Pattern Recognition 53 (2016) 229–237

Contents lists available at ScienceDirect

Pattern Recognition

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

The Kolmogorov–Sinai entropy in the setting of fuzzy sets for image texture analysis and classification

Tuan D. Pham*

Department of Biomedical Engineering, Linköping University, 58183 Linköping, Sweden

ARTICLE INFO

Article history:
Received 3 September 2014
Received in revised form
17 December 2015
Accepted 19 December 2015
Available online 29 December 2015

Keywords:
Chaos
Kolmogorov–Sinai entropy
Fuzzy sets
Uncertainty
Texture analysis
Image classification

ABSTRACT

The Kolmogorov–Sinai (K–S) entropy is used to quantify the average amount of uncertainty of a dynamical system through a sequence of observations. Sequence probabilities therefore play a central role for the computation of the entropy rate to determine if the dynamical system under study is deterministically non-chaotic, deterministically chaotic, or random. This paper extends the notion of the K–S entropy to measure the entropy rate of imprecise systems using sequence membership grades, in which the underlying deterministic paradigm is replaced with the degree of fuzziness. While constructing sequential probabilities for the calculation of the K–S entropy is difficult in practice, the estimate of the K–S entropy in the setting of fuzzy sets in an image is feasible and can be useful for modeling uncertainty of pixel distributions in images. The fuzzy K–S entropy is illustrated as an effective feature for image analysis and texture classification.

1. Introduction

Texture is an important feature to describe an attribute of an image, which is a phenomenon existing in natural scenes, physical and biological appearances, and art work. Mathematical methods such as measures of smoothness, coarseness, and spatial regularity of distributions of pixels are often used to quantify and classify the texture content of different types of images. Although texture is inherently present in images, it is easy to recognize but difficult to define [1]. This is because texture is subject to human perception, and therefore there is no single precise mathematical definition of texture [2]. In general, there are three main approaches to the analysis of texture: statistical, structural, and spectral [3]. Statistical approaches characterize textures as smooth, coarse, grainy, and so on. Structural techniques deal with the arrangement of image primitives, such as the description of texture based on regularly spaced parallel lines. Spectral techniques are based on properties of the frequency content of an image to detect global periodicity by taking into account high-energy, narrow peaks in the spectrum. It is known that texture analysis has a long and rich development in image processing and computer vision [4–7] as there is a textbook devoted to the topic of image texture [8]. In fact, the extraction of texture features continues receiving considerable attention to both technical developments [9–14] and applications [15–20].

It has been pointed out that texture analysis is difficult because of the lack of effective methods for characterizing texture at different scales and in terms of spatial multiresolution; therefore, techniques such as wavelet transform and Gabor filter can be useful for texture analysis [6,21]. Although there are many developments of methods and their applications for texture analysis, the challenge of the classification of texture in images still remains. This is because the properties of texture are ill-defined and complex in nature, being due to the high degree of local spatial variations of image intensity in orientation and scale. Such complexity makes it a burdensome task for current mathematical models to effectively discriminate various types of textural information.

As a matter of fact, information and uncertainty are a coupled entity that exists and needs to be addressed in many complex problems of pattern recognition [22]. It is the uncertainty in information, which makes real-life patterns be both predictable and unpredictable at once. For example, in performance art, we know the pianist will play all the notes that Beethoven wrote, but the performance seems like it can go anywhere spontaneously. This is known as “the uncertainty principle and pattern recognition” [23]. A major school of thought for quantifying the predictability or uncertainty in complex information processing is the theory of chaos and nonlinear dynamics. Chaos is the study of surprises, the nonlinear and the unpredictable; while traditional science deals with supposedly predictable phenomena. In spite of its wide applications in many fields of science and engineering, techniques developed for measuring nonlinear behavior in chaotic

* Tel.: +46 13 286778.
E-mail address: tuan.pham@liu.se

http://dx.doi.org/10.1016/j.patcog.2015.12.012
0031-3203/

signals are mainly concerned with time-series data. Little effort has been spent on the formulation and extension of chaos analysis for quantifying the unpredictability of the nature of texture.

The study presented in this paper was motivated by the utilization of the theories of chaos and fuzzy sets for addressing dynamic uncertainty to characterize the spatial arrangement of the distribution of intensities in textural images. In fact, it has been pointed out that the theory of chaos and fuzzy logic are among the most interesting fields of mathematical research, and these two theories were applied to study the semantic dynamics of self-reference [24]. The rest of this paper is organized as follows. Section 2 briefly describes the concept of the Kolmogorov–Sinai (K–S) entropy. The notion of the fuzzy K–S entropy is introduced in Section 3. Section 4 presents the estimate of the fuzzy K–S entropy in images. Experimental results on different types of textural images and benchmark data are discussed in Section 5. Finally, Section 6 is the conclusion of the research findings.

2. K–S entropy

The theory of chaos and the notion of entropy measure in information theory have been found useful for solving various problems in complex systems [25–27], ranging from electromagnetic transmission of information to medicine and biology [28–30]. The three most well-known quantitative measures of chaos are the Lyapunov exponents, Kolmogorov–Sinai (K–S) entropy, and mutual information [31,32]. This paper particularly focuses on the K–S entropy [33] to extend its notion to measure the average amount of uncertainty rate under imprecise information. Although the Shannon entropy cannot be used to identify chaotic systems, its concept forms the basis for the formulation of the K–S entropy [34]. In the analysis of time series, the K–S entropy [35–37], which is thought to be a number indicating the average time rate of newly created information induced by the time evolution of chaotic trajectories, has also been known as the measure-theoretic entropy or metric entropy, which has three important properties: (1) sequence probabilities, (2) an entropy rate, and (3) limits. To describe the state-space characteristics of a dynamical system, consider a two-dimensional state-space region represented with a box that is divided into smaller boxes or cells, of which each side has length ϵ. As the system evolves with time, the trajectories of the dynamical system will propagate over a number of cells covered by the state-space region. Regarding the first property, the K–S entropy measures the uncertainty of a system associated with a sequence of outcomes or observations of the trajectories after m units of time as follows:

$$S_m = -k \sum_{i=1}^{N_m} p_i \log(p_i), \tag{1}$$

where S_m is the Shannon entropy calculated at time m, p_i is the probability of a trajectory in the ith cell after m units of time, N_m is the number of possible phase-space routes that the system visits at time m, k is a constant and will be taken as 1 to simplify the mathematical expression, and $p_i \log(p_i) = 0$ when $p_i = 0$.

The second property of the K–S entropy is characterized by the rate of change of entropy as the system travels over time. The K–S entropy H_m after m units of time that has length τ is expressed as

$$H_m = \frac{1}{\tau}(S_{m+1} - S_m), \tag{2}$$

where H_m is the rate of change of the entropy resulting in going from time $t = m\tau$ to $t = (m+1)\tau$.

Thus, the average K–S entropy, denoted as H_{KS}, when the number of time steps, denoted as N_t, approaches infinity to cover the entire attractor is

$$H_{KS} = \lim_{N_t \to \infty} \frac{1}{N_t \tau} \sum_{m=0}^{N_t - 1} S_{m+1} - S_m = \lim_{N_t \to \infty} \frac{1}{N_t \tau}(S_{N_t} - S_0). \tag{3}$$

The third property of the K–S entropy requires two limits: one limit takes the time interval to zero, that is $\lim_{\tau \to 0}$; the second limit takes the bin or cell size ϵ to zero, that is $\lim_{\epsilon \to 0}$. Taking together the above three properties, the complete definition of the K–S entropy is said to be the average entropy per unit time at the limit of time approaching infinity, and at the limits of the cell size and of the time interval taking to zero. In mathematical notation,

$$H_{KS} = \lim_{\tau \to 0} \lim_{\epsilon \to 0} \lim_{N_t \to \infty} \frac{1}{N_t \tau}(S_{N_t} - S_0). \tag{4}$$

For discrete systems or when H_{KS} is applied to iterated maps, τ is set to 1 and $\lim_{\tau = 0}$ is dropped out in Eq. (4) [31,32]; thus the K–S entropy becomes

$$H_{KS} = \lim_{\epsilon \to 0} \lim_{N_t \to \infty} \frac{1}{N_t}(S_{N_t} - S_0). \tag{5}$$

It should be noted that either time steps or space steps are of step sizes to refer to either time evolution or spatial evolution, respectively. To maintain a coherent mathematical presentation of the K–S entropy for time series and images, which are interchangeably addressed in following sections, the term "time step" will be used to indicate either time or space.

3. K–S entropy of fuzzy sets

Let $X = \{x\}$ be a collection of points. A fuzzy set A in X is defined as a set of ordered pairs: $A = \{(x, \mu_A(x)) \mid x \in X\}$, where $\mu_A(x)$ is called the fuzzy membership function for the fuzzy set A, which maps each element of X to a real number in the interval $[0, 1]$ [38]. The entropy of the fuzzy set A, denoted by $D(A)$, is a measure of the degree of its fuzziness, which has the following three properties [39]:

1. $D(A) = 0$ if A is a crisp (non-fuzzy) set, that is, if $\mu_A(x) \in \{0, 1\} \forall x$;
2. $D(A)$ is maximum $\Leftrightarrow \mu_A(x) = 0.5$ (most fuzzy) $\forall x \in X$; and
3. $\forall x, D(A) \geq D(A^*)$ where $\mu_{A^*}(x)$ is any "sharpened" (less fuzzy) version of $\mu_A(x)$, such that $\mu_{A^*}(x) \geq \mu_A(x)$ if $\mu_A(x) \geq 0.5$ and $\mu_{A^*}(x) \leq \mu_A(x)$ if $\mu_A(x) \leq 0.5$.

The above properties formally define a measure of fuzziness that is related to the vague distinction between a set X and its negation $\overline{X}$: the closer X and $\overline{X}$, the more fuzzy X. Motivated by the concept of the Shannon entropy [40], the entropy of a discrete fuzzy set A is defined as [39]

$$D(A) = -\sum_{i}^{n} \mu_A(x_i) \log(\mu_A(x_i)) - (1 - \mu_A(x_i)) \log(1 - \mu_A(x_i)). \tag{6}$$

Based on the definition of the entropy of a fuzzy set, the uncertainty of a fuzzy system in the context of the K–S entropy measured with a sequence of observations after m units of time can be defined as

$$D_m = k \sum_{i=1}^{N_m} F_i, \tag{7}$$

where k is previously defined as a constant and taken to be 1, and F_i is the Shannon's function [39,40] of a fuzzy trajectory in the ith cell:

$$F_i = -\mu_i \log(\mu_i) - (1 - \mu_i) \log(1 - \mu_i), \tag{8}$$

where $\mu_i \in [0, 1]$ is the fuzzy membership grade of a trajectory in the ith cell (it is not necessary that $\sum_i \mu_i = 1$ because the notion of