Generalized recurrence plots for the analysis of images from spatially distributed systems

Angelo Facchini *, Chiara Mocenni, Antonio Vicino

Department of Engineering Information, University of Siena, Via Roma, 56, I-53100, Siena, Italy
Center for the Study of Complex Systems, University of Siena, Via Tommaso Pendola, 37, I-53100 Siena, Italy

ABSTRACT

We propose a new method for the analysis of images showing patterns emerging from the evolution of spatially distributed systems. The generalized recurrence plot (GRP) and the generalized recurrence quantification analysis (GRQA) are exploited for the investigation of such patterns. We focus on snapshots of spatio-temporal processes such as the formation of Turing structures and traveling waves in the Belousov–Zhabotinsky reaction, satellite images of spatial chlorophyll distribution in seas and oceans (similar to turbulent flows), colonies of Dycistelium discoideum, fractals, and noise. The method is based on the GRP and GRQA and particularly on the measures determinism (DET) and entropy (ENT), providing a new criterion for the assessment and classification of images based on the simultaneous evaluation of their global and local structure. The DET–ENT diagram is introduced and compared with the classical image analysis entropy defined on the pixels’ values. The method proposed provides appealing performances in the case of images showing complex spatial patterns.

1. Introduction

The Recurrence Plot (RP) is a visual tool for the investigation of temporal recurrences in phase space. It was initially designed to display recurring patterns and to investigate nonstationarity in time series [1]. Recurrence is the most important feature of chaotic systems, while nonstationarity may arise from other reasons such as parameter drifting, time varying driving forces, sudden changes in dynamics etc. In recent years, RPs found a wide range of applications in the time series analysis of nonstationary phenomena. For instance, RPs have been used in the analysis of biological systems including neuronal spike trains [2], electromyographic data [3], intercranial EEG recordings [4], electrocardiograms recording [5], protein folding [6], DNA sequences [7], nonlinear phenomena in voice production [8]. The popularity of RPs lies in the fact that their structures are visually appealing, and that they allow for the investigation of high dimensional dynamics by means of a simple two-dimensional plot.

For a better understanding and quantification of the recurrences, Webber and Zbiluth have proposed a set of quantification measures, which are mainly based on the statistical distribution of the line structures in the RP. By means of the Recurrence Quantification Analysis [9], the RP can be used as a tool for the exploration of bifurcation phenomena and dynamics changes also in nonstationary and short time series. Some authors have also related the quantification measures to the invariants of the phase space, like the correlation dimension and the largest Lyapunov exponent [10].

In the field of biological physics, spatially distributed systems play an important role in the understanding of phenomena such as chemotaxis, distribution of species, growth phenomena and reaction-diffusion systems like the prototypical Belousov–Zhabotinsky reaction [11]. In the modeling of ecological aquatic systems, understanding the spatial dynamics is crucial for determining and forecasting the evolution of the ecosystem, e.g. the spatial distribution of microalgae and nutrients, which are highly informative about the ecosystem health status. It has been shown [12,13] that the dynamics of dissolved oxygen can be investigated by means of time series analysis of values recorded in specific points or by means of models based on reaction diffusion mechanisms [14,15]. Both approaches are based on direct and reliable measures, which in many cases are very hard to obtain. An important open problem is that of reconstructing or understanding the spatial dynamics by remote sensing techniques and satellite images. For example, the chlorophyll distribution in water is the result of the spatial and temporal evolution of the drift reaction-diffusion system governing the dynamics of the area. The problem
is then to retrace important information about the system by analyzing the patterns showed by the images.

Recently, extensions of the RP to spatial systems have been proposed [16,17]. In particular, Marwan et al. [16] have assessed the trabecular bone structures under different health conditions.

The aim of this paper is to apply the GRP to the analysis of images showing patterns resulting from the evolution of a spatially distributed dynamical system, trying to catch some signatures of it in the image. Here we use the GRP to extract information about the system from the image by introducing Determinism and Entropy, two indicators based on the analysis of patterns recurrent in the image.

In fact, the main issue of the paper is to offer a simultaneous evaluation of the image both on the global and local scales. In this sense, the aim of the paper is different from that of classical image analysis approaches based on segmentation [18], feature extraction [19], or use of complexity measures related to the semantics of the image [20]. The paper is organized as follows: in Section 2, we describe the classical and Generalized Recurrence Plot and the Recurrence Quantification Analysis. A new formulation of the generalized structures in the GRP is also introduced. In Section 3, we describe our image database and the DET–ENT diagram for the analysis of images. In Section 4, the results of the analysis are reported, and we test the method by analyzing prototypical images of fractals, noise, and periodic structures. After that, we focus on “biologically inspired” images, such as chlorophyll distribution in seas and oceans, colonies of Dictyostelium discoideum, and reaction-diffusion waves and Turing structures in chemical reactions [21]. In Section 5, we state our conclusions.

2. Recurrence plots and generalized recurrence plots

The Recurrence Plot is essentially a two dimensional binary diagram indicating the recurrence that occur in an m-dimensional phase space within an arbitrarily defined threshold ε at different times i, j. The RP is easily expressed as a two dimensional square matrix with ones and zeros representing the occurrence (ones) or not (zeros) of states \( \hat{x}_i \) and \( \hat{x}_j \) of the system:

\[
R_{ij} = \Theta(\varepsilon - \|\hat{x}_i - \hat{x}_j\|)
\]

where \( N \) is the number of the measured states \( \hat{x}_i \), \( \Theta(\cdot) \) is the step function, and \( \| \cdot \| \) is a norm. In the graphical representation, each non-zero entry of \( R_{ij} \) is marked by a black dot in the position \( (i, j) \). Since any state recurrs with itself, the RP matrix fulfills \( R_{ii} = 1 \) and hence it contains the diagonal Line of Identity (LOI).

To compute an RP, a norm must be defined. Usually the \( l_{\infty} \) norm is used, because it is independent of the phase space dimension and no rescaling of \( \varepsilon \) is required. Furthermore, special attention must be paid to the choice of the threshold \( \varepsilon \). Although there is not a general rule for the estimation of \( \varepsilon \), the noise level of the time series plays an important role in its choice. Usually, \( \varepsilon \) is chosen as a percentage of the diameter of the attractor, not greater than 10\% [22].

An RP is characterized by typical patterns, whose structure is helpful for understanding the underlying dynamics of the system investigated. These patterns can be classified according to two features: typology and textures. Typology provides a global view of the RP, and allows a first understanding of the RP. A homogeneous distribution of points is usually associated with stationary stochastic processes, e.g. gaussian or uniform white noise. Periodic structures, such as long diagonal lines parallel to the LOI indicate periodic behaviors, while drifts in the structure of the recurrences are often due to a slow variation of some parameter of the system and white areas or bands indicate non stationarity and abrupt changes in the dynamics. Recently, curved macrostructures have been related to very small frequency variations in periodic signals [23].

The textures consist of the local structures forming the patterns in the RP. They may be: (a) Single points, if the state does not persist for a long time; (b) Diagonal lines of length \( l \), indicating that the trajectory visits the same portion of the phase space at different times; (c) Vertical and horizontal lines, indicating that the state changes very slowly in time. Because of the screen resolution and the length of the time series, it is difficult to analyze the RP only by means of visual inspection (which is anyway useful to detect, e.g. simple non stationarities). To cope with this problem, the RQA offers a set of indicators computed on the structures of the RP (isolated points, vertical and horizontal lines). In the following we will briefly describe and consider only a subset of the RQA measures (for an extensive discussion of other RQA measures the reader may refer to [22]).

2.1. Image analysis: The generalized recurrence plot

Following Marwan [16], we define the RP for a d-dimensional data-set as the 2d-dimensional RP specified by:

\[
R_{ij} = \Theta(\varepsilon - \|\hat{x}_i - \hat{x}_j\|)
\]

where \( \hat{i} = i_1, i_2, \ldots, i_d \) is the d-dimensional coordinate vector and \( \hat{x}_i \) is the associated phase-space vector. This RP, called Generalized Recurrence Plot, accounts for recurrences between the d-dimensional state vectors. Although it can no longer be visualized, its quantification is still possible. Furthermore, as in the one-dimensional case, the LOI is replaced by a linear manifold of dimension \( d \) for which \( R_{ij} = 1, \forall i = j \).

An image is a two-dimensional cartesian object composed of scalar values and in this special case the GRP reads:

\[
R_{i_1,j_1,j_2} = \Theta(\varepsilon - \|x_{i_1,j_1} - x_{i_2,j_2}\|)
\]

where each black dot represents a spatial recurrence between two pixels, and every pixel is identified by its coordinates \( (i_1, i_2) \), being \( i_1 \) and \( i_2 \) the row and the column index respectively. In this case, the recurrence plot is a four-dimensional RP and contains a two-dimensional identity plane, defined by setting \( i_1 = j_1 \) and \( i_2 = j_2 \).

2.2. Generalized recurrence quantification analysis

Since the GRP of an image is four dimensional, its visual inspection is possible only by projections in three or two dimensions. Although this is possible (see e.g. [16], page 548), relevant information is hard to extract, and one must cope with the fact that Generalized Recurrence Plots lose their visual appeal. Despite this drawback, we can still talk about recurrence plots, since our four dimensional space is filled by black dots associated with spatial recurrences between the image’s pixels. Furthermore, RQA is allowed, since the structures described before (as isolated points and lines parallel to the LOI) are present or generalized. In the following we will describe how to generalize the structures formed by the recurrences.

In order to perform RQA, a new definition of the structures must be adopted (details are given in [16]). Here we are interested only in the distribution of the diagonal lines, which have an equivalent in the diagonal patches of length \( l \), defined as follows:

\[
\prod_{k_1, k_2 = 0}^{l - 1} R_{i_1 + k_1, i_2 + k_2} \equiv 1.
\]
In this paper we introduce a different definition of structure of length l: instead of looking for two-dimensional patches, we look for the distribution of line segments in the four dimensional GRP. This is done by sampling the patch structures with lines. In the image, this can be retraced by looking for the recurrences only in the diagonal direction (k₁ = k₂). Our formulation of diagonal line reads:

\[
(1 - R_{i1,i2,j1,j2})(1 - R_{i1+1,i2+j1,j2+1}) \times \prod_{k=0}^{l} R_{i1+k,j1+k,j2+k} = 1.
\]

By applying this new definition, denoting by l the length of a line structure, i.e. the number of recurrent points in it, we build the histogram P(l) of the line lengths and define the GRQA measures as in the one dimensional case. In particular, we focus on Recurrence Rate (RR), Determinism (DET) and Entropy (ENT), defined as follows:

\[
RR = \frac{1}{N^2} \sum_{i,j} P(i,j),
\]

\[
DET = \frac{\sum_{l=\min} P(l)}{N^2},
\]

\[
ENT = - \sum_{l=\min} P(l) \log P(l), \quad p(l) = \frac{P(l)}{\sum_{l=\min} P(l)}.
\]

where \(\min\) is the minimum length considered for the diagonal structures. As for the one dimensional case, the RR (Recurrence Rate) is the fraction of recurrent points with respect to the total number of possible recurrences. The DET (Determinism) is the fraction of recurrent points forming diagonal structures with respect to all the recurrences, and the ENT (Entropy) is a complexity measure of the distribution of the diagonal line in the GRP. The RR is a density measure of the RP, a typical value ranges between 10 and 20\% [22]. The measure DET is introduced as a measure of the predictability of the system, because it accounts for the diagonal structures in the RP. In the one dimensional case, i.e. time series, a line of length l indicates that, for l time steps, the trajectory in the phase space has visited the same region at different times. High values of DET (60\%–70\% or more) indicate that the recurrence points are mainly organized in lines, therefore, in its evolution, the phase space trajectories visit mainly the same regions. The measure ENT is a complexity measure of the RP. It refers to the Shannon entropy with respect to the probability to find a diagonal line of exactly length l. For periodic signal or uncorrelated noise the value is small (∼0.2–0.8); while for chaotic systems, e.g. Lorenz, ENT ∼ 3–4. The computation of the measures was casted in the floating point format and normalized between 0 and 1. To guarantee the homogeneity of the ENT and DET measures, we varied \(\varepsilon\) in order to obtain a RR ∼ 20\% and the DET and ENT were computed considering a minimum length for the diagonal structures of \(\min = 4\).

### 3.1. Description of the images

Our image database consists of eight classes of images showing different patterns: (1) Fractals: large (500 × 500 pixels) and (2) Fractals: small (150 × 150 pixels), (3) Chlorophyll distribution in seas and oceans (different sizes, freely available on the web: www.oceancolor.org), (4) Uniform white noise (100 × 100 pixels), (5) the Belousov–Zhabotinsky reaction [24], showing Turing structures [25] and (6) Diffusion waves [26] (different sizes, courtesy of F. Rossi), (7) Aggregates of Dictyostelium discoideum colonies (different sizes, freely available on the web), and (8) periodic distribution of horizontal, vertical and diagonal lines (100 × 100 pixels). A part of the images, namely fractals, noise, and periodic images are artificially generated, while the other (except the one reported in Fig. 2(b)) come from experiments showing the behavior of different spatially distributed systems. The images of diffusion waves, Turing structures, and Dictyostelium discoideum are snapshots of different movies showing the temporal evolution of the system. Fig. 1 shows a sample of the images’ database: (a) Chlorophyll distribution, (b) Turing patterns in the BZ reaction, (c) a fractal, and (d) diffusion waves in the BZ reaction.

Before starting the analysis, every image has been converted to grayscale and in the jpg format. The level of gray of each pixel was casted in the floating point format and normalized between 0 and 1. To guarantee the homogeneity of the ENT and DET measures, we varied \(\varepsilon\) in order to obtain a RR ∼ 20\% and the DET and ENT were computed considering a minimum length for the diagonal structures of \(\min = 4\).

### 3.2. Example of the analysis method

We now briefly show the steps of the GRP and GRQA analysis. We consider two artificially generated images (size 500 × 500 pixels) extracted from the evolution of a reaction diffusion system showing Turing patterns. Fig. 2(a) shows the initial condition set to uniform random distribution, while in Fig. 2(b) the state of the system with Turing structures is displayed.

First we set the threshold \(\varepsilon\) in order to obtain a Recurrence Rate of about 20\%, after that we set the minimum line length \(\min = 4\) and compute the histogram of the line length distribution. Fig. 3(a) shows the histogram relative to the white uniform noise: as expected, and analogously to the one dimensional case, only few and short lines are present (\(\max = 13\)) and the line length is exponentially distributed. The histogram relative to the Turing pattern is shown in Fig. 3(b); in the first part (4 < l < 20) the line length is exponential but within the remaining part of the histogram is more complex. The visual inspection of the histograms is confirmed by the values of the quantification measures: white noise has DET = 0.61\% and ENT = 0.61, the Turing patterns have DET = 15.35\% and ENT = 1.457.

---

1 All the images are available at the url http://csc.unisi.it/PHYSIO/.
4. Results

Following the method described in Section 3.2, we have computed the DET and ENT values of all the images in our database, and in the beginning the DET–ENT diagram is filled without distinguishing among the different image types. Fig. 4(a) shows the position of the images in the DET–ENT diagram. As one can note, the images form different clusters denoted by the capital letters A, B, C, D, E, F. In Fig. 4(b) we distinguish the different classes of the images in the DET–ENT diagram and we analyze the content of the clusters. The cluster A (ENT = 0.75–1.25, DET = 64%–78%) contains only periodic images. This is consistent with the properties of the DET and ENT described in Section 3: high DET indicates a strong periodic structure, low ENT is related to a trivial pattern. Cluster B (ENT = 0.55–0.75, DET = 0.3%–1.6%) contains only images of uniform white noise: this is again consistent with our hypothesis, the images are globally unstructured (low DET) and locally the patterns are trivial (low ENT). The red dot in the cluster indicates the position of the image shown in Fig. 2(a). Cluster C (ENT = 0.9–1.25, DET = 1.5%–5.5%) mainly contains images of...
The histograms of the line length distribution: (a) White noise: In analogy with the one-dimensional case, the length are exponentially distributed and the maximum length is short. (b) Turing patterns: In the beginning an exponential distribution is found, while in the remaining part the histogram is more complex.

The ENT–DET diagram: (a) The position of all the analyzed images and the clusters. (b) The position of the images is indicated by the different symbols. Except for some isolated cases, the clusters contain: A: Periodic images, B: Uniform white noise, C: Small fractals, D: Large fractals, E: Reaction diffusion systems showing Turing structures, chlorophyll distribution in sea and oceans, colonies of Dictyostelium discoideum, F: Diffusion waves.

A large fraction of the images coming from experiments and showing patterns formed by different reaction-diffusion systems are contained in the last two clusters. In particular Cluster E (ENT = 2.4–2.85, DET = 26%–38%) contains the images of large fractals: the self similarity range increases (higher ENT) and the image shows a macroscopic structure (higher DET, see e.g. Fig. 1(c)).

4.1. Comparison with the pixel based entropy

We have compared our method with a classical measure of image complexity: the entropy based on the color distribution of the pixels:

$$S_{pix} = - \sum_{i=1}^{N} p(x_i) \log p(x_i) \quad (7)$$

where $p(x_i)$ is the i-th component of the pixel values histogram. The x axis of Fig. 5(a) reports the $S_{pix}$ values of all the images in the database. The data are randomly spread along the y direction for the sake of readability. As in the case of the DET diagram the points have some analogies. Even if cluster F contains the diffusion waves in the BZ reaction, the structure of the patterns is different, and under the point of view of the model, in this case the formulation of the equation is simpler, since the diffusion waves are only cylindrical.
Fig. 5. (Color online) The values of $S_{\text{pix}}$ (the y axis spreads uniformly the points for the sake of the readability). Except for F, the different clusters do not homogeneously group the image types.

Cluster A ($S_{\text{pix}} = 3–4$) contains only periodic images, cluster B ($S_{\text{pix}} = 4–5$) contains both periodic images and Turing structures, cluster C ($S_{\text{pix}} = 5–6$) contains small fractals and diffusion waves, cluster D ($S_{\text{pix}} = 6.3–7.2$) groups Dictostelium discoideum, small fractals, and some chlorophyll images. Finally clusters E contains chlorophyll distribution, small fractals, and big fractals, while the white noise is contained in the cluster F.

The pixel color based entropy fails to discriminate the different patterns. This is because of its definition: it is based only on the histogram of the pixel color distribution, and its value is invariant with respect to the structure of the image. Furthermore, the white
noise is found to have the highest entropy: the more the image is rich in colors, the more the entropy is high. Therefore, the clusters shown by Fig. 5 are formed by images having similar tones and colors, no matters about the structures.

The values of DET and ENT have been plotted with respect to the pixels' based entropy in order to check if the diagrams ENT–$S_{pix}$ and DET–$S_{pix}$ show performance similar to that of the DET–ENT diagram. Fig. 6(a) shows that ENT and $S_{pix}$ are not correlated, while Fig. 6(b) shows that images in the DET–$S_{pix}$ diagram cluster poorly if compared with the DET–ENT diagram. The red dot and the red square refer to the patterns of Fig. 2(a) and Fig. 2(b) (uniform noise and Turing patterns), respectively. It is worth noticing that they fall outside the respective clusters in both the ENT–$S_{pix}$ and DET–$S_{pix}$ diagrams.

4.2. Limitations of the method and effect of noise

The main limitation lies essentially in the fact that one loses the information on colors and semantic of the image. Indeed, feature extraction and pattern recognition do not appear feasible by means of DET and ENT indicators.

With reference to the effect of noise on the image, the level of noise may affect strongly both the local and the global scales. In order to show how the method performs with noisy images, an increasing level of noise ranging from 10% to 100% has been added to the image of Fig. 2(b). Fig. 7 shows how the image moves on the DET–ENT diagram by increasing the noise level. The insets display a portion (100 × 100 pixels) of the image and show how the pattern is modified by the noise. As expected, the corresponding position of the noisy images in the DET–ENT diagram moves towards the region of noise (cluster B of Fig. 4).

5. Conclusions

We have generalized the properties of the standard one dimensional RQA by extending the RQA measures to two dimensional data sets. To better characterize the spatial systems, we have proposed the DET–ENT diagram, based on the recurrence quantification measures of the generalized recurrence plot.

Both indicators are based on the quantification of the global/local structures present in the image, rather than on the values of the colors. Indeed, DET and ENT are not related to statistical information measures based on the image color distribution. For this reason they allow an overall assessment of the image by a simultaneous evaluation of the global and local scales.

We have analyzed images showing artificially generated patterns (fractals, noise, periodic structures) and images representing the snapshots of the evolution of spatially distributed systems of different origin: chlorophyll in seas and ocean, Turing patterns and diffusion waves in the BZ reaction, and colonies of Dictyostelium discoideum showing diffusion waves.

The results show that the images cluster in the DET–ENT diagram, allowing one to distinguish the patterns generated by different systems. In particular, the DET–ENT diagram identifies images with different patterns and physical nature, such as periodic patterns, noise, diffusion waves, and small fractals. The remaining patterns, except for large fractals, are mainly contained in the same cluster.

This clustering “feature” of the method can be exploited in real applications for analyzing medical, biological, or chemical images emerging from unknown spatially distributed systems. The comparison of our measure with the classical image complexity indicator $S_{pix}$ (the entropy defined on the color distribution) shows that the DET–ENT diagram provides appealing clustering performances. Furthermore, the method can be easily automatized and extended to the analysis of a set of images or a sequence of movie frames.

A deeper understanding of the cluster overlapping phenomenon due to artifacts is subject of current work through a sensitivity analysis of specific parameters of the algorithm such as, e.g., RR and size/resolution of the image. Moreover, further investigation is needed for understanding how the method can be exploited to detect the spatio-temporal evolution and possible structural changes in the dynamics of distributed systems.

Acknowledgements

The authors are grateful to the referees for their important suggestions and comments. The authors thank F. Rossi for providing the images and movies of the BZ reaction and N. Marwan for his comments.

References


