Information Sciences 459 (2018) 36–52

Contents lists available at ScienceDirect

Information Sciences

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

A Gaussian pyramid approach to Bouligand–Minkowski fractal descriptors

João Batista Florindo [c,a,*], Dalcimar Casanova [b], Odemir Martinez Bruno [c]

[a] Institute of Mathematics, Statistics and Scientific Computing - University of Campinas, Rua Sérgio Buarque de Holanda, 651, Cidade Universitária "Zeferino Vaz" - Distr. Barão Geraldo, Campinas, SP CEP 13083-859, Brazil

[b] Department of Informatics, Federal University of Technology – Paraná, Via do Conhecimento, Km 1, Pato Branco, PR, CEP 85503-390, Brazil

[c] Scientific Computing Group – São Carlos Institute of Physics – University of São Paulo, Av. Trabalhador São Carlense, 400, São Carlos, SP 13560-970, Brazil

A R T I C L E I N F O

Article history:
Received 23 May 2016
Revised 8 February 2018
Accepted 13 May 2018
Available online 14 May 2018

Keywords:
Fractal descriptors
Gaussian pyramid
Texture analysis
Pattern recognition

A B S T R A C T

This work proposes a method to extract features from texture images by applying a Gaussian pyramid multiscale approach to the Bouligand–Minkowski fractal descriptors. The proposal starts from the texture image and computes the stack of multi-resolution images that compose the pyramid, in both directions, of reduction and expansion. In the following, each image in the stack is mapped onto a surface, which is dilated by spheres with variable radii and the dilation volumes are used to compute the Bouligand–Minkowski fractal descriptors for each level. Both the descriptors of each level and combinations with descriptors from the original image are verified in the classification of well-known databases of textural images. The proposed method outperformed other classical and state-of-the-art descriptors with a significant advantage in most cases, including situations where random noise is added to the images.

1. Introduction

Since the seminal work of Mandelbrot [16], the fractal geometry has demonstrated to be a powerful tool for the analysis of objects in the real world. Particularly, with the dissemination of computers, a large number of works have proposed to analyze digital images representing such objects through fractal-based methods [2,6,11,15,26,29,32–34].

In this way, an important objective is to use fractal geometry to find meaningful features to describe the object of interest in the image. The most simple and widely used fractal feature is the fractal dimension [19,20]. Nevertheless, despite being a descriptive metric, this is only a global value and cannot faithfully express important details of an image observed at different scales. Furthermore, the fractal dimension of a real-world structure changes depending on the scale of analysis and thus it is not robust enough to characterize an image. Another limitation of fractal dimension is to describe objects by a single fractal value whereas in most cases they exhibit a multifractal behavior [15].

Several methods have been proposed to address such limitations of fractal dimension. In multifractal theory [13,15], the analysis is performed in two steps: first, a local measure of regularity (Hölder exponent) is assigned to each pixel in the

* Corresponding author.
E-mail addresses: jbflorindo@ime.unicamp.br (J.B. Florindo), casanova@utfpr.edu.br (D. Casanova), bruno@ifsc.usp.br (O.M. Bruno).
URL: http://scg.ifsc.usp.br (J.B. Florindo), http://scg.ifsc.usp.br (D. Casanova), http://scg.ifsc.usp.br (O.M. Bruno)

https://doi.org/10.1016/j.ins.2018.05.037
0020-0255/© 2018 Published by Elsevier Inc.

image; second, the features are obtained from the fractal dimension of each subset of the image whose regularity values are similar. Different methods to estimate the Hölder exponent have been proposed, for example, using Gaussian filters [33] or wavelets [34]. The multifractal spectrum has also been associated to other texture approaches, like SIFT and textons in [32], for example. A related texture descriptor is the local fractal dimension [26,29]. Such alternative employs methods especially adapted to estimate the "fractality" of the neighborhood of a pixel and extract statistics from the distribution of these measures along the image. The local values can also be used in other frameworks like textons [29] or binary patterns [26]. Alternatively, in multiscale fractal dimension (MFD) [4,18], the power-law relation typically associated to the computation of the fractal dimension is explored to estimate the dimension at specific scales. This is done by computing the derivative of the power-law curve and the features of interest are provided by properties of the derivative curve, such as minima, maxima, area under the curve graph, etc.

More recently, fractal descriptors have been proposed in [2] as a generalization of MFD to provide a meaningful set of features capable of describing the most complex structures. The basic idea is to use all the values in the power-law curve and how they are inter-related to provide the image features. Among the methods to compute fractal descriptors, Bouligand–Minkowski has been successful in a number of theoretical and applied works [2,6,10,11]. In this method, the texture image is mapped onto a three-dimensional cloud of points and after that each point is dilated by spheres. The descriptors are computed by using the volumes of the dilated structure when the sphere radius is varied. Such descriptors quantify how and when the wave-fronts of growing spheres behave within the cloud and, as a consequence, it end up being a complete and straightforward measure of pixel arrangements within the image. Although such method has demonstrated its efficiency, the descriptors do not take into account that a real-world image is intrinsically multiscale and relevant information may be enclosed at different resolutions. Gaussian pyramid theory confirms that features extracted from different levels of sampling convey uncorrelated information that can be quite useful for the image representation [25,28,35].

Based on these assumptions, we propose a method that computes the fractal descriptors from different resolution levels in the Gaussian pyramid [5]. At first, the pyramid is constructed in two directions: reduction and expansion. In the reduction, the image is down-sampled recursively and convolved with a Gaussian filter. In the expansion, the image is recursively up-sampled using an interpolation process. In the following, the Bouligand–Minkowski fractal descriptors are computed from each level in the pyramid (in both directions). The features are obtained either from the fractal descriptors at each level or from a combination of descriptors from different levels and from the original texture image.

The performance of the proposal is assessed in a task of classification of three well-known databases of texture images, namely, Outex [22], USPTex [3] and UIUC [14]. The success rate in the classification is compared to other well-known texture descriptors. The proposed method achieved the best results, outperforming state-of-the-art and classical methods widely used in the literature.

2. Gaussian pyramid

Gaussian pyramids are a well-known multiscale approach to image analysis employed in a number of works in the literature [25,28,35]. Roughly speaking, it consists of applying successive down or up-samplings to the original image followed by a convolution with a Gaussian kernel. The use of up or down-sample determines what is called the direction of the operation. Thus the reduction direction decreases the resolution of the image (down-sample), while the expansion acts in the opposite way.

The reduction acts as an averaging operation where, in the lth level, I_l is given by

$$I_l^R(i,j) = \sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n) I_{l-1}^R(2i+m, 2j+n), \tag{1}$$

where w is the weight matrix, typically a 5×5 matrix of real values obtained from the two-dimensional Gaussian function in discrete domain. Similarly, the expansion at the level l is expressed by

$$I_l^E(i,j) = \sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n) I_{l-1}^E\left(\frac{i-m}{2}, \frac{j-n}{2}\right) \tag{2}$$

In practice, the first step when obtaining Gaussian pyramids is to compute a discrete filter to represent the Gaussian kernel. Given the separability of this kernel, all the process is significantly simplified by considering only one dimension (signals). As it will be seen later, the generalization to images (two dimensions) is straightforward. The Gaussian filter is approximated by the N-tap Gaussian in [5], where N is the number of weights. Such weights may be obtained from Pascal triangle (Fig. 1), as the Gaussian function describes the normal distribution, which is the continuous form of the binomial distribution represented by the coefficients in that triangle. The motivation for using N-tap filtering instead of the classical Gaussian smoothing mask in [5] is mainly for computational efficiency. In fact, the authors in [5] show that the second is closely approximated by the first and the tap coefficients are encoded by the sum of signed powers of two. In this way, the operation of multiplication is replaced by simple addition and shifting, what significantly improves the computational performance. This strategy makes the algorithm faster in general without any significant loss of precision. Another characteristic of this operation is that the original image can be exactly reconstructed, implying that besides allowing the analysis of the image from different scale viewpoints, the original information is also preserved. Finally, the smoothing parameter σ