From fuzzy recurrence plots to scalable recurrence networks of time series

TUAN D. PHAM

Department of Biomedical Engineering, Linkoping University - 581 83 Linkoping, Sweden

The extraction of useful features of nonlinear dynamic systems in time series with the use of network analysis has been increasingly attracting the attention of various research communities. Methods for network analysis of univariate time series have been widely applied and recently reviewed in [1,2], including the pioneering work of complex network analysis of pseudoperiodic time series [3], visibility graphs [4], recurrence networks [5], and ordinal partition networks [6]. Applications of network analysis of time series have covered many areas, typically such as stock markets [7], autistic spectrum disorder [8], traffic management and control in intelligent transportation systems [9], climate change [9–12], geoscience [13], rock fracture mechanics [14], and information retrieval for the diagnostics of dry-eye disease [15].

Among major methods for network analysis of nonlinear time series, the recurrence network approach is an important tool for studying this type of complex systems [2]. However, a problem for the analysis of recurrence networks is the handling of very large networks with respect to the computational requirements in terms of computing time and memory storage. Therefore, it is of interest to develop some theoretical model that enables scalable recurrence networks. The construction of recurrence networks from fuzzy recurrence plots [16], which is to be presented in this paper, can provide a solution to this demand.

The notion of a fuzzy recurrence plot is introduced with the definition of a recurrence plot (RP) as follows. Let \( X = \{x\} \) be a set of phase-space states, in which \( x \) is the \( i \)-th state of a dynamical system in \( m \)-dimensional space. A recurrence plot (RP) is an \( N \times N \) matrix in which an element \( (i, j) \), \( i = 1, \ldots, N \), \( j = 1, \ldots, N \), is represented with a black dot if \( x_i \) and \( x_j \) are considered to be closed to each other [17]. A symmetrical RP can be defined as [18]

\[
R_{ij} = H(\epsilon - \|x_i - x_j\|),
\]

where \( R_{ij} \) is an element \( (i, j) \) of the recurrence matrix \( R \), \( \epsilon \) is a similarity threshold, \( H(\cdot) \) is the Heaviside step function that yields either 0 or 1 if \( (\epsilon - \|x_i - x_j\|) < 0 \) or otherwise, respectively.

A recurrence network can be defined in terms of its adjacency matrix as follows [11]:

\[
A_{ij} = R_{ij} - \delta_{ij},
\]

where \( A_{ij} \) is an element \( (i, j) \) of the binary adjacency matrix, and \( \delta_{ij} \) is the Kronecker delta. The adjacency matrix of a recurrence network can be alternatively expressed as

\[
A = R - I,
\]

where \( A \) and \( I \) are the adjacency and identity matrices, both of size \( N \times N \), respectively.

A fuzzy recurrence plot (FRP) [16] is described as follows. Let \( X = \{x\} \), and \( V = \{v\} \) be the sets of phase-space states and fuzzy clusters of the states, respectively. Fuzzy clusters are defined here as classes or prototypes to which the phase-space states belong with different
grades of membership (technical details about how the fuzzy membership grades can be computed by minimizing the objective function of the fuzzy c-means algorithm (FCM) [19] were described in [16]). A fuzzy relation $R$ from $X$ to $V$ is a fuzzy subset of $X \times V$ characterized by a fuzzy membership function $\mu \in [0, 1]$. This fuzzy membership grade expresses the degree of similarity of each pair $(x, v)$ in $R$ that has the following properties [20]:

- $\mu(x, x) = 1, \forall x \in X$ (reflexive);
- $\mu(x, v) = \mu(v, x), \forall x \in X, \forall v \in V$ (symmetrical);
- $\mu(x, z) = \vee_{v}[\mu(x, v) \wedge \mu(v, z)], \forall x \in X, \forall z \in Z$ (transitive), where the symbols $\vee$ and $\wedge$ stand for max and min, respectively.

The similarity between the phase-space states and their fuzzy cluster centers can be expressed as the fuzzy membership grades indicating different degrees at which the phase-space states belong to different fuzzy cluster centers, computed using the FCM algorithm, based on which the similarity between the pairs of the phase-space states can be inferred using the max-min operator (transitivity) of a fuzzy relation to construct the fuzzy recurrence matrix (FRM). It should be pointed out that the construction of an FRP is different from the recently developed coarse graining of time series by introducing the recurrence plot of recurrence plots [21], where the latter method uses the distance plot to define unthresholded short-term RPs at the lower layer, and then computes the distances between every pair of the short-term RPs to construct a recurrence plot at the higher layer.

An FRM displays the recurrences of the phase-space trajectory as a grayscale image whose intensity values are the complements of the fuzzy membership grades, where darker pixels indicate stronger magnitudes of recurrence. An $\alpha$-cut FRP yields a defuzzified or binary recurrence matrix, denoted as $FRM_\alpha$, which can be mathematically expressed as

$$FRM_\alpha = \begin{cases} 0, & \mu(x, z) \geq \alpha, \\ 1, & \text{otherwise}, \end{cases}$$

\hspace{1cm} (4)

where $\alpha \in [0, 1]$, and 0 and 1 indicate black and white pixels, respectively.

When being represented as a grayscale image, several features of an FRP can be readily extracted in the context of texture analysis, such as the gray-level co-occurrence matrix (GLCM) [22]. The value of a component of the GLCM as a function of distance $h$ is defined as

$$c_h(p, q) = \sum_{(u, v) | u = p \oplus v = q} n(h), \forall p, q \in \mathcal{L},$$

\hspace{1cm} (5)

in which $f_u$ and $f_v$ are pixels at locations $u$ and $v$ and have intensity values $p$ and $q$, respectively, which are separated by the lag $h$, $\oplus$ stands for the logical AND operator, $\mathcal{L}$ is the set of the image intensity levels, and $n(h)$ is the total number of pairs of pixels offset by $h$.

The probability of the co-occurrence of $p$ and $q$ with respect to $h$ is

$$p_h(p, q) = \frac{c_h(p, q)}{n(h)},$$

\hspace{1cm} (6)

where $n(h)$ is the total number of pairs of pixels offset by $h$.

Using the image intensities and probabilities, GLCM-based texture features can be extracted from an FRP, including contrast, correlation, energy, homogeneity, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measures of correlation, autocorrelation, dissimilarity, cluster prominence, cluster shade, maximum probability, inverse difference, inverse difference normalized, and inverse difference moment normalized [23].

As another measure of texture, the experimental semi-variogram for an image is defined as [24]

$$\gamma(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} [f(x_i) - f(x_i + h)]^2,$$

\hspace{1cm} (7)

where $f(x_i)$ is the image intensity at $x_i$, $h$ is a distance, and $m(h)$ is the total number of pairs of pixels separated by $h$.

Being analogous to the definition of the adjacency matrix of an undirected and unweighted recurrence network [5], the adjacency matrix of an undirected and unweighted $\alpha$-cut recurrence network, which is a defuzzified (binary) adjacency matrix and denoted as $R_\alpha$, can be defined in terms of an $FRM_\alpha$ as

$$R_\alpha = FRM_\alpha - I.$$

\hspace{1cm} (8)

The notion of an undirected and unweighted $\alpha$-cut (defuzzified) recurrence network has been defined from an FRP. In other words, a defuzzified recurrence network is a binary recurrence network that is constructed from the $\alpha$-cut FRM, which no longer holds the fuzzy membership values between 0 and 1. We next discuss its scalability issue. Let $\tau$ and $m$ be the time lag or reconstruction delay, and the embedding dimension, respectively. The number of the phase-space state of the dynamical system of $N$-point time series is $G = N - (m - 1)\tau$. It is obvious that, for a large value of $N$, a recurrence network will have a large number of vertices that is proportional to $G$ if the adjacency matrix is not sparse. Such a situation makes it computationally expensive or impractical for computing the recurrence network properties with respect to the computational complexity. The scalability of a defuzzified adjacency matrix of an undirected $\alpha$-cut recurrence network can be obtained by replacing the phase-space states with a number of phase-space prototypes, denoted as $c$, that are the number of fuzzy cluster centers and can be much smaller than $G$. Again by using the fuzzy relation
Scalable recurrence networks

Fig. 1: Fuzzy recurrence plots of the logistic map with \( a = 3.83 \) ((a) and (b)), \( a = 3.679 \) ((c) and (d)), \( a = 3.791 \) ((e) and (f)), and \( a = 4 \) ((g) and (h)).

to infer the fuzzy membership grades of similarity between the cluster pairs \((\phi, \theta) \in V\), we have

\[ \mu(\phi, \phi) = 1, \quad \forall \phi \in V; \]
\[ \mu(\phi, x) = \mu(x, \phi), \quad \forall \phi, x \in V; \]
\[ \mu(\phi, \theta) = \vee_x [\mu(\phi, x) \wedge \mu(x, \theta)], \quad \forall \phi, \theta \in V. \]

Fig. 2: A defuzzified, undirected, and unweighted recurrence network of the logistic map with \( a = 3.791 \) (laminar), constructed from a \( 15 \times 15 \) adjacency matrix scaled down from a \( 1000 \times 1000 \) adjacency matrix with \( \alpha = 0.3 \), and \( \beta = 0.005 \).

A defuzzified \( \beta \)-cut prototype recurrence matrix of size \( c \times c \) \((1 < c < G)\), denoted as \( M_\beta \), can be obtained by

\[ M_\beta = \begin{cases} 1, & \mu(\phi, \theta) \geq \beta, \\ 0, & \text{otherwise}, \end{cases} \tag{9} \]

where \( \beta \in [0, 1] \).

Finally, the adjacency matrix of an undirected and unweighted \( \beta \)-cut prototype recurrence network, denoted as \( B_\beta \), is defined as follows:

\[ B_\beta = M_\beta - I. \tag{10} \]

An example to show the potential application of the FRP and its defuzzified recurrence networks is given with the use of the logistic map and its parameters presented in [5]. The logistic equation is

\[ x(n+1) = ax(n)(1 - x(n)), \]

where \( x(1) = 0.1 \), with the number of iterations = 2000. Four different dynamical regimes of the logistic map were generated with \( a = 3.830 \) (period-3), \( 3.679 \) (band merging), \( 3.791 \) (laminar), and \( 4.000 \) (outer crisis). To avoid transient responses of the logistic map [5], the last 1000 data points were used for the construction of the FRP with time delay \( \tau = 1 \), embedding dimension \( m = 1 \). Figure 1 shows the FRPs of the four regimes for \( c = 4 \) and 15 with 200 data points. It can be observed that the smaller number of clusters allows more phase-space states to be close to each other, which results in a plot denser with darker pixels. To illustrate the scalability of the defuzzified recurrence networks, the number of fuzzy clusters \( c = 15 \), \( \alpha = 0.3 \), and \( \beta = 0.005 \) were used to reduce the adjacency matrix whose original size is of \( 1000 \times 1000 \) to \( 15 \times 15 \). Figure 2 shows the graph of a defuzzified recurrence network of the logistic map with \( a = 3.791 \) (laminar), constructed from a \( 15 \times 15 \) adjacency matrix.

To extract GLCM-based textures of the FRPs, the distance \( h \) of the GLCMs for the grayscale images of the FRPs was based on the horizontal proximity of the pixels, which is the pixel next to the pixel of interest on the
same row of the image. Four GLCM features were computed, which are the contrast, correlation, energy, and homogeneity.

The GLCM-based contrast ∈ [0, (S−1)²], where S is the length of the GLCM, is a measure of the intensity contrast between a pixel and its neighbor over the whole image:

\[ \text{Contrast} = \sum_{p,q} |p-q|^2 p_h(p,q), \]  
(11)

which is zero for a constant image.

The GLCM-based correlation ∈ [−1, 1] is defined as

\[ \text{Correlation} = \sum_{p,q} \frac{(p - \mu_p)(q - \mu_q)p_h(p,q)}{\sigma_p \sigma_q}, \]  
(12)

where \( \mu_p \) and \( \sigma_p \) are the mean and standard deviation of \( p \), respectively.

The GLCM-based energy ∈ [0, 1] is the sum of squared elements in the GLCM:

\[ \text{Energy} = \sum_{p,q} p_h(p,q)^2, \]  
(13)

which is 1 for a constant image.

The GLCM-based homogeneity ∈ [0, 1] is the measure of the closeness of the distribution of elements in the GLCM to the GLCM diagonal:

\[ \text{Homogeneity} = \sum_{p,q} \frac{p_h(p,q)}{1 + |p-q|}. \]  
(14)

The semivariogram (SV) of the FRPs was extracted with \( h = 1, \ldots, 4 \) in the horizontal and vertical directions of the FRPs.

Two popular properties of recurrence network measures were calculated, which are the average clustering coefficient, denoted as \( C \), and the characteristic path length, denoted as \( L \). The average clustering coefficient for the adjacency matrix \( A \) of an undirected, unweighted graph is defined as [25]

\[ C = \frac{1}{N} \sum_{i=1}^{N} C_i, \]  
(15)

where

\[ C_i = \frac{\sum_{k=1}^{N} A_{ij} A_{jk} A_{ki}}{l_i(l_i - 1)}, \]  
(16)

where

\[ l_i = \sum_{j=1}^{N} A_{ij}. \]  
(17)

The characteristic path length, which is the average shortest path, for the adjacency matrix of an undirected, unweighted graph is defined as follows:

\[ L = \frac{1}{N(N-1)} \sum_{i,j=1, i \neq j}^{N} d_{ij}, \]  
(18)

where \( d_{ij} \) is the length of the shortest path between nodes \( i \) and \( j \).

Tables 1 and 2 show the GLCM-based and SV-based texture features extracted from the four dynamical regimes of the logistic map, respectively. Table 3 shows the two recurrence network measures of the defuzzified non-prototype \( R_\alpha \) and prototype \( B_\beta \) for the four dynamical regimes of the logistic map. As being observed from fig. 1, the texture of period-3 regime is distinct from the other three dynamical regimes (band merging, laminar, and outer crisis) of the logistic map. The GLCM-based and SV-based features of the FRP of the period-3 are mostly different from the other three regimes.

<table>
<thead>
<tr>
<th>Regime</th>
<th>( a )</th>
<th>Contrast</th>
<th>Correlation</th>
<th>Energy</th>
<th>Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period-3</td>
<td>3.830</td>
<td>32.6667</td>
<td>−0.5000</td>
<td>0.3333</td>
<td>0.4167</td>
</tr>
<tr>
<td>Band merging</td>
<td>3.679</td>
<td>4.4473</td>
<td>0.0848</td>
<td>0.5985</td>
<td>0.8385</td>
</tr>
<tr>
<td>Laminar</td>
<td>3.791</td>
<td>4.3769</td>
<td>−0.0183</td>
<td>0.5843</td>
<td>0.8289</td>
</tr>
<tr>
<td>Outer crisis</td>
<td>4.000</td>
<td>4.5494</td>
<td>0.0066</td>
<td>0.6082</td>
<td>0.8390</td>
</tr>
</tbody>
</table>

Table 1: GLCM-based texture features of fuzzy recurrence plots of the logistic map.

<table>
<thead>
<tr>
<th>Regime</th>
<th>( a )</th>
<th>( b )</th>
<th>( C )</th>
<th>( L )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period-3</td>
<td>3.830</td>
<td>0.9502</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Band merging</td>
<td>3.679</td>
<td>0.9504</td>
<td>0.9253</td>
<td>5.7222</td>
</tr>
<tr>
<td>Laminar</td>
<td>3.791</td>
<td>0.9439</td>
<td>0.9037</td>
<td>5.9646</td>
</tr>
<tr>
<td>Outer crisis</td>
<td>4.000</td>
<td>0.9502</td>
<td>0.9060</td>
<td>6.4113</td>
</tr>
</tbody>
</table>

Table 2: SV-based texture features of fuzzy recurrence plots of the logistic map.

<table>
<thead>
<tr>
<th>Regime</th>
<th>( a )</th>
<th>( b )</th>
<th>( R_\alpha )</th>
<th>( B_\beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period-3</td>
<td>3.830</td>
<td>0.9502</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Band merging</td>
<td>3.679</td>
<td>0.9504</td>
<td>0.9253</td>
<td>5.7222</td>
</tr>
<tr>
<td>Laminar</td>
<td>3.791</td>
<td>0.9439</td>
<td>0.9037</td>
<td>5.9646</td>
</tr>
<tr>
<td>Outer crisis</td>
<td>4.000</td>
<td>0.9502</td>
<td>0.9060</td>
<td>6.4113</td>
</tr>
</tbody>
</table>

Table 3: Defuzzified non-prototype and prototype recurrence network measures for different dynamical regimes of the logistic map.
Likewise, both non-prototype and prototype (scalable) recurrence network measures of the period-3 are also readily distinguishable from those of the other three regimes. Because of scalability, the characteristic path lengths of the prototype defuzzified recurrence networks of the dynamic regimes, except for the period-3 where both cases are of minimum length being equal to 1, are smaller than the non-prototype (original) defuzzified recurrence networks. In general, the texture features and recurrence network measures of the band merging, laminar, and outer crisis regimes are also different, suggesting the sensitivity of the texture features of the FRPs and their defuzzified recurrence network measures in relation to the changes in the dynamical behaviors of the logistic map.

Being analogous to the quantification analysis of RPs, which tries to overcome the difficulty in visualizing an RP with limited computer resolution by introducing several quantitative measures of the point and line distributions of the RP, such as the recurrence rate, entropy, and determinism [18], the texture analysis of the grayscale images of the FRPs is able to differentiate the properties of the four dynamical regimes of the logistic map in terms of contrast, correlation, energy, homogeneity, and spatial statistics. To compare the current results with those obtained from the recurrence networks studied in [5], although the values for the clustering coefficients and the characteristic path lengths of the four dynamical regimes of the logistic map provided by the two methods are different due to different generated textures, the two sets of results have similar orders of magnitude. For the clustering coefficients, the ascending order given by the defuzzified recurrence network is: laminar, band merging, outer crisis, and period-3; and by the recurrence networks it is: laminar, outer crisis, band merging, and period-3. For the characteristic path lengths, both methods have the same ascending order of magnitude: period-3, band merging, laminar, and outer crisis.

As another example to illustrate the potential application of FRPs and defuzzified recurrence networks, the photoplethysmography (PPG) signals of an elderly participant with dementia and a caregiver reported in [26] were used in this study. The original purpose was to measure the therapeutic communications between the caregiver and elderly participants with dementia via wireless sensor measurements of physiological signals. Figure 3 shows the PPG signals of 2500 instances measured on the elderly and caregiver and their associated FRPs, using the same parameters specified for the logistic map. Tables 4 and 5 show the GLCM-based and SV-based texture features of the FRPs of the elderly subject and caregiver. While it is difficult to discern the difference between the patterns of the two PPG signals, the FRP of the PPG signal of the elderly subject can be seen to exhibit smaller texture structure along the diagonal than that of the caregiver. The GLCM-based features in terms of contrast, correlation, energy, and homogeneity have different values for the two subjects, particularly the contrast in image intensity. The SV-based features between the two subjects are most different for $h = 3$ (0.0207 vs. 0.0170). Table 6 shows the defuzzified non-prototype and prototype recurrence network measures of the two subjects in terms of the clustering coefficient and characteristic path length, where the clustering coefficients are more distinctive for the prototype that enables the scaling of the original adjacency matrix from $2500 \times 2500$ to $15 \times 15$.

Table 4: GLCM-based texture features of fuzzy recurrence plots of PPG signals.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Contrast</th>
<th>Correlation</th>
<th>Energy</th>
<th>Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>0.2798</td>
<td>0.9290</td>
<td>0.7316</td>
<td>0.9554</td>
</tr>
<tr>
<td>Caregiver</td>
<td>0.2507</td>
<td>0.9358</td>
<td>0.7455</td>
<td>0.9618</td>
</tr>
</tbody>
</table>

Table 5: SV-based texture features of fuzzy recurrence plots of the PPG signals.

<table>
<thead>
<tr>
<th>Subject</th>
<th>$h$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Elderly</td>
<td>0.0044</td>
<td>0.0125</td>
<td>0.0207</td>
<td>0.0284</td>
</tr>
<tr>
<td>Caregiver</td>
<td>0.0040</td>
<td>0.0105</td>
<td>0.0170</td>
<td>0.0232</td>
</tr>
</tbody>
</table>

Table 6: Defuzzified non-prototype and prototype recurrence network measures for the PPG signals.

<table>
<thead>
<tr>
<th>Subject</th>
<th>$C_1$</th>
<th>$B_3$</th>
<th>$R_0$</th>
<th>$B_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>0.9382</td>
<td>0.8874</td>
<td>5.0808</td>
<td>1.0286</td>
</tr>
<tr>
<td>Caregiver</td>
<td>0.9391</td>
<td>0.9758</td>
<td>5.2284</td>
<td>1.2000</td>
</tr>
</tbody>
</table>

Fig. 3: PPG signals and associated fuzzy recurrence plots.
The mental characteristics in terms the above example of the synchronized PPG-signal analysis of the patients with dementia and the caregiver have important clinical implications in that if the therapeutic communication between patients and caregivers can be quantified, then the cognitive stimulation can be accurately assessed if it can effectively help patients improve their quality of life, and if it is necessary to change the current psychosocial planning and management for a better therapeutic intervention [26]. In addition, PPG is a low-cost physiological measurement technique for studying age-related health conditions [27]; therefore, if being feasible, the PPG-based cognitive stimulation therapy would be more cost-effective than conventional treatments.

When the number of the phase-space states $N$ gets very large, the reduction in the computational requirement using scalable recurrence networks can be appreciated by the complexity analysis of the clustering coefficient and the characteristic path length. For the clustering coefficient, the complexity of the computational time using the mathematical Big $O$ notation (https://en.wikipedia.org/wiki/Big_O_notation), which is adopted to analyze the complexity of computer algorithms with respect to the growth of their running time or space storage as the input size grows, is $O(N^3)$ and the space complexity is $O(N^2)$, using matrix multiplication [28]. For the characteristic path length, using the breadth-first-search algorithm [29], the time complexity for unweighted networks is $O(N^2 + NE)$, where $E$ is the number of edges in the graph, while the space requirement using the adjacency matrix representation is $O(N^2)$.

In summary, image texture analysis of FRPs and measures of defuzzified recurrence networks have been presented, which have potential applications for the analysis of complex signals. The defuzzified recurrence network analysis can be extended to address directed and weighted recurrence networks, as well as multivariate recurrence networks [30].

**Additional remark:** The Matlab code for constructing defuzzified and scalable recurrence networks is available at the author’s personal homepage: https://sites.google.com/site/professorTuAnPham/codes.

**REFERENCES**