an unbiased estimator for $\log f(x_i)$, it would reach an unbiased estimator for the entropy. We will estimate $\log f(x_i)$ by considering the probability density function $P_{ik}(\varepsilon)$ for the distance between x_i and its k-th nearest neighbour. The probability $P_{ik}(\varepsilon)\mathrm{d}\varepsilon$ is equal to the chance that there is one point within distance $r \in [\varepsilon,\ \varepsilon + \mathrm{d}\varepsilon]$ from x_i, there are numbers of $k-1$ other points at smaller distances, and that the remaining $n-k-1$ points have larger distances from x_i. Denote the mass of the ε-ball centred at x_i by $p_i(\varepsilon)$:

$$p_i(\varepsilon) = \int_{\|x - x_i\| < \varepsilon} f(x)\mathrm{d}x. \quad (7)$$

Applying the trinomial formula, $P_{ik}(\varepsilon)$ is obtained as:

$$\begin{aligned} P_{ik}(\varepsilon) &= \frac{(n-1)!}{(k-1)!(n-k-1)!} \\ &\quad \times \frac{\mathrm{d}p_i(\varepsilon)}{\mathrm{d}\varepsilon} p_i^{k-1}(1-p_i)^{n-k-1}. \end{aligned} \quad (8)$$

It can be easily verified that $\int P_{ik}(\varepsilon)\mathrm{d}\varepsilon = 1$. According to the distribution $P_{ik}(\varepsilon)\mathrm{d}\varepsilon$, the expected

value of the function $\log p_i(\varepsilon)$ is

$$E(\log p_i(\varepsilon)) = \int_0^\infty P_{ik}(\varepsilon)\log p_i(\varepsilon)\mathrm{d}\varepsilon = k\binom{n-1}{k} \times \int_0^1 p^{k-1}(1-p)^{n-k-1}\log p\,\mathrm{d}p = \psi(k)-\psi(n). \quad (9)$$

For the last equality,

$$\int_0^1 x^{a-1}(1-x)^{b-1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (10)$$

$$\psi(x) = \Gamma'(x)/\Gamma(x), \quad (11)$$

where $\Gamma(x)$ is the digamma function. The expectation is taken over the positions of all other $n-1$ points, with x_i fixed. Assuming that $f(x)$ is almost constant in the entire ε-ball around x_i, we obtain:

$$p_i(\varepsilon) \approx c_m\varepsilon^m f(x_i) = \frac{\pi^{m/2}}{\Gamma(1+m/2)}\varepsilon^m f(x_i), \quad (12)$$

where m is the dimension and c_m is the volume of the m-dimensional unit ball. Substituting Eq. (12) into Eq. (9), we obtain

$$-\log f(x_i) \approx mE(\log(\varepsilon)) + \log(c_m) + \psi(n) - \psi(k), \quad (13)$$

which finally leads to the unbiased kNN estimator for the differential entropy:

$$H_k(x) = \frac{m}{n}\sum_{i=1}^n \log\varepsilon_{i,k} + \log(c_m) + \psi(n) + \psi(k), \quad (14)$$

where $\varepsilon_{i,k}$ is the distance from x_i to its k-th nearest neighbour.

For $k=1$, the estimator was given by Kozachenko and Leonenko[12]

$$H(x) = \frac{m}{n}\sum_{i=1}^n \log\varepsilon_i + \log(c_m) + \log(n-1) + \gamma, \quad (15)$$

where γ is Euler's constant and $\gamma \approx 0.5772$, ε_i is the distance of x_i from its nearest neighbour. In this paper, we adopt a modified estimator[13]

$$H(x) = \frac{1}{n}\sum_{i=1}^n \ln[n\cdot\varepsilon_i] + \ln 2 + \gamma. \quad (16)$$

For a given time series, the estimation of the differential entropy for time delay embedded versions with an embedding dimension m and a time delay τ is denoted as $H(x,\, m,\, \tau)$, which can be computed by Eq. (16), but $H(x,\, m,\, \tau)$ is not robust with respect to dimensionality. In order to highlight the robustness, we can use an ensemble of surrogates of the original signal to standardize $H(x,\, m,\, \tau)$. Surrogates $x_{s,i}$ are best generated by iterative amplitude adjusted Fourier transform (IAAFT),[14] which retains both the signal distribution and the autocorrelation structure of the original signal. The estimation for the surrogates $x_{s,i}$ is denoted as $H(x_{s,i},\, m,\, \tau)$ and the standardized result is represented by entropy ratio $I(m,\tau)$:

$$I(m,\tau) = \frac{H(x,m,\tau)}{\langle H(x_{s,i},m,\tau)\rangle_i}, \quad (17)$$

where i is the number of surrogates, $\langle\cdot\rangle_i$ denotes the average over all values of i. Employing the minimum description length[15] to limit the higher embedding dimensions, we find that the entropy ratio $I(m,\tau)$ becomes $R_{\mathrm{ent}}(m,\, \tau)$:

$$R_{\mathrm{ent}}(m,\tau) = I(m,\tau)\left(1+\frac{m\ln N}{N}\right), \quad (18)$$

where N is the number of delay vectors.

The variation of the reconstructed parameters leads to the folding and stretching behaviour of the attractor in the reconstruction. Differential entropy can be used to characterize the disorder degree of the attractor, so when the differential entropy is at a minimum, the attractor has the best regular structure.

4. Simulation experiments

Numerical time series generated from the Lorenz system, Duffing system and Rossler system respectively are used for simulation to illustrate the efficiency and the applicability of the above method.

The Lorenz system[16] is a computationally simple chaotic system of three variables, given by

$$\begin{cases} \mathrm{d}x/\mathrm{d}t = \sigma(y-x), \\ \mathrm{d}y/\mathrm{d}t = rx - y - xz, \\ \mathrm{d}z/\mathrm{d}t = -bz + xy, \end{cases} \quad (19)$$

where σ, b and r take the standard values 16, 4 and 45.92. For these parameters, the Lorenz system exhibits nonperiodic and chaotic motion. The first state variable x is chosen as the observable and a total of 7000 points are generated by using the fourth-order Runge–Kutta method; the time step h in the Runge–Kutta procedure is 0.01 and the first 2000 points are omitted before the time series is acquired.

For the obtained time series, 10 sets of surrogates are generated by the IAAFT method. The entropy estimate $H(x,\, m,\, \tau)$ and $H(x_{s,i},\, m,\, \tau)$ is computed

using Eq. (16), then the entropy ratios are evaluated for $m = 2, \ldots, 10$ and $\tau = 0, \ldots, 10$. The plot of entropy ratios is shown in Fig. 1(a), and the minimum of the plot ($R_{\text{ent}} = 0.9460$) yields the optimum set of reconstructed parameters $m_{\text{opt}} = 4$ and $\tau_{\text{opt}} = 5$.

The AMI for the time series is shown in Fig. 1(b). The first minimum of this function occurs at $\tau = 10$; this is the value of the time delay. In Fig. 1(c), we display the percentage of global FNN for this Lorenz data as a function of dimension. The percentage of FNN goes to zero at $m = 3$, and $m = 3$ is the necessary integer dimension for unfolding the attractor from time series. Through a comparison between the differential entropy method and AMI/FNN, we clearly find that there is an incongruence.

To further verify the correctness and feasibility of the method, reconstructions are performed by these two methods with the results shown in Figs. 1(d) to 1(f). Figure 1(d) shows the two-dimensional projection of the actual attractor while figures 1(e) and 1(f) show the reconstructions based on AMI/FNN and the differential entropy method respectively. It is clear that the feature of the attractor is more accurately captured by the differential entropy method.

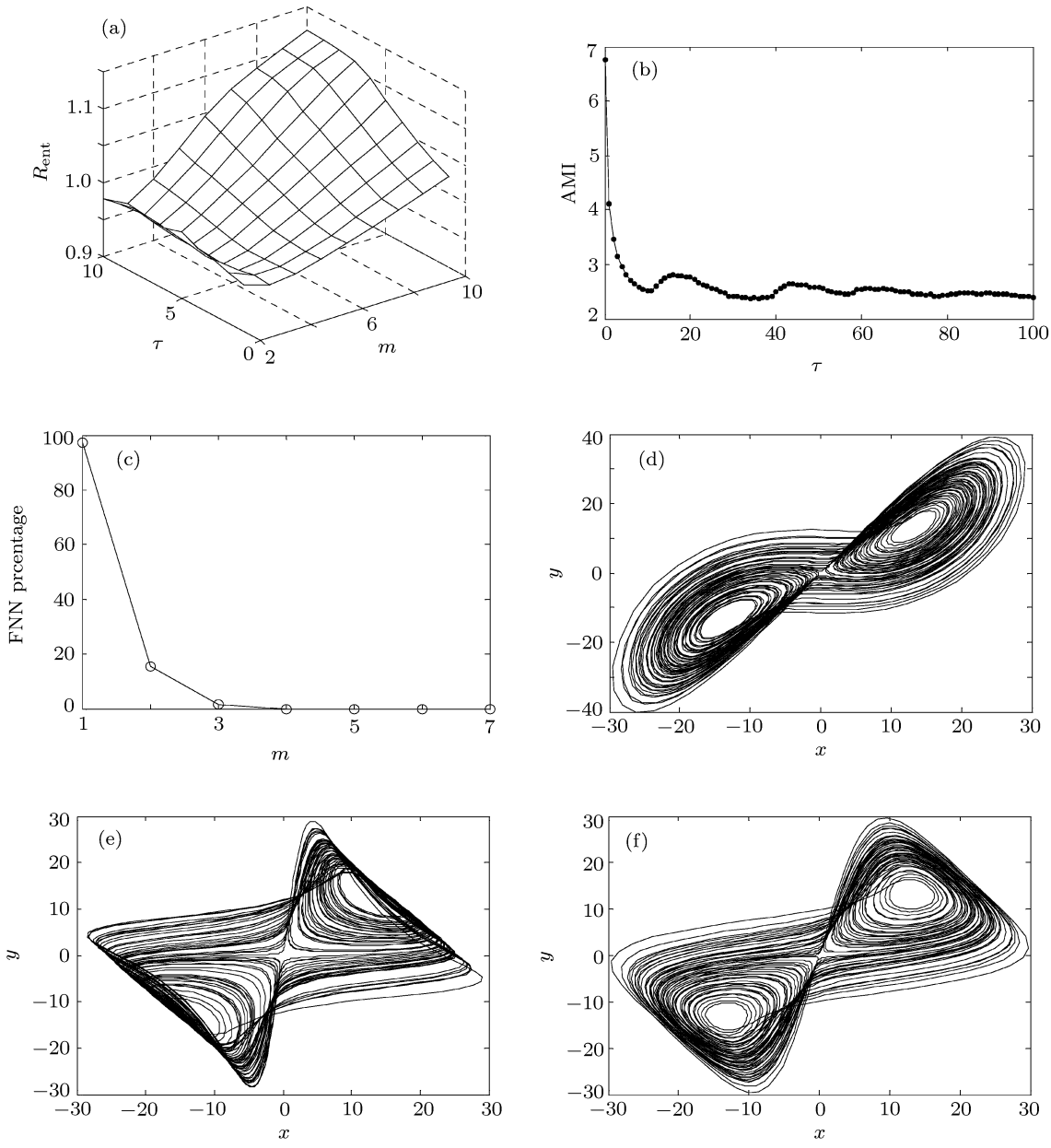


Fig. 1. Lorenz system ($\sigma = 16$, $b = 4$, $r = 45.92$): (a) the plot of differential entropy ratios; (b) the AMI for the Lorenz time series; (c) the percentage of FNN for the Lorenz time series; (d) actual attractor; (e) reconstructed attractor based on AMI/FNN; (f) reconstructed attractor based on differential entropy.

Consider a form of the forced Duffing system[17] given by:

$$\begin{cases} \mathrm{d}x/\mathrm{d}t = y, \\ \mathrm{d}y/\mathrm{d}t = -\sigma y - ax(1+x^2) + F\cos(z), \\ \mathrm{d}z/\mathrm{d}t = \Omega, \end{cases} \tag{20}$$

which becomes chaotic when $\sigma = 0.06$, $a = 0.7$, $F = 7.5$, $\Omega = 1$. Equation (20) is integrated by the fourth-order Runge–Kutta method with time step $h = 0.05$; 7000 points are obtained and the first 2000 points are omitted.

Ten sets of surrogates are generated by the IAAFT method. The differential entropy of the original time series and its surrogates are estimated. Figure 2 (a) shows the entropy ratio plot for $m = 2, \ldots, 10$ and $\tau = 0, \ldots, 10$. The minimum of the entropy ratio occurs at $R_{\mathrm{ent}} = 0.9527$ and the optimum set of reconstructed parameters $m_{\mathrm{opt}} = 3$ and $\tau_{\mathrm{opt}} = 6$ are found.

For comparison, those parameters are obtained by the AMI/FNN method. The AMI for the time series is shown in Fig. 2(b) and the time delay is $\tau = 11$. From Fig. 2(c), the FNN percentage indicates a minimum embedding dimension $m = 3$. It is clear that the obtained parameters are dramatically different from those obtained by the differential entropy method.

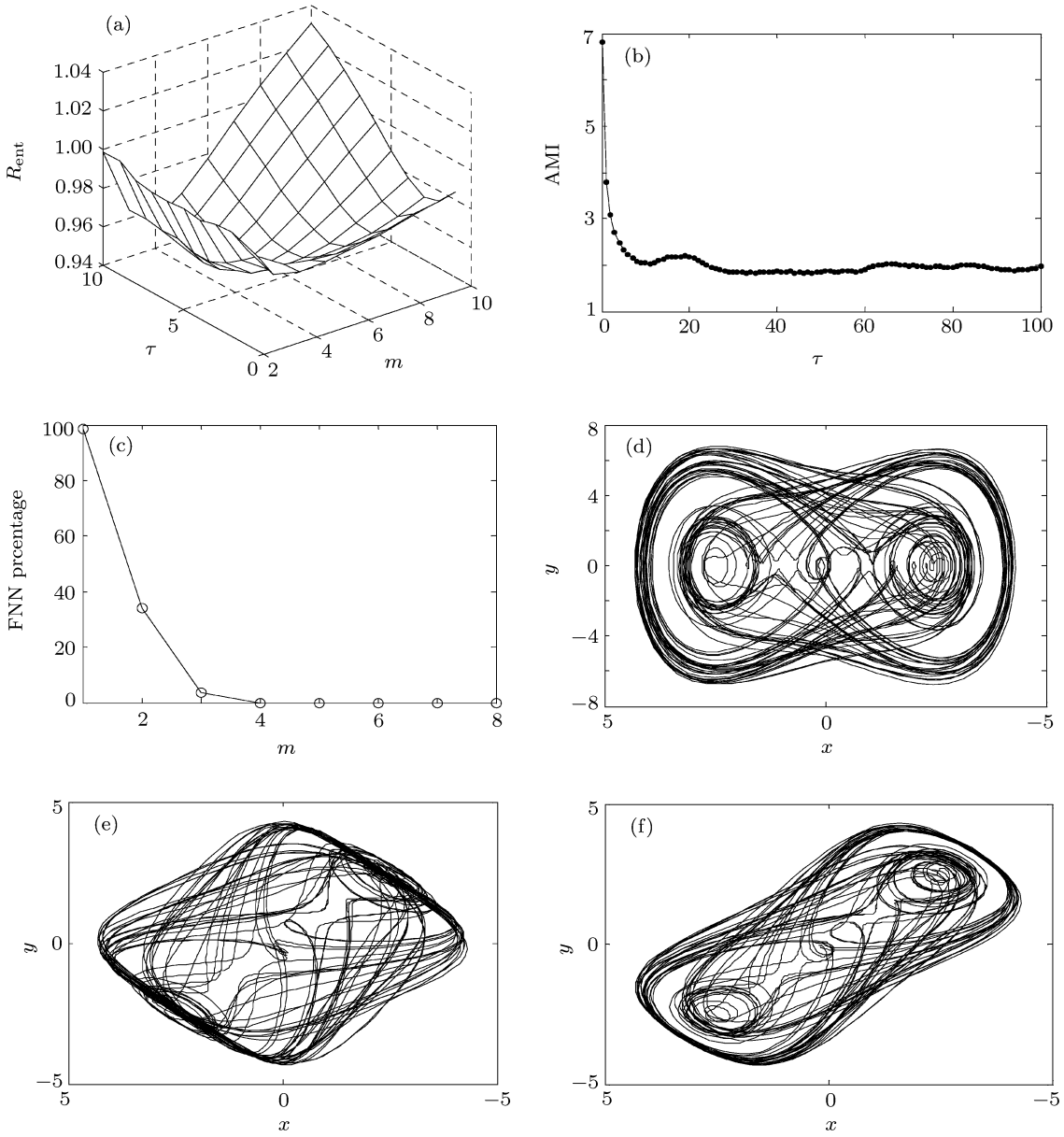


Fig. 2. Duffing system ($\sigma = 0.06$, $a = 0.7$, $F = 7.5$): (a) the plot of differential entropy ratios; (b) the AMI for the Duffing time series; (c) the percentage of FNN for the Duffing time series; (d) actual attractor; (e) reconstructed attractor based on AMI/FNN; (f) reconstructed attractor based on differential entropy.

For showing two-dimensional projections of the phase trajectory, figure 2(d) shows a projection of the actual attractor while figures 2(e) and 2(f) show reconstructions based on AMI/FNN and the differential entropy method. Clearly, the reconstruction is better by using the differential entropy method.

The third is the Rossler system:[18]

$$\begin{cases} \mathrm{d}x/\mathrm{d}t = -(y+z), \\ \mathrm{d}y/\mathrm{d}t = x + \mathrm{d}y, \\ \mathrm{d}z/\mathrm{d}t = e + z(x-f), \end{cases} \tag{21}$$

where $d = 0.2$, $e = 0.2$, $f = 5$. For these parameters, the Rossler system exhibits broadband chaos. Equation (21) is solved and 7000 points are acquired by the fourth-order Runge–Kutta method with time step $h = 0.05$; the first 2000 points are omitted.

Ten sets of surrogates are generated by the IAAFT method. $H(x,\ m,\ \tau)$ and $H(x_{s,i},\ m,\ \tau)$ are computed by using Eq. (16). Figure 3(a) shows the plot of entropy ratios for $m = 2, \ldots, 10$ and $\tau = 0, \ldots, 15$. The minimum of the plot ($R_{\mathrm{ent}} = 0.9946$) corresponds to the optimum set of reconstructed parameters $m_{\mathrm{opt}} = 4$ and $\tau_{\mathrm{opt}} = 7$.

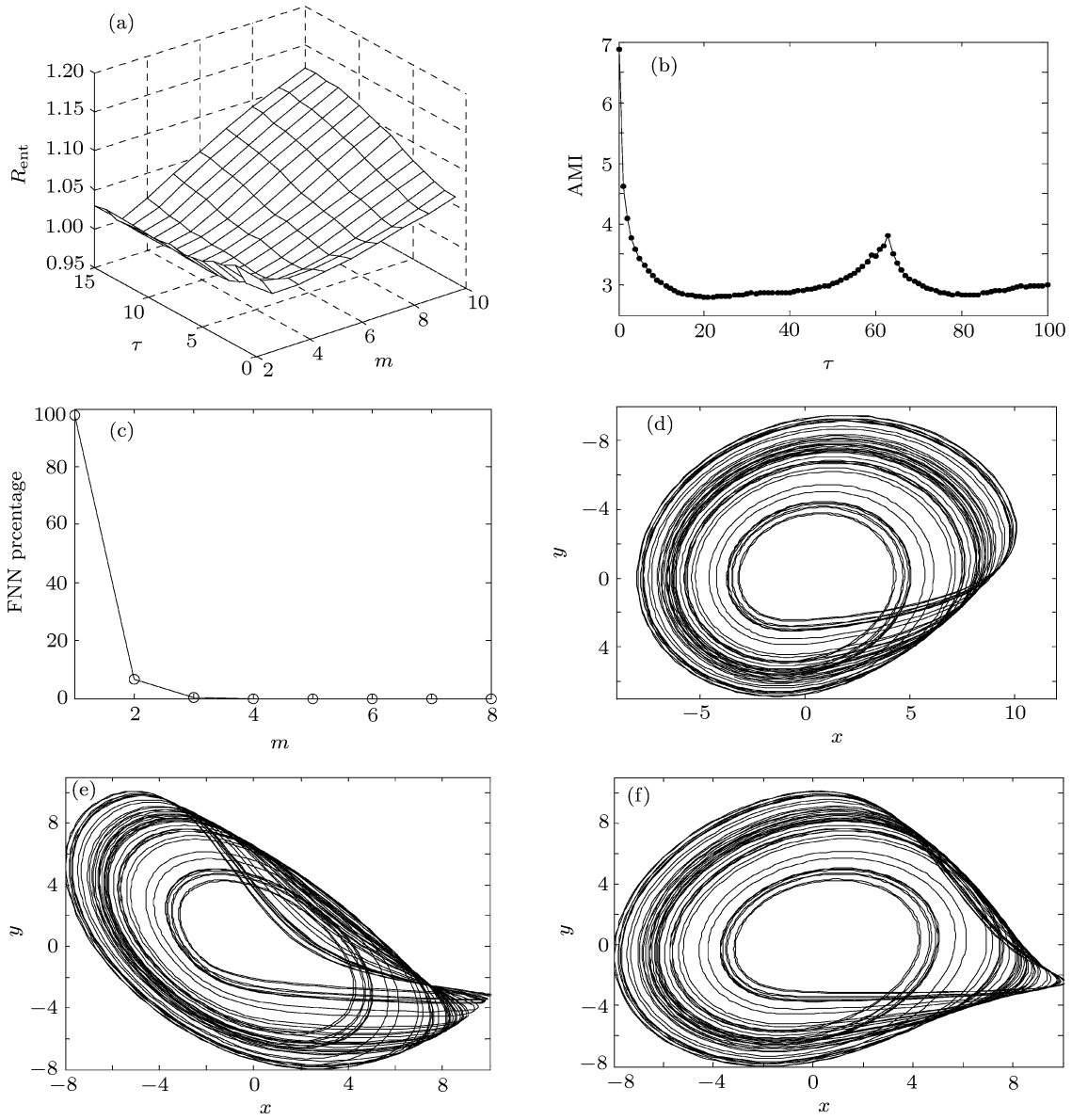


Fig. 3. Rossler system ($d = 0.2$, $e = 0.2$, $f = 5$): (a) the plot of differential entropy ratios; (b) the AMI for the Rossler time series; (c) the percentage of FNN for the Rossler time series; (d) actual attractor; (e) reconstructed attractor based on AMI/FNN; (f) reconstructed attractor based on differential entropy.

To obtain embedding parameters by the AMI/FNN method, we find that the first minimum of AMI appears at $\tau = 22$ as depicted in Fig. 3(b). In Fig. 3(c), the FNN percentage indicates a minimum embedding dimension $m = 3$. Obviously, the obtained parameters are inconsistent with those obtained by the differential entropy method.

Figure 3(d) shows a projection of the actual attractor, figures 3(e) and 3(f) show projections based on AMI/FNN and differential entropy method respectively. Note that the features of the attractor are more accurate by using the differential entropy method.

The Lyapunov exponent and the correlation dimension are invariable for a chaotic attractor, so the reconstruction quality is evaluated by these two parameters. In this paper, the G–P algorithm is employed to compute the correlation dimension, and the Lyapunov exponent based on Rosenstein's method[19] is estimated. For simplification, the calculations for reconstruction based only on the differential entropy method are shown in Figs. 4 and 5.

The calculated results and errors for AMI/FNN and the differential entropy method are provided in Table 1. It is clear that the reconstruction based on the differential entropy method can effectively decrease the calculation errors of the correlation dimension and the largest Lyapunov exponent, which is superior to the AMI/FNN method.

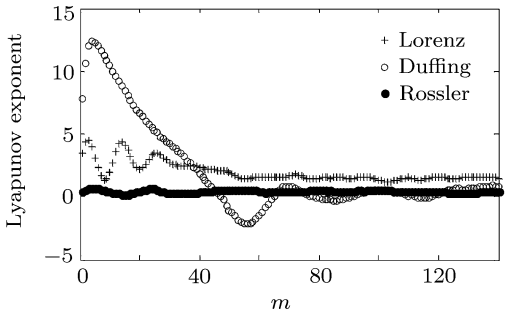


Fig. 4. Lyapunov exponent.

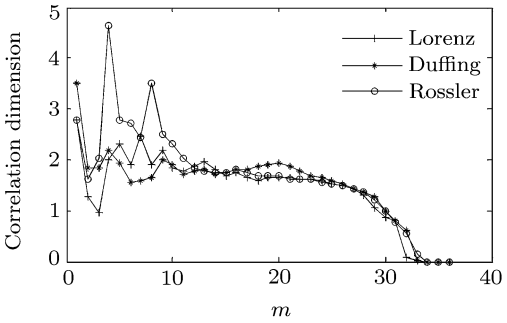


Fig. 5. Correlation dimension.

Table 1. Calculation result of the Lyapunov exponent and correlation dimension.

system	Lyapunov exponent					correlation dimension				
	actual value	AMI/FNN	error	DE	error	actual value	AMI/FNN	error	DE	error
Lorenz	1.4976	1.4571	–0.0405	1.4725	–0.0251	20377	2.0879	+0.0502	2.0564	+0.0187
Duffing	0.2748	0.2073	–0.0675	0.3300	+0.0552	1.8201	1.8456	+0.0255	1.8310	+0.0109
Rossler	0.4013	0.4376	+0.0363	0.4273	+0.0260	1.9544	1.9683	+0.0139	1.9598	+0.0054

5. Conclusion

In order to extract the chaotic characteristic accurately, this paper presents an efficient method based on the differential entropy for phase space reconstruction. The optimum set of reconstructed parameters is given by the minimum of the differential entropy. Compared with conventional methods, we find that a clear advantage of the differential entropy method is that both embedding dimension and delay time can be determined simultaneously. The simulation results show that the estimated embedding parameters by the differential entropy method can capture the potential dynamical information from the observed chaotic time series better than those by the AMI/FNN method. Furthermore, the accuracy of invariants in phase space reconstruction is improved by means of estimating the correlation dimension and the largest Lyapunov exponent. All the experiments prove that the differential entropy method is practicable for determining appropriate reconstructed parameters.

References

[1] Song Q S, Feng Z R and Li R H 2009 *Acta Phys. Sin.* **58** 5057 (in Chinese)

[2] Liu J H, Zhang H G and Feng J 2008 *Acta Phys. Sin.* **57** 6868 (in Chinese)

[3] Meng Q F, Peng Y H and Sun J 2007 *Chin. Phys.* **16** 3220

[4] Small M and Tse C K 2004 *Physica* D **194** 283

[5] Cao L Y 1997 *Physica* D **110** 43

[6] Yang S Q and Jia C Y 2002 *Acta Phys. Sin.* **51** 2452 (in Chinese)

[7] Kim H S, Eykholt R and Salas J D 1999 *Physica* D **127** 48

[8] Meng Q F, Peng Y H and Sun J 2007 *Chin. Phys.* **16** 1252

[9] Cong R, Liu S L and Ma R 2008 *Acta Phys. Sin.* **57** 7487 (in Chinese)

[10] Gautama T and Danilo P M 2003 *IEEE* **6** 29

[11] Xiao F H, Yan G R and Han Y H 2005 *Acta Phys. Sin.* **54** 550 (in Chinese)

[12] Kozachenko L F and Leonenko N N 1987 *Problems of Information Transmission* **23** 95

[13] Beirlant J, Dudewicz E J, Gyorfi L and Vander M 1997 *Int. J. Math. Stat. Sci.* **6** 17

[14] Schreiber T and Schmitz A 2000 *Physica* D **142** 346

[15] Zhao Y and Michael S 2006 *IEEE* **53** 722

[16] You R Y, Chen Z, Xu S C and Wu B X 2004 *Acta Phys. Sin.* **53** 2882 (in Chinese)

[17] Liu B Z and Peng J H 2004 *Nonlinear Dynamics* (Vol. 1) (Beijing: Higher Education Press) pp. 132–143 (in Chinese)

[18] Wang A L and Yang C X 2002 *Acta Phys. Sin.* **51** 2719 (in Chinese)

[19] Rosenstein M T, Collins J J and Deluca C J 1993 *Physica* D **65** 117