Brachiaria species identification using imaging techniques based on fractal descriptors

http://www.producao.usp.br/handle/BDPI/50555

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo
Brachiaria species identification using imaging techniques based on fractal descriptors

João Batista Florindo,a Núbia Rosa da Silva,a,b Liliane Maria Romualdo,c Fernanda de Fátima da Silva,c Pedro Henrique de Cerqueira Luz,c Valdo Rodrigues Herling,c Odemir Martinez Bruno,a,b,*

a Scientific Computing Group, São Carlos Institute of Physics, University of São Paulo (USP), cx 369, 13560-970 São Carlos, São Paulo, Brazil
b Institute of Mathematics and Computer Science, University of São Paulo (USP), Avenida Trabalhador São-carlense, 400, 13566-590 São Carlos, São Paulo, Brazil
c Department of Animal Science, College of Animal Science and Food Engineering, University of São Paulo, Avenida Duque de Caxias Norte 225, Caixa Postal 23, 13635-900 Pirassununga, São Paulo, Brazil

Article info

Article history:
Received 20 September 2013
Received in revised form 2 January 2014
Accepted 9 February 2014

Keywords:
Fractal descriptors
Texture analysis
Brachiaria species identification

Abstract

The use of a rapid and accurate method in diagnosis and classification of species and/or cultivars of forage has practical relevance, scientific and trade in various areas of study, since it has broad representation in grazing from tropical regions. Nowadays it occupies about 90% of the grazing area along Brazil and, besides the grazing areas to feed ruminants, Brachiaria also corresponds to about 80% of seeds being traded in all the world, bringing a large amount of money to Brazil. To identify species and/or cultivars of this genus is of fundamental importance in the fields that produce seeds, to ensure varietal purity and the effectiveness of improvement programs. Thus, leaf samples of fodder plant species Brachiaria were previously identified, collected and scanned to be treated by means of artificial vision to make the database and be used in subsequent classifications. Forage crops used were: Brachiaria decumbens cv. IPEAN; Brachiaria ruziziensis Germain & Evrard; Brachiaria brizantha (Hochst. ex. A. Rich.) Stapf; Brachiaria arrecta (Hack.) Stent. and Brachiaria spp. The images were analyzed by the fractal descriptors method, where a set of measures are obtained from the values of the fractal dimension at different scales. Therefore such values are used as inputs for a state-of-the-art classifier, the Support Vector Machine, which finally discriminates the images according to the respective species. The proposed method outperforms other state-of-the-art image analysis methods and makes possible the correct prediction of species in more than 93% of the samples. Such remarkable result is consequence of the better suitability of representing complex structures like those arising in the plant leaves by measures of complexity from fractal geometry. Finally, this high correctness rate suggests that the fractal method is an important tool to help the botanist.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The knowledge and understanding of functional properties of plants make possible to develop advances in several areas like medicine to cure diseases, produce and improve species to feed people and animals (Camargo and Smith, 2009). Involving this last topic, the analysis of consumption by animals is very important because the animal production can be improved from grazed pastures. Specifically, ruminants have their amount of feeding directly linked with the processes of particle-size reduction during the feeding. Due to the physical strength of grasses, ruminant animals consume larger quantities of forages with lower resistance to breakdown (Herrero et al., 2001). Since the grass is extremely important for animal food, it becomes the object of study here being one of the main forms of ruminant feeding is through grazing Brachiaria.

The genus Brachiaria consists of herbaceous, perennial or annual, erect or decumbent. Belonging to the grass family, it presents approximately one hundred species, and therefore their correct classification is of great importance for the genetic improvement of forage species and purity of the species in the field of seed production (Parsons, 1972; Wenzl et al., 2000; Arroyave et al., 2013). Grouping plants into genus is a way of facilitating the understanding of the diversity of the grasses, according to the particularities of each species (Parsons, 1972).

* Corresponding author at: Scientific Computing Group, São Carlos Institute of Physics, University of São Paulo (USP), cx 369, 13560-970 São Carlos, São Paulo, Brazil. Tel.: +55 16 3373 8728; fax: +55 16 3372 2218.
E-mail address: bruno@scg.ifsc.usp.br (O.M. Bruno).
The study of techniques to classify species and/or cultivars of the genus *Brachiaria* has practical, scientific and commercial relevance, since it has broad representation in grazing from tropical regions. Originating in Africa, it has been easily adapted to the climatic conditions in Brazil and nowadays it occupies about 90% of the grazing area along the entire country. In economic terms, besides the grazing areas to feed ruminants, *Brachiaria* also responds to about 80% of seeds being traded in all the world, bringing a large amount of money to Brazil. To identify species and/or cultivars of this genus is of fundamental importance for practical issues of Biology, Agronomy and Animal Science, as well as in the fields that produce seeds, to ensure varietal purity and the effectiveness of improvement programs.

The classification of grasses is mainly based on the characters of the spikelet structure and its arrangement. The main taxonomic feature of the genus *Brachiaria*, despite not being present in many species, is the reversed or adaxial position of the spikelet and of ligule. This spikelet is relatively large, oval or oblong and it is arranged regularly in a row along one side of the rachis. However, the taxonomy of this genus is not satisfactory, both in terms of species composition and in their inter-relation with other genus. Problems related to incorrect classifications often occur among *Brachiaria* species commonly used in pastures, as well as among accessions of germplasm collections.

In this context, this work proposes a computational imaging technique to address this taxonomical problem. The literature on applications of image analysis has shown a number of works classifying plant species based on the image of structures like the leaves. For example, in [Mokhtarian and Abbasi (2004)](#) a shape analysis method (considering only information from the contour) based on Curvature Scale Space is applied to the identification of species of *Chrysanthemum* achieving a success rate close to 100%. Another shape-based approach is proposed in [Neto et al. (2006)](#), to classify some species of weed. They use a variant of Fourier transform and obtain about 90% of success. In [Wang et al. (2008)](#), a variant of *Brachiaria* species is used to classify some species of weed. They use a variant of Fourier transform and obtain about 90% of success. In [Wang et al. (2008)](#), the focus is on the pre-processing of leaf images with complicated background. Again, a shape analysis is carried out and a general database of plants is classified with success rate close to 90%. Finally, in [Backes et al. (2009)](#), the authors propose to use information from the tonalities of all the pixels in the image inside the leaf, instead of using only information from the contour. They propose a method based on fractal geometry and achieve a great success rate on a complicated database of plants from the Brazilian flora.

Since there is a great variability among natural species of *Brachiaria*, to identify really discriminant characters becomes a difficult task so that seeking for techniques that improve the identification will contribute to studies within this theme, as well as provide a reasonable system of classification, since there is no such system for the genus *Brachiaria*. Based on this and taking into account the great results of fractal-based approaches on leaf images, as demonstrated in works like [Backes et al. (2009)](#), the objective of this study was to take the volumetric Bouligand–Minkowski and Probability fractal descriptors associated with the Canonical Transform to classify samples from five species of *Brachiaria* cultivars. This methodology provides a set of coefficients for each image that will characterize it. The tests were performed in a large database with almost ten thousand of samples including the superior and inferior face of leaves obtaining 92.84% of correctness rate in the classification (of all leaves). The text in this paper is organized as follows. In Sections 2, 2.2 and 3 the theory of the methods is explained. In Section 4 the description of the method Bouligand–Minkowski with probability dimension applied to texture characterization. Section 5 shows the experiments in a database of *Brachiaria* leaves and in Section 6 the results are analyzed. The paper is concluded in Section 7.

### 2. Fractal geometry

Fractal geometry ([Mandelbrot, 1968](#)) is the area of Mathematics which deals with fractal objects. These are geometrical structures characterized by two main properties: the infinite self-similarity and infinite complexity. In other words, these elements are recursively composed by similar structures. In addition, they exhibit a high level of detail on arbitrarily small scales.

In the same way as in the Euclidean geometry, fractal objects are described by numerical measures. The most widespread of such measures is the fractal dimension. Given a geometrical set $X$ (set of points in the $N$-dimensional space), the fractal dimension of $D(X)$ is expressed in the following equation:

$$D(X) = N - \lim_{\epsilon \to 0} \frac{\log(\Psi(\epsilon))}{\log(\epsilon)},$$

where $\Psi$ is a fractality measure and $\epsilon$ is the scale parameter. The literature presents various definitions for the fractality measure ([Tricott, 1995; Russ, 1994](#)). The following sections describe two of such approaches.

#### 2.1. Bouligand–Minkowski

One of the best-known methods for estimating the fractal dimension of an object is the Bouligand–Minkowski approach ([Tricott, 1995](#)). In this solution, the grayscale image $I : [1 : M] \times [1 : N]$ is mapped onto a surface $S$, using the following relation:

$$S = [(i,j,k) \mid (i,j) \in [1 : M] \times [1 : N], k = I(i,j)].$$

Then, each point having co-ordinates $(x,y,z)$ is dilated by a sphere with variable radius $r$. Therefore, the dilation volume $V(r)$ may be computed by the following expression:

$$V(r) = \sum \chi_{[r]}(i,j,k),$$

where $(i,j,k)$ are points in the surface $S$, $\chi$ is the characteristic function and $\chi(r)$ refers to the following set:

$$\chi(r) = \left\{ (x,y,z) \mid (x-P_1)^2 + (y-P_2)^2 + (z-P_3)^2 \leq r \right\},$$

in which $(P_1, P_2, P_3)$ is in $S$. In practice, the Euclidean Distance Transform ([Fabbri et al., 2008](#)) is used to determine the value of $V(r)$.

Finally, the fractal dimension $D_{BM}$ itself is given by:

$$D_{BM} = 3 - \lim_{r \to 0} \frac{\log(V(r))}{\log(r)}.$$  

The limit in the above expression is calculated by plotting the values of $V(r)$ against $r$, in log–log scale, and the limit is the slope of a straight line fitting the log–log curve. [Fig. 1](#) exemplify the process.

#### 2.2. Probability dimension

Also referred to as Voss dimension, this method obtains the fractal dimension from the statistical distribution of pixel intensities within the image ([Voss, 1986](#)). Like in the Bouligand–Minkowski method, the image analyzed is converted into a three-dimensional surface $S$. Hence, the surface is surrounded by a grid of cubes with side $\delta$. By varying the value of $\delta$, the information function $N_\delta$ is provided through:

$$N_\delta(\delta) = \sum_{m=1}^{N} \frac{1}{m} P_m(\delta),$$

where $N$ is the maximum possible number of points within a single cube and $P_m(\delta)$ is the probability of $m$ points in $S$ belonging to the same cube.
Finally, the fractal dimension $D$ is estimated by the following relation:

$$D = \lim_{\delta \to 0} \frac{\ln N_\delta}{\ln \delta}.$$  

(1)

As with the Bouligand–Minkowski approach, the limit is computed by a least squares fit.

3. Fractal descriptors

Fractal descriptors are a methodology that extracts meaningful information of an object of interest by means of an extension of the fractal dimension definition (Bruno et al., 2008; Florindo and Bruno, 2012; Florindo et al., 2011). Instead of using the dimension for describing the object, the fractal descriptors $u$ use the entire set of values in the fractality curve:

$$u : \log(e) \to \log(N).$$  

(2)

The values of $u$ can be used directly to compose the feature vector (Florindo and Bruno, 2012) or after some sort of transform (Florindo et al., 2011). In the present work, we use the raw data of $\log(N)$.

The fractal descriptors extract significant information from the object at different scales. The values of the fractality for larger radii measure the global aspect of the structure. In the case of plant leaves like those analyzed here, these data concern important information regarding the general aspect of the nervure distribution. On the other hand, the smallest radii provide essential information on the variability of the pixel intensities inside a local neighborhood. Biologically, these micro-patterns are tightly related to the constitution of the plant tissue.

4. Proposed method

The proposed methodology combines the previously described approaches to compose a precise and robust tool to identify plant species from an image of the leaf of *Brachiaria*. The feature vector is obtained by concatenating the Bouligand–Minkowski and Probability descriptors and then applying a dimensionality reduction procedure based on a canonical transform to the merged descriptors.

Let the Bouligand–Minkowski descriptors be represented by the vector $D_1$:

$$D_1 = \{x_1, x_2, \ldots, x_n_1\}$$

and the Probability features expressed through $D_2$:

$$D_2 = \{y_1, y_2, \ldots, y_n_2\}.$$  

Here, we have $n_1 = 85$ and $n_2 = 40$.

Then, in a problem of species discrimination over a database of leaf images, we can define two feature matrices, $M_{1n_1}$ and $M_{2n_2}$, one for each descriptor, where the rows correspond to the descriptor vectors for each image to be analyzed. After that, the matrices are concatenated horizontally, giving rise to the matrix $M$. Next, this matrix is transformed into one matrix $\hat{M}$:

$$\hat{M} = M_{\text{inter}}M_{\text{intra}}^{-1},$$  

(3)

where $M_{\text{intra}}$ and $M_{\text{inter}}$ are, respectively, the inter and intra-class matrix. The intra-class matrix is defined by:

$$S_{\text{intra}} = \sum_{i=1}^{K} \sum_{c_i} (X(i, \cdot) - \overline{C}_i)(X(i, \cdot) - \overline{C}_i)^T,$$  

(4)

where $C_i$ is the $i$th class in $M$. $K$ is the total number of classes, $M(i, \cdot)$ expresses the $i$th row (sample) of $M$ and $\overline{C}_i$ is a row vector representing the average descriptors of each class $C_i$. On its turn, the inter-class matrix is provided by the following expression:

$$S_{\text{inter}} = \sum_{i=1}^{K} N_i (\overline{C}_i - \overline{M})(\overline{C}_i - \overline{M})^T,$$  

(5)

in which $N_i$ is the number of samples of the $i$th class.

The concatenation process ensures that the resulting descriptors emphasize the best discriminative properties of each fractal approach. In this case, both descriptors provide a different perspective of the object. While the Bouligand–Minkowski capture a multiscale mapping of the texture morphology, the Probability method gives a detailed description of the statistical distribution of the pixels along the gray-level image, as those descriptors are computed considering the density of pixel intensities at each region of the image, resulting in the probability of a pixel tonality arising at a particular point. The sum of these viewpoints makes possible a detailed description of the patterns within the image, at different scales.

In the present study, the concatenated descriptors are obtained from the windows extracted from the scanned image of the analyzed grass. Fig. 2 illustrates the process involved in computing the descriptors, since the original image until the descriptors themselves.

5. Experiments

The leaf samples were collected in the agrostologic field at the Faculdade de Zootecnia e Engenharia de Alimentos (FZEA-USP), which is located in Pirassununga city at state of São Paulo, Brazil. The leaf images were collected manually, directly from live plants, with extreme carefulness in not damaging the leaf surface. All plants grew with ideal conditions of nutrients and lighting. In this study, five species were taken: *Brachiaria decumbens* Stapf. cv. Ipean, *Brachiaria ruziziensis* Germain & Evrard, *Brachiaria brizantha* (Hochst. ex. A. Rich.) Stapf., *Brachiaria arrretcha* (Hack.) Stent. and *Brachiaria spp.*

After collecting the leaves, they were submitted to a scanning procedure, where the superior and inferior face of the leaf was
scanned in 1200 dpi (dots-per-inch) resolution and saved in a lossless image format with no compression. It was obtained 5 sheets with 10 different tillers, totalling 100 samples (50 images from the superior faces of the leaves and 50 images from the inferior faces).

As the leaves were scanned manually, they were not properly aligned. Therefore, the images were vertically aligned according to the central axis using the Radon transform (Deans, 1993). Subsequently, for mounting the database, it was randomly obtained about 20 sub-images of 200 × 200 pixels without overlapping and considering all the leaf surface, avoiding stains margins and allowing the central vein (Fig. 3). At Fig. 3 it can also be seen the preparation of the signature for each sample. The method is applied to each sub-image from the sample (superior and inferior face of the leaf) to obtain a feature vector \( F \) with features \( \{F_1, F_2, F_3, \ldots, F_n\} \) where \( n \) is the number of features. Afterward, the signature from the superior face of the image is concatenated with the signature from the inferior face to generate the final signature of the sample. Therefore, the final database is composed by 9832 images, 4916 images from superior face and 4916 from inferior face of leaves. Fig. 4 shows some samples for each class of the database.

The proposed method is applied to compute descriptors from the images of the analyzed grass species. Therefore, the obtained descriptors are used as the input of a classifier, in this case, the Support Vector Machine method (Vapnik, 1999). The classification is carried out in a 10-fold cross-validation scheme (Vapnik, 1999).

6. Results

The graph in Fig. 5 shows the success rate of each compared texture descriptors when the number of descriptors is varied. The first descriptors are selected as those with the highest correlation among themselves, by following the order of the eigenvectors on the transform in Section 4. We notice that even with only 4 elements, the proposed method achieves a correctness rate close to 90%. Gabor and Fourier have a close behavior, while the fractal descriptors presents an outstanding performance for any number of elements greater than 3.

Table 1 shows the best success rate achieved by each method by using an optimal number of descriptors. The success rate is accompanied by other important measures related to the performance of each approach applied to discriminate the grass species. In this table, \( ND \) is the number of descriptors, \( CR \) is the correctness rate, \( \kappa \) is the \( \kappa \) index and \( AE_1 \) and \( AE_2 \) are the type 1 and 2 errors. \( \kappa \) index measures how much the prediction is better than a simple random guess for the classes. In practice, it follows the correctness rate when the measures have normal behavior. More details can be seen in Fleiss (1981) The number of descriptors ranges from 1 to 30, as after this point the performance tends to stabilize. The greatest success rate and smallest errors were obtained by the combined fractal descriptors. Such outcome confirms the effectiveness of the proposed methodology in this classification task, providing an excellent categorization of the plant species.

Tables 2–4 exhibit the confusion matrices for each compared descriptor. These tables are helpful to describe either the correctly classified samples as well as the false negatives and false positives, outside the main diagonal. We observe that, despite some minor differences among the methods, the methodology developed in the present study demonstrated to be the most reliable solution to identify the grass samples. Reinforcing the values in Table 1, the type 1 and 2 errors are very similar, expressing the homogeneity of the distinguished classes.

Finally, Table 5 shows the ratio of success in the identification of the species when combining the proposed fractal approach with other descriptors. The results indicate that a high precision (more than 97%) in the automatic identification of the species is achieved by a combination with classical methods and using only 10 descriptors in this task. Such remarkable result confirms that more than only an efficient individual method, fractal descriptors have a great potential to be combined into a more general computer vision system to provide a highly accurate outcome.

The great performance achieved by the proposed fractal descriptors is due to the nature of the fractal modelling, as mathematical fractals and natural objects have a great deal in common. The similarities are related to the high complexity usually found in the nature as well as the self-similarity property, which also appears often in parts of plants, like leaves, flowers, etc. Actually,
the fractal descriptors are tightly related to important physical attributes of the leaf, such as roughness, reflectance and distribution of colors and brightness levels. In turn, this set of properties is capable of identify plant species faithfully, using their digital image representation, as demonstrated in Bruno et al. (2008). The present study confirms such reliability and robustness of fractal descriptors and shows that this is a powerful tool for the species categorization of the grasses analyzed here.

Fig. 3. Samples acquisition. Sub-images of 200 × 200 pixels size obtained from samples. A signature $F = \{F_1, F_2, F_3, \ldots, F_n\}$ is extracted for each sub-image and, afterward, the signature from the superior and inferior face of the leaf are concatenated to obtain the final signature.
The studied species have different sizes, shapes, colors and amount of piles over the foliar surface, which interfere in the results provided by the computational analysis. 

Brachiaria decombens Stapf. is the most widespread in areas of beef cattle farms, mainly in the middle-west and north of Brazil. Its most remarkable feature is the high density of piles over the foliar surface, differing from the remaining studied species, with variants without piles (glabra) or slightly hairy (pubescent). Brachiaria brizantha, Brachiaria arrecta and Brachiaria spp. have leaf fronds brighter and greener than the others. Other properties that

The studied species have different sizes, shapes, colors and amount of piles over the foliar surface, which interfere in the results provided by the computational analysis. 

Brachiaria decombens Stapf. is the most widespread in areas of beef cattle farms, mainly in the middle-west and north of Brazil. Its most remarkable feature is the high density of piles over the foliar surface, differing from the remaining studied species, with variants without piles (glabra) or slightly hairy (pubescent). Brachiaria brizantha, Brachiaria arrecta and Brachiaria spp. have leaf fronds brighter and greener than the others. Other properties that
are not taken into consideration in this work, such as shapes and serrated edges can be worthy in the classification, however, the mentioned features, piles, color and brightness, are sufficient for a precise classification.

Below, some characteristics of the foliar surface that are discriminant elements among the classes:

- **Brachiaria decumbens** Stapf cv. IPEAN: Lanceolate or linear-lanciolate leaves, 10–15 cm length and 15 mm wide, soft and densely hairy.
- **Brachiaria ruziziensis** Germain & Evrard: The leaves do not have any pile or are slightly hairy, linear and lanceolate, 10–30 cm length and 0.8–2.0 cm wide.
- **Brachiaria brizantha** (Hochst. ex. A. Rich.) Stapf cv. Marandu: The leaves can be naked or covered by fine and short piles (pubescent), with roll-shaped sheaths; sharp leaf fronds, 5–30 cm length and 0.6–1.6 cm wide, cartilaginous-margined, ciliated.
- **Brachiaria arrecta** (Hack.) Stent.: Leaves are lanciolate, heart-shaped, 70–150 mm length and 12–25 mm wide, naked, bright, with juicy appearance and dark green color.
- **Brachiaria spp.** This is an inter-species hybrid of **Brachiaria arrecta** and **Brachiaria purpurascens**. It has purple-green stem and inner face with a few piles; few piles in the sheath and nodes.

Like in other works applying imaging techniques to analyse plant leaves, here, the fractal descriptors quantify all these characteristics of each specie of *Brachiaria* by describing multiscale patterns of pixel colors and brightnesses at each region of the leaf. Such arrangements of pixels are tightly related to properties like density of piles, level of brightness or darkness on the leaf frond, shape of sheaths, etc. These properties are difficult or even unfeasible to be quantified manually or by simple statistical approaches and the fractal descriptors make possible a reliable and precise automatic classification of the *Brachiaria* samples.

7. Conclusions

The present study proposed a combination of fractal descriptor approaches to discriminate among species of *Brachiaria* grass, based on the digital images from their leaves. The proposed solution achieved a high success rate even using a low number of features. Such result confirms the effectiveness and reliability of fractal descriptors in this kind of task.

This result is also remarkable from a biological perspective, as *Brachiaria* grasses are one of the most important foods for animals that are used for the labor and human consumption. The precise discrimination of species makes possible to better understand the distribution of species in a region and, as a consequence, to optimize the necessary attention for that region.

Despite the importance of this study, the literature shows very few works on *Brachiaria* classification and the present is the first to obtain such a great effectiveness. Such so good results suggest the use of fractal descriptors as a powerful method to identify these species and markedly help the taxonomy specialist.

Acknowledgments

Odemir Martinez Bruno gratefully acknowledges the financial support of CNPq (National Council for Scientific and Technological Development, Brazil) (Grants 308449/2010-0 and 473893/2010-0) and FAPESP (The State of São Paulo Research Foundation) (Grant 2011/01523-1). Núbia Rosa da Silva acknowledges support from FAPESP (Grant 2011/21467-9). João Batista Florindo acknowledges support from FAPESP (Grant 2012/19143-3).

References