Physics Letters A 381 (2017) 604–615

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Robust reconstruction of a signal from its unthresholded recurrence plot subject to disturbances

Aloys Sipers [a,*], Paul Borm [a], Ralf Peeters [b]

[a] Department of Bèta Sciences and Technology, Zuyd University, The Netherlands
[b] Department of Data Science and Knowledge Engineering, Maastricht University, The Netherlands

ARTICLE INFO

Article history:
Received 14 March 2016
Received in revised form 10 December 2016
Accepted 12 December 2016
Available online 15 December 2016
Communicated by C.R. Doering

Keywords:
Recurrence plot
Graph theory
Robust reconstruction
Disturbances
Singular value decomposition

ABSTRACT

To make valid inferences from recurrence plots for time-delay embedded signals, two underlying key questions are: (1) to what extent does an unthresholded recurrence (URP) plot carry the same information as the signal that generated it, and (2) how does the information change when the URP gets distorted. We studied the first question in our earlier work [1], where it was shown that the URP admits the reconstruction of the underlying signal (up to its mean and a choice of sign) if and only if an associated graph is connected. Here we refine this result and we give an explicit condition in terms of the embedding parameters and the discrete Fourier spectrum of the URP. We also develop a method for the reconstruction of the underlying signal which overcomes several drawbacks that earlier approaches had. To address the second question we investigate robustness of the proposed reconstruction method under disturbances. We carry out two simulation experiments which are characterized by a broad band and a narrow band spectrum respectively. For each experiment we choose a noise level and two different pairs of embedding parameters. The conventional binary recurrence plot (RP) is obtained from the URP by thresholding and zero-one conversion, which can be viewed as severe distortion acting on the URP. Typically the reconstruction of the underlying signal from an RP is expected to be rather inaccurate. However, by introducing the concept of a multi-level recurrence plot (MRP) we propose to bridge the information gap between the URP and the RP, while still achieving a high data compression rate. We demonstrate the working of the proposed reconstruction procedure on MRPs, indicating that MRPs with just a few discretization levels can usually capture signal properties and morphologies more accurately than conventional RPs.

1. Introduction

One of the well-established techniques for analyzing a complex dynamical system is the recurrence plot (RP), see [2]. For a dynamical system a recurrence is said to occur whenever its trajectory in phase space is (nearly) self-intersecting. Examination of the recurrence pattern that occurs over an observed time interval, often provides useful information – not just when trajectories are somehow smooth and well predictable, but also in the case of complex dynamics and chaos. In these cases, recurrence quantification analysis (RQA), see [3], offers various tools to extract such information from RPs.

RPs can also be used to analyze an observed *scalar* signal $x(t)$. To do so, it is common to first embed it into some M-dimensional space by constructing a vector of time-delayed copies $X(t) := (x(t), x(t+\tau), \ldots, x(t+(M-1)\tau))^T$. This is known as the time-delay embedding method, first introduced in [4]. Note that the vector $X(t)$ carries information on the recent history of the signal $x(t)$ and captures its local morphology. In the literature there exist several other modeling and signal processing techniques, such as AR-modeling [5] and singular spectral analysis [6], which also interpret $X(t)$ as a state vector of the dynamical process generating $x(t)$, but we will not go into them here.

The *unthresholded* recurrence plot (URP) is defined as the bivariate distance function $\mathrm{URP}_X(u,v) := \|X(u) - X(v)\|$, for some chosen norm $\|\cdot\|$. From it, the (binary) RP is obtained by choosing a threshold value $\varepsilon > 0$ and defining $\mathrm{RP}^{\varepsilon}_X(u,v) := \Theta(\varepsilon - \mathrm{URP}_X(u,v))$. Here $\Theta(\cdot)$ denotes the Heaviside step function, given by $\Theta(x) = 1$ for $x \geq 0$ and $\Theta(x) = 0$ for $x < 0$. By construction, $\mathrm{RP}^{\varepsilon}_X(u,v) = 1$ (and it is zero otherwise) if and only if the vectors $X(u)$ and $X(v)$ differ by at most ε, as measured by the norm $\|\cdot\|$. This qualifies $X(v)$ as a near recurrence of $X(u)$ and motivates the terminology used.

* Corresponding author.
E-mail addresses: aloys.sipers@zuyd.nl (A. Sipers), paul.borm@zuyd.nl (P. Borm), ralf.peeters@maastrichtuniversity.nl (R. Peeters).

http://dx.doi.org/10.1016/j.physleta.2016.12.028
0375-9601/

When inferences are made about a signal $x(t)$ from its RP, it is important to understand to what extent an RP carries information that is unique for the underlying signal. This has been studied by several authors; see, e.g., [7], [8] and [9]. In [1], we have investigated this question in detail for the URP of a time-delay embedded signal. There we used the Euclidean norm to define the URP and we restricted the discussion to zero mean periodic signals which admit a Fourier series representation and have finite power. We will do the same in the present paper. This is a fairly large class of signals which includes many signals encountered in practice, such as digitally sampled measurement signals. The zero mean requirement is natural, as the URP does not carry any information on the mean of $x(t)$. It is easy to see that it does not carry information on the sign of $x(t)$ either. Periodicity (with a known period, which we normalize to be 1) allows us to deal with finite interval effects in a convenient way.

In [1, Theorem 3.4] we characterized uniqueness of the underlying signal $x(t)$ for a given URP (up to a sign ± 1) by means of connectedness of a simple undirected graph $G_X = (K_X, E)$. For this graph, the node set K_X corresponds to the discrete Fourier frequencies that occur in the signal $x(t)$. The edge set E is determined by the choice of embedding dimension M and the time-delay τ. It was found that the (discrete Fourier) power spectrum of $x(t)$ can always be reconstructed from URP_X, but it generally depends on the embedding parameters M and τ which other information can also be retrieved. Several special cases were analyzed, see [1, Corollary 3.5]. However, the approach has two main shortcomings: (1) The characterization of unique reconstructibility of $x(t)$ from URP_X (up to a sign) by connectedness of G_X, is implicit. An explicit characterization in terms of K_X, M and τ is lacking for the general case but would be useful to have: to be able to assess whether a given URP is maximally informative (i.e., the URP determines the underlying de-averaged signal up to a sign) about its underlying signal, and to help selecting embedding parameters M and τ to ensure such a property when generating URPs and RPs. (2) If G_X is connected, then the actual reconstruction of $x(t)$ from URP_X in [1] is not very practical – nor recommendable from a computational point of view. It involves several double integrals which may be hard to compute accurately, and in general it requires the search for suitable paths in the graph G_X to connect selected nodes.

In Section 2 we address and resolve the first issue. We present a new and explicit characterization of connectedness of G_X and, in the course of deriving it, we establish that the diameter of a connected graph G_X always equals 1 or 2. This new characterization shows for each choice of M precisely which values of τ will cause disconnectedness. For such values, the URP will not be maximally informative and in choosing the embedding parameters we typically want to avoid them. We illustrate this result with a couple of examples.

In Section 3 we address the second issue. The 2D-Fourier transform of URP_X^2 is found to have coefficients with a special structure that can be exploited to compute $x(t)$. The coefficients are used to calculate a square matrix $\tilde{I}_X$ which is the Hadamard (i.e., entry-wise) product $\tilde{I}_X = W \odot T$ of a rank-one matrix W (determined by the non-zero Fourier coefficients of $x(t)$ only) and a known matrix T (depending entirely on M and τ). The matrix T allows for the determination of connectedness of G_X. If connectedness holds, then, using the property that the diameter is at most 2, T can be factored from $\tilde{I}_X$ to obtain W. With singular value decomposition (SVD), the Fourier coefficients of the signal $x(t)$ can then be computed from W in a robust way.

In Section 4 we present a couple of simulation experiments to demonstrate this procedure. We analyze robustness against noise of the reconstruction procedure, by adding noise of a few levels to a URP before reconstruction. We also investigate sensitivity of the reconstruction near values of τ for which disconnectedness of G_X occurs.

Third, in Section 5 we perform experiments to study the impact of truncation on URPs, such as applied when computing RPs. This fosters the idea of using 'multi-level recurrence plots' (MRPs) instead of RPs, preserving more information while still achieving a high compression rate compared to URPs.

Section 6 concludes the paper, with a discussion of the results obtained. All the proofs are collected in Appendix A.

2. Connectedness of the graph G_X

We consider the class of real zero mean periodic signals $x(t)$ with period 1, which are square integrable on $[0, 1)$. Any such signal admits a Fourier series representation, see [10, Sect. 4.26], denoted by:

$$x(t) = \sum_{k \in \mathbb{Z}} c_k e^{2\pi k t i}. \tag{1}$$

As the mean is zero $c_0 = 0$, and as $x(t)$ is real the complex-valued Fourier coefficients c_k satisfy $c_{-k} = \overline{c_k}$ for all $k \in \mathbb{Z}$. Square integrability gives $\sum_{k \in \mathbb{Z}} |c_k|^2 < \infty$. Convergence of the Fourier series on the right hand side to the function $x(t)$ happens in the L^2-sense; it happens pointwise under additional smoothness conditions, see [10, Ch. 5]. Note that discrete-time (regularly sampled) signals with a finite number of observations can also be included in this set-up, as they admit finite Fourier series representations and can be associated with continuous-time periodic signals through interpolation and periodic extension.

For a given choice of embedding dimension M and time-delay $\tau \in (0, 1)$, let $X(t)$ be the corresponding time-delay embedded trajectory in $\mathbb{R}^M$ and $\text{URP}_X(u, v) = \|X(u) - X(v)\|$ the unthresholded recurrence plot, in which $\|\cdot\|$ denotes the Euclidean norm. Note that $X(t)$ is also periodic with period 1 and given by:

$$X(t) = \sum_{k \in \mathbb{Z}} c_k e^{2\pi k t i} T_k, \tag{2}$$

in which

$$T_k = \begin{pmatrix} 1 \\ z^k \\ \vdots \\ z^{(M-1)k} \end{pmatrix}, \quad \text{with } z = e^{2\pi \tau i}. \tag{3}$$

As in [1], the associated (simple, undirected) graph G_X is defined to have the integer labeled nodes $K_X = \{k \in \mathbb{Z} \mid c_k \neq 0\}$. Two distinct nodes $p, q \in K_X$ are defined to be adjacent if and only if the L^2-inner product $\langle T_p, T_q \rangle$ is nonzero. The latter is equivalently characterized by $[(p-q)\tau \in \mathbb{Z}] \vee [(p-q)M\tau \notin \mathbb{Z}]$; see [1, Lemma 3.1].

From [1, Theorem 3.4], we have that the signal $x(t)$ can be reconstructed (up to a sign ± 1) from URP_X if and only if the graph G_X is connected. In [1, Corollary 3.5] several special cases are addressed for which it is possible to quickly decide on connectedness of the graph G_X. Below, we extend these results by providing a complete and explicit characterization of connectedness of G_X. We also show that if G_X is connected, then the diameter of G_X is either 1 or 2. This means that a connected graph G_X is either complete ($\text{diam}(G_X) = 1$) or every two nodes are connected by a path of length at most 2.

Theorem 2.1. *Let $x(t)$ be a non-zero signal from the class introduced above, and $G_X = (K_X, E)$ the associated graph for a given choice of embedding dimension M and time-delay $\tau \in (0, 1)$. For the associated set of node indices K_X, define the set of ordered node pairs as*