Texture descriptors by a fractal analysis of three-dimensional local coarseness

Florindo, Joao Batista; Landini, Gabriel; Bruno, Odemir Martinez

DOI:
10.1016/j.dsp.2015.03.013

Document Version
Peer reviewed version

Citation for published version (Harvard):
Florindo, JB, Landini, G & Bruno, OM 2015, 'Texture descriptors by a fractal analysis of three-dimensional local coarseness', *Digital Signal Processing*, vol. 42, pp. 70-79. https://doi.org/10.1016/j.dsp.2015.03.013

Texture descriptors by a fractal analysis of three-dimensional local coarseness

Joao Batista Florindo[a,*], Gabriel Landini[a], Odemir Martinez Bruno[b]

[a] *Oral Pathology Unit, School of Dentistry, University Of Birmingham, St Chad's Queensway, Birmingham, B4 6NN, United Kingdom*
[b] *Instituto de Física de São Carlos (IFSC), Universidade de São Paulo Av. Trabalhador São Carlense, 400, 13560-970 - São Carlos, SP, Brasil*
phone/fax: +55 16 3373 8728 / +55 16 3373 9879

Abstract

This work proposes a new method of extracting texture descriptors from digital images based on local scaling properties of the greyscale function using constraints to define connected local sets. The texture is first mapped onto a three-dimensional cloud of points and the local coarseness under different scales is assigned to each point p. This measure is obtained from the size of the largest "connected" set of points within a cube centred at p. Here, the "connected set" is defined as the set of points such that for each point in the local domain there is at least one other point at a distance smaller than a threshold t. Finally, the Bouligand-Minkowski fractal descriptors of the local coarseness of each pixel are computed. The classificatory power of the descriptors on the Brodatz, Vistex, UIUC and UMD databases showed an improvement over the results obtained with other well-known texture descriptors reported in the literature. The performance achieved also suggests possible applications to real-world problems where the images are best analysed as textures.

Keywords: Pattern Recognition, Fractal Descriptors, Local Connectivity, Image Analysis

1. Introduction

There have been several fractal-based methods proposed for the analysis of complexity in images, including the analysis of image textures [1, 2, 3, 4, 5, 6, 7].

*Corresponding author
Email addresses: jbflorindo@gmail.com (Joao Batista Florindo), G.Landini@bham.ac.uk (Gabriel Landini), bruno@ifsc.usp.br (Odemir Martinez Bruno)

Fractal geometry in image analysis has also found practical applications in a number of areas [8, 9, 10, 44, 11, 12, 13, 14].

In a "texture image", the analysis is typically focused on the statistical and geometrical relations amongst pixel intensity patterns in different regions and scales. In this context, fractal analysis becomes a powerful tool to address the problem of measuring the complexity or the homogeneity of the texture across scales. A pattern can be expressed by certain type of homogeneity at a particular resolution and here is where fractal geometry provides a straightforward procedure to detect and relate such properties. Moreover, real world objects commonly have some degrees of intrinsic self-similarity and therefore they might be more suitably represented as approximations to fractal objects rather than to regular Euclidean ones. Some of the successful fractal-based approaches tested in this field include "multiscale fractal dimension" [15], multifractals [16] and "fractal descriptors" [17]. Particularly, the latter has demonstrated to be highly efficient for the discrimination of general textures [18, 19, 20, 13, 14, 21].

Here, we propose an alternative way of extracting texture descriptors based on fractal geometry. The texture descriptors are computed using the Bouligand-Minkowski fractal descriptors [20] based on the local coarseness of each pixel [22, 23] rather than on the image intensity values [20, 17, 19]. The proposed procedure performs a two-level complexity analysis. In the local domain, the coarseness describes the clustering (or homogeneity) of the pixel neighbourhood, while globally, the fractal descriptors represent the distribution of connectivities across the image. These two complementary types of information appear to be fundamental in describing and discriminating texture patterns at varying scales in a more straightforward way than other statistical or geometrical solutions proposed. Some advantages of the suggested approach are:

- A large number of real-world scenes (like those described by the textures analysed here) have fractal-like characteristics;

- The "fractal properties" of textures are most often not homogeneous across scales as expected with ideal fractal objects but the fractal descriptors

enable the quantification of such variability across the image;

- The local coarseness provides information about pixel neighbourhoods, which is rather richer than pixel intensity alone.

The method performance was assessed on four well-known image databases (Brodatz [24], Vistex [25], UIUC [33] and UMD [16]) and the results were compared to other texture descriptors reported in the literature (Gabor [26], Fourier [27], Grey-Level Cooccurrence Matrix (GLCM) [28], Multifractal [16], Local Binary Patterns (LBP) [29], Soft-LBP [35], Fuzzy-LBP [36] and textons (VZ) [34]).

2. Related Works

Since the seminal work of Mandelbrot [42], several fractal-based methods have been proposed in the literature to analyse texture images, and particularly for the problem of texture classification. The most state-of-the-art and successful approaches can be essentially divided into three categories: texton-based, multifractals and fractal descriptors.

Texton (also called bag-of-words) methods follow the general scheme presented in [34]. The basic idea in this approach is to associate a vector of measures (texton or "word") to each pixel or region of interest in the image and cluster them into a number of groups (dictionary). Therefore, for each image in the training and testing database, a model is built by using the histogram of pixels whose corresponding textons are closer to a particular group in the dictionary.

An example of texton-based fractal method is illustrated in [37], where the image is submitted to a filter bank as in [34] and the textons are estimated by the local fractal dimension of each filter response, computed by using the exponential relation between the sum of pixel intensities within a neighbourhood of the reference pixel and the radius of such neighbourhood. The remaining proce-

dures are classical in texton-based methods [34], involving K-means clustering of textons and classification by nearest neighbours with χ^2 distance.

The second group of fractal-based methods includes the multifractal spectrum, which quantifies the distribution of a regularity parameter (the Holder exponent) within local neighbourhoods of each image pixel. In [40] and [16] the neighbourhood regularity is quantified by using a similar procedure to that employed in [37], that is, the power-law relation between the sum of the pixel intensities after Gaussian filtering within a neighbourhood and the radius of the neighbourhood. The image is therefore partitioned into subsets, based on the values of the local Holder exponents, and the texture features are given by the box-counting dimension of each subset.

A more elaborated method to compute the local regularity is described in [41], using wavelet leaders. A wavelet leader is the maximum response of a wavelet decomposition inside a scale-space neighbourhood, that is, the neighbourhood in this case is three-dimensional, including not only all the adjacent points in the decomposition level as usual, but also those correspondent points in neighbour scales in the wavelet pyramid. The use of wavelet leaders attenuates one of the main problems with wavelet transforms, which is the large number of small coefficients obtained for natural images. Another novelty in this method, compared to [16], is the use of a multi-orientation approach to the image, to overcome the orientation sensitivity of the wavelet transform. A more complete version, including an adaptation to dynamic textures can be found in [38].

Multifractal methods can be further divided into dense and sparse approaches. In dense methods, the Holder exponent is computed at each pixel in the image. This is the case of the above methods. Another strategy is the sparse approach, where the dimension is computed only over particular regions rather than over all the pixels as proposed in [39]. In that case, the orientation histograms of the scale-invariant gradient of the image is used to partition the image for the posterior computation of the multifractal spectrum.

Finally, the third category of fractal methods in texture classification is composed of the fractal descriptors, originally proposed in [17]. They employ

the values of the power-law curve associated to the fractal modelling to provide the image features in a straightforward manner. One of the most investigated approaches in this category are the Bouligand-Minkowski descriptors [20], where the texture is mapped onto a three-dimensional cloud of points and all the points are simultaneously dilated by spheres with radius r. The descriptors are obtained from the total volumes of the dilted cloud with various values of r. More details are given in Section 4.2. These features have been used either directly [20] or after some post-processing procedure [19]. Another variant is described in [18], where the entropy of the fractal descriptors are computed under different scales to accomplish the classification task.

The proposed method can be considered as part of the third group, although the local connectivity can also be associated to the concept of local regularity as employed in the multifractal spectrum. In a sense, it is similar to the multifractal approach, as both rely on a two-layer analysis: first, a local quantification of regularity (here expressed by the connectivity dimension), and second, a global distribution of such property. However, our proposal replaces the box-counting dimension of partition sets by the Bouligand-Minkowski descriptors. Such descriptors provide a more descriptive representation as, more than estimating the local dimension in the spatial domain of the image, they also reveal how the local regularity behaves across multiple scales. Moreover, the connectivity is also locally scale-invariant, which makes it a type of density function and allows the generation of descriptors robust to invariances in illumination changes, as discussed in [16].

3. Fractal Geometry

A fractal is a mathematical object with self-similarity (i.e. parts are similar to the whole) and typically high complexity (i.e. persistence of distinguishable details at various scales).

Same as Euclidean geometry defines fundamental measures such as area and perimeter, fractal geometry defines its own measures, one of the most important being the "fractal dimension". This measures how the complexity (or space

occupation) of the object changes with changes of observational scale. Its value can be obtained from the general expression:

$$D = \lim_{\epsilon \to 0} \frac{\log(\mathfrak{M}(\epsilon))}{\log(\epsilon)}, \tag{1}$$

where $\mathfrak{M}(\epsilon)$ is a self-similarity measure which grows with the scale ϵ following a power law.

Despite mathematical fractals being ideal constructs, many seemingly self-similar and complex objects are easily found in the real world. In this context, there is vast literature on modelling real-world problems through fractal geometry and quantifying important properties using fractal dimensions. Applying fractal theory to digitised images requires to redefine the fractal dimension in a discrete and finite space, and several methods have been developed for this purpose [30, 31]. Two of these will be discussed below, the Bouligand-Minkowski and local connected dimension.

3.1. Bouligand-Minkowski

This method estimates the dimension of binary objects, but it can be straightforwardly extended to grey-level textures [20].

The grey-level image is considered as a cloud of points in the three-dimensional Euclidean space, such that a pixel in the coordinate (x, y) and with intensity z is mapped onto a point with coordinates (x, y, z). Then, each point in the cloud is morphologically dilated by spheres with radius r and the volume $V(r)$ of the dilated cloud is computed. The dimension is given by:

$$D = 3 - \alpha, \tag{2}$$

where α is the slope of the linear regression of $\log(V(r))$ on $\log(r)$.

3.2. Local Connected Dimension

The Local Connected Dimension [22] has been applied to binary images. As shown in Section 4 we use an adapted version for grey-scale images. This consist of computing a local dimensional value relative to each possible position of the

analysis, i.e. for each image pixel. Given a pixel p, within a local neighbourhood of size r_{max}, its local dimension $D_C(p)$ can be obtained by counting the number N of pixels connected (using, for example, 8-neighbour pixel connectivity in the plane) to p within a square window with side-length r always centred at p. When r is varied, the dimension is given by:

$$D_C(p) = \alpha_C, \tag{3}$$

where α_C is the slope of the linear regression of $\log(N(r))$ on $\log(r)$.

3.3. Bouligand-Minkowski Fractal Descriptors

While the fractal dimension has been shown to be useful in a variety of applications, it still is a single number, and this might be insufficient to model complex or heterogeneous objects. To take advantage of fractal geometry without being bound to a single number, the authors in [17] proposed using "fractal descriptors", i.e. to use all the values from $\log(\mathfrak{M}(\epsilon))$ rather than their scaling relation. More specifically, they used the $\log(V(r))$ curve from Bouligand-Minkowski dimension to provide powerful features for texture image. These features can be used directly, after transformation, or combined (for instance, using different colour channels [21] or image windows [18]) and have been successfully used in texture analysis [18, 19, 20, 13, 14, 21].

4. Proposed Method

We propose a new method to compute fractal descriptors based on adapting the local connectivity concept to grey-level images to describe local patterns in images. Such adaptation essentially consists of mapping the image onto a three-dimensional cloud (considering the grey-level of the pixel as the third dimension) and replacing the concept of "adjacency" by that of "an Euclidean distance smaller than a threshold".

4.1. Local Connectivity

The local fractal dimension has been used before to estimate the structure of binary sets. This is based on the local scaling of mass (binary pixels) using the mass-radius dimension and repeating the analysis to small regions, relative to a centre point position which can be repeated for many (or all) possible positions in an image. This method can be constrained to "local connected set" rather than just local sets by considering only the mass that is "locally connected" to the origin of the analysis. Such an approach was successfully applied in [22] to characterise the structure of retinal vessel patterns and to cancer and pre-cancer invasive patterns [9]. That approach, however, can be modified to accommodate non-binary sets (i.e. other image types). Here we consider the scaling of pixel intensities in grey-scale images as if they were embedded in a three-dimensional Euclidean space defined by the image coordinates and the intensity scale.

Based on this approach, the connectivity concept can be redefined by replacing the adjacency contact relation with a new connectedness property based on a three-dimensional Euclidean distance satisfying a predefined threshold.

To achieve this, the grey-level image $I : [1, M] \times [1, N] \rightarrow \Re$ is mapped into a cloud of points S in a three-dimensional Euclidean space, such that each x and y coordinate pairs are the coordinates of each pixel and the z coordinate is the respective pixel intensity.

$$S = \{(x, y, z)|I(x, y) = z\} \tag{4}$$

Following the mapping, a connectivity measure is computed for each point p in the cloud by considering a cube C_p^r with side-length r centred at the point:

$$C_p^r = \{(x, y, z)|x \in [x_p - r, x_p + r], y \in [y_p - r, y_p + r], \in [z_p - r, z_p + r]\} \tag{5}$$

The connectivity measure $M_p^{r,t}$ is given by the number of points inside C_p^r and connected to p. Given that the concept of connectivity in three dimensions is not so straightforward as in two, here a connectivity based on Euclidean distances is employed. In this context, the set $\mathfrak{C}_p^{r,t}$ of points connected to p is a

subset of points within C_p^r such that each element has at least one other point at a distance smaller than t within the cube.

$$\mathfrak{C}_p^{r,t} = \{p\} \cup \{p_i | p_i \in C_p^r \text{ and } \exists p_j \in C_p^r | D(p_i, p_j) < t\}, \tag{6}$$

where D is the three-dimensional Euclidean distance.

An optimized algorithm to find the connected components maps the points in the cube into a non-weighted graph $G(V, E)$ such that:

$$v \in V \text{ iff } p_v \in C_p^r, \tag{7}$$

$$e_{i,j} \in E \text{ iff } \{p_i, p_j\} \subseteq C_p^r \text{ and } D(p_i, p_j) < t, \tag{8}$$

and then searches for the connected component of $G(V, E)$ that includes p.

The connectivity measure $M_p^{r,t}$ is given by the size (number of vertices) of the connected component $\mathfrak{C}_p^{r,t}$. Figure 1 illustrates the connected components in two dimensions to facilitate visualization. Figure 1(a), (b), (c) and (d) show the steps for a window (two-dimensional version of the cube) with growing length r. In each iteration the border of the current window is highlighted in black. Inside the current window all the points at a distance smaller than t are connected by an edge. Such process generates a graph with multiple connected components. The points taking part into the connected component that contains the centre pixel p are painted red. The connectivity of p (for each r) is given by the number of red points inside the respective window.

In a similar way to the analysis of binary images, to characterise the scaling of local connectivity in grey-scales, a local Holder exponent can be computed by:

$$\alpha(p) = \lim_{r \to 0} \frac{\log(M_p^{r,t})}{log(r)}. \tag{9}$$

Given that the range of r cannot be made large enough to estimate the relation with confidence, the limit loses part of its significance and the coarse (or coarse-grained) Holder exponent [23] arises as a more interesting measure. This is given by:

$$\alpha_r(p) = \frac{\log(M_p^{r,t})}{log(r)}. \tag{10}$$

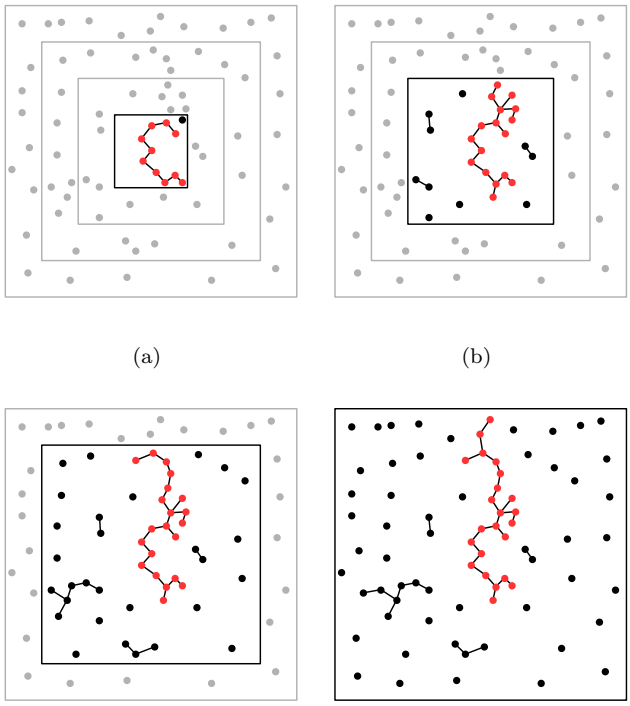

Figure 1: Computing the connectivity of a pixel p (centre point). (a), (b), (c) and (d) exemplify the process for different window lengths. Inside the current window (highlighted in black) all the points at a distance smaller than t are connected by an edge resulting in a multi-component graph. The connectivity is given by the number of points in the connected component containing the centre point (these are painted red).

The threshold t is an empirically pre-defined value. After computing α_r for each pixel, the procedure ends up with a matrix of values M_{α_r} for each r. Here, the values of r were varied between 1 and 6, with t ranging between 1 and 10.

Finally, the Bouligand-Minkowski fractal descriptors are computed over each one of the six matrices M_{α_r} (one for each value of r) and concatenated to provide the final texture descriptors.

4.2. Bouligand-Minkowski Descriptors

The procedure to compute the Bouligand-Minkowski descriptors of each matrix consists of two steps. First, each point in the matrix is mapped onto a three-dimensional cloud of points. Second, all the points in the cloud are simultaneously dilated by a sphere with radius r and the total volume $V(r)$ of the dilated cloud is computed. The descriptors correspond to the values of $V(r)$ when r ranges within a pre-defined range. More details are given in the following.

At first, the real-valued matrix $M_{\alpha_r} : [1, W] \times [1, H] \rightarrow \Re$, where W is the width and H the height of the matrix, is mapped onto a three-dimensional cloud of points C. Such mapping is obtained by associating each point in M_{α_r} with coordinate $(x, y) \in ([1, W] \times [1, H])$ and such that $M(x, y) = z$, $z \in \Re$, with a point with coordinates $(x, y, z) \in C$.

In the following, each point in C is dilated by a radius r, and the total volume $V(r)$ of the dilated cloud is computed. In practice, all the points are replaced by a sphere with radius r and, depending on the value of r and the distance among the points in the cloud, such spheres can touch and merge. Therefore, the total dilated volume is not just the sum of the volumes of the individual spheres, but it encloses relevant information about the distribution of points in the cloud and, as a consequence, the distribution of α_r in M.

Particularly, when $V(r)$ is analysed within a range of values of r, it provides a useful insight about the homogeneity of α_r. If such distribution is homogeneous, the cloud C is regular and the curve $\log r \times \log V(r)$ is similar to a straight line. If, on the other hand, such coefficients are distributed in an irregular way on M,

there is a larger number of values of r for which new collisions arise and thus the $\log - \log$ curve is more irregular as well. Such behaviour of $V(r)$ is what makes it appealing to summarize the information expressed by the local coarseness.

The total volume $V(r)$ of the dilated cloud corresponds to the number of points pertaining to the union of spheres $B(p, r)$ centred at each point $p \in C$ with radius r:

$$V(r) = \sum_{p'} \mathbf{1}_U(p'), \tag{11}$$

being $\mathbf{1}$ the indicator function ($\mathbf{1}_U(p') = 1$ if $p' \in U$ and 0, otherwise) and

$$U = \bigcup_{p \in C} B(p, r). \tag{12}$$

In practice, an efficient way to compute $V(r)$ is by using the Euclidean Distance Transform (EDT), given that the set $B(p, r)$ in the above expression contains the points at a distance at most r from p. In a three-dimensional space, the EDT of a point p' is provided by:

$$EDT(p') = \min_{p \in C}(\mathbf{dist}(p, p')), \tag{13}$$

where $\mathbf{dist}(p, p')$ is the Euclidean distance between the points. More details on methods to compute the EDT efficiently can be found in [43].

The Bouligand-Minkowski descriptors $\mathfrak{D}(u)$ are obtained by computing the EDT over all points within a region of interest around C. Those EDT values are thus increasingly sorted into a vector o and the descriptors correspond to the logarithm of the cumulated number of points p' such that $EDT(p') \leq o(u)$:

$$\mathfrak{D}(u) = \log \left[\sum_{i=1}^{u} \delta(EDT(p') - o(u)) \right], \tag{14}$$

where δ is the unit response function ($\delta(x) = 1$, if $x = 0$, and $\delta(x) = 0$, otherwise). The index u corresponds to the non-negative values of the EDT within the region of interest. For example, for $r \leq 2$ there are 4 possible values for u (1, $\sqrt{2}$, $\sqrt{3}$ and 2), and thus 4 descriptors are computed. Here we use $r \leq 10$, providing 85 descriptors.

As the number of descriptors can become very large in most cases, a Principal Component Analysis is performed after the concatenation. The following algorithm express each step in a pseudo-code language, while Figure 2 synthesizes these steps on a diagram. That diagram shows the steps involved in the proposed method, sequentially from top to bottom. First of all, a grey level image is depicted, following by the matrices M_{α_r}, represented as intensity images for r between 1 and 6. After that, the Bouligand-Minkowski descriptors are computed for each previous matrix, providing the exhibited $\log - \log$ curves. Finally, at the bottom, the Bouligand-Minkowski descriptors are concatenated and submitted to a Principal Component Analysis with the aim of reducing the dimensionality. It is worth noting that M_{α_r} behaves like a multiscale transform over the texture where increasing values of r gradually "smooth" details, since points falling inside a same cube are expected to have similar Holder coarseness.

for $r = 1$ until 6 **do**
  for all $p \in I$ **do**
    for all $q_1 \in C_r^p$ **do**
      for all $q_2 \in C_r^p \wedge q_2 \neq q_1$ **do**
        if $distance(q_1, q_2) \leq t$ **then**
          $add(G, \{q_1, q_2\})$
        end if
      end for
    end for
    $\mathfrak{C}_p^{r,t} \leftarrow$ findConnectedComponent(G,p)
    $M_{\alpha_r}(p) \leftarrow \frac{\log(|\mathfrak{C}_p^{r,t}|)}{\log(r)}$
  end for
  $D^r \leftarrow$ BouligandMinkowskiDescriptors(M_{α_r})
end for
$D \leftarrow$ concatenate($D^1, D^2, D^3, D^4, D^5, D^6$)
descriptors $\leftarrow$ PCA(D)

Figure 3 shows how the proposed descriptors can precisely classify some textures from Brodatz database. Even with only a few PCA scores the classes can be distinguished. It is still worth stressing that although using more components causes the curves approximate each other, as they less contribute to the variability, those components when put together allow higher precision when managed by efficient classifiers. Such promising results are consequence of combining the measure of complexity under different perspectives: the first (Holder coarseness) more local and focused on the pixel neighbourhood, the second (Bouligand-Minkowski) more global, expressing how the topology of the texture is defined. The complementary information given by both provides a richer method to well characterize even the most complex and irregular textures.

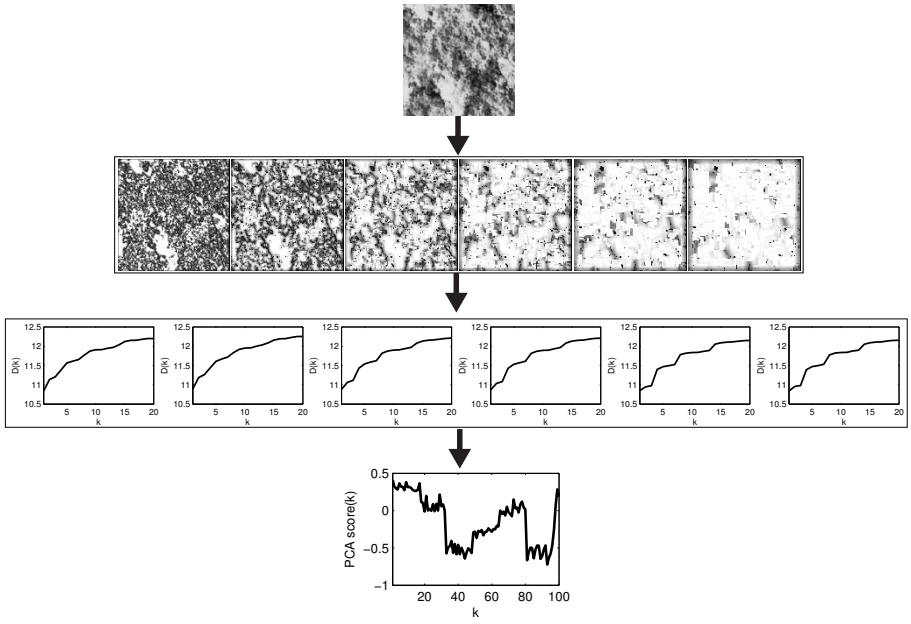

Figure 2: Steps in the proposed method. At the top, the grey-level texture to be analysed. Below, the M_{α_r} matrices expressing the coarseness at each pixel in the original image, with $r \in [1, 6]$, represented in an intensity image. In the following row, the respective Bouligand-Minkowski curves computed over each M_{α_r} is exhibited and finally the concatenated descriptors after applying a Principal Component Analysis are shown at the bottom of the diagram.

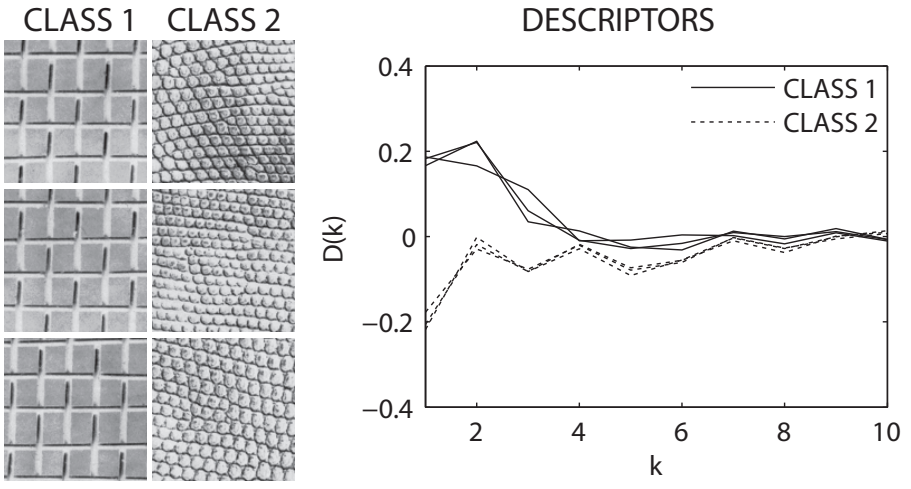

Figure 3: First PCA scores (D(k)) of the proposed descriptors from images of two classes (objects). The groups can be identified by the descriptors even visually.

5. Experiments

The performance of our approach was tested against other well-known texture descriptors to classify four benchmark databases (Brodatz [24], Vistex [25], UIUC [33] and UMD [16]).

The Brodatz database is a collection of grey-scale images from photos of an architectural textures book [24]. One hundred and eleven images were used, with each one being split into 16 non-overlapping windows, resulting in 111 classes with 16 samples each.

The Vistex database is a classical texture image collection [25] composed by colour images with different resolutions. Images of size 512×512 pixels were converted to grey-level textures and split into 16 non-overlapping windows, resulting in 54 classes with 16 samples each.

For the UIUC database we use a more recent version employed in [33], composed by 25 classes with 40 grey-level samples per class and each sample has a resolution of 256×256 pixels.

The forth database is UMD [16], composed by 25 classes with 40 grey-level images in each one and each image has a high resolution of 1280×960. To speed up the computation, here a downsampled version was employed and each image has a 256×192 resolution.

The total number of descriptors depends on the maximum radius of the Bouligand-Minkowski dilation. Table 1 shows this relation, although in most cases, a reduced number of descriptors were enough to obtain the success rates reported here. The numbers used in practice are discussed in Section 6.

The descriptors were computed for each image and the results were compared (in terms of precision over a cross-validation classification of each database) with other classical and state-of-the-art approaches, to know, (Gabor [26], Fourier [27], Grey-Level Cooccurrence Matrix (GLCM) [28], Multifractal [16], Local Binary Patterns (LBP) [29], Soft-LBP [35], Fuzzy-LBP [36] and textons (VZ) [34]).

The classification was carried out through a Linear Discriminant Analysis

Table 1: Total number of descriptors for each radius used by the Bouligand-Minkowski dilation.

Dilation radius	Number of descriptors
1	6
2	24
3	48
4	84
5	132
6	186
7	252

using a 10-fold scheme for cross-validation [32]. The precision (named "Success Rate" in Results section), in this case, corresponds to the average percentage of images correctly classified in each round of the cross-validation procedure, according to the ground-truth provided by each database.

6. Results and Discussions

6.1. Parameter settings

The cube side-lengths considered to compute the connected components and therefore the Holder exponent were fixed between 1 and 6, as greater values became computationally unfeasible on standard hardware. For the threshold t, larger values were thought to provide more information as they produced larger components, however, distances greater than 10 were, again, computationally costly. Moreover, the use of larger values of t tends to make the coarseness measures quite similar along the neighbourhood and such local homogeneity would impair the ability of fractal descriptors to detect small-scale patterns in the texture.

The remaining variable to establish was the dilation radius in the Bouligand-Minkowski analysis. Table 2 shows the success rates of classification (as a percentage) for a number of dilation radii and the number of fractal descriptors

generated for Brodatz data set. This was done to identify radii values providing best performance with a minimum number of descriptors (and avoid over-training issues such as dimensionality curses).

Table 2: Success rates of classification and number of descriptors for different dilation radii on the Brodatz database.

Dilation radius	Success rate (%)	Number of descriptors
1	69.93	6
2	87.67	23
3	91.16	48
4	91.95	73
5	92.12	105
6	92.40	138
7	92.74	242

The same test and results on the Vistex database are shown in Table 3, for UIUC in Table 6.1 and for UMD in Table 6.1.

Table 3: Success rates of classification and number of descriptors for different dilation radii on the Vistex database.

Dilation radius	Success rate (%)	Number of descriptors
1	62.74	6
2	89.59	23
3	91.44	33
4	93.87	73
5	95.03	132
6	95.14	118
7	95.60	142

From the results above, we concluded that 6 was a reasonable value for the dilation radius as it produces a number of descriptors similar to those used in other established methods while outperforming them in the rates of classification obtained.

Table 4: Success rates of classification and number of descriptors for different dilation radii on the UIUC database.

Dilation radius	Success rate (%)	Number of descriptors
1	59.10	6
2	73.30	24
3	81.70	48
4	86.80	65
5	87.90	112
6	89.50	124
7	90.10	229

Table 5: Success rates of classification and number of descriptors for different dilation radii on the UMD database.

Dilation radius	Success rate (%)	Number of descriptors
1	66.90	5
2	81.50	19
3	86.40	47
4	90.80	83
5	93.10	116
6	94.00	150
7	93.90	178

6.2. Classification

The graphs and tables below show the results obtained by the proposed method as well as the performance of other approaches. Figure 4 shows the success rates on the Brodatz database according to the number of descriptors used. Most methods show a rapid increase in the correct classification rates as the number of descriptors increase, then reaching a stability level. While our method does not produce the best performance with few descriptors it shows the best performance when these are 25. Furthermore the performance increase appears to be smoother than for the other methods, too.

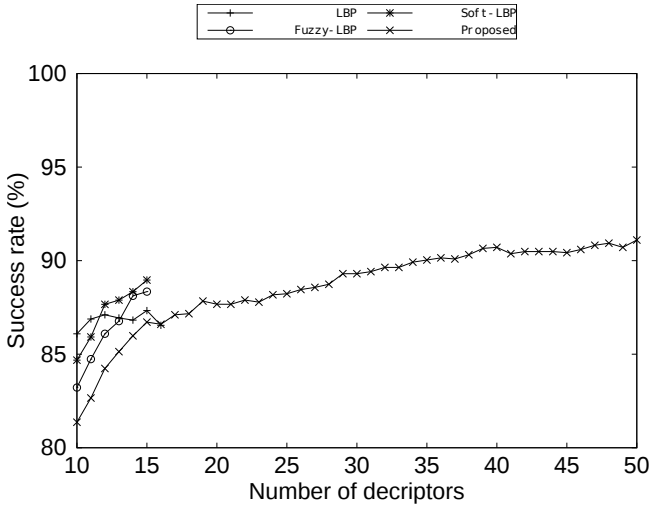

Figure 4: Success rates of classification and number of descriptors for various methods applied to the Brodatz database.

Table 6 shows the best classification results achieved for each method in the previous graph, the number of descriptors necessary to reach such rate and the associated cross-validation error. Except for the classical Fourier approach, the other methods have similar performances (around 86%), while however our proposed method achieved nearly 92% (a notable improvement, given the size and complexity of the textures database).

Figure 5 shows another very useful and hepful way of evaluating the performance of a classifier, i.e, the confusion matrix. This diagram aims to show

Table 6: Success classification rates (with the respective errors) and number of descriptors for various methods applied to the Brodatz database.

Method	Success rate (%)	Number of descriptors
LBP	87.33±0.02	15
GLCM	86.48±0.02	70
Multifractal	85.64±0.03	70
Gabor	85.42±0.02	19
Fourier	78.71±0.03	15
Fuzzy LBP	88.34±0.02	15
Soft LBP	88.96±0.02	15
Textons	81.47±0.02	97
Proposed	91.84±0.01	96

the number of elements from the class A (expected) that were assigned to the class B (predicted). The elements that are correctly classified are represented in the diagonal, while the number of misclassified samples can be infered from the outside. Here the values on the diagonal are expressed by grey-levels (the darker the point, the higher the number of samples), whilst those outside are depicted in red levels only to facilitate the visualization. In this type of representation, an ideal classifer is expected to have the maximum possible of dark points on the diagonal and the minimum in the outside. In Figure 5, as supposed, the best methods in this visual sense are LBP and the proposed descriptors. However, although they behave in a similar way in some cases like the confused samples around the class 40, the proposed fractal descriptors exhibits less dark red points confirming the higher correctness rate and identifying the classes where more or less samples can be distinguished.

Figure 6 shows the classification rates for the Vistex database. The relative results are similar to those for Brodatz, although the curves here are more irregular; this is likely due to a more pronounced variation in patterns amongst the images.

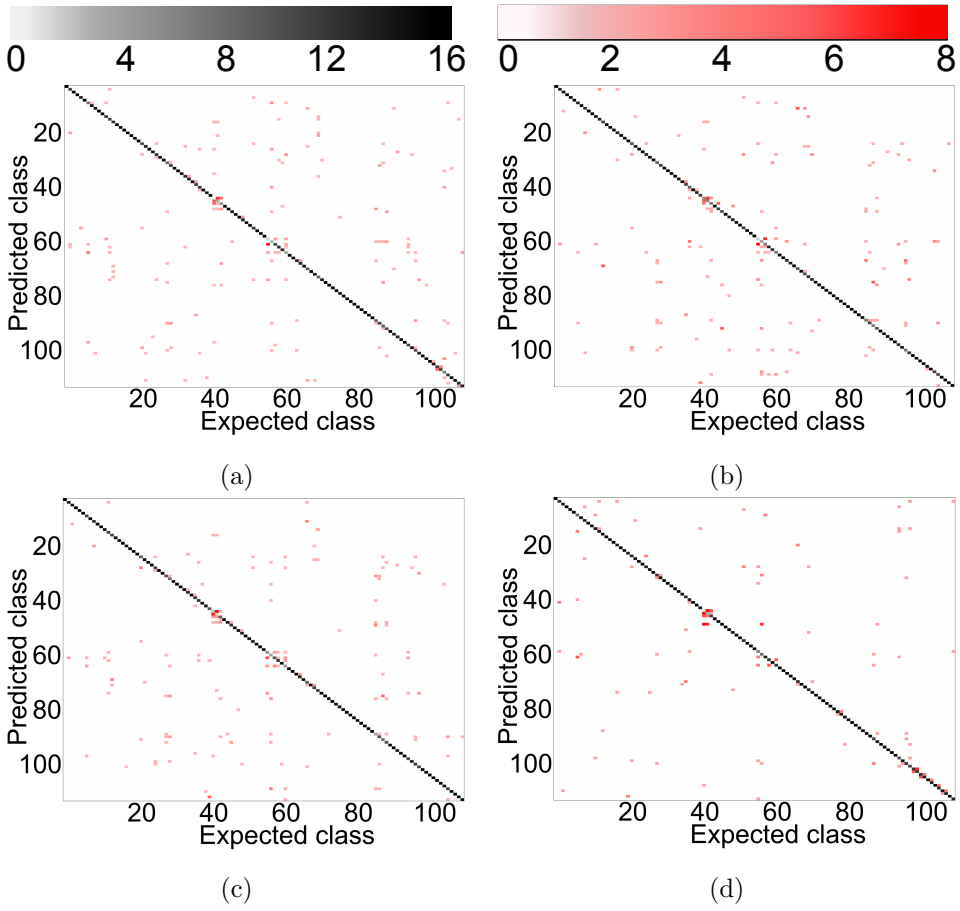

Figure 5: Confusion matrices for the methods on Brodatz data set. (a) LBP. (b) Fuzzy-LBP. (c) Soft-LBP. (d) Proposed method.

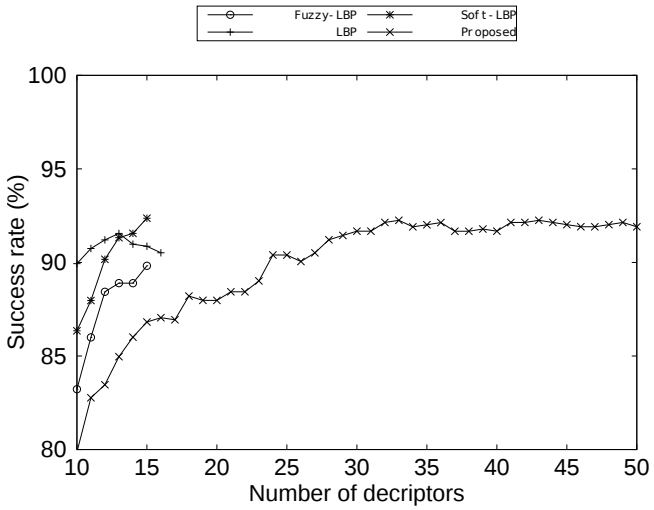

Figure 6: Success classification rates and number of descriptors for the Vistex database.

Table 7 shows the percentage of images correctly classified in Vistex database. The reduced number of classes appears to benefit the performance of the methods, as the rates are greater than for the Brodatz database results. Again, the proposed method demonstrates its potential in the analysis of complex textures; the Vistex images contain a high level of heterogeneity caused by shadows, orientation, scale, etc. However, the combination of local and global multiscale analysis in the the proposed method achieved about 94.5% correct classification.

Figure 7 shows the confusion matrices for Vistex in the same scheme used for Brodatz. Despite the difference in appearance with the matrices for the analysis of the Brodatz textures (which have fewer classes, Figure 5), the proposed descriptors correctly identifies more classes. Apart from the reduced number of red points, this also can be observed by a diagonal more continuous, with less grey gaps. Moreover, the classification errors are not concentrated at any specific classes, ensuring higher reliability to the method for a practical application.

Figure 8, Table 8 and Figure 9 show the results for the classification of UIUC database, following the same scheme adopted for the previous databases,

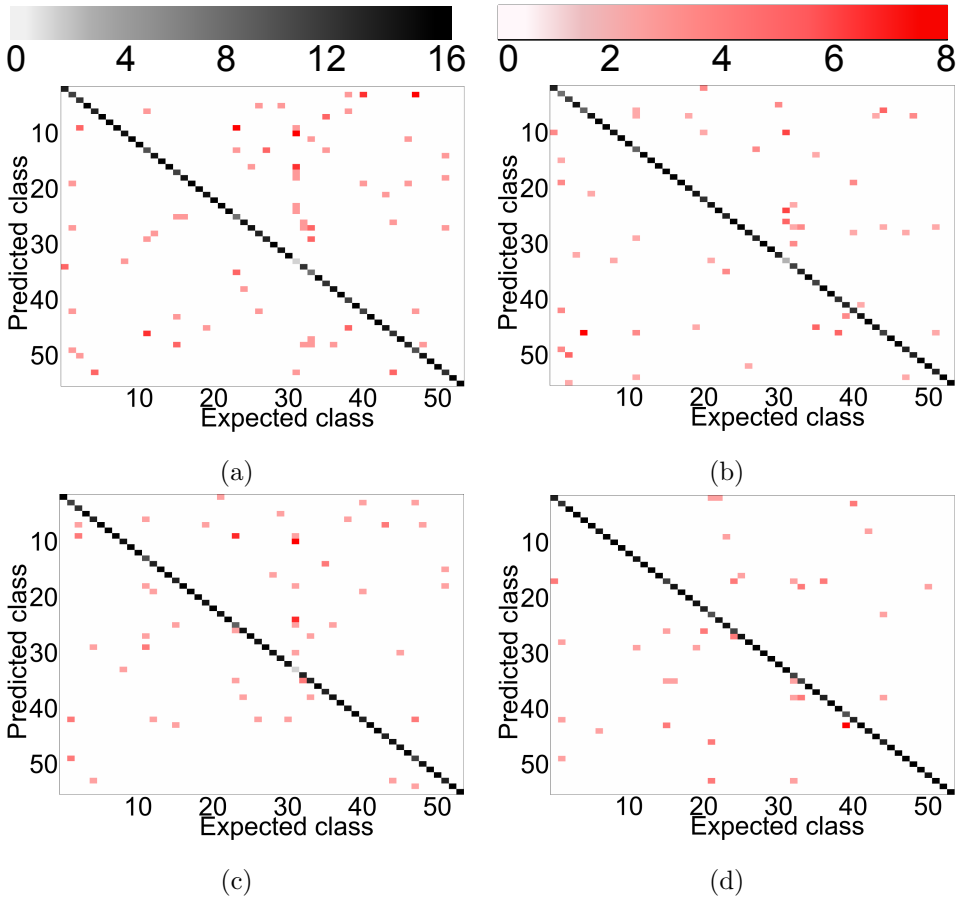

Figure 7: Confusion matrices for the methods on Vistex data set. (a) Fuzzy-LBP. (b) LBP. (c) Soft-LBP. (d) Proposed method.

Table 7: Success classification rates (with respective errors) and number of descriptors for various methods applied to the Vistex database.

Method	Success rate (%)	Number of descriptors
LBP	91.55±0.03	13
GLCM	88.21±0.03	70
Multifractal	88.31±0.03	76
Gabor	90.39±0.01	17
Fourier	84.49±0.02	15
Fuzzy LBP	89.82±0.04	15
Soft LBP	92.36±0.03	15
Textons	86.00±0.02	98
Proposed	94.45±0.02	74

that is, success rate against number of descriptors, highest success rates and confusion matrices. An interesting point to be observed in this case is that methods like LBP and variants, which provided good results in the classification of Vistez and Brodatz, now present results below the average. The main cause of such discrepancy are the significant changes in viewpoint, scale and illumination conditions on UIUC samples. Approaches like classical LBP (and its variants) and GLCM focus their analysis on grey-levels and local neighbourhood, whereas their global descriptors are not complex and precise enough to identify a global change in albedo for example. On the other hand, multifractals and textons, using, respectively, local measures invariant to illumination and multiple types of filters, were capable of identify samples even when they are presented under different perspectives and distances from the observer. Finally, the proposed method again achieved the highest percentage of images correctly classified. Such performance is a consequence, as stated before, of the efficiency of combining two complementary local and global fractal analysis, addressing the respective variations amongst samples from the same class by means of descriptors that quantify the local regularity as well as the texture

patterns at each scale instead of using only the pixel intensities or ambiguous global measures like histograms or fractal dimension.

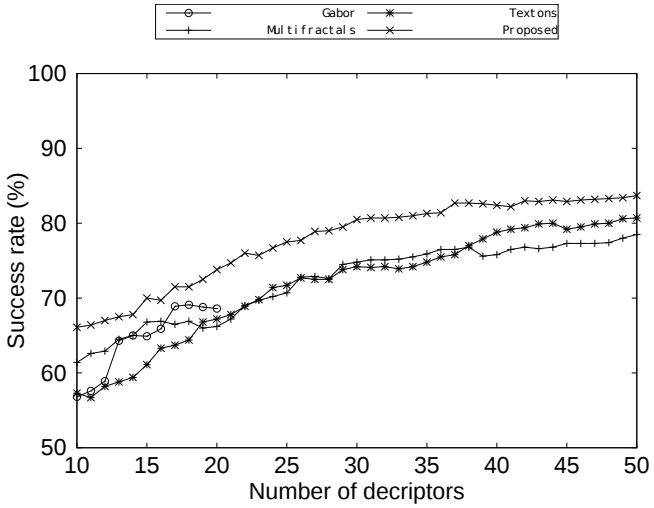

Figure 8: Success classification rates and number of descriptors for the UIUC database.

Table 8: Success classification rates (with respective errors) and number of descriptors for various methods applied to the UIUC database.

Method	Success rate (%)	Number of descriptors
LBP	57.80±0.05	14
GLCM	58.70±0.03	13
Multifractal	82.40±0.03	70
Gabor	69.10±0.03	18
Fourier	67.30±0.03	10
Fuzzy LBP	51.60±0.06	15
Soft LBP	54.90±0.05	15
Textons	86.70±0.03	97
Proposed	88.00±0.02	85

Finally, Figure 10, Table 9 and Figure 11 exhibit the results for the classification of UMD by the compared descriptors. A similar discussion to that for UIUC database is also valid here. In fact, UMD shares similarities with UIUC,

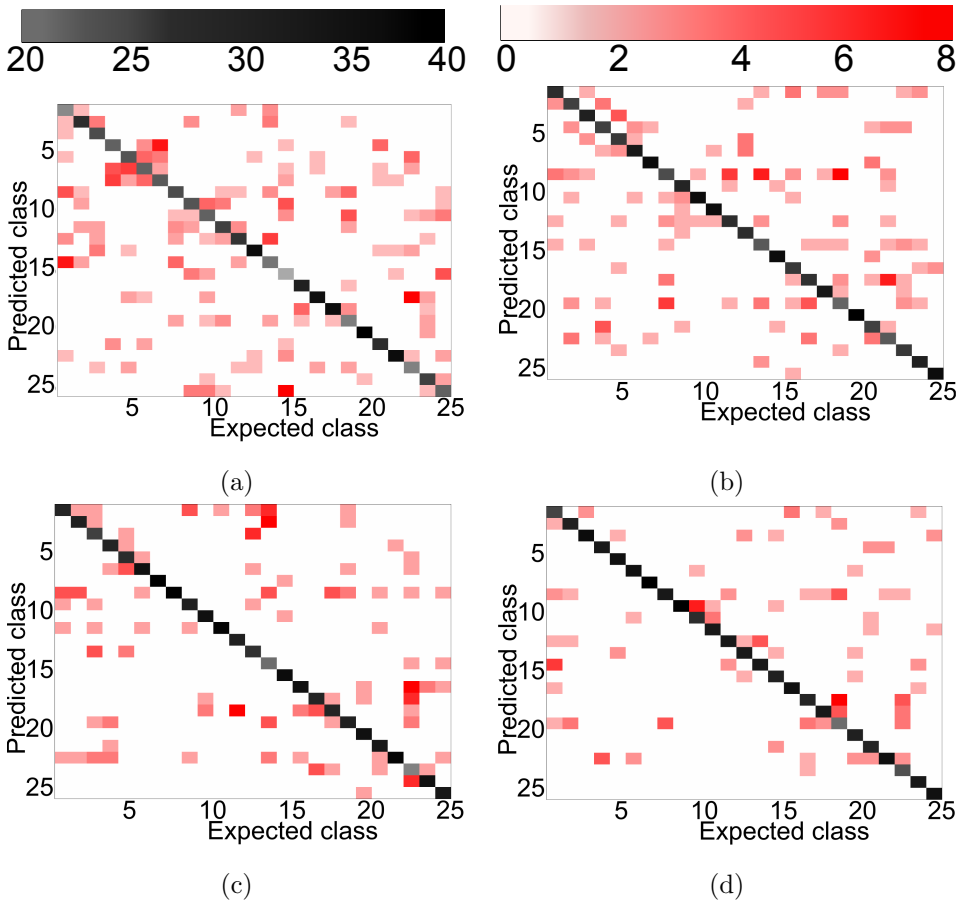

Figure 9: Confusion matrices for the methods on UIUC data set. (a) Gabor. (b) Multifractals. (c) Textons. (d) Proposed method.

mainly in their high variance of viewpoints and scales, in addition to the uncontrolled illumination conditions. The good results for multifractals was also expected since the database was proposed and employed in the corresponding paper ([16]). Another observation is that in Figure 10 multifractal descriptors outperformed the proposed method for a number of descriptors smaller than 50. This can also be explained by characteristics of the database and, particularly, by the complexity of the samples, containing, for instance, dozens of packets of wall anchors or tins of joint compound, each unit with complex labels attached. Even in this case, however, the proposed method provided the highest success rate when more than 80 descriptors were employed.

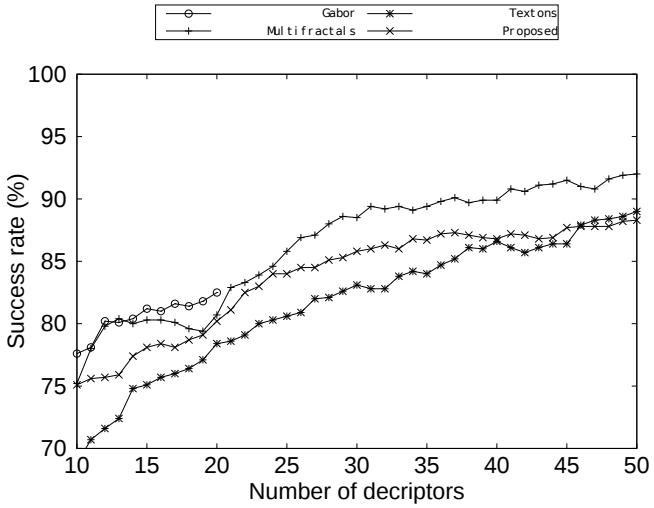

Figure 10: Success classification rates and number of descriptors for the UMD database.

Based on these results, we conclude that the combination of fractal descriptors and local coarseness Holder exponent allows a rich and precise description of complex and heterogeneous textures. One advantage of our method is that it analyses two domains. The first one is local and provides measures of the clustering of pixel neighbourhoods while the second, deals with the distribution of the clustering across the image, giving a measure of its homogeneity. Moreover, computing the coarseness at different cube sizes makes this procedure a multiscale analysis. Furthermore, the fractal descriptors provide a detailed analysis

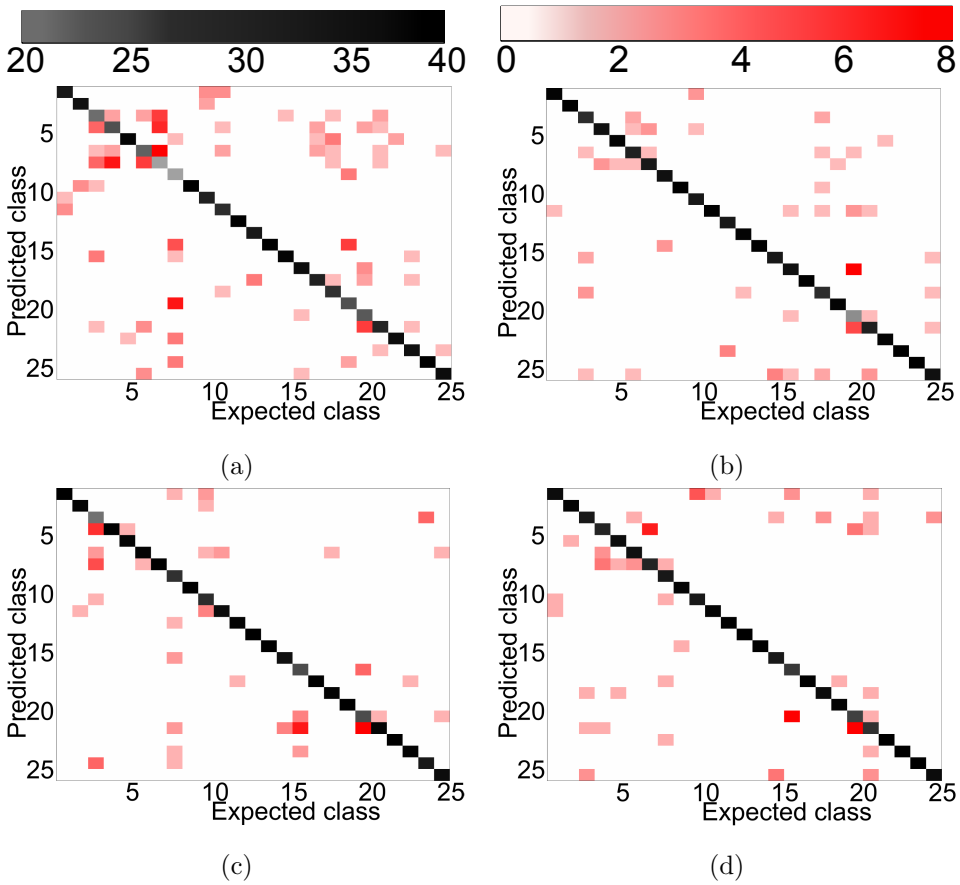

Figure 11: Confusion matrices for the methods on UMD data set. (a) Gabor. (b) Multifractals. (c) Textons. (d) Proposed method.

Table 9: Success classification rates (with respective errors) and number of descriptors for various methods applied to the UMD database.

Method	Success rate (%)	Number of descriptors
LBP	77.90±0.03	12
GLCM	78.20±0.04	16
Multifractal	92.10±0.03	72
Gabor	82.50±0.03	20
Fourier	75.50±0.02	15
Fuzzy LBP	78.20±0.03	15
Soft LBP	77.80±0.03	14
Textons	92.50±0.02	100
Proposed	93.00±0.02	100

in terms of the spatial distribution of such clustering measure.

We empirically identified that with regards to the local fractal measure of natural images, the connectivity itself can also be characterised as self-similar (as shown in [22]) while the Bouligand-Minkowski descriptors summarise how the variation of such self-similarity occurs at the different scales analysed [20, 17].

The result of this double-level multiscale and fractal analysis (through fractal descriptors and the local coarseness, respectively) leads to a set of quantifiers that are perhaps more robust to abrupt texture variations. Such robustness is a consequence of two main points: firstly, the local dimension is taken over a neighbourhood, which attenuates the effect of a punctual irregularity (noise, for example) in the pixel and, secondly, the fractal descriptors capture relevant information at different scales and, hence, localised variations would not compromise the global performance of the descriptor to a large extent. The inherent multiscale procedure also retains information about different levels of details in the images, making possible a more precise and reliable classification.

7. Conclusions

We proposed a new method to extract texture descriptors from grey-level images by computing the Bouligand-Minkowski fractal descriptors from a matrix containing the coarseness Holder exponent (logarithm of the number of points connected to each pixel in the original image).

The method was applied to classify well-known databases and the performance compared to other classical and state-of-the-art texture analyses published in the literature. Our method outperformed all other analyses for the compared databases.

The results suggest that fractal descriptors and local coarseness exponent provide complementary information about the textures. While the coarseness index measures pixel clustering and consequently the neighbourhood homogeneity, the fractal descriptors provide a measure of the regularity of the distribution of Holder exponents, and consequently the distribution of patterns along scales. After removal of redundancies through PCA, the descriptors become a powerful tool to represent and describe complex textures.

Acknowledgments

Odemir Martinez Bruno gratefully acknowledges the financial support of CNPq (National Council for Scientific and Technological Development, Brazil) (Grant Nos. 308449/2010-0 and 473893/2010-0) and FAPESP (Grant No. 2011/01523-1). Joao Batista Florindo acknowledges support from FAPESP (The State of São Paulo Research Foundation) (Grant No. 2013/22205-3).

References

[1] J. Wang, N. Zheng, Y. Liu, G. Zhou, Parameter analysis of fractal image compression and its applications in image sharpening and smoothing, Signal Processing: Image Communication 28 (6) (2013) 681 – 687.

[2] H. Xiaoqing, Z. Qin, L. Wenbo, A new method for image retrieval based on analyzing fractal coding characters, Journal of Visual Communication and Image Representation 24 (1) (2013) 42 – 47.

[3] C.-W. Shih, H.-C. Chu, Y.-M. Chen, C.-C. Wen, The effectiveness of image features based on fractal image coding for image annotation, Expert Systems with Applications 39 (17) (2012) 12897 – 12904.

[4] X.-Y. Wang, Y.-X. Wang, J.-J. Yun, An improved no-search fractal image coding method based on a fitting plane, Image and Vision Computing 28 (8) (2010) 1303 – 1308.

[5] A. Napryushkin, V. Kibitkin, V. Pleshanov, Linear transformation based error correction algorithm for fractal dimension estimation of images, Signal Processing 90 (6) (2010) 2094 – 2101.

[6] M. Keyvanpour, F. Merrikh-Bayat, An effective chaos-based image watermarking scheme using fractal coding, Procedia Computer Science 3 (0) (2011) 89 – 95.

[7] J. Velzquez-Garca, K. Oleschko, J. A. Muoz-Villalobos, M. Velsquez-Valle, M. M. Menes, J.-F. Parrot, G. Korvin, M. Cerca, Land cover monitoring by fractal analysis of digital images, Geoderma 160 (1) (2010) 83 – 92.

[8] G. Landini, Fractals in microscopy, Journal of Microscopy 241 (1) (2011) 1–8.

[9] G. Landini, J. W. Rippin, How important is tumour shape? quantification of the epithelial-connective tissue interface in oral lesions using local connected fractal dimension analysis, The Journal of Pathology 179 (2) (1996) 210–7.

[10] P.-L. Lin, P.-W. Huang, C.-H. Lee, M.-T. Wu, Automatic classification for solitary pulmonary nodule in {CT} image by fractal analysis based on fractional brownian motion model, Pattern Recognition 46 (12) (2013) 3279 – 3287.

[11] J. xiang Du, C.-M. Zhai, Q.-P. Wang, Recognition of plant leaf image based on fractal dimension features, Neurocomputing 116 (0) (2013) 150 – 156.

[12] Y. Wu, Q. Lin, Z. Chen, W. Wu, H. Xiao, Fractal analysis of the retrogradation of rice starch by digital image processing, Journal of Food Engineering 109 (1) (2012) 182 – 187.

[13] J. B. Florindo, M. S. Sikora, E. C. Pereira, O. M. Bruno, Characterization of nanostructured material images using fractal descriptors, Physica A: Statistical Mechanics and its Applications 392 (7) (2013) 1694 – 1701.

[14] J. B. Florindo, N. R. da Silva, L. M. Romualdo, F. de Fátima da Silva, P. H. de Cerqueira Luz, V. R. Herling, O. M. Bruno, Brachiaria species identification using imaging techniques based on fractal descriptors, Computers and Electronics in Agriculture 103 (0) (2014) 48 – 54.

[15] E. T. M. Manoel, L. da Fontoura Costa, J. Streicher, G. B. Müller, Multiscale fractal characterization of three-dimensional gene expression data, in: SIBGRAPI, IEEE Computer Society, 2002, pp. 269–274.

[16] Y. Xu, H. Ji, C. Fermüller, Viewpoint invariant texture description using fractal analysis, International Journal of Computer Vision 83 (1) (2009) 85–100.

[17] O. M. Bruno, R. de Oliveira Plotze, M. Falvo, M. de Castro, Fractal dimension applied to plant identification, Information Sciences 178 (12) (2008) 2722–2733.

[18] J. B. Florindo, O. M. Bruno, Texture analysis by multi-resolution fractal descriptors, Expert Systems with Applications 40 (10) (2013) 40224028.

[19] J. B. Florindo, A. R. Backes, M. de Castro, O. M. Bruno, A comparative study on multiscale fractal dimension descriptors, Pattern Recognition Letters 33 (6) (2012) 798–806.

[20] A. R. Backes, D. Casanova, O. M. Bruno, Plant leaf identification based on volumetric fractal dimension, International Journal of Pattern Recognition and Artificial Intelligence 23 (6) (2009) 1145–1160.

[21] A. R. Backes, D. Casanova, O. M. Bruno, Color texture analysis based on fractal descriptors, Pattern Recognition 45 (5) (2012) 1984 – 1992.

[22] G. Landini, P. I. Murray, G. P. Misson, Local connected fractal dimensions and lacunarity analyses of 60 degrees fluorescein angiograms, Investigative Ophthalmology and Visual Science 36 (13) (1995) 2749–2755.

[23] H. Peitgen, H. Jürgens, D. Saupe, Chaos and Fractals: New Frontiers of Science, Springer, 2004.

[24] P. Brodatz, Textures: A photographic album for artists and designers, Dover Publications, New York, 1966.

[25] Vistex, Vistex. vision texture database (2009).
URL `http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html`

[26] B. Manjunath, W. Ma, Texture features for browsing and retrieval of image data, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (8) (1996) 837–842.

[27] R. C. Gonzalez, R. E. Woods, Digital Image Processing (2nd Edition), Prentice Hall, Upper Saddle River, 2002.

[28] R. M. Haralick, Statistical and structural approaches to texture, Proceedings of the IEEE 67 (5) (1979) 786–804.

[29] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on featured distributions, Pattern Recognition 29 (1) (1996) 51–59.

[30] K. Falconer, Fractal Geometry: Mathematical Foundations and Applications, Wiley, Chichester,UK, 2003.

[31] J. C. Russ, Fractal Surfaces, Plenum Press, New York, 1994.

[32] R. O. Duda, P. E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973.

[33] K. Kilic, R. Abiyev, Exploiting the synergy between fractal dimension and lacunarity for improved texture recognition, Signal Processing 91 (10) (2011) 2332–2344.

[34] M. Varma, A. Zisserman, A Statistical Approach to Texture Classification from Single Images, International Journal of Computer Vision 62 (1-2) (2005) 61 – 81.

[35] T. Ahonen, M. Pietik, M. Pietikäinen, Soft histograms for local binary patterns, in: Proceedings of the Finnish signal processing symposium, IEEE, 2007, pp. 1–4.

[36] D.K. Iakovidis, E.G. Keramidas, D. Maroulis, Fuzzy Local Binary Patterns for Ultrasound Texture Characterization, in: Image Analysis and Recognition, Springer Berlin / Heidelberg, 2008, pp. 750–759.

[37] M. Varma, R. Garg, Locally Invariant Fractal Features for Statistical Texture Classification, in: IEEE 11th International Conference on Computer Vision, IEEE, 2007, pp. 1–8.

[38] H. Ji, X. Yang, H. Ling, Y. Xu, Wavelet Domain Multifractal Analysis for Static and Dynamic Texture Classification, IEEE Transactions on Image Processing 22 (1) (2013) 286 – 299.

[39] Y. Xu, S. Huang, H. Jib, C. Fermüller, Scale-space texture description on SIFT-like textons, Computer Vision and Image Understanding 116 (1) (2012) 999 – 1013.

[40] Y. Xu, H. Ji, C. Fermüller, A projective invariant for textures, in: Computer Vision and Pattern Recognition, IEEE, 2006, pp. 1932–1939.

[41] Y. Xu, X. Yang, H. Ling, H. Ji, A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid, in: Computer Vision and Pattern Recognition, IEEE, 2010, pp. 161–168.

[42] B. B. Mandelbrot, The Fractal Geometry of Nature, Freeman, New York, 1982.

[43] R. Fabbri, L. Costa, J. Torelli, O. Bruno, 2D Euclidean distance transform algorithms: A comparative survey, ACM Computer Surveys 40 (1) (2008) 1 – 44.

[44] R. A. Eid, G. Landini, Quantification of the global and local complexity of the epithelial-connective tissue interface of normal, dysplastic, and neoplastic oral mucosae using digital imaging, Pathology Research & Practice 199 (7) (2003) 475 – 482.