Chaos, Solitons and Fractals 91 (2016) 679–688

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Discriminating image textures with the multiscale two-dimensional complexity-entropy causality plane

Luciano Zunino [a,b,*], Haroldo V. Ribeiro [c]

[a] Centro de Investigaciones Ópticas (CONICET La Plata - CIC), C.C. 3, 1897 Gonnet, Argentina
[b] Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
[c] Departamento de Física, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil

ARTICLE INFO

Article history:
Received 4 December 2015
Revised 6 July 2016
Accepted 6 September 2016

Keywords:
Texture images
Roughness
Entropy
Complexity
Ordinal patterns probabilities
Multiscale analysis

ABSTRACT

The aim of this paper is to further explore the usefulness of the two-dimensional complexity-entropy causality plane as a texture image descriptor. A multiscale generalization is introduced in order to distinguish between different roughness features of images at small and large spatial scales. Numerically generated two-dimensional structures are initially considered for illustrating basic concepts in a controlled framework. Then, more realistic situations are studied. Obtained results allow us to confirm that intrinsic spatial correlations of images are successfully unveiled by implementing this multiscale symbolic information-theory approach. Consequently, we conclude that the proposed representation space is a versatile and practical tool for identifying, characterizing and discriminating image textures.

1. Introduction

The development of complexity measures for two or higher dimensional data has been recognized as a long-standing goal [1]. Several approaches were introduced during the last two decades for a quantitative distinction between different types of ordering or pattern in two-dimensional signals, such as images [2,3]. In particular, techniques to detect fractal and multifractal features have been shown to be useful for dealing with the characterization of self-similar and extended self-similar objects [4–8]. These approaches have their roots in the seminal work of Mandelbrot [9], who just introduced fractal geometry to mimic natural textured patterns. Cloudy textures, such as those associated with mammographic, terrain, fire, dust, cloud, and smoke images can be suitably described by these scaling and multiscaling analysis [5,6]. Actually, recent effective applications in heterogeneous fields confirm that these fractal techniques are highly valuable tools, *e.g.* identification of lesion regions of crop leaf affected by diseases [10], Hurst exponent estimation performed on satellite images to measure changes on the Earth's surface [11], and determination of scaling properties in encrypted images [12]. Despite all these significant efforts, the development of a robust methodology to detect and quantify spatial structures in images still represents an open and subtle problem. Along this research direction, we have previously introduced an extension of the complexity-entropy causality plane to more than one dimension [13]. It has been shown that the two-dimensional version of this information-theory-derived tool is very promising for distinguishing between two-dimensional patterns. Motivated by this fact, in the present paper, we implement the two-dimensional complexity-entropy causality plane in different numerical and experimental contexts with the aim of testing its potentiality as a texture image quantifier. Furthermore, a multiscale generalization of the original recipe is proposed for characterizing the dominant textures at different spatial scales. As it will be shown, this multiscale approach offers a considerable improvement to the original tool introduced in Ref. [13]. Since any image corresponds to a two-dimensional ordered array, we conjecture that the proposed multiscale ordinal symbolic approach can be a useful alternative for an efficient and robust characterization of its features, offering deeper insights into the understanding of the underlying phenomenon that governs the spatial dynamics of the system at different resolution scales.

This paper is organized as follows. In the next section, a brief review of the two-dimensional complexity-entropy causality plane is given. Besides, its generalization to multiple spatial scales is also described. In Section 3, we have included several numerical and experimental applications. More precisely, in Section 3.1, an initial periodic ornament is carefully analyzed when adding a variable

* Corresponding author.
E-mail addresses: lucianoz@ciop.unlp.edu.ar (L. Zunino), hvr@dfi.uem.br (H.V. Ribeiro).

http://dx.doi.org/10.1016/j.chaos.2016.09.005

degree of randomness by changing the color of each pixel with a given probability. A numerically controlled example to illustrate the importance of implementing a multiscale analysis is detailed in Section 3.2. The normalized Brodatz texture database is studied in Section 3.3 and results obtained from the characterization of some real images of interest are included in Section 3.4. Finally, the main conclusions of this research are summarized in the last Section 4.

2. Complexity-entropy causality plane for two-dimensional patterns

A two-dimensional symbolization procedure, following the encoding scheme introduced by Bandt and Pompe (BP) [14], is applied to the image under study. Given a $N_x \times N_y$ image (2D array), the symbolic sequences are obtained by considering the spatial ranking information (ordinal or permutation patterns) associated with overlapping subarrays of size $D_x \times D_y$. This procedure can be better introduced with a simple example; let us assume that we start with the 3×3 array given below

$$A = \begin{pmatrix} 3 & 4 & 8 \\ 5 & 6 & 7 \\ 2 & 8 & 9 \end{pmatrix}.$$

Four parameters, the embedding dimensions $D_x, D_y \geq 2$ ($D_x, D_y \in \mathbb{N}$, the number of symbols that form the ordinal pattern in the two orthogonal directions) and the embedding delays τ_x and τ_y ($\tau_x, \tau_y \in \mathbb{N}$, the spatial separation between symbols in the two orthogonal directions) are chosen. The matrix is partitioned into overlapping subarrays of size $D_x \times D_y$ with delays τ_x and τ_y in the horizontal and vertical directions, respectively. The elements in each new partition are replaced by their ranks in the subset. For instance, if we set $D_x = D_y = 2$ and $\tau_x = \tau_y = 1$, there are four different partitions associated with A. The first subarray $A_1 = \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}$ is mapped to the ordinal pattern (0123) since $a_0 \leq a_1 \leq a_2 \leq a_3$. The second partition is $A_2 = \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 6 & 7 \end{pmatrix}$, and (0231) will be its related ordinal motif because $a_0 \leq a_2 \leq a_3 \leq a_1$. The next subarray $A_3 = \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix} = \begin{pmatrix} 5 & 6 \\ 2 & 8 \end{pmatrix}$ is associated with the ordinal pattern (2013), and the last one $A_4 = \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix} = \begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix}$ is also mapped to the motif (0123). Subarrays with consecutive elements are taken in the above example because the embedding delays τ_x and τ_y are fixed equal to one. However, non-consecutive elements of the original array can be considered by changing the embedding delays. For instance, by choosing $\tau_x = 2$ and $\tau_y = 1$ only two partitions are obtained from array A, namely $A_1 = \begin{pmatrix} 3 & 8 \\ 5 & 7 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 5 & 7 \\ 2 & 9 \end{pmatrix}$. Their related ordinal permutations will be (0231) and (2013), respectively. Finally, in the case $\tau_x = \tau_y = 2$ only one subarray, $A_1 = \begin{pmatrix} 3 & 8 \\ 2 & 9 \end{pmatrix}$, with motif (2013) can be obtained. It is worth remarking that different spatial resolution scales are taken into account by changing the embedding delays.

An ordinal pattern probability distribution

$$P_{BP} = \{p(\pi_i), i = 1, \ldots, (D_x D_y)!\}, \tag{1}$$

is subsequently obtained by computing the relative frequencies of the $(D_x D_y)!$ possible ordinal patterns π_i. For a reliable estimation of this distribution, the image size must be at least an order of magnitude larger than the number of possible ordinal states, *i.e.* $N_x N_y \gg (D_x D_y)!$. It is clear that the 2D encoding scheme previously described is not univocally defined. Actually, instead of ordering the elements row-by-row, an alternative column-by-column ordering recipe could be proposed. However, the BP probability distribution (Eq. (1)) would remain unchanged since only the label given to the accessible states would change by implementing this alternative definition.

Once the BP probability distribution has been obtained, any information-theory-derived quantifier can be estimated. In particular, and in order to introduce the complexity-entropy diagram, the two involved measures—entropy and complexity—need to be defined. Around a decade ago, Rosso et al. [15] proposed to use the normalized Shannon entropy and the normalized Jensen-Shannon complexity for such a purpose. It has been shown that chaotic and stochastic time series are located at different regions of this representation space, thus allowing an efficient discrimination between these two kinds of dynamics that are commonly very hard to distinguish. Given any arbitrary discrete probability distribution $P = \{p_i, i = 1, \ldots, M\}$, the Shannon's logarithmic information measure is given by

$$S[P] = -\sum_{i=1}^{M} p_i \ln p_i \,. \tag{2}$$

The Shannon entropy $S[P]$ is regarded as a measure of the uncertainty associated to the physical processes described by the probability distribution P. It is equal to zero when we can predict with full certainty which of the possible outcomes i whose probabilities are given by p_i will actually take place. Our knowledge of the underlying process described by the probability distribution is maximal in this instance. In contrast, this knowledge is minimal and the entropy (ignorance) is maximal ($S_{\max} = S[P_e] = \ln M$) for the equiprobable distribution, *i.e.* $P_e = \{p_i = 1/M, i = 1, \ldots, M\}$. The Shannon entropy is a quantifier for randomness. It is well-known, however, that the degree of structure present in a process is not quantified by randomness measures and, consequently, measures of statistical or structural complexity are necessary for a better understanding of complex dynamics [16]. As stated by Lange et al. [17]: *One would like to have some functional C[P] adequately capturing the "structurednes" in the same way as Shannon's entropy captures randomness*. There is not a universally accepted definition of *complexity* [18]. In this work, we have implemented the effective statistical complexity measure (SCM) introduced by Lamberti et al. [19], following the seminal notion advanced by López-Ruiz et al. [20], through the product

$$\mathcal{C}_{JS}[P] = \mathcal{Q}_J[P, P_e]\; \mathcal{H}_S[P] \tag{3}$$

of the normalized Shannon entropy

$$\mathcal{H}_S[P] = S[P]/S_{\max} \tag{4}$$

and the disequilibrium $\mathcal{Q}_J[P, P_e]$ defined in terms of the Jensen-Shannon divergence. That is,

$$\mathcal{Q}_J[P, P_e] = \mathcal{J}[P, P_e]/\mathcal{J}_{\max} \tag{5}$$

with

$$\mathcal{J}[P, P_e] = S[(P + P_e)/2] - S[P]/2 - S[P_e]/2 \tag{6}$$

the above-mentioned Jensen-Shannon divergence and $\mathcal{J}_{\max}$ the maximum possible value of $\mathcal{J}[P, P_e]$. Being more precise, $\mathcal{J}_{\max} = -\frac{1}{2}[\frac{M+1}{M}\ln(M+1) - 2\ln(2M) + \ln(M)]$ is obtained when one of the components of P, say p_m, is equal to one and the remaining p_i are equal to zero. The Jensen-Shannon divergence quantifies the difference between two (or more) probability distributions. For further details about this information-theory divergence measure please see Refs. [21,22]. Note that the above introduced SCM depends on two different probability distributions, the one associated to the system under analysis, P, and the uniform distribution, P_e.