Fractal measures of image local features: an application to texture recognition

Pedro M. Silva[1] · Joao B. Florindo[2]

Abstract

Here we propose a new method for the classification of texture images combining fractal measures (fractal dimension, multifractal spectrum and lacunarity) with local binary patterns. More specifically we compute the box counting dimension of the local binary codes thresholded at different levels to compose the feature vector. The proposal is assessed in the classification of three benchmark databases: KTHTIPS-2b, UMD and UIUC as well as in a real-world problem, namely the identification of Brazilian plant species (database 1200Tex) using scanned images of their leaves. The proposed method demonstrated to be competitive with other state-of-the-art solutions reported in the literature. Such results confirmed the potential of combining a powerful local coding description with the multiscale information captured by the fractal dimension for texture classification.

Keywords Fractal geometry · Multifractals · Texture classification · Box counting · Local binary patterns

1 Introduction

Texture analysis, and in particular texture recognition, has been one of the most important tasks in computer vision. Despite most of the initial applications of this area being focused

✉ Joao B. Florindo
   jbflorindo@ime.unicamp.br

   Pedro M. Silva
   icmpedro@gmail.com

[1] Federal Institute of Education, Science and Technology of Espírito Santo, Rodovia Governador José Sete, nº 184 - Itacibá, CEP 29150-410, Cariacica, ES, Brasil

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

on material sciences, during the last decades it has been applied to problems in a wide range of research areas, such as Medicine [8, 21], Biology [34, 37], Engineering [12, 43], Physics [14, 45], and many others.

Although texture images do not have a consensual formal definition, they are usually associated to structures commonly found in nature and composed by elements whose complexity makes them difficult to be described by the classical Euclidean geometry. This observation naturally opened space for the possibility of analyzing such objects by employing techniques developed in fractal geometry [34, 38, 39]. Indeed, a fractal is a geometrical set characterized by the self-repetition of basic patterns and by a high degree of complexity. It was in fact described by Mandelbrot in his seminal book [26] as the "geometry of nature".

Despite the recent popularization of deep learning / neural networks methods in texture analysis, the intrinsic connection between the way that natural structures are composed and the self-similar replication of fractals allows the development of specific purpose models, which present some advantages over deep learning methods, like not requiring so much annotated data for training and the possibility of more straightforward interpretation of the obtained results. The difficulty in obtaining large amounts of annotated data and importance of result interpretation are key aspects, for example, in the analysis of medical images.

However, the theoretical interest in fractals still struggles with some practical limitations. The first and most obvious one is that there is no mathematical fractal in nature. Fractals are defined in infinite scales and this is obviously infeasible in the real world. In this way we need to be aware that the fractal geometry representation is an approximation, a modeling of a real system. A second point is that fractal geometry deals with geometrical sets embedded within a continuous N-dimensional space, whereas the most practical way of analyzing a real world object is by inspecting its digital image representation. This issue was addressed for example in [29], where a relation between the "fractality" of a surface and its respective photography was established. That work also accomplished a psychological experiment that established connections between the fractal dimension and the natural visual perception of attributes such as smoothness and roughness.

Since then, a number of works applying fractal geometry to texture recognition have been presented in the most diverse areas of applications. Nevertheless, a third point still remains without a convincing solution: this is the ideal representation of the image to estimate measures like the fractal dimension.

Among the attempts we have seen the fractional Brownian motion (fBm) model [44] (and similar variations like the Fourier-based fractal dimension [32]) and the geometrical approaches [9]. fBm has achieved considerable success in texture synthesis, but there is no relevant success reported in recognition. A possible explanation for this is the high degree of specialization of the model, which is the reason for its success in texture synthesis, but at the same time also makes it highly prone to over-fitting in texture recognition. On the other hand, the geometrical approaches rely on interpreting the image as a surface (or cloud of points) embedded in the three-dimensional Euclidean space. This approach has achieved interesting results in texture classification [11]. However, it clearly lacks an effective explanation for using the pixel spatial domain (its localization) and the intensity domain (its gray value) to compose a unique three-dimensional space. Questions like "how a homogeneous resolution scale can be defined in such conditions?" are definitely not well answered.

The multifractal theory [38, 39, 41, 42] represented a noticeable attempt to address this issue with an advanced mathematical framework based on Measure Theory. It also achieved promising results on classification problems [38, 39], even though not too much of these results have been reported on more modern and challenging texture databases.

Finally, a simple solution would be to take a family of images resulting from thresholding the original texture image at different levels. The direct application of this strategy to the texture image is not feasible as the relation between the gray values of neighbor pixels is lost and this is known to be a fundamental characteristic to distinguish textures of different materials. The authors in [31] elegantly solved this issue by applying the threshold strategy to the local binary pattern (LBP) [28] map of the texture. This is a representation well known to capture the relations between neighbor pixels [24]. This allows the preservation of such locality information even in the thresholded image.

Based on this context, we propose here the combination of numerical techniques to estimate the fractal dimension of binary images with a well known local encoding of texture images, namely, LBP. We use the fractal dimension of the encoded image thresholded at a range of levels to compose a vector of image descriptors, which are employed for texture classification. To compose the descriptors we verify the use of box-counting and Bouligand-Minkowski fractal dimension as well as the lacunarity measure and the multifractal spectrum.

Two main novelties are presented by the proposed method: the combined use of numerical estimators of fractal dimension, lacunarity and multifractal spectrum to summarize the information expressed by an LBP map, replacing the basic histogram typically used, and a statistical model explaining how the numerical estimators of fractal dimension analyzed here are capable of quantifying and distinguishing complex patterns arising in natural textures.

The proposed descriptors are compared with other fractal-based texture features, namely, invariant multifractals [38], wavelet multifractals [39], and pattern lacunarity spectrum [31]. Other classical and state-of-the-art texture descriptors are also compared, such as VZ-MR8 [35], local binary patterns [28], convolutional neural networks [6], and others. Comparative tests were carried out over three well known texture benchmark data sets, to know, KTH-TIPS2b, UIUC and UMD. The descriptors were also applied to the identification of species of Brazilian plants (database 1200Tex). The proposed method outperformed the compared approaches in terms of classification accuracy and the results confirmed the potential of such a strategy to provide rich and meaningful texture descriptors.

2 Fractal geometry

Roughly speaking, a fractal is a mathematical object defined at infinite scales and characterized by self-similarity, i.e., repetition of geometrical patterns at different scales and high degree of complexity, which basically means that a hypothetical observer would never see the same object at different scales. It is also commonly distinguished from Euclidean elements for presenting non-integer dimension. Its formal definition is in fact stated in terms of a dimension, technically called Hausdorff dimension.

2.1 Hausdorff dimension

Given a set X, its Hausdorff measure is defined as:

$$\mathcal{H}^s(X) = \lim_{\sigma \to 0} H^s_\sigma(X) \text{ where } H^s_\sigma(X) = \inf \left\{ \sum_{i=1}^{\infty} |X_i|^s : X_i \text{ is a } \sigma\text{-cover of } X \right\}, \quad (1)$$

where $|X_i|$ is the diameter of X_i given by $|X_i| = \sup\{d(x, y) : x, y \in X_i\}$, being $d(x, y)$ a metric. We say that a family of sets $\{X_i\}$ is a σ-cover of a set X if:

$$\begin{cases} X \subset \bigcup_{i=1}^{\infty} X_i \\ 0 < |X_i| < \sigma. \end{cases} \tag{2}$$

If X is a fractal structure, it can be demonstrated that there exists a real and non-negative value D such that $H_{\sigma}^s = \infty$ for $s < D$ and $H_{\sigma}^s = 0$ for $s > D$. Then D is defined as the *Hausdorff dimension* of X. Formally:

$$D = \inf\{s : \mathcal{H}^s(F) = 0\} = \sup\{s : \mathcal{H}^s(F) = \infty\}. \tag{3}$$

To gain some intuition on this formalism, we illustrate it with a rather simplified example. Let us suppose that X is a hyper-cube corresponding to the interval $[0, 1]^D$ and we are trying to fill it with s-dimensional cubes with side-length $\frac{1}{n}$. This corresponds to a σ-cover of X with $\sigma = \frac{1}{n}$. If D is integer, it is easy to see that we need n^D of such $\frac{1}{n}$-length cubes to perfectly cover X. As we are assuming that all cubes have the same side-length, the sum of diameters, which is the idea of $\sum_{i=1}^{\infty} |X_i|^s$ in (1), would roughly be the number of cubes multiplied by the side-length, i.e., $n^D \cdot \frac{1}{n^s} = n^{D-s}$. A cover with $\sigma \to 0$ corresponds in this case to $n \to \infty$. But if $n \to \infty$, we have that $n^{s-D} \to \infty$ if $s < D$ and $n^{D-s} \to 0$ if $s > D$. Therefore D is the dimension of X. The Hausdorff dimension is basically the extension of such idea to non-integer values of D. The reader can refer to the specialized literature, e.g. in [9], for more details.

The most widely accepted definition of fractal is that of a geometrical set whose Hausdorff dimension strictly exceeds its Euclidean (topological) dimension.

2.2 Numerical estimates in fractal geometry

Although objects with fractal characteristics can be easily found in nature, such real-world structures differ from mathematical fractals in two crucial points: first, they do not have infinite self-similarity; second, the rules dictating the formation of the structure are usually not known. The analytical calculation of the Hausdorff dimension in this scenario is infeasible [26]. In this context, a large number of numerical values were developed to estimate to which extent these objects can be approximated by a fractal set. Especial interest has been devoted to methods that estimate the fractal dimension of these structures.

Essentially, the computation of the Hausdorff measure involves an infinite covering by units with diameter smaller than σ. The diameters of these elements are therefore raised to an exponent s and summed up. A discrete approximation of this operation can be accomplished with the aid of an exponential function:

$$M_{\sigma} \propto \sigma^s, \tag{4}$$

where M_{σ} is a measure of the object at the scale σ, in such a way that any detail larger than σ is not counted. By changing the definition of M and σ, we have definitions of the fractal dimension that are alternative to Hausdorff formulation. These alternative definitions may (and often do) assume values that do not coincide with the Hausdorff dimension, but they preserve the idea of measuring the complexity and spatial occupation of the object: the most complex the structure, the highest the dimension. Many of these alternative definitions are more suitable for numerical computation. Among the most popular definitions possessing this property we have box-counting and Bouligand-Minkowski dimension and lacunarity [9, 40].

2.2.1 Box-counting

Mathematically, if X is a non-empty bounded subset of $\mathbb{R}^n$ and $N_\delta(X)$ is the smallest number of sets with diameter at most δ covering X, then we define the *lower* and *upper* *box-counting dimension* of X respectively by

$$\underline{\dim}_B(X) = \liminf_{\delta \to 0} \frac{\log N_\delta(X)}{-\log \delta}$$
$$\overline{\dim}_B(X) = \limsup_{\delta \to 0} \frac{\log N_\delta(X)}{-\log \delta}. \tag{5}$$

If both limits have the same value, we simply call it the *box-counting* dimension:

$$\dim_B(X) = \lim_{\delta \to 0} \frac{\log N_\delta(X)}{-\log \delta}. \tag{6}$$

In fact, this definition has strong connections with the Hausdorff definition (1). It basically means that $N_\delta(X) \simeq \delta^{-\dim_B(X)}$. In this way and considering that $\delta \ll 1$, we have $N_\delta(X)\delta^s \to \infty$ if $s < \dim_B(X)$ and $N_\delta(X)\delta^s \to 0$ if $s > \dim_B(X)$. But $N_\delta(X)\delta^s$ corresponds to $n^D \cdot \frac{1}{n^s}$ in our intuition on the Hausdorff dimension and that was also the sum of diameters in a δ-cover of X in (1). More details can be found in [9].

However, the most interesting equivalent definition of box-counting dimension for practical purposes is that defined under a mesh of hyper-cubes. In $\mathbb{R}^n$ we define a δ-mesh of hypercubes by

$$[m_1\delta, (m_1 + 1)\delta] \times \cdots \times [m_n\delta, (m_n + 1)\delta], \tag{7}$$

where $m_1, \cdots, m_n$ are integer numbers. The measure N_δ in (6) can be replaced by the number of δ-mesh cubes intersecting X as they can be demonstrated to be equivalent measures [9].

2.2.2 Bouligand-Minkowski

The first step in the calculus of the Bouligand-Minkowski dimension of $X \in \mathbb{R}^n$ is to dilate X by balls in $\mathbb{R}^n$ with radius δ forming the structure $D(\delta)$ according to:

$$D(\delta) = \{p : \|p - p'\|_2 \leq \delta, \forall p' \in X\}. \tag{8}$$

The hypervolume of $D(\delta)$ is obtained by simply counting the number of points in $D(\delta)$, i.e,

$$V(\delta) = \sum_{p \in \mathbb{R}^n} \chi_{D(\delta)}(p), \tag{9}$$

where χ is the characteristic function: $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ otherwise.

Finally, the dimension $\dim_M$ is provided by

$$\dim_M(X) = n - \lim_{\delta \to 0} \frac{\log V(\delta)}{\log \delta}. \tag{10}$$

The relation with Hausdorff dimension can be justified in a similar way to that employed for the box-counting dimension.

2.2.3 Lacunarity

For the calculus of the lacunarity of X a hypercube with side-length δ glides through X and a histogram $H_\delta(k)$ stores the number of hypercubes intersecting exactly k points of X. The lacunarity at scale δ is computed by

$$\Lambda_\delta(X) = \frac{E[(H_\delta)^2]}{(E[H_\delta])^2}, \tag{11}$$

where E is the expected value for the distribution represented by H. Finally, a scale-independent lacunarity is provided by

$$\Lambda(X) = \lim_{\delta \to 0} \frac{\log \Lambda_\delta(X)}{\log \delta}. \tag{12}$$

Roughly speaking, $\Lambda_\delta(X)$ captures the idea of a ratio between the variance of the number of hypercubes (assuming mean zero) and that number on its own (the square plays the role of a normalizer). Such variance represents, for example, the heteogeneity of a binary image for each δ. $\Lambda(X)$ measures how this measures varies with scale. The reader can find more details in [26].

2.2.4 Multifractals

For the multifractal spectrum we follow the method in [30]. First, the image is partitioned into boxes of sizes L. Therefore a local measure μ_i is defined for a range of q values as

$$\mu_i(q, L) = \frac{\mathfrak{p}_i^q(L)}{\sum_{i=1}^{N(L)} \mathfrak{p}_i^q(L)}, \tag{13}$$

where $\mathfrak{p}_i(L)$ is the ratio (probability) of white points (points of interest of the analyzed object) falling inside the ith box of size L and $N(L)$ is the number of L-sized boxes.

The spectrum $f(q)$ is obtained by

$$f(q) = \lim_{L \to 0} \frac{\sum_{i=1}^{N(L)} \mu_i(q, L) \log[\mu_i(q, L)]}{\log L}. \tag{14}$$

The essential idea of a multifractal spectrum is to capture "fractality" at two levels, locally and globally. Locally we define a measure μ, whereas globally the spectrum expresses the fractal dimension of the set of points whose local measure is α. Here, we employ the moments method as in [30] and it can be demonstrated that the exponent q is tightly related to α in the original definition.

3 Proposed method

We propose here the combined use of box-counting, Bouligand-Minkowski, lacunarity and multifractal spectrum to quantify the fractality of binary images with a well known local encoding of texture images, namely, LBP codes. We use the fractal measure of the encoded image thresholded at a range of levels to compose a vector of image descriptors, which are employed for texture classification.

In practice, for a binary image I like that used here, δ values follow an exponential sequence $\delta = 2, 4, 8, 16, \cdots, \log(M)$ where M is the smallest dimension of the image. In the calculus of box-counting dimension, for each δ the number of squares intersecting the object of interest is accumulated in the variable N_δ and the numerical dimension is estimated by

$$\dim_B(I) = -\alpha, \tag{15}$$

where α is the slope of the straight line fit by minimum least squares to the curve $\log \delta \times \log N_\delta$. We also employ the linear coefficient of the straight line to compose the feature vector. Figure 1 illustrates the procedure.
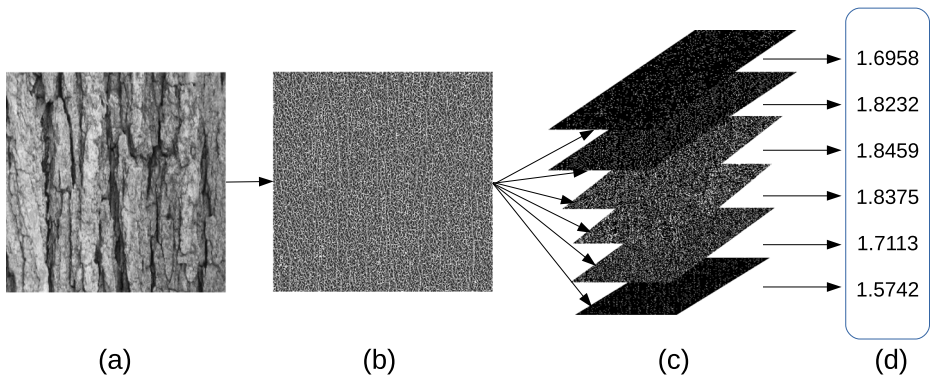
Fig. 1 Box counting descriptors for LBP mappings. To simplify the visualization we use $P = 4$ and $R = 1$ as LBP parameters. **a** Texture. **b** LBP codes. **c** LBP thresholds. **d** Box-counting dimension

The procedure to compute Bouligand-Minkowski dimension is similar and we use Euclidean distance transform to optimize the calculus of $V(\delta)$ and we employ a maximum radius of 9, as recommended in [3]. Again, we use the slope and linear coefficient of a straight line fit to $\log \delta \times \log V(\delta)$.

For the lacunarity, the radius δ ranges between 2 and 14, as recommended in [31] and we also employ both linear and angular coefficients to compose the feature vector.

Finally, the multifractal spectrum is computed according to the description in Section 2.2.4 and using $L = 2, 3, 5, 10, 25, 50, 100, 125, 250$ as in [30].

3.1 Motivation

Fractal dimension and other fractal measures in their strict sense, as in [9], are usually not suitable for the analysis of digital images. To start with, digital images are discrete and as a consequence their geometrical representation is countable. Fractal dimension of countable sets is, by definition, zero [9] and hence such images cannot be identified using fractal dimension. One could construct a continuous model based on the image, but that would introduce artificial data to the process.

Based on this context we develop here a statistical model to investigate how measures like box counting or Bouligand-Minkowski dimension behave in the analysis of digital images. For that purpose, LBP codes can be assumed to follow a uniform distribution, which substantially simplifies the statistical analysis. It is also worth to mention that the other measures explored here, i.e., lacunarity and multifractal spectrum, can be employed in a similar analysis, but the conclusions are expected to be similar, given that both also share the same main objective of measuring the fractality of the image.

3.1.1 Box counting

For box counting, we can employ a model similar to that developed in [20]. Let us suppose that the probability of a white point in the binary image (point of interest) is p and we have a total of N points. Let us also suppose that in $\mathbb{R}^1$ the data is enclosed within the interval $I = [0, 1]$. The probability that a box with size s (subinterval of I) does not contain any point is

$$p_0 = (1 - s)^{Np}. \tag{16}$$

Correspondingly, the probability of the same box to contain at least one point of interest is the complement of p_0:

$$p_1 = 1 - (1 - s)^{Np}. \tag{17}$$

The expected number of nonempty boxes used in the calculus of box counting dimension is given by

$$B(s) = [1 - (1 - s)^{Np}]/s. \tag{18}$$

Likewise, in $\mathbb{R}^2$ we work on an interval $I^2 = [0, 1] \times [0, 1]$ and have

$$B(s) = [1 - (1 - s^2)^{Np}]/s^2. \tag{19}$$

The trick in [20] is that for a self-similar set with self-similar dimension d_S we would have

$$B(s) = [1 - (1 - s^{d_S})^{Np}]/s^{d_S}. \tag{20}$$

Box counting dimension is obtained by fitting a straight line to the curve $\log s \times \log B(s)$ using least squares. Variable s assumes value within a predefined range (typically linear or exponential) $s_1, s_2, \cdots, s_n$. To simplify the mathematical notation we define $x_i = \log s_i$ and $y_i = \log B(s_i)$. Least squares have a statistical interpretation in which the straight line equation $y = \alpha x + \beta$ is defined in terms of variances and covariances.

At first, we define the variances of x and y and the covariance between x and y:

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \overline{x})^2}{n} \quad \sigma_y^2 = \frac{\sum_{i=1}^n (y_i - \overline{y})^2}{n} \quad \text{cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{n}, \tag{21}$$

where $\overline{x}$ and $\overline{y}$ denote the mean of x and y respectively. Then it is well known from least squares theory that

$$\alpha = \frac{\text{cov}(x, y)}{\sigma_x^2} \quad \text{and} \quad \beta = \overline{y} - \alpha \overline{x}. \tag{22}$$

Such relations express how the measure (number of boxes) is statistically associated to the scale s. Figure 2 shows α and β for different numbers of points N_p and dimension d_S. For different similarity dimensions, the box dimension (α) and linear coefficient (β) present different behaviors when plotted against the distribution of points (Np). In terms of image descriptors, this corresponds to a more complete representation than the sole dimension or the probability p, even though those important features continue to determine the shape of the curve.
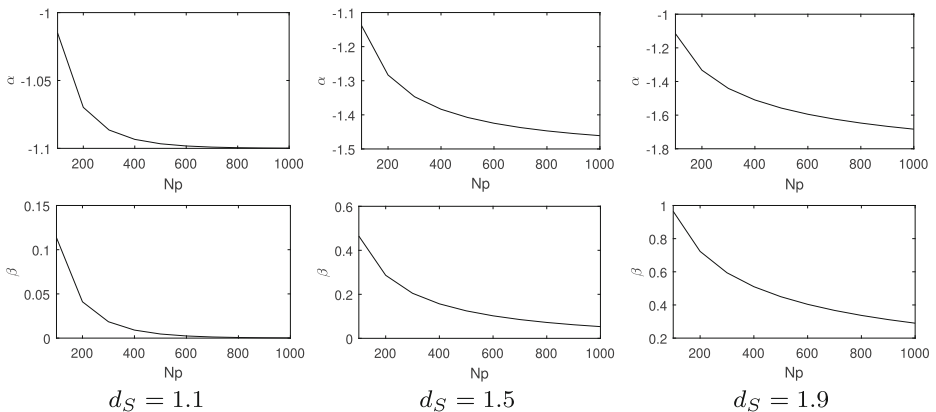

Fig. 2 α and β in box counting for different numbers of points Np and dimension d_S

3.2 Bouligand-Minkowski

A formal analysis of the dilation area of circles is significantly more complicated than for box counting and the interested reader can have some idea of that in [33]. Nevertheless, for a statistical study, we can reduce the problem to a one-dimensional binary "image" in $\mathbb{R}^1$, more specifically over the [0, 1] interval. Now the morphological dilation of Bouligand-Minkowski corresponds to placing bars with length s ($s < 1$) centered at points (white points) randomly dropped over the interval. The aim is to compute the expected length covered by at least one bar.

As we need to take care of the boundaries of [0, 1], it turns out that for a single bar, the probability p_1 of a point at the position x being covered can assume three possible values depending on the value of x:

$$p_1 = \begin{cases} x + \frac{s}{2} & \text{if } s \in [0, \frac{r}{2}] \\ s & \text{if } x \in [\frac{s}{2}, 1 - \frac{s}{2}] \\ -x + \frac{s}{2} + 1 & \text{if } x \in [1 - \frac{s}{2}, 1]. \end{cases} \tag{23}$$

The probability of x being uncovered is $1 - p_1$. In a similar way to what we did in box counting, we suppose that our "image" has N points and the probability of occurrence of a white point is p. Therefore the expected number of white points is Np. The probability of x being uncovered after n bars is randomly placed in [0, 1] is $(1 - p_1)^{Np}$. Hence the probability of x being covered by at least one bar is

$$p_n = 1 - (1 - p_1)^{Np}. \tag{24}$$

Finally the expected length of the covered region is given by

$$\langle L \rangle = \int_0^1 p_n dx. \tag{25}$$

In practice we work with dilation radii much smaller than the size of the image, which makes possible to disregard the effect of boundaries and focus on the region $x \in [\frac{s}{2}, 1 - \frac{s}{2}]$. This simplifies the integral to

$$\langle L \rangle = 1 - (1 - s)^{Np}. \tag{26}$$

This expression can be rewritten as

$$\langle L \rangle = 1 - \left(1 - \frac{Nps}{Np}\right)^{Np}, \tag{27}$$

which for large N allows for the use of the exponential limit:

$$\lim_{n \to \infty} \langle L \rangle = 1 - e^{-Nps}. \tag{28}$$

This limit confirms the power law relation naturally appearing in fractal-like phenomenons. However, for our purposes, it is also worthwhile to investigate (26), which determines the behavior of $\langle L \rangle$ for not so large values of N.

Using a strategy similar to that in [20], in two dimensions we have circles with radius s and (26) is rewritten as

$$\langle L \rangle = 1 - (1 - \pi s^2)^{Np}. \tag{29}$$

For a self-similar structure with self-similar dimension d_S we can resort to the formula for the volume of the n-dimensional ball, yielding

$$\langle L \rangle = 1 - \left(1 - \frac{\pi^{d_S/2}}{\Gamma\left(\frac{d_S}{2} + 1\right)} s^{d_S}\right)^{Np}, \tag{30}$$

where Γ is the Euler's gamma function. α and β are provided by (22) as usual. Figure 3 shows α and β for different numbers of points N_p and self-similar dimension d_S. As in Fig. 2 we have different curves for each similarity dimension.

In general, all this analysis corroborates the interest in analyzing local characteristics of the image like LBP codes by inspecting properties of the $\log - \log$ dimension curve. Here we use the slope and linear coefficient of such curve and verify that such potential theoretically predicted is also confirmed in practical situations.

4 Experiments

The performance of the proposed descriptors was evaluated on three benchmark databases widely used in the recent literature of texture classification methods. We also applied the same method to a practical problem, namely the identification of species of Brazilian plants using scanned images of leaves.

The first benchmark data set is KTHTIPS-2b [16], a database comprising 4752 images uniformly divided into 11 material classes. An important characteristic of this data is its focus on the material represented in the image rather than on the instance of the photographed object. In each material class the images can still be divided into 4 samples. Each sample follows a particular scheme of scale, pose and illumination. The validation protocol is the most typically employed in the literature, where 1 sample is used for training and the remaining 3 samples are used for testing. The accuracy (percentage of images correctly classified) and respective standard deviation are obtained by averaging out the results for the 4 possible combinations of training/testing.

The second database is UIUC [23], containing 1000 images evenly divided into 25 texture categories. The images were collected under non-controlled conditions and contain variation in albedo, perspective, illumination and scale. For the validation split we randomly select 20 images of each texture for training and the remaining 20 images for testing. This is repeated 10 times to provide the average accuracy and deviation.

The third data set is UMD [38]. This is composed by a collection of 1000 high-resolution images collected by a family camera without any illumination control. The images are uncalibrated and unregistered and are categorized into 25 classes, each one with 40 images.
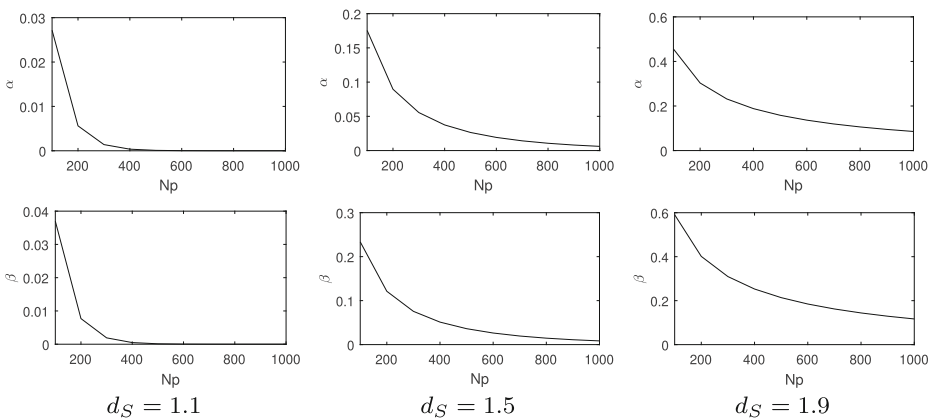

Fig. 3 α and β in Bouligand-Minkowski for different numbers of points Np and self-similar dimension d_S

Each image has a resolution of 1280×960. The database is characterized by high variance in viewpoint and scale, what makes its classification a more challenging task.

To reduce dimensionality and attenuate the effect of redundant information, the proposed descriptors are subject to principal component analysis [18] before the classification. For the classification, we tested four possibilities: support vector machines (SVM) [7] with a configuration similar to what is employed in [6], i.e., linear kernel, $C = 1$ and L^2 normalization, random forests (RF) [17] with a number of trees empirically determined from the training set within the [100, 500] interval, artificial neural networks (ANN) [27] using one hidden layer with the number of neurons defined empirically (around 10 worked fine) and trained by scaled conjugate gradient, and linear discriminant analysis (LDA) [22].

5 Results and discussion

An initial test was accomplished to verify the performance of the proposed method for the four compared classifiers, i.e., SVM, RF, ANN, and LDA. For this test we employed box counting descriptors as those have more straightforward interpretation and allow more unbiased evaluation of the classifier. Table 1 lists the percentage of images correctly classified in each compared database as well as the respective standard deviation. LDA provided higher accuracies in all tested data. The success of linear approaches can be justified by the relatively large number of input features, which causes non-linear approaches to overfit. Based on this and on the fact that LDA has no parameter to be adjusted, which makes it more controllable, the remaining experiments are carried out using this classification scheme.

In Table 2 we analyze the performance of the individual fractal descriptors: box counting (BC), Bouligand-Minkowski (BM), lacunarity (L), and multifractal (MF). From that table it is possible to state that each fractal metric can be more or less recommended depending on particularities of the data set. This also suggests that combining different fractal features could provide even better classification results.

Table 3 shows the accuracy when some of the investigated fractal features are combined into a single vector of descriptors. Other combinations were verified but at the end those ones presented in Table 3 had the most competitive performances. In general, some subtle improvement over the individual features was obtained in most data sets when combined descriptors were used.

Table 4 lists the accuracy performance of the proposed descriptors in KTHTIPS-2b, UMD and UIUC, compared with other results published in the literature. Details concerning parameters and other implementation details for each result can be obtained in the respective references. Here the fractal features outperformed advanced approaches like SIFT + VLAD or SIFT + BoVW in KTHTIPS-2b. Methods based on automatic (deep) learning

Table 1 Classification accuracy using SVM and LDA classifier

Database	SVM	RF	ANN	LDA
KTHTIPS-2b	63.4±3.4	63.5±3.8	59.7±7.6	**65.5±3.5**
UIUC	96.3±0.5	95.3±0.8	89.1±1.9	**97.3±0.6**
UMD	98.0±0.7	97.4±1.0	95.3±1.3	**98.6±0.6**
1200Tex	81.9±1.7	81.2±0.6	62.1±3.7	**86.6±1.1**

Bold values correspond to the highest accuracies

Table 2 Classification accuracy of the individual fractal descriptors: BC (box counting), BM (Bouligand-Minkowski), L (lacunarity), and MF (multifractal)

Database	BC	BM	L	MF
KTHTIPS-2b	65.5±3.5	**66.3±3.1**	65.9±3.6	58.9±3.3
UIUC	97.4±0.5	97.6±0.4	**97.8±0.5**	86.7±1.4
UMD	98.6±0.6	99.0±0.6	**99.1±0.4**	97.2±0.6
1200Tex	**86.4±1.1**	85.9±1.3	84.9±1.1	67.2±1.5

Bold values correspond to the highest accuracies

like FC-CNN were also outperformed in UIUC and UMD. These are canonical examples of what could be defined as "textures in their strict sense" and the results confirmed the potential of fractal-based methods to analyze such types of images. Finally, MFS and PLS are examples of fractal-based approaches for texture recognition. Both were also outperformed by the proposed descriptors in UIUC and UMD.

Figure 4 shows the confusion matrices for the benchmark databases. Box counting descriptors were used to generate these pictures. Such representations essentially confirm the accuracies in Table 2, but they also provide information regarding the accuracy in each class, what opens possibility for a more elaborate analysis on the classification outcomes. In Fig. 4a we see KTHTIPS-2b with lower accuracy in classes 3 ("corduroy") and 5 ("cotton"). These are actually materials frequently confused as they are types of clothing fabrics and possess similar texture patterns. UIUC, on the other hand, yields a nearly perfect classification result, with some significant misclassification only in class 8 - granite - (confused with 18 - carpet). Despite being different materials, they are both characterized by a granular appearance, which poses some difficulties for the automatic discrimination.

5.1 Identification of plant species

Table 5 lists the accuracy of the fractal-LBP descriptors in 1200Tex database [3], compared with a few state-of-the-art results published in the literature. This is a set of images of plant leaves for 20 Brazilian species collected *in vivo*. For each species 20 samples were collected, washed, aligned with the vertical axis and photographed by a scanner. The image of each sample was split into 3 non-overlapping windows with size 128×128. Those windows were extracted from regions of the leaf presenting less texture variance and were converted into gray scale images, resulting in a database with 1200 images. From each species 30 images were randomly selected for training and the remaining images for testing. This procedure was repeated 10 times, which allowed the computation of the average accuracy and respective deviation (in Tables 2 and 3).

Table 3 Classification accuracy for some combinations of fractal descriptors

| Database	BC+BM	BC+L	BC+MF	BM+L	BM+MF	L+MF	BC+BM+L	BC+BM+L+MF
KTHTIPS-2b	67.5±2.1	**67.6±2.1**	65.0±2.1	67.4±2.6	65.4±2.2	66.7±3.1	66.4±2.0	66.4±2.0
UIUC	97.9±0.5	**98.1±0.5**	96.4±0.7	98.0±0.4	96.3±0.8	97.0±0.6	**98.1±0.6**	97.6±0.5
UMD	98.9±0.6	99.1±0.5	99.1±0.5	99.1±0.4	**99.3±0.2**	99.1±0.3	99.0±0.5	99.2±0.3
1200Tex	86.2±1.0	86.3±0.8	82.2±1.6	85.9±1.1	81.6±1.8	80.5±1.6	**86.3±0.6**	82.5±1.4

Bold values correspond to the highest accuracies

Table 4 Accuracy of the proposed descriptors compared with other texture descriptors in the literature

Method	KTH-TIPS2b	UIUC	UMD
VZ-MR8 [35]	46.3	92.9	–
LBP [28]	50.5	88.4	96.1
VZ-Joint [36]	53.3	78.4	–
BSIF [19]	54.3	73.4	96.1
LBP-FH [1]	54.6	–	–
CLBP [15]	57.3	95.7	98.6
SIFT+LLC [6]	57.6	96.3	98.4
ELBP [25]	58.1	–	–
SIFT + KCB [5]	58.3	91.4	98.0
LBP_{riu2}/VAR [28]	58.5[1]	84.4	95.9
SIFT + BoVW [5]	58.4	96.1	98.1
PCANet (NNC) [4]	59.4[1]	57.7	90.5
SIFT + VLAD [5]	63.1	96.5	99.3
RandNet (NNC) [4]	60.7[1]	56.6	90.9
SIFT+IFV [5]	58.2	97.0	99.2
ScatNet (NNC) [2]	63.7[1]	88.6	93.4
DeCAF [5]	70.7	94.2	96.4
FC-CNN VGGM [6]	71.0	94.5	97.2
MFS [38]	–	92.7	93.9
PLS [31]	–	96.6	99.0
Proposed	67.6	98.1	99.3

All the results except for the proposed method were obtained from the literature. A '-' indicates that no result was published for that method on that database. A superscript [1] denotes a slightly different protocol in KTH-TIPS2b where 3 samples are used for training and 1 sample for testing

Figure 5 complements Table 5 by exhibiting the confusion matrix for the proposed method (box counting features). Generally speaking, the accuracies in all classes are high and the most critical situation occurs in class 8, which is confused for example with class 6. These correspond to samples with similar textures, especially with regards to the arrangement of nervures and leaf microtexture, which are prominent elements for the process of distinguishing among samples from different species.

Generally speaking, fractal descriptors still demonstrate competitiveness when compared with several state-of-the-art approaches for texture recognition. It is well known that characteristics intrinsic to the way that these materials are formed in nature contribute to this relation to a significant extent. This can be even more easily observed in data sets of "pure" textures, like UIUC and UMD, as well as in practical problems where such types of images naturally arise. This is the case in many biological applications and here is illustrated by the foliar surface images. To summarize, the most excelling advantages of the proposed method are its high accuracy without requiring neither large amounts of data nor specific computational resources for training, meaningfulness of fractal characteristics, and an implementation based on simple numerical methods. As for shortcomings, we can mention the relatively large number of features, even though techniques like regularization or principal component analysis, as used here, can attenuate that point and prevent any potential overfit.
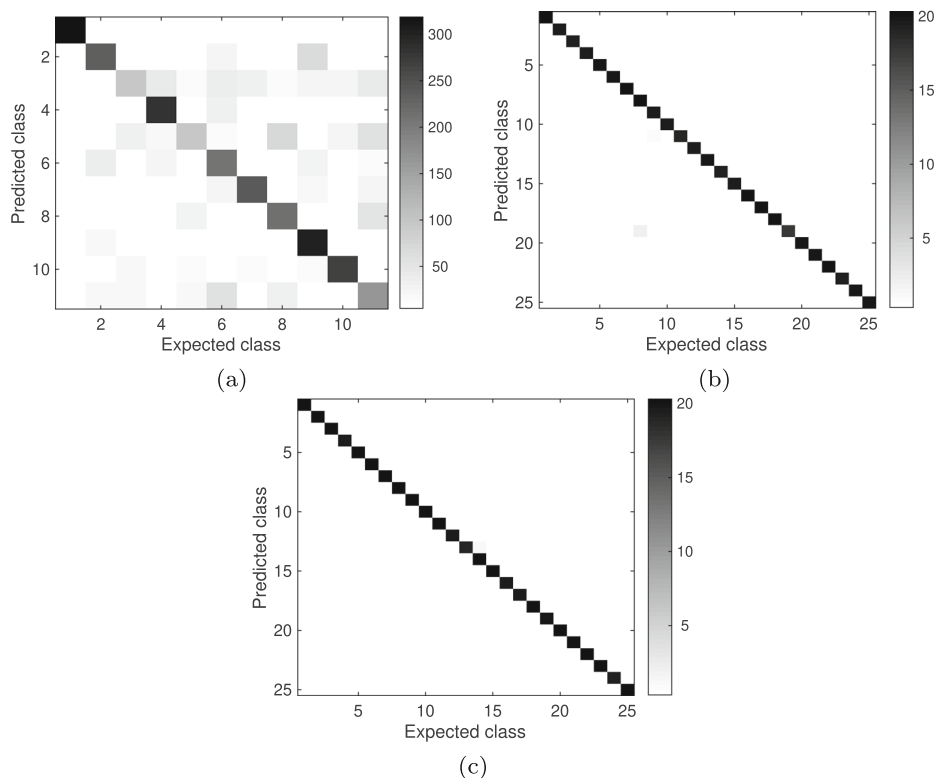
Fig. 4 Confusion matrices. **a** KTHTIPS-2b. **b** UIUC. **c** UMD

The results encourage more research on this topic at the same time that it presents fractal descriptors as an alternative that should be verified in practical problems, as they can achieve competitive performance, for example, without requiring large amounts of training data and they also provide more natural interpretation for the obtained results as fractal sets have been classically associated with a mathematical model of nature structures.

Table 5 Accuracy of the fractal-LBP descriptors compared with other results in the identification of plant species

Method	Accuracy (%)
LBPV [15]	70.8
Network diffusion [13]	75.8
FC-CNN VGGM [6]	78.0
Gabor [3]	84.0
FC-CNN VGGVD [6]	84.2
Schroedinger [10]	85.3
SIFT + BoVW [5]	86.0
Proposed	86.3
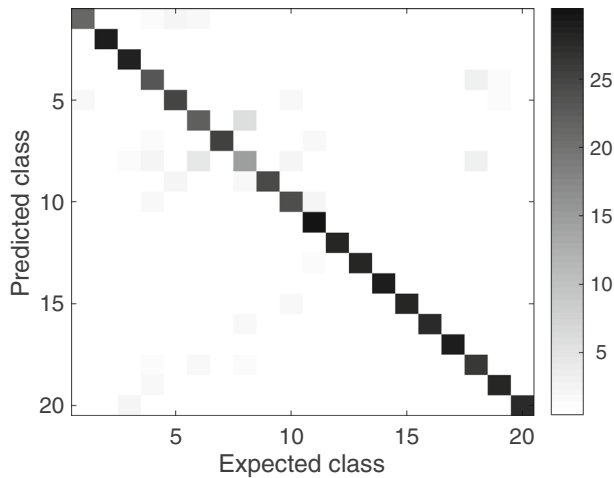
Fig. 5 Confusion matrix for the plant database

6 Conclusions

Here we developed a combination of local binary patterns with numerical estimates of fractal dimension. More specifically, we compute the dimension of the LBP codes thresholded at different levels to compose the image feature vector.

The performance of our proposal was assessed in the classification of benchmark databases typically used in the literature. We also employed such descriptors in a practical problem with significant importance in botany and related areas, namely, the identification of species of Brazilian plants. In both cases our method obtained promising results, comparable to state-of-the-art results recently published in the literature.

The results presented here suggest that the combination of fractal geometry (and potentially other fractal measures) with a local encoding like LBP can be rather useful to represent all the rich information conveyed by a texture image.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

References

1. Ahonen T, Matas J, He C, Pietikäinen M (2009) Rotation invariant image description with local binary pattern histogram fourier features. In: Salberg AB, Hardeberg JY, Jenssen R (eds) Image analysis. Springer, Berlin, pp 61–70
2. Bruna J, Mallat S (2013) Invariant scattering convolution networks. IEEE Trans Pattern Anal Mach Intell 35(8):1872–1886
3. Casanova D, de Mesquita Sá Junior JJ, Bruno OM (2009) Plant leaf identification using gabor wavelets. Int J Imaging Syst Technol 19(3):236–243
4. Chan T, Jia K, Gao S, Lu J, Zeng Z, Ma Y (2015) Pcanet: A simple deep learning baseline for image classification? IEEE Trans Image Process 24(12):5017–5032
5. Cimpoi M, Maji S, Kokkinos I, Mohamed S, Vedaldi A (2014) Describing textures in the wild. In: Proceedings of the 2014 IEEE Conference on computer vision and pattern recognition. IEEE Computer Society, Washington. 3606–3613

6. Cimpoi M, Maji S, Kokkinos I, Vedaldi A (2016) Deep filter banks for texture recognition, description, and segmentation. Int J Comput Vis 118(1):65–94
7. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
8. Dhal KG, Galvez J, Ray S, Das A, Das S (2020) Acute lymphoblastic leukemia image segmentation driven by stochastic fractal search. Multimed Tools Appl 79(17-18):12227–12255
9. Falconer K (2004) Fractal geometry: mathematical foundations and applications. Wiley, New York
10. Florindo JB, Bruno OM (2017) Discrete schroedinger transform for texture recognition. Inform Sci 415:142–155
11. Florindo JB, Casanova D, Bruno OM (2018) A gaussian pyramid approach to bouligand-minkowski fractal descriptors. Inform Sci 459:36–52
12. Gao TJ, Zhao D, Zhang TW, Jin T, Ma SG, Wang ZH (2020) Strain-rate-sensitive mechanical response, twinning, and texture features of NiCoCrFe high-entropy alloy: experiments, multi-level crystal plasticity and artificial neural networks modeling. J Alloys Compd, 845
13. Gonçalves WN, da Silva NR, da Fontoura Costa L, Bruno OM (2016) Texture recognition based on diffusion in networks. Inform Sci 364(C):51–71
14. Grochalski K, Wieczorowski M, Pawlus P, H'Roura J (2020) Thermal sources of errors in surface texture imaging. Materials 13(10)
15. Guo Z, Zhang L, Zhang D (2010) A completed modeling of local binary pattern operator for texture classification. IEEE Trans Image Process 19(6):1657–1663
16. Hayman E, Caputo B, Fritz M, Eklundh JOPajdla T, Matas J (eds) (2004) On the significance of real-world conditions for material classification. Springer, Berlin
17. Ho TK (1995) Random decision forests. In: Proceedings of the third international conference on document analysis and recognition (Volume 1) - Volume 1. IEEE Computer Society, Washington, p 278. ICDAR '95
18. Jolliffe I (2002) Principal component analysis. Springer Series in Statistics. Springer, Berlin
19. Kannala J, Rahtu E (2012) Bsif: Binarized statistical image features. In: ICPR. IEEE Computer Society, pp 1363–1366
20. Kenkel N (2013) Sample size requirements for fractal dimension estimation. Community Ecol 14(2):144–152
21. Krishnamoorthi N, Chinnababu VK (2019) Hybrid feature vector based detection of glaucoma. Multimedi Tools Appl 78(24):34247–34276
22. Krzanowski WJ (ed) (1988) Principles of multivariate analysis: a user's perspective. Oxford University Press, Inc., New York
23. Lazebnik S, Schmid C, Ponce J (2005) A sparse texture representation using local affine regions. IEEE Trans Pattern Anal Mach Intell 27(8):1265–1278
24. Liu J, Chen Y, Sun S (2019) A novel local texture feature extraction method called multi-direction local binary pattern. Multimed Tools Appl 78(13):18735–18750
25. Liu L, Zhao L, Long Y, Kuang G, Fieguth P (2012) Extended local binary patterns for texture classification. Image Vision Comput 30(2):86–99
26. Mandelbrot BB (1983) The fractal geometry of nature, 3rd edn. W. H. Freeman and Comp., New York
27. McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5(4):115–133
28. Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
29. Pentland AP (973) Fractal-based description. In: Proceedings of the eighth international joint conference on artificial intelligence - volume 2. Morgan Kaufmann Publishers Inc., San Francisco. IJCAI'83
30. Posadas A, Gimenez D, Bittelli M, Vaz C, Flury M (2001) Multifractal characterization of soil particle-size distributions. Soil Sci Soc Am J 65(5):1361–1367
31. Quan Y, Xu Y, Sun Y, Luo Y (2014) Lacunarity analysis on image patterns for texture classification. In: 2014 IEEE conference on computer vision and pattern recognition, pp 160–167
32. Russ J (1994) Fractal surfaces. Fractal Surfaces. Springer, Berlin
33. da S Oliveira MW, Casanova D, Florindo JB, Bruno OM (2014) Enhancing fractal descriptors on images by combining boundary and interior of minkowski dilation. Physica A 416:41–48
34. Taraschi G, Florindo JB (2020) Computing fractal descriptors of texture images using sliding boxes: An application to the identification of Brazilian plant species. Physica A 545
35. Varma M, Zisserman A (2005) A statistical approach to texture classification from single images. Int J Comput Vis 62(1):61–81
36. Varma M, Zisserman A (2009) A statistical approach to material classification using image patch exemplars. IEEE Trans Pattern Anal Mach Intell 31(11):2032–2047

37. Verma G, Luciani ML, Palombo A, Metaxa L, Panzironi G, Pediconi F, Giuliani A, Bizzarri M, Todde V (2018) Microcalcification morphological descriptors and parenchyma fractal dimension hierarchically interact in breast cancer: A diagnostic perspective. Comput Biol Med 93:1–6
38. Xu Y, Ji H, Fermüller C (2009) Viewpoint invariant texture description using fractal analysis. Int J Comput Vis 83(1):85–100
39. Xu Y, Yang X, Ling H, Ji H (2010) A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid. In: CVPR, IEEE Computer Society, pp 161–168
40. Yang Q, Peng F, Li JT, Long M (2016) Image tamper detection based on noise estimation and lacunarity texture. Multimed Tools Appl 75(17):10201–10211
41. Zaghloul R (2019) Hiary H. A multifractal edge detector (online). Multimedia Tools and Applications, Al-Zoubi, MB
42. Zaghloul R, Hiary H, Al-Zoubi MB (2020) A multifractal edge detector. Multimed Tools Appl 79(9-10):5807–5828
43. Zhang J, Liu Y, Yan K, Fang B (2019) A fractal model for predicting thermal contact conductance considering elasto-plastic deformation and base thermal resistances. J Mech Sci Technol 33(1):475–484
44. Zhang P, Barad H, Martinez A (1990) Fractal dimension estimation of fractional brownian motion. In: IEEE Proceedings on Southeastcon, vol 3, pp 934–939
45. Zheng Q, Fan J, Li X, Wang S (2018) Fractal model of gas diffusion in fractured porous media. Fractals 26(03):1850035