NOLTA, IEICE

Paper

Criterion for determining the optimal delay of attractor reconstruction using persistent homology

Shotaro Tsuji [1 a)] *and Kazuyuki Aihara* [1,2 b)]

[1] *Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo*
7-3-1 Hongo Bunkyo-Ku, Tokyo 113-8656, Japan

[2] *Institute of Industrial Science, University of Tokyo*
4-6-1 Komaba Meguro-ku, Tokyo 153-8505, Japan

[a)] *tsuji@sat.t.u-tokyo.ac.jp*
[b)] *aihara@sat.t.u-tokyo.ac.jp*

Received June 4, 2018; Revised October 3, 2018; Published January 1, 2019

Abstract: Persistent homology computes the number and the widths of holes of a shape. The delay-coordinate embedding reconstructs an attractor from time series data. Recently the combination of persistent homology and attractor reconstruction is proposed for clustering time series data. We propose a criterion to determine the delay of attractor reconstruction for periodic and chaotically periodic signals. Our criterion chooses the delay that maximizes the hole width of reconstructed attractors. We compare it to the criterion with mutual information and that by Perea and Harer. It is found that the combination of our criterion with Perea and Harer's criterion works well.

Key Words: attractor reconstruction, persistent homology, optimal delay

1. Introduction

There is a topological criterion for choosing the optimal delay for attractor reconstruction. It was proposed in the study of Perea and Harer [1] for periodic signals. Their criterion is examined numerically for periodic and chaotically periodic signals.

1.1 Problem statement

In delay-coordinate embedding [2, 3], the appropriate delay should be determined, as it controls the shape of the reconstructed attractors and influences the quality of time series analysis.

The delay-coordinates obtained from a time series $x(t)$ are defined to be the following n-dimensional vector:

$$y(t) = (x(t), x(t+a), x(t+2a), \ldots, x(t+(n-1)a)), \tag{1}$$

where the parameter a is the delay, and n is the embedding dimension.

Nonlinear Theory and Its Applications, IEICE, vol. 10, no. 1, pp. 74–89 DOI: 10.1587/nolta.10.74

Takens' theorem [2, 3] ensures that the shape reconstructed by this mapping is homeomorphic to the original attractor under certain conditions. Let A be an attractor of a dynamical system and let $\xi(t)$ be an orbit in A. For a measurement function $h : A \to \mathbb{R}$, let $x(t) = h(\xi(t))$ and Φ be a mapping from original m-dimensional state space to n-dimensional Euclidean space such that $\Phi(\xi(t)) = y(t) = (h(\xi(t)), h(\xi(t+a)), \ldots, h(\xi(t+(n-1)a)))$. For $a > 0$ and $n > 2m$, Takens' theorem states that A is homeomorphic to the reconstructed attractor $\tilde{A} = \Phi(A)$. That is, the delay-coordinate embedding induces a dynamical system conjugate to the dynamical system that $\xi(t)$ obeys.

Although the original attractor A and the reconstructed attractor $\tilde{A}$ are homeomorphic, it is known that the shape of the reconstructed attractor depends on the delay a. There are two related phenomena called redundancy and irrelevance [4]. Redundancy refers to the concentration of the reconstructed attractor on the diagonal set. This occurs when a is overly small because the values of the components in the delay coordinates are near each other. Irrelevance refers to chaotic systems in which the reconstructed attractor has considerably more complicated shape compared with the original attractor. This occurs when a is overly large because the causality between the point $y(t)$ and the point $y(t+a)$ disappears. Therefore, the choice of the delay is highly important in the analysis of time series data with delay-coordinate embedding. Delay-coordinate embedding has been used for calculating the correlation dimension and the Lyapunov exponents of time series. These are defined for an orbit in phase space; thus, delay-coordinate embedding is required.

Recently, combination of delay-coordinate embedding and persistent homology has been proposed for the classification of time series data [5–9]. Persistent homology can count the number of holes in a given shape and estimate the width of each hole. In the aforementioned classification method, persistent homology is applied to the attractors reconstructed from time series data and is subsequently used for machine learning.

1.2 Criterion for optimal delay

The optimal delay a should be chosen as the minimum value that maximizes the hole width of reconstructed attractors with holes. If we consider the attractors with a hole, such attractors are generated from either periodic or chaotically periodic trajectories.

In the aforementioned applications, time series data are classified with persistent homology. Persistent homology computes the number and widths of holes. It is needed to precisely detect holes in the given shape. When the delay a is too small, the reconstructed attractor comes close to the diagonal set and the hole becomes too narrow. If the trajectory is added by noise, the hole is drowned out by the noise and it cannot be detected with persistent homology. Even if there is no noise, tiny holes appear in the gaps within the trajectory. The hole that represents the periodicity cannot be distinguished from these tiny holes. Consequently we cannot classify the time series data with too small delay. Thus the delay should be chosen to maximize the hole width of reconstructed attractors. In order to measure the width of a hole, we take the death filtration value of the persistent homology generator with the maximum persistence. We call this value the most significant death value. We cannot choose the excessively large delay because of the irrelevance. Excessively large delay deforms the shape of the reconstructed attractor. Although we cannot clearly say how the reconstructed attractor is deformed, this deformation should be avoided. Thus we regard the first value of the delay that maximizes the most significant death value as the optimal delay.

1.3 Literature review

A large number of criteria for determining the delay of attractor reconstruction have been proposed. In most of them, geometric–heuristic approaches are adopted.

In 1986, Fraser and Swinney [10] proposed using the delay that attains the first local minimum of the mutual information between $x(t)$ and $x(t+a)$. For sufficiently small mutual information, the values of $x(t)$ and $x(t+a)$ can be as independent as possible, so that the reconstructed attractor may be expanded. This criterion is still widely used.

Liebert, Pawelzik, and Schuster [11] proposed the wavering product for simultaneously determining the delay and the embedding dimension. The wavering product measures the extent to which

the neighboring points in the attractor remain in the neighborhood when the embedding dimension increases.

Buzug and Pfister [12, 13] proposed the fill-factor and the integral local deformation. The fill-factor determines the optimal delay so that redundancy may be avoided by measuring the capacity of the space that the reconstructed attractor fills. The integral local deformation measures the divergence of neighboring trajectories with time.

Rosenstein, Collins, and de Luca [14] proposed the average displacement. The average displacement measures the distance of the reconstructed trajectory from the diagonal set. They chose the delay as the first value for which the slope of the delay versus the average displacement plot is 40% of the first value of the slope.

There is an another approach that uses statistical models and the minimum description length principle.

Judd and Mees [15] proposed combining the problem of determining the optimal parameter of the delay-coordinate embedding and the modeling of the observed dynamical system. Their method determines the non-uniform embedding that minimizes the description length of the estimated model.

Small and Tse [16] used the method proposed by Judd and Mees for determining the delay and the embedding dimension of uniform embedding. They suggested that the embedding window, i.e., the product of the delay and the embedding dimension, is more important than the delay alone.

A topological criterion has also been proposed. Perea and Harer [1] proposed that the optimal delay for periodic signals is $an = T/2$, where T is the period. They analyzed the behavior of delay-coordinate embedding by embedding a circle and calculated the value of the delay that realizes the roundest reconstructed circle. Moreover, they combined delay-coordinate embedding and persistent homology to detect the periodicity of signals.

Mutual information as well as the method of Perea and Harer will be discussed later.

1.4 Contributions of this study

A new measure for determining the optimal delay for attractor reconstruction was proposed, which is called the most significant death value, in order to examine the criterion of Perea and Harer. A reconstructed attractor whose hole is most distinguishable can be obtained by maximizing the most significant death value. The criterion of Perea and Harer was compared with mutual information using numerical experiments. It was demonstrated that mutual information does not always determine the optimal delay.

1.5 Structure of the paper

In Section 2, persistent homology and Takens' theorem are introduced. In Section 3, two methods for determining the delay time are reviewed, and their limitations are discussed. In Section 4, the most significant death value is introduced, and an analysis of the proposed criterion is provided. In Section 5, the materials and the procedures of the experiment are explained. In Section 6, the results of the experiment are presented. In Section 7, the relationship between the results and the proposed criterion is discussed. Moreover, the proposed criterion is compared with the mutual information method. Section 8 concludes this paper.

2. Background

In this section, persistent homology and Takens' theorem are briefly introduced. Then, two criteria for determining the optimal delay are reviewed, namely, mutual information and the method of Perea and Harer.

2.1 Persistent homology

Persistent homology [17] is a technique in computational topology. Edelsbrunner, Letcher, and Zomorodian [18] introduced the concept of persistent homology to analyze shapes of proteins. In mathematics, homology groups are used to analyze the shape of continuous spaces; however, on computers, spaces are discrete. Thus, homology groups cannot directly be used on computers. Then, the

filtration of spaces is introduced to imitate a continuous space, and the homology group applied to a filtration is called persistent homology.

The homology group [19, 20] counts the number of holes of arbitrary dimension in a given shape. The meaning of a "hole of arbitrary dimension" should be clarified: An n-dimensional hole is the space obtained by removing an $(n+1)$-dimensional ball. For example, a 1-dimensional hole is an ordinary hole and a 2-dimensional hole is a void. Exceptionally, a connected component is regarded as a 0-dimensional hole.

Homology groups are defined for simplicial complexes. A simplicial complex is converted into a chain complex C, which is a collection of free modules $\{C_n\}$. A homology group is the quotient group of a cycle group $Z_n(C) = \mathrm{Ker}\ \partial_n$ and a boundary group $B_n(C) = \mathrm{Im}\ \partial_{n+1}$, where ∂_n is the boundary operator. The homology group is denoted by $H_n(C) = Z_n(C)/B_n(C)$. A generator of a homology group corresponds to a non-bounding cycle in a simplicial complex.

Vietoris–Rips complexes [21] are simplicial complexes defined for a set of points in a metric space. For any points in the set whose pairwise distances are at most r, a simplex is formed that consists of these points, and thus a complex is obtained. The Vietoris–Rips complex of a point cloud X is denoted by $\mathrm{VR_r(X)}$.

The concept of filtration is introduced to avoid determining the threshold of a Vietoris–Rips complex. A filtration of simplicial complexes is a sequence whose complexes are preceded by their subcomplexes. For two thresholds $r < r'$, the Vietoris–Rips complex for the threshold r is a subcomplex of that for the threshold r', i.e., $\mathrm{VR_r(X)} \subset \mathrm{VR_{r'}(X)}$.

A persistent homology group [19] is defined for a filtration of simplicial complexes. By considering the homology group of each complex in a filtration, the homomorphisms between the resulting homology groups are also obtained. These homology groups and homomorphisms are called the persistent homology. The persistent homology of a point cloud X is denoted by $PH_*(X)$.

A generator of a persistent homology group has a lifetime, or persistence. The homomorphisms in the persistent homology describe how a generator at a filtration value changes at another filtration value. The filtration value at which a generator is born is called the birth time or the birth filtration value. The filtration value where a generator dies is called the death time or the death filtration value.

The persistent homology of a Vietoris–Rips complex of a point cloud can be used to estimate the width of holes in the point cloud. When a cycle in a persistent homology dies, it is filled up in the filtration of the Vietoris–Rips complex. This death filtration value is nearly equal to the width of the hole represented by the cycle.

2.2 Takens' theorem

As mentioned above, Takens' theorem ensures that the reconstructed attractor is homeomorphic to the original attractor under certain conditions.

Let $\xi(t)$ be a trajectory of a dynamical system on an Euclidean space $\mathbb{R}^m$. Let $h : \mathbb{R}^m \to \mathbb{R}$ be a continuous function and a be a positive real number. Then, the delay-coordinate mapping $\Phi : \mathbb{R}^m \to \mathbb{R}^n$ is defined by

$$\Phi(\xi(t)) = (h(\xi(t), h(\xi(t+a)), \ldots, h(\xi(t+(n-1)a))). \tag{2}$$

Assuming that $\xi(t)$ belongs to an attractor A in $\mathbb{R}^m$ and $n > 2m$, this mapping is generically an embedding.

As the mapping Φ is only generically an embedding, there are cases where Φ is not an embedding. For example, if $\xi(t)$ is periodic with period T, the mapping Φ is not an embedding for $a = T, 2T, 3T, \ldots$ [2]. Then, the points of $\Phi(\xi(t))$ collapse on the diagonal set because $\xi(t+a) = \xi(t)$.

If Φ is an embedding, the original attractor A and the reconstructed attractor $\tilde{A} = \Phi(A)$ are homeomorphic. However, as explained in Introduction, when the resulting delay a is excessively small, the reconstructed attractor $\tilde{A}$ concentrates around the diagonal set. This phenomenon is known as redundancy. Hence, the delay time should be properly chosen, and a criterion for determining it is necessary.

Excessively small as well as large delay should be avoided. An excessively large delay a may result in a reconstructed attractor of complicated shape. This phenomenon is called irrelevance. Although the reconstructed attractor and the original attractor are homeomorphic, the complicated shape of the former is not suitable for processing time series data, as it would yield unsatisfactory results.

3. Related Criteria

The typical criterion as well as criteria related to the present study will now be reviewed.

3.1 Mutual information

A widely used criterion for the optimal delay is mutual information [10]. For example, TISEAN [22], a software package for nonlinear time analysis, implements this method.

The first minimum of the mutual information between $x(t)$ and $x(t+a)$ is chosen when a increases from 0. Then, $x(t)$ and $x(t+a)$ are the most independent with respect to a. Such $x(t)$ and $x(t+a)$ are well distinguishable and expand the shape of the reconstructed attractor. Choosing a minimum of the mutual information avoids redundancy. The first minimum is selected to avoid irrelevance.

The mutual information between $x(t)$ and $x(t+a)$ is defined as follows. Let S be the set of the values of $x(t)$, and Q be the set of the values of $x(t+a)$. Then, the product of S and Q is defined as the set of pairs $(x(t), x(t+a))$. The probability distribution of S, Q, and (S,Q) is denoted by P_S, P_Q, and P_{SQ}, respectively. Then, the mutual information of S and Q is defined by

$$I(S,Q) = \int P_{SQ}(s,q) \log \left(\frac{P_{SQ}(s,q)}{P_S(s) P_Q(q)} \right) dsdq. \tag{3}$$

3.2 Method of Perea and Harer

Perea and Harer [1] analyzed the behavior of the circles embedded in the delay-coordinate space. Their aim was to detect periodic signals from time series data. They assumed that the periodic signals can be expanded into Fourier series. Before analyzing such periodic signals they mapped a circle into the delay-coordinate space for simplicity.

They found that the embedded circle is roundest when the delay satisfies $an = T/2$. The embedded circle becomes an ellipse in the delay-coordinate space. In their study, "roundest" implies that the major axis and the minor axis of the ellipse have the same length.

It should be noted that when the embedded circle is "roundest", it has the maximum width.

3.3 Limitations

The mutual information method does not appear to be suitable for expanding reconstructed attractors. It does not directly imply the expansion of the reconstructed space although the values of $x(t)$ and $x(t+a)$ may not be correlated.

The problem is that the delay determined by mutual information does not depend on the embedding dimension. Essentially, this method is suitable only for the case where the embedding dimension is two, because it only calculates the mutual information between $x(t)$ and $x(t+a)$.

The criterion suggested by Perea and Harer can be used only for signals of a certain class. It is restricted to periodic signals that can be expanded in Fourier series. The simplest example of such signals is a sine wave. It is not known whether this criterion can be applied to arbitrary periodic or chaotically periodic signals.

4. Proposed criterion

Perea and Harer [1] combined delay-coordinates and persistent homology in order to recognize the given time series as periodic as summarized below. They introduced the maximum persistence as an index of periodicity. It is defined as

$$\mathrm{mp(A)} = \max_{\mathrm{x} \in \mathrm{PH_1(A)}} \mathrm{pers(x)},$$

where A is the reconstructed attractor and pers(x) is the persistence of the homology generator x.

They supposed that the time series data can be written as

$$x(t) = \cos(Lt),$$

and embedded this into the delay coordinate space of dimension n with the delay a. They obtained an ellipse in the delay coordinate space. They investigated when the delay makes the ellipse "roundest". The semi-major axis and the semi-minor axis of this ellipse can be calculated by hand. The thought was that the longer the semi-minor axis the rounder the ellipse. They found that the delay that makes the ellipse "roundest" is given by

$$a = \frac{1}{n}\frac{2\pi}{L}.$$

Moreover, they investigated the delay coordinate embedding when the time series data can be written as

$$x(t) = \sum_{k=0}^{N}(a_k \cos kt + b_k \sin kt).$$

It means that the time series data are represented by the partial sum of the Fourier series. They analyzed the condition on the dimension of embedding and they obtained the following inequality:

$$n > 2N.$$

There are concerns on the analysis of Perea and Harer. How large N should we choose? Furthurmore, the optimal delay is different for each higher harmonics. This can be an obstacle for choosing the delay.

Let us see the frequency spectrum of some real data. Figure 2 shows the FFT spectrum of the Japanese vowel /a/. Even if we count the number of prominent harmonics we should set $N = 30$. In Fig. 4 the FFT spectrum of the limit cycle of van der Pol equations has a lot of higher harmonics. Such data require large N and it makes us to face the difficulty to choose the optimal delay because there are many harmonics. We will give another analysis of delay coordinate embedding for this problem.

In addition, the maximum persistence is slightly smaller than the width of the hole because the birth time of the homology generator is not zero since the points in the data are apart from each other. Let us propose another index to measure the periodicity of a signal from the point of view of the hole width. The most significant death value MSDV(A) is defined as the death value of a, where $a \in PH_1(A)$ and $\mathrm{pers(a)} = \max_{\mathrm{c}\in \mathrm{PH}_1(\mathrm{A})} \mathrm{pers(c)}$.

We evaluate the width of the hole of the reconstructed attractor under some assumptions, without Fourier series. Consider a function $x : \mathbb{R} \to \mathbb{R}$. We denote the value of x at t by $x(t)$. We assume that the function x is continuous and has the period of T. Without loss of generality, we can assume that $x(0) = x(T) = 0$. We also assume that $x(T/2) = 0$ holds and $x(t) > 0$ for $0 < t < T/2$ and $x(t) < 0$ for $T/2 < t < T$ hold.

We consider the following set:

$$A = \{x(t) \mid t \in I,\, I \subset \mathbb{R}\}, \tag{4}$$

where I is an interval. We embed this set into the delay coordinate space of dimension n with delay a. Let $y_a(t)$ be the vectors of the embedded set of A:

$$y_a(t) = (x(t), x(t+a), \ldots, x(t+(n-1)a)). \tag{5}$$

Let $\tilde{A}$ be the set of $y_a(t)$: $\tilde{A} = \{y_a(t) \mid t \in I\}$. The set $\tilde{A}$ becomes a closed curve because the function x is periodic if we choose the length of I that is sufficiently larger than the value of T.

We denote the hole width of $\tilde{A}$ by $w_a(\tilde{A})$. Intuitively, as follows the minimum of the distance between the origin and the point $y_a(t)$ is smaller than the hole width:

$$\min_t \|y_a(t)\| \le w_a(\tilde{A}). \tag{6}$$

Let ξ be a trigonometric function with the period of T:

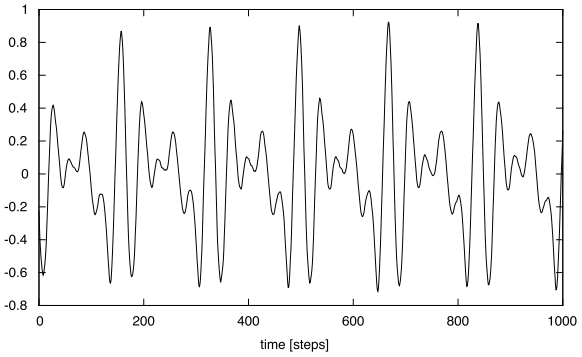

Fig. 1. The signal of the Japanese vowel /a/.

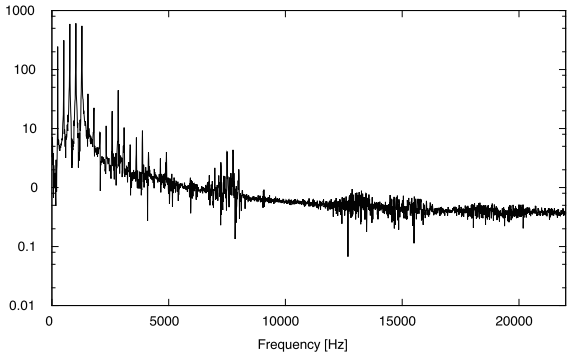

Fig. 2. The FFT spectrum of the Japanese vowel /a/.

$$\xi(t) = C \sin\left(\frac{2\pi}{T}t\right), \tag{7}$$

where C is the constant that satisfies $|\xi(t)| \leq |x(t)|$ for all t. We embed $\xi(t)$ into the delay-coordinate space:

$$\eta_a(t) = (\xi(t), \xi(t+a), \ldots, \xi(t+(n-1)a)). \tag{8}$$

Then we obtain the following inequalities:

$$\min_t \|\eta_a(t)\| \leq \min_t \|y_a(t)\| \leq w_a(\tilde{A}). \tag{9}$$

Furthermore, we calculate the first term of the equation above and get the following inequalities:

$$C\left(\frac{n}{2} - \frac{1}{2}\sqrt{\frac{1-\cos(4an\pi/T)}{1-\cos(4a\pi/T)}}\right) \leq \min_t \|y_a(t)\| \leq w_a(\tilde{A}). \tag{10}$$

Although we cannot know whether $w_a(\tilde{A})$ attains the maximum when $\eta_a(t)$ attains the maximum with varying the value of a, at least the value of $w_a(\tilde{A})$ may be sufficient large and it is larger than $\max_a \min_t \|\eta_a(t)\|$. The first term of Eq. (10) attains the maximum when $a = \frac{1}{n}\frac{T}{2}$ holds and then we get the following inequalities:

$$C\frac{n}{2} \leq \min_t \|y_a(t)\| \leq w_a(\tilde{A}). \tag{11}$$

Further we consider whether we can apply the analysis above to several time series data. We can apply the analysis to the van der Pol system because we can see that the data in Fig. 3 satisfies the assumption. Figure 5 shows the x-coordinate of the Rössler attractor. Though the signal is not strictly periodic, we are interested in whether our analysis works in this case. Figure 7 shows the x-coordinate of the Lorenz attractor. The Lorenz attractor has two holes and each hole corresponds to the positive value and negative value of Fig. 7 respectively. Although the Lorenz attractor is not strictly periodic but has two holes, our analysis may be able to roughly applied to each hole. But we cannot choose the period of the signal from the spectrum the Lorenz attractor shown in Fig. 8. Thus, instead of the spectrum, we use the median intervals of crossing some positive and negative values. Here we choosed 7.5 and -7.5 respectively.

5. Methods

Numerical experiments were conducted to test the proposed criterion. Then, the results were compared with those obtained by the mutual information method and the criterion of Perea and Harer.

5.1 Materials

The harmonic oscillator, the van der Pol system, the Rössler system, and the speech signal of a Japanese vowel /a/ were used in the experiment. The definitions of these systems as well as the rationale for their choice will now be provided.

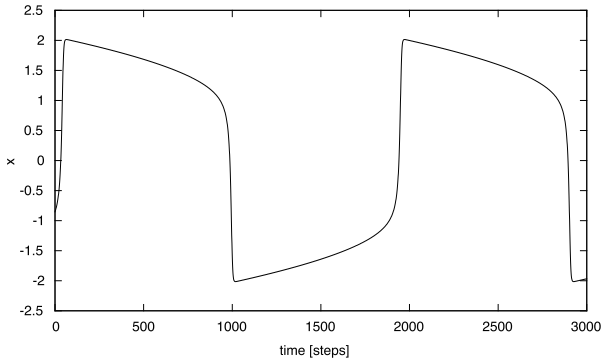

Fig. 3. The x coordinate of the limit cycle of the van der Pol system.

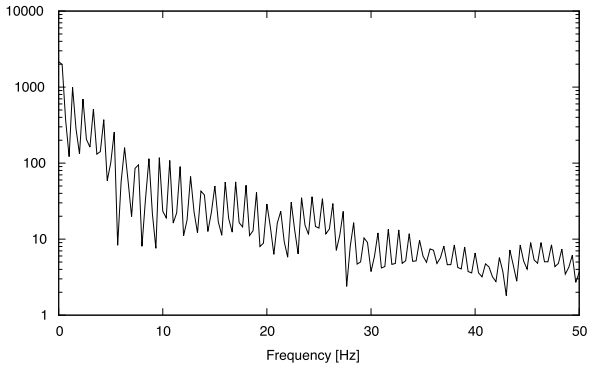

Fig. 4. The FFT spectrum of the limit cycle of van der Pol system.

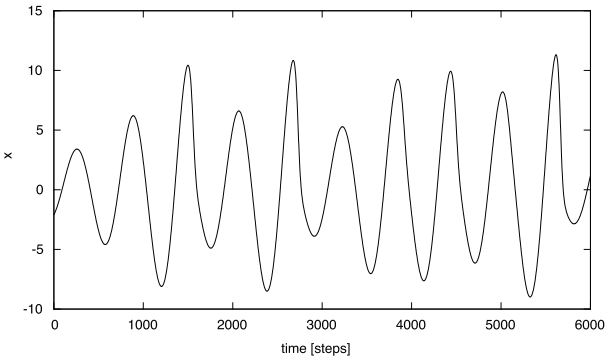

Fig. 5. The x coordinate of the Rössler attractor.

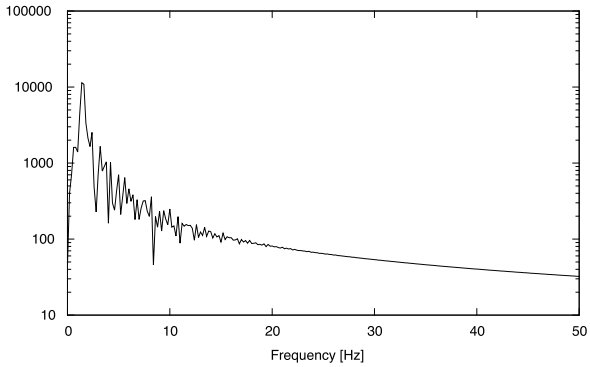

Fig. 6. The FFT spectrum of the Rössler attractor.

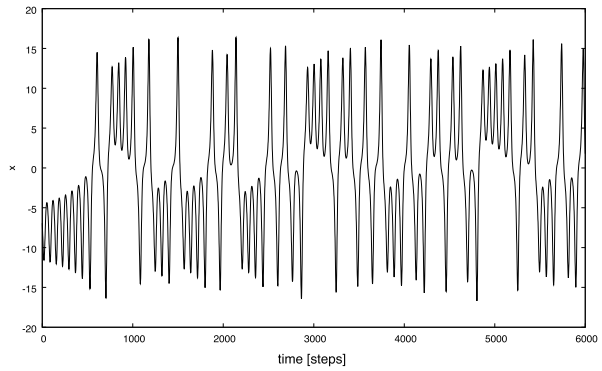

Fig. 7. The x coordinate of the Lorenz attractor.

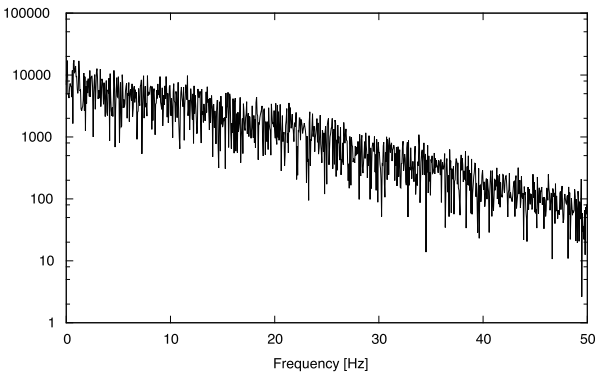

Fig. 8. The FFT spectrum of the Lorenz attractor.

The harmonic oscillator was chosen for investigating the relationship between the proposed criterion and the criterion of Perea and Harer. The harmonic oscillator is defined as

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -x, \end{cases} \tag{12}$$

where the period of the solution is 628 steps.

The van der Pol system [23] was chosen as an example of a periodic signal with high harmonics that can be expressed by Fourier series. The van der Pol system is defined as

$$\begin{cases} \dot{x} = y, \\ \dot{y} = \mu(1 - x^2)y - x, \end{cases} \tag{13}$$

where $\mu = 10$, and the period of the solution is 1908 steps.

The Rössler system [24] was chosen as an example of a chaotically periodic signal. The Rössler system is defined as

$$\begin{cases} \dot{x} = -y - z, \\ \dot{y} = x + ay, \\ \dot{z} = b + z(x - c), \end{cases} \tag{14}$$

where $a = 0.2$, $b = 0.2$, and $c = 5.7$. The solution converges to a strange attractor. The inverse of the strongest frequency of the FFT spectrum was used as the characteristic period, as the trajectory of the Rössler attractor has no strict period; however, it is nearly periodic. The characteristic period is 625 steps.

The Lorenz system [25] was chosen as another example of a chaotically periodic signal. The Lorenz system is defined as

$$\begin{cases} \dot{x} = s(y - z), \\ \dot{y} = x(r - z) - y, \\ \dot{z} = xy - bz, \end{cases} \tag{15}$$

where $r = 28$, $s = 10$, and $c = 8/3$. The solution converges to a strange attractor. The median of the interval that the x-coordinate value crosses 7.5 and that that value crosses -7.5 was used as the half of the characteristic period. The characteristic period is 58 steps.

The speech signal was chosen to test the application of the proposed criterion to a real world signal. The signal for the Japanese vowel /a/, spoken by a female speaker, was used. The data were obtained from the Vowel Database: Five Japanese Vowels of Males, Females, and Children Along with Relevant Physical Data (JVPD) [26]. This database is maintained by the Speech Resources Consortium of the National Institute for Informatics (`http://research.nii.ac.jp/src/en/JVPD.html`). The signal for Japanese vowel /a/ is shown in Fig. 1, and its FFT spectrum is shown in Fig. 2.

The solution of each dynamical system was calculated by the fourth order Runge-Kutta method with step size 0.001. It was evolved for sufficiently large time. The first 1000 samples of the solutions were dropped because the solutions that are not in the attractor should be discarded.

The first projection map, that is the map taking the first coordinate of the signal, was used as an observation function. The resulting time series data were used as input to the experiment.

5.2 Procedure

Each input time series data were first embedded into a delay-coordinate space to obtain a reconstructed attractor. The dimension of the delay-coordinate spaces ranged from 2 to 10. The delay time varied from one step to the step that was equal to half the period of the input signal. The first 5000 points of each attractor were used.

Subsequently, the first persistent homology of each reconstructed attractor was computed. The Ripser [27] software application was used for computation. The computer that was used had four Intel Xeon E5-4640 CPUs (2.40GHz, 8 cores) and 1TB RAM. The computation time for one attractor ranged from a few seconds to a day.

Finally, the graph of the most significant death value versus the delay was plotted for each embedding dimension of each reconstructed attractor. The most significant death value was extracted from the first persistent homology of each reconstructed attractor. The delay time was normalized by the period or the characteristic period.

6. Results

In this section, the results of the numerical experiments are presented.

Figure 9 shows the most significant death values of the harmonic oscillator. The first peaks are achieved at $a = \frac{1}{n}\frac{T}{2}$. This result coincides with the prediction of Perea and Harer.

Figure 10 shows the most significant death values of the limit cycle of the van der Pol system with parameter $\mu = 10$. It is clear that the first peaks are achieved at $a = \frac{1}{n}\frac{T}{2}$ except for the case of embedding dimension three, where the point around $a = \frac{1}{3}\frac{T}{2}$ is flattened; however, this point can be considered the maximum.

Figure 11 shows the most significant death values of the Rössler attractor. The ravine in the graphs is realized at a point slightly smaller than 1/2. At the points $a = \frac{1}{n}\frac{T}{2}$, the graphs exhibit saddle

points instead of peaks. The graphs increase after the point $a = \frac{1}{n}\frac{T}{2}$.

Figure 12 shows the most significant death values of the Lorenz attractor. There is no ravine such as in the case of the Rössler attractor. At the points $a = \frac{1}{n}\frac{T}{2}$, each value of the graphs is near to the first peak.

Figure 13 shows the most significant death values of the Japanese vowel /a/. The graphs have several ravines between $a/T = 0$ and $a/T = 0.5$. The points where $a = \frac{1}{n}\frac{T}{2}$ do not correspond to peaks. In fact, in the case $n = 3$, they fall near the minimum.

7. Discussion

The relationship between the proposed criterion and that of Perea and Harer is herein discussed in light of the experimental results.

7.1 Dynamical systems

The most significant death value yields results that coincide with those of Perea and Harer for the harmonic oscillator. Figure 9 clearly shows that the first peaks are attained at $a = \frac{1}{n}\frac{T}{2}$. This fact implies that the word "roundest" in Perea and Harer's study is equivalent to the maximum of the width of a circle.

The criterion suggested by Perea and Harer is appropriate for the limit cycle of the van der Pol system. This system is an example of a nonlinear periodic signal. As the period of zero-crossing points is half the period of the limit cycle, the given signal can be made to correspond to a sine wave with the period of the given signal. The absolute value of this sine wave is bounded by that of the given signal.

The reconstructed attractor does not always recover its original shape when the delay time maximizes the most significant death value. Figure 14 shows a reconstructed trajectory of the harmonic oscillator embedded in the delay-coordinate space with delay equal to 63 and embedding dimension equal to five, which makes the equation $a = \frac{1}{n}\frac{T}{2}$ hold. This reconstructed trajectory recovers its original shape. By contrast, Fig. 15 shows a reconstructed limit cycle of the van der Pol system embedded in the delay-coordinate space with delay equal to 191 and embedding dimension equal to five, which makes the equation $a = \frac{1}{n}\frac{T}{2}$ hold. This reconstructed trajectory does not recover its original shape. We have to explain why the reconstructed trajectory of the harmonic oscillator recovers its original shape but that of the limit cycle of the van der Pol system does not. The harmonic oscillator is a special case. It can recovers its original shape because the reconstructed ellipse that has the widest hole width is the circle.

For the Rössler attractor, it should be considered whether to use or not the criterion $a = \frac{1}{n}\frac{T}{2}$. In the results, the first peak of the graphs turns into a saddle. The maximum of the most significant death value is attained for larger delay time. Considering irrelevance, excessively large delay time results in unsatisfactory reconstruction. The reconstructed attractors are checked for irrelevance.

Two figures are examined to compare the delay satisfying the criterion with the delay attaining the maximum. Figure 16 shows a plot of the reconstructed Rössler attractor with embedding dimension equal to five and delay equal to 62. This delay time satisfies the equation $a = \frac{1}{n}\frac{T}{2}$. Figure 17 shows a plot of the reconstructed Rössler attractor with embedding dimension equal to five and delay equal to 198. This delay time attains the maximum of the graph shown in Fig. 11. It can be seen that Fig. 16 shows the reconstructed attractor whose shape resembles that of the original attractor. The trajectory of the reconstructed attractor shown in Fig. 17 is considerably tangled. Thus, it can be concluded that the delay satisfying the criterion $a = \frac{1}{n}\frac{T}{2}$ is preferable to the larger delay. This is natural because the aim was to maximize the hole width, and the original shape cannot be obtained from the observed signals. The reason why the MSDV of the Rössler attractor became large when the delay was large may be the irrelevance. The irrelevance deformed the reconstructed attractor and the hole width may be enlarged.

For the Lorenz attractor, the criterion $a = \frac{1}{n}\frac{T}{2}$ is almostly appropriate. The most significant values where the criterion holds are slightly smaller than the peaks in Fig. 12. This may be because the crossing intervals vary around the characteristic period. Moreover, the MSDVs get greater where

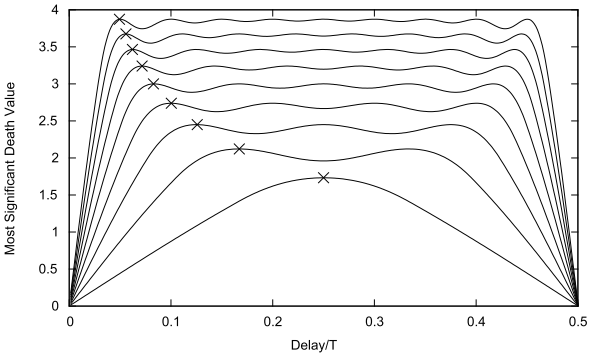

Fig. 9. Most significant death values of the harmonic oscillator. The horizontal axis represents the delay divided by the period $T = 628$, and the vertical axis the most significant death values. The lowest line is the case of embedding dimension two. The embedding dimension is in an increasing order. The x marks plotted in the graph represent the points where $a = \frac{1}{n}\frac{T}{2}$.

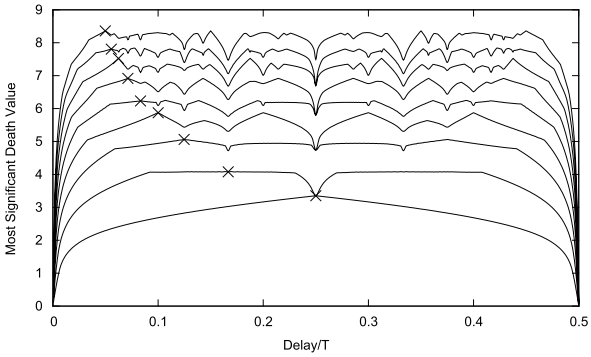

Fig. 10. Most significant death values of the van der Pol system with $\mu = 10$. The horizontal axis represents the delay divided by the period $T = 1908$, and the vertical axis the most significant death values. The lowest line is the case of embedding dimension two. The embedding dimension is in an increasing order. The x marks plotted in the graph represent the points where $a = \frac{1}{n}\frac{T}{2}$.

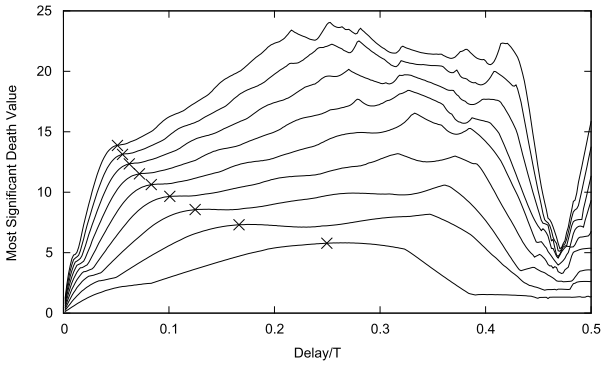

Fig. 11. Most significant death values of the Rössler attractor. The vertical axis represents the delay divided by the characteristic period $T = 625$, and the horizontal axis the most significant death values. The lowest line is the case of embedding dimension two. The embedding dimension is in an increasing order. The x marks plotted in the graph represent the points where $a = \frac{1}{n}\frac{T}{2}$.

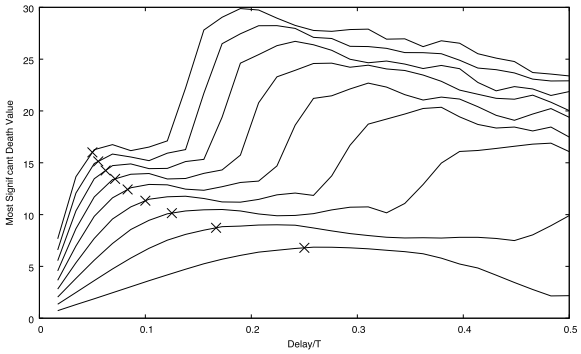

Fig. 12. Most significant death values of the Lorenz attractor. The vertical axis represents the delay divided by the characteristic period $T = 58$, and the horizontal axis the most significant death values. The lowest line is the case of embedding dimension two. The embedding dimension is in an increasing order. The x marks plotted in the graph represent the points where $a = \frac{1}{n}\frac{T}{2}$.

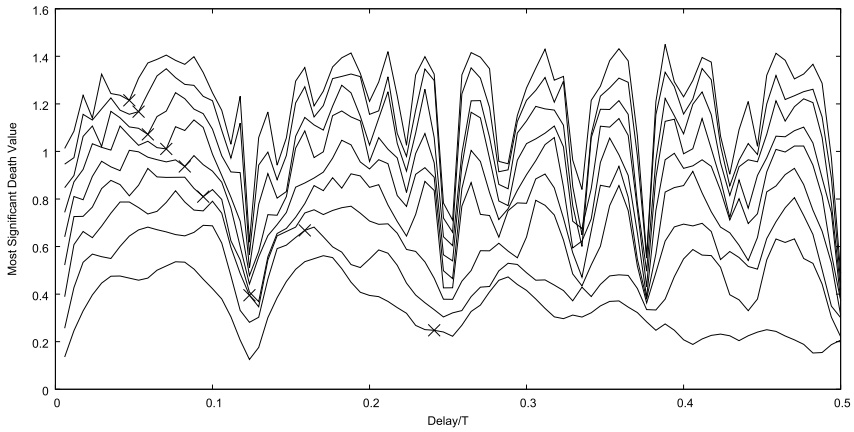

Fig. 13. Most significant death values of the Japanese Vowel /a/. The horizontal axis is the delay divided by the characteristic period $T = 170$, and the vertical axis the most significant death values. The lowest line is the case of embedding dimension two. The embedding dimension is in an increasing order. The x marks plotted in the graph represent the points where $a = \frac{1}{n}\frac{T}{2}$.

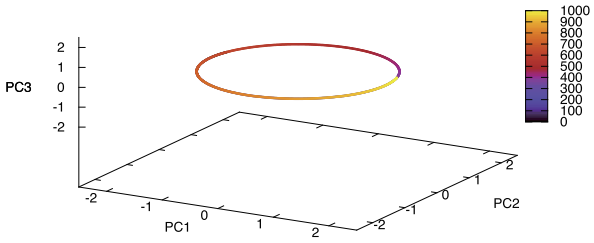

Fig. 14. A reconstructed trajectory of the harmonic oscillator with embedding dimension equal to five and delay equal to 62. The axes of this plot represent principal components. This delay was chosen to satisfy the equation $a = \frac{1}{n}\frac{T}{2}$.

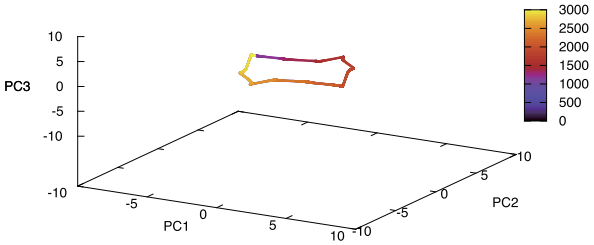

Fig. 15. A reconstructed limit cycle of the van der Pol system with embedding dimension equal to five and delay equal to 191. The axes of this plot represent principal components. This delay was chosen to satisfy the equation $a = \frac{1}{n}\frac{T}{2}$.

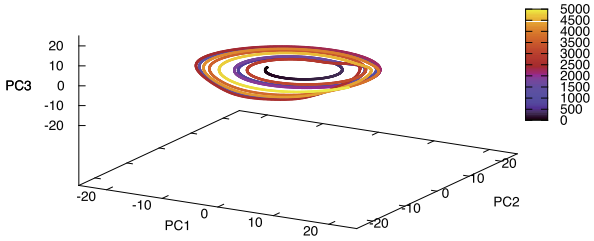

Fig. 16. A reconstructed Rössler attractor with embedding dimension equal to five and delay equal to 62. The axes of this plot represent principal components. This delay was chosen to satisfy the equation $a = \frac{1}{n}\frac{T}{2}$.

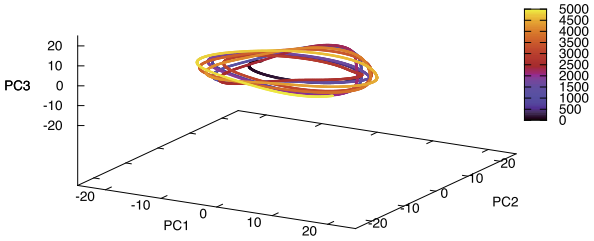

Fig. 17. A reconstructed Rössler attractor with embedding dimension equal to five and delay equal to 198. The axes of this plot represent principal components. This delay attains the first peak of the graph.

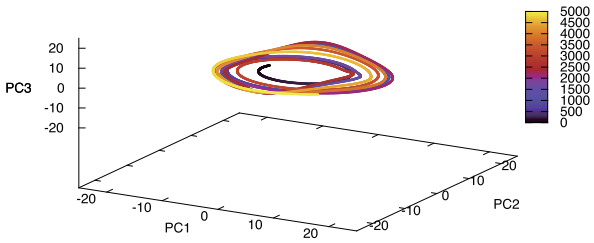

Fig. 18. A reconstructed Rössler attractor with embedding dimension equal to five and delay equal to 125. The axes of this plot represent principal components.

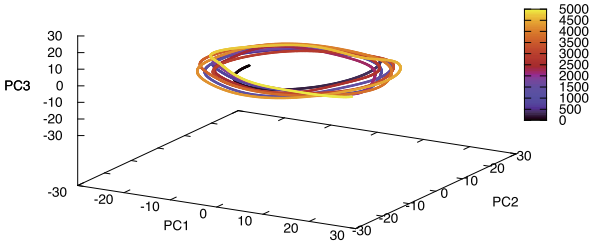

Fig. 19. A reconstructed Rössler attractor with embedding dimension equal to ten and delay equal to 125. The axes of this plot represent principal components.

the delay has greater values, except the $n = 2$ case. We check what happens in the case of larger delays. Figure 20 shows a plot of the reconstructed Lorenz attractor with embedding dimension equal to five and delay equal to 6. This delay time satisfies the equation $a = \frac{1}{n}\frac{T}{2}$. Figure 21 shows the reconstructed attractor where the delay attains the maximum of the MSDV of the $n = 5$ case. The graph of the MSDV is shown in Fig. 12 and this delay is 22. The shape shown in Fig. 20 resembles the original attractor, but the reconstructed attractor shown in Fig. 21 is extremely deformed. Thus, it can be concluded that the delay satisfying the criterion $a = \frac{1}{n}\frac{T}{2}$ is also preferable. In the case of the Lorenz attractor, the third hole appears when the delay is large. This is seen in Fig. 21. This third hole makes the MSDV higher. This may be caused because the number of the sampled points is small. The third hole will be filled up with the infinite length of samples.

7.2 Japanese vowel /a/

The graph of the most significant death values of the signal for the Japanese vowel /a/ has several ravines, as shown in Fig. 13. This is due to the fact that the human voice resonates in the mouth and the speech signal has higher harmonics. The eighth harmonic contributes to the first ravine in the graph. As the speech signal has several harmonics, it is difficult to determine a single period T for the criterion of Perea and Harer. By contrast, the most significant death value can determine the

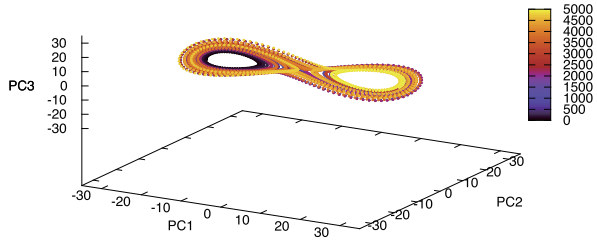

Fig. 20. A reconstructed Lorenz attractor with embedding dimension equal to five and delay equal to 6. The axes of this plot represent principal components. This delay was chosen to satisfy the equation $a = \frac{1}{n}\frac{T}{2}$.

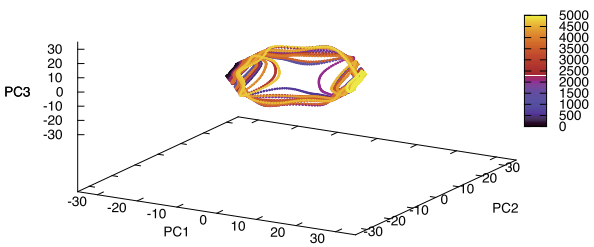

Fig. 21. A reconstructed Lorenz attractor with embedding dimension equal to five and delay equal to 22. The axes of this plot represent principal components. This delay was chosen to the delay attains the maximum of the MSDV.

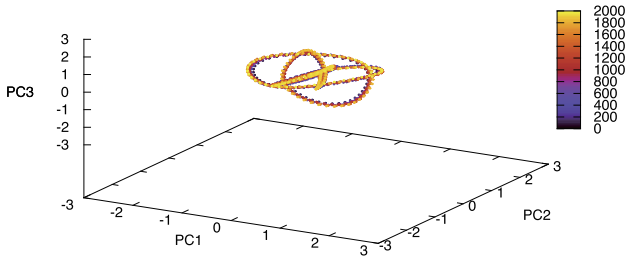

Fig. 22. A reconstructed attractor of Japanese vowel /a/. The embedding dimension was 10 and the delay was 12. The axes of this plot represent principal components. This delay was chosen to maximize the most significant death value.

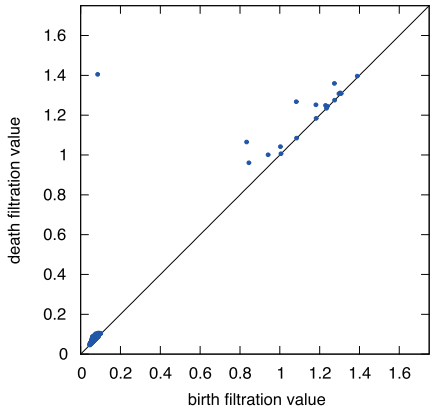

Fig. 23. First persistence diagram of a reconstructed attractor of the Japanese vowel /a/. The embedding dimension was 10 and the delay was 12. This delay was chosen to maximize the most significant death value. The horizontal axis represents the birth filtration value and the vertical axis the death filtration value.

delay time.

An example of reconstruction is shown in Fig. 22, and its persistence diagram is shown in Fig. 23. The embedding dimension is 10 and the delay is 12. This delay attains the maximum of the most significant death value in Fig. 13. It cannot be determined whether the reconstruction is satisfactory or not because the original is unknown. The persistence diagram suggests that the reconstructed attractor has one large hole.

7.3 Comparison with mutual information

The most significant death value is now compared with mutual information. Figure 24 shows the plot of mutual information versus delay for the harmonic oscillator. The first local minimum is marked on the graph. If the first local minimum is chosen as the delay, the reconstructed attractor collapses owing to redundancy. Similarly, the delay attaining the first local minimum of the mutual information is small for the limit cycle of the van der Pol system. Figure 25 shows the plot of the mutual information for this system. However, the global minimum of the mutual information of both systems is about $0.25T$ and this delay is suitable for the delay coordinate embedding where $n = 2$. Thus, it can be suggested that mutual information is not necessarily suitable for determining the delay of periodic signals in the cases where the embedding dimension is large.

As seen in Fig. 26, the mutual information suggests that the optimal delay is approximately $0.2T$. Figure 11 shows that this delay may be suitable for attractor reconstruction for embedding dimension equal to two or three. However, this delay may cause irrelevance for embedding dimension greater than three. We show the reconstructed attractors in higher dimension and with the delay $0.2T$. Figure 18 shows the reconstructed attractor of embedding dimension 5 with the delay 125, and Fig. 19 shows

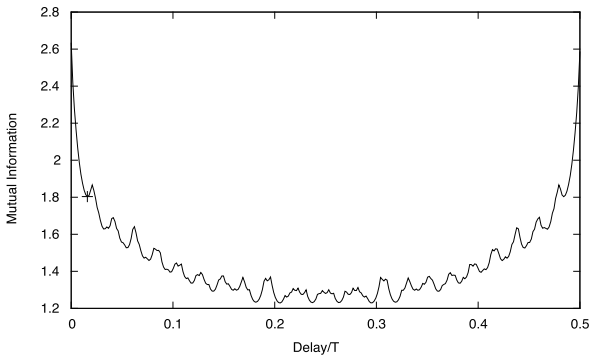

Fig. 24. Plot of mutual information versus delay for the harmonic oscillator. The horizontal axis represents the delay normalized by the period and the vertical axis the mutual information between $x(t)$ and $x(t+a)$. A "+" mark represents the first minimum on the graph.

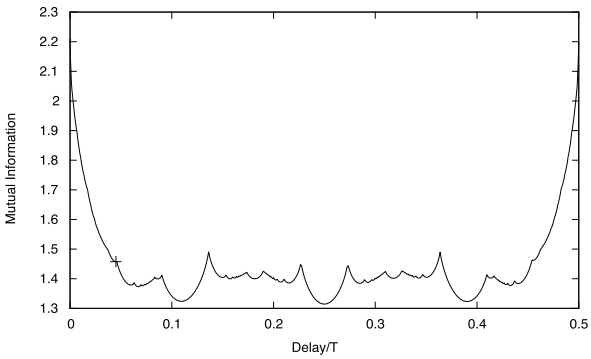

Fig. 25. Plot of mutual information versus delay for the van der Pol system with parameter $\mu = 10$. The horizontal axis represents the delay normalized by the period and the vertical axis the mutual information between $x(t)$ and $x(t+a)$. A "+" mark represents the first minimum on the graph.

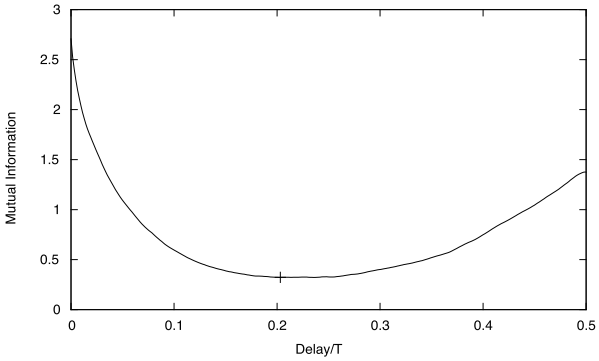

Fig. 26. Plot of mutual information versus delay for the Rössler attractor. The horizontal axis represents the delay normalized by the period and the vertical axis the mutual information between $x(t)$ and $x(t+a)$. A "+" mark represents the first minimum on the graph.

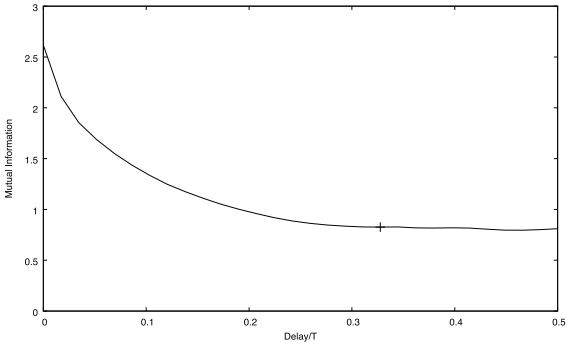

Fig. 27. Plot of mutual information versus delay for the Lorenz attractor. The horizontal axis represents the delay normalized by the period and the vertical axis the mutual information between $x(t)$ and $x(t+a)$. A "+" mark represents the first minimum on the graph.

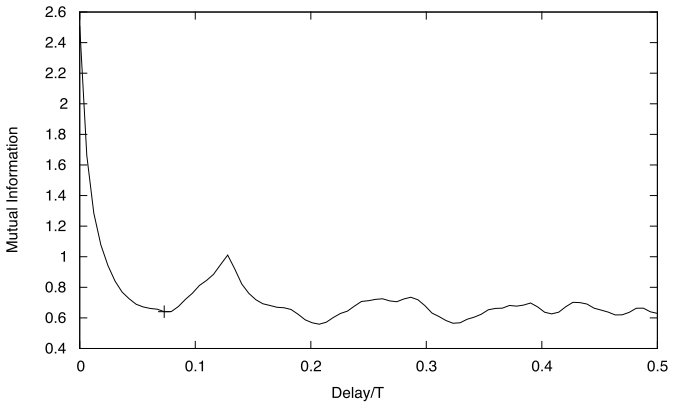

Fig. 28. Plot of mutual information versus delay for the Japanese vowel /a/. The horizontal axis represents the delay normalized by the period and the vertical axis the mutual information between $x(t)$ and $x(t+a)$. A "+" mark represents the first minimum on the graph.

the reconstructed attractor of embedding dimension 10 with the delay 125. This delay equals to $0.2T$. Altough in the five dimensional case the reconstructed attractor seems to be formed well, in the ten dimensional case the reconstructed attractor tangles and there occurs the irrelevance.

Figure 27 shows the mutual information of the Lorenz attractor. The first local minimum is about $0.34T$. This value is too large even if the embedding dimension is two. For these two systems, the

mutual information cannot estimate the appropriate delay for delay coordinate embeddings.

It can be seen in Fig. 28 that the optimal delay for the Japanese vowel /a/, as determined by the mutual information, is approximately $0.75T$. It is can be said that this delay is appropriate for attractor reconstruction, because the most significant death value is sufficiently large in Fig. 13.

Considering the discussion above, mutual information does not yield consistent results. The delay determined by mutual information is satisfactory for some systems and unsatisfactory for others. Although the mutual information is not so good, the maximum of the most significant death value yields too large delay for chaotic systems as seen in previous sections.

8. Conclusion

The most significant death value can determine the delay that maximizes the hole width of reconstructed attractors of several examples of periodic signals shown in this paper. However, for chaotically periodic signals, the maximum of the most significant death value yields the irrelevance. The criterion of Perea and Harer, that the delay a satisfies $a = \frac{1}{n}\frac{T}{2}$, is also good for the Rössler attractor and the Lorenz attractor. The variable T is the period or a characteristic period of the given signal. Nevertheless, the criterion of Perea and Harer is not suitable for the signal of Japanese vowel /a/, which is periodic with harmonics. For this signal, the maximum of the most significant death value yields good delay for attractor reconstruction. Compared with the most significant death value, mutual information was found not good for attractor reconstruction. It sometimes returns too small delay, and the delay is only suitable for too low-dimensional embeddings.

Multiple delays may be suitable for practical applications, as the optimal delay depends on the period of the signals, and the period cannot always be assumed. The vineyards [28], which is a continuous set of persistence diagrams, may be useful for applications. However, there is the problem of the computational cost of the persistent homology.

Acknowledgments

This work was supported by the University of Tokyo–NEC Future AI scholarship, and this work was supported by JST CREST, Grant Number JPMJCR14D2, Japan.

References

[1] J.A. Perea and J. Harer, "Sliding windows and persistence: an application of topological methods to signal analysis," *Foundations of Computational Mathematics*, vol. 15, no. 3, pp. 799–838, June 2015.

[2] F. Takens, "Detecting strange attractors in turbulence," *Dynamical Systems and Turbulence, Warwick 1980*, pp. 366–481, 1981.

[3] T. Sauer, J.A. Yorke, and M. Casdagli, "Embedology," *Journal of Statistical Physics*, vol. 65, no. 3, pp. 579–616, November 1991.

[4] M. Casdagli, S. Eubank, J.D. Farmer, and J. Gibson, "State space reconstruction in the presence of noise," *Physica D*, vol. 51, no. 1, pp. 52–98, August 1991.

[5] C.M.M. Pereira and R.F. de Mello, "Persistent homology for time series and spatial data clustering," *Expert Systems with Applications*, vol. 42, no. 15, pp. 6026–6038, September 2015.

[6] L.M. Seversky, S. Davis, and M. Berger, "On time-series topological data analysis: new data and opportunities," *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1014–1022, July 2016.

[7] V. Venkataraman, K.N. Ramamurthy, and P. Turaga, "Persistent homology of attractors for action recognition," *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 4150–4154, September 2016.

[8] J. Garland, E. Bradley, and J.D. Meiss, "Exploring the topology of dynamical reconstructions," *Physica D*, vol. 334, no. 1, pp. 49–59, November 2016.

[9] S. Maletić, Y. Zhao, and M. Rajković, "Persistent topological features of dynamical systems," *Chaos*, vol. 26, no. 5, 053105, May 2016.

[10] A.M. Fraser and H.L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 32, pp. 1134–1140, February 1986.
[11] W. Liebert, K. Pawelzik, and H.G. Schuster, "Optimal embeddings of chaotic attractors from topological considerations," *Europhysics Letters*, vol. 14, no. 6, pp. 521–526, March 1991.
[12] T.M. Buzug, T. Reimers, and G. Pfister, "Optimal reconstruction of strange attractors from purely geometrical arguments," *Europhysics Letters*, vol. 17, no. 7, pp. 605–610, December 1990.
[13] T.M. Buzug and G. Pfister, "Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors," *Physical Review A*, vol. 45, no. 10, pp. 7073–7084, May 1992.
[14] M.T. Rosenstein, J.J. Collins, and C.J. de Luca, "Reconstruction expansion as a geometry-based framework for choosing proper delay times," *Physica D*, vol. 73, no. 1, pp. 82–98, May 1994.
[15] K. Judd and A. Mees, "Embedding as a modeling problem," *Physica D*, vol. 120, no. 3, pp. 273–286, September 1998.
[16] M. Small and C.K. Tse, "Optimal embedding parameters: a modeling paradigm," *Physica D*, vol. 194, no. 3, pp. 283–296, July 2004.
[17] H. Edelsbrunner and J. Harer, "Persistent homology—a survey," *Contemporary Mathematics*, vol. 453, pp. 257–282, 2008.
[18] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," *Discrete & Computational Geometry*, vol. 28, no. 4, pp. 511–533, November 2002.
[19] H. Edelsbrunner and J.L. Harer, *Computational Topology: An Introduction*, American Mathematical Society, 2010.
[20] I. Tamura, *Topology*, Iwanami shoten (published in Japanese), 1972.
[21] A. Zomorodian, "Fast construction of the Vietoris-Rips complex," *Computer & Graphics*, vol. 34, no. 3, pp. 263–271, June 2010.
[22] R. Hegger, H. Kantz, and T. Schreiber, "Practical implementation of nonlinear time series methods: The TISEAN package," *Chaos*, vol. 9, no. 2, pp. 413–435, June 1999.
[23] S.H. Strogatz, *Nonlinear Dynamics and Chaos*, Westview Press, 2014.
[24] O.E. Rössler, "An equation for continuous chaos," *Physics Letter A*, vol. 57, no. 5, pp. 397–398, July 1976.
[25] E.N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 5, pp. 130–141, March 1963.
[26] G. Ohyama, T. Deguchi, and H. Kasuya, "Construction of Japanese vowel database uttered by native speakers over a wide range of age", *Proc. Spring Meeting of ASJ*, 2–P–15(a) 2011, [in Japanese]
[27] U. Bauer, *Ripser*, Visit https://github.com/Ripser/ripser for the latest version of Ripser.
[28] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov, "Vines and vineyards by updating persistence in linear time," *Proceedings of the Twenty-second Annual Symposium on Computational Geometry*, pp. 119–126, June 2006.