Visualizing high-dimensional time series data

Yoshito Hirata and Kazuyuki Aihara

Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
Email: yoshito@sat.t.u-tokyo.ac.jp

Abstract—We propose a method for visualizing high-dimensional time series data. In this method, we pile recurrence plots for various observables, and project this pile along a time axis and/or the axis of observables. We call this method layered recurrence plot. We show that the proposed method is useful for observing transitions in high-dimensional dynamics.

1. Introduction

Although the developments of measurement techniques now enable us to produce a lot of high-dimensional time series data, it is still difficult to understand intuitively the underlying dynamics. A common set of techniques for such a purpose include dimension reduction [1, 2, 3] and clustering [4]. Because the types of such techniques are currently limited, we would like to add another type of methods, which is the visualization of high-dimensional data.

In this manuscript, we propose how to use recurrence plots [5, 6] for visualizing high-dimensional time series data. First, we pile recurrence plots obtained from different observables. Then, we project this pile along a time axis or the axis of observables. We demonstrate using numerical examples that the proposed approach shows the properties of the underlying dynamics.

The rest of this manuscript is organized in the following way: In Section 2, we introduce recurrence plots more formally. In Section 3, we propose the pile of recurrence plots as a layered recurrence plot. In Section 4, we show some numerical examples. In Section 5, we conclude this manuscript.

2. Recurrence plots

2.1. Recurrence plots

A recurrence plot [5, 6] is a two-dimensional plot originally proposed for visualizing a time series data. Suppose that a time series $\{x(i) : i = 1, 2, \ldots, I\}$ is given. Let us denote a threshold by ϵ. Then a recurrence plot is defined as

$$R(i, j) = \begin{cases} 1, \text{if } d(x(i), x(j)) < \epsilon, \\ 0, \text{otherwise}. \end{cases} \quad (1)$$

If $R(i, j) = 1$, then we plot a point at (i, j). If $R(i, j) = 0$, then we do not plot anything at (i, j). This simple plot can show a lot of things. For example, we can calculate correlation dimension and correlation entropy by using recurrence plots [7, 8]; In addition, we learned that we can reproduce a rough shape of the original time series from a recurrence plot even if the original time series is given in a time series whose dimension is more than one [9]. Therefore, a recurrence plot eventually contains almost all information for the underlying dynamics except for the spatial scale. From this viewpoint, a recurrence plot is a nice tool for visualizing time series data.

The portion of places where a dot is plotted is called the recurrence rate [10]. We choose the threshold such that the recurrence rate becomes 20%.

2.2. Extensions of recurrence plots

There are two known extensions of recurrence plots to multivariate time series: The first extension is called cross recurrence plot, while the second extension is called joint recurrence plot.

Suppose that there are two time series $\{x_k(i) : i = 1, 2, \ldots, I\}$ and $\{x_l(j) : j = 1, 2, \ldots, J\}$ given. Then, a cross recurrence plot [11] is defined as

$$CR_{k,l}(i, j) = \begin{cases} 1, \text{if } d(x_k(i), x_l(j)) < \epsilon_{k,l}, \\ 0, \text{otherwise}. \end{cases} \quad (2)$$

To use the cross recurrence plot, two observables should have the same dimension and the similar values.

Suppose that there are two time series $\{x_k(i) : i = 1, 2, \ldots, I\}$ and $\{x_l(i) : i = 1, 2, \ldots, I\}$. This time, the length of two time series should be the same. Then, the other existing extension, a joint recurrence plot [12], is defined as

$$R_k(i, j) = \begin{cases} 1, \text{if } d(x_k(i), x_k(j)) < \epsilon_k, \\ 0, \text{otherwise}, \end{cases} \quad (3)$$

$$R_l(i, j) = \begin{cases} 1, \text{if } d(x_l(i), x_l(j)) < \epsilon_l, \\ 0, \text{otherwise}, \end{cases} \quad (4)$$

$$JR_{k,l}(i, j) = R_k(i, j)R_l(i, j). \quad (5)$$

In Ref. [6], the joint recurrence plot is further extended to a multivariate time series whose dimension is greater than 2 in the following way:

$$JR(i, j) = \prod_{k=1}^{K} R_k(i, j). \quad (6)$$

The problem of this definition is that $JR(i, j)$ becomes very sparse when K is large.

3. Layered recurrence plots

Alternatively, we define a layered recurrence plot as

$$R_k(i, j) = \begin{cases} 1, \text{if } d(x_k(i), x_k(j)) < \epsilon_k, \\ 0, \text{otherwise}, \end{cases} \tag{7}$$

for $k \in \{1, 2, \ldots, K\}$. The difference between recurrence plots and the layered recurrence plot is that we pile the set of recurrence plots on the top of them successively.

For the visualization, we apply some operation to marginalize j or k. One of such operations can be the product like defined in Eq. (6) (see Ref. [6]). Another possibility could be a sum over j or k. When we take a sum over k, then we call it a layered recurrence plot summed over space:

$$S(i, j) = \sum_k R_k(i, j). \tag{8}$$

Using this $S(i, j)$, we can show which time is close to which time, namely the strength of recurrence given a pair of times. The better point compared to Eq. (6) is that we can see some non-zero value if some of observables in some time are close to those in another time.

When we take a sum over j, then we call it a layered recurrence plot summed over time:

$$T(i, k) = \sum_j R_k(i, j). \tag{9}$$

Using this $T(i, k)$, we can show the relationship between time and observables, namely the strength of recurrence given a pair of time and spatial index. This way of looking at a time series is new for this manuscript.

4. Results

We show two examples of layered recurrence plots in this Section.

The first example is a set of coupled logistic maps [13]:

$$\begin{aligned} & y_n(t+1) \\ = \; & (1-2\eta)(3.8 y_n(t)(1-y_n(t))) \\ & +(\eta-\xi)(3.8 y_{n+1}(t)(1-y_{n+1}(t))) \\ & +(\eta+\xi)(3.8 y_{n-1}(t)(1-y_{n-1}(t))), \end{aligned} \tag{10}$$

where we define $y_{n+100}(t) = y_n(t)$, $\eta = 0.05$ and $\xi = 0.01$. We generated a time series of length 200 from this system.

A time series looks like one shown in Fig. 1.

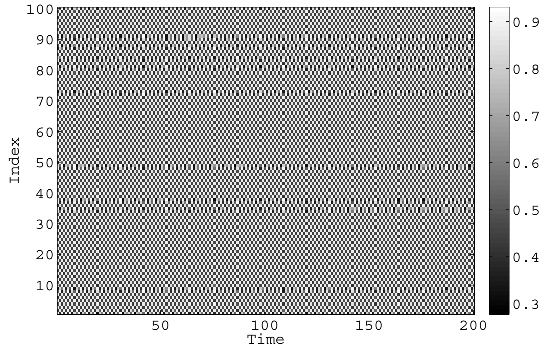

Figure 1: A time series of coupled logistic maps.

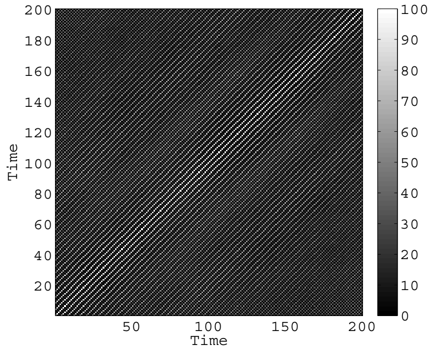

Figure 2: The layered recurrence plot summed over space for the time series of the coupled logistic maps shown in Fig. 1.

First, we apply Eq. (8). The results are shown in Fig. 2. Here, whiter points have recurrences for more observables. We can clearly see the pseudo-periodicity of the underlying dynamics although this pseudo-periodicity is not apparent in the time series shown in Fig. 1. This tendency is shown more clearly in Fig. 2 than the recurrence plot for the original high-dimensional time series (see Fig. 3). We can see from Fig. 2 that the time series is nonstationary because the graph is darker at the top left and the bottom right corners.

Second, we apply Eq. (9). The results are shown in Fig. 4. We can observe in Fig. 4 that there are some travelling waves from smaller indices to the larger around the indices of 20, 65, and 95. These travelling waves cannot be seen clearly in the time series shown in Fig. 1.

Thus, this example of the coupled logistic maps demonstrated that the layered recurrence plot is good at emphasizing small tendencies hidden within the given time series.

We look at the Lorenz'96 model [14, 15] as our second example. The Lorenz'96 model is described as follows:

$$\frac{du_i}{dt} = u_{i-1}(u_{i+1} - u_{i-2}) - u_i + F - \frac{h_u c}{b} \sum_{j=1}^{J} v_{j,i}, \tag{11}$$

$$\frac{dv_{j,i}}{dt} = cbv_{j+1,i}(v_{j-1,i} - v_{j+2,i}) - cv_{j,i} + \frac{h_v c}{b} u_i, \tag{12}$$

$$u_{I+i} = u_i, v_{j+J,i} = v_{j,i+1}, v_{j-J,i} = v_{j,i-1}, \tag{13}$$

where we set $I = 40$, $J = 5$, $F = 8$, $b = 10$, $c = 10$, $h_u = 1$, and $h_v = 1$. We assume that we can observe $v_{j,i}$ for $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$. Hence, the time series is now 200-dimensional.

A time series is shown in Fig. 5. We can sightly see the travelling waves in this figure.

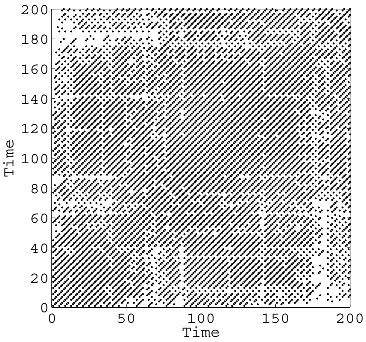

Figure 3: The recurrence plot for the original high-dimensional time series.

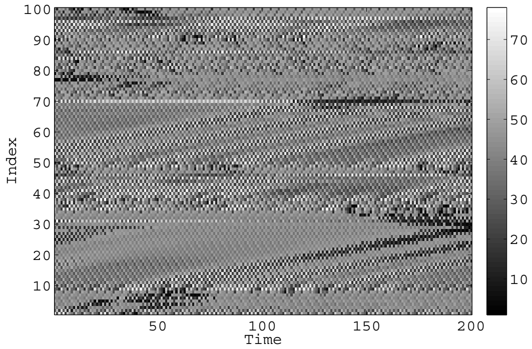

Figure 4: The layered recurrence plot summed over time for the time series of the coupled logistic maps shown in Fig. 1.

A layered recurrence plot summed over space is shown in Fig. 6. We can also see in this figure that there is the pseudo-periodicity. But, this time, diagonal whiter lines tend to be interrupted by blacker spaces. This interruption corresponds to sensitive dependence on initial conditions [16]. The similar dynamical tendency is also seen in a recurrence plot for the original 200-dimensional time series (Fig. 7).

A layered recurrence plot summed over time is shown in Fig. 8. In this figure, we can more clearly see that there are some travelling waves than Fig. 5.

5. Conclusions

We have proposed a way to visualize a high-dimensional time series using recurrence plots, which is called a layered recurrence plot. When we pile recurrence plots obtained from various observables and apply some operation to marginalize the axis of space, then we can see the pseudo-periodicity and sensitive dependence on initial conditions of given time series data. When we apply some operation to marginalize the axis of time, we can show the relationship between the time and the observables. Especially, travelling waves can be easily observed. We believe that the proposed method help to understand the complicated dynamics hidden in a high-dimensional time series.

The methods seem to work for high-dimensional time series obtained from wide contexts, although we should accumulate emperical evidnece for their efficacy in the future. We also will investigate the dependences on the parameters we used in our future communication.

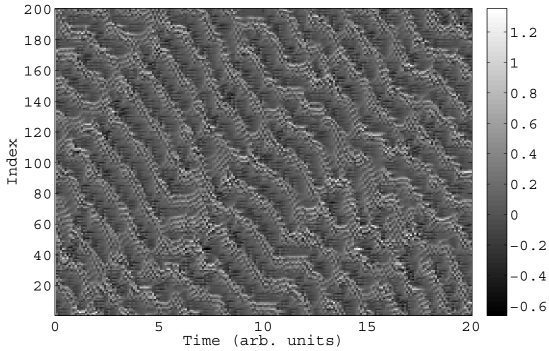

Figure 5: A time series $v_{j,i}$ of Lorenz'96 model.

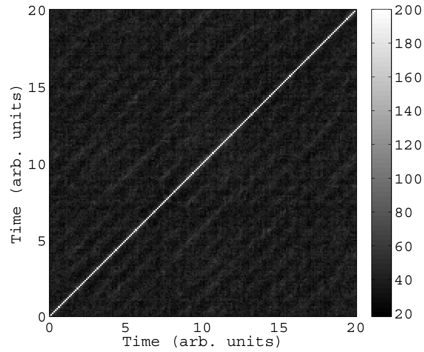

Figure 6: A layered recurrence plot summed over space for the time series of Lorenz'96 model shown in Fig. 5.

Acknowledgments

This work was supported by Platform for Dynamic Approaches to Living System from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] J. C. Gower, "Some distance properties of latent root and vector methods used in multivariate analysis," *Biometrika*, vol.53, pp.325–338, 1966.

[2] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol.290, pp.2319–2323, 2000.

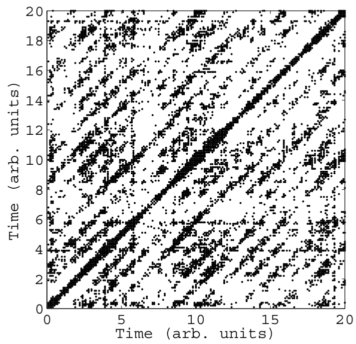

Figure 7: A recurrence plot for the original 200-dimensional time series of Lorenz'96 model.

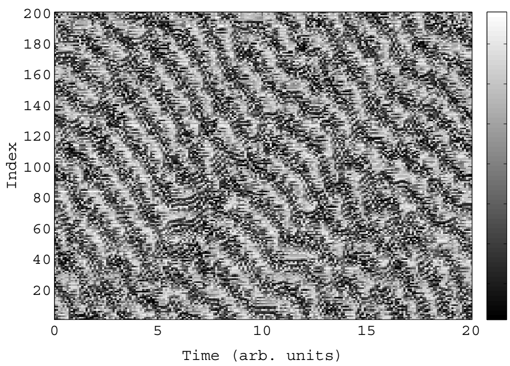

Figure 8: A layered recurrence plot summed over time for the time series of Lorenz'96 model shown in Fig. 5.

[3] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol.290, pp.2323–2326, 2000.

[4] J. MacQueen, "Some methods for classification and analysis of multivariate observations," In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (University of California Press, Berkeley, USA, 1967), pp.281–297.

[5] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhys. Lett.*, vol.5, pp.973–977, 1987.

[6] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Phys. Rep.*, vol.438, pp.237–329, 2007.

[7] P. Faure and H. Korn, "A new method to estimate the Kolmogorov entropy from recurrence plots: its application to neuronal signals," *Physica D*, vol.122, pp.265–279, 1998.

[8] M. Thiel, M. C. Romano, P. L. Read, and J. Kurths, "Estimation of dynamical invariants without embedding by recurrence plots," *Chaos*, vol.14, pp.234–243, 2004.

[9] Y. Hirata, S. Horai, and K. Aihara, "Reproduction of distance matrices and original time series from recurrence plots and their applications," *Eur. Phys. J. Special Topics*, vol.164, pp.13–22, 2008.

[10] J. P. Zbilut and C. L. Webber Jr., "Embeddings and delays as derived from quantification of recurrence plots," *Phys. Lett. A*, vol.171, pp.199–203, 1992.

[11] J. P. Zbilut, A. Giuliani, and C. L. Webber Jr. "Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification," *Phys. Lett. A*, vol.246, pp.122–128, 1998.

[12] M. C. Romano, M. Thiel, J. Kurths, and W. von Bloh, "Multivariate recurrence plots," *Phys. Lett. A*, vol.330, pp.214–223, 2004.

[13] K. Kaneko, "Period-doubling of kink-antikink patterns, quasiperiodicity in antiferro-like structures and spatial intermittency in coupled logistic lattice: Towards a prelude of a 'Field Theory of Chaos'," *Progress of Theoretical Physics*, vol.72, pp.480–486, 1984.

[14] E. N. Lorenz, "Predictability: A problem partly solved," in *Proc. ECMWF Seminar on Predictability*, Reading, UK, 1995, pp.1–18.

[15] J. A. Hansen and L. A. Smith, "The role of operational constraints in selecting supplementary observations," *J. Atmos. Sci.*, vol.57, pp.2859–2871, 2000.

[16] Y. Hirata and K. Aihara, "Devaney's chaos on recurrence plots," *Phys. Rev. E*, vol.82, art no.036209, 2010.