

Fusion of time series representations for plant recognition in phenology studies[☆]



Fabio A. Faria^{a,b,*}, Jurandy Almeida^{a,b}, Bruna Alberton^c, Leonor Patricia C. Morellato^c, Ricardo da S. Torres^b

^aInstitute of Science and Technology, Federal University of São Paulo – UNIFESP, São José dos Campos 12247-014, SP, Brazil

^bInstitute of Computing, University of Campinas – UNICAMP, Campinas 13083-852, SP, Brazil

^cDept. of Botany, Sao Paulo State University – UNESP Rio Claro, 13506-900, SP, Brazil

ARTICLE INFO

Article history:

Available online 12 March 2016

Keywords:

Plant species identification
Classifier fusion
Diversity measures

ABSTRACT

Nowadays, global warming and its resulting environmental changes is a hot topic in different biology research area. Phenology is one effective way of tracking such environmental changes through the study of plant's periodic events and their relationship to climate. One promising research direction in this area relies on the use of vegetation images to track phenology changes over time. In this scenario, the creation of effective image-based plant identification systems is of paramount importance. In this paper, we propose the use of a new representation of time series to improve plants recognition rates. This representation, called recurrence plot (RP), is a technique for nonlinear data analysis, which represents repeated events on time series into a two-dimensional representation (an image). Therefore, image descriptors can be used to characterize visual properties from this RP images so that these features can be used as input of a classifier. To the best of our knowledge, this is the first work that uses recurrence plot for plant recognition task. Performed experiments show that RP can be a good solution to describe time series. In addition, in a comparison with visual rhythms (VR), another technique used for time series representation, RP shows a better performance to describe texture properties than VR. On the other hand, a correlation analysis and the adoption of a well successful classifier fusion framework show that both representations provide complementary information that is useful for improving classification accuracies.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Global warming and its resulting environmental changes have raised important research topics of different disciplines. Among those is phenology that studies recurrent life cycles events and its relationship to climate [35]. To increase the range of study sites and species and the effectiveness of phenological observations, technological devices (e.g., multi-channel imaging sensors) have been successfully applied to provide metrics for estimating changes on phenological events, such as leaf development and senescence [3,4,33].

Plant species recognition in the digital images is not a trivial task, especially in tropical vegetations, where one single image

may include a huge number of species [4]. This task is very time consuming since it has to be done in the field, first by matching each crown in the image to the tree in the soil and then by identifying the tree at species level.

Our goal in this work is to support automatic plant species recognition tasks based on phenological pattern information. Different patterns correspond to different species, as well as similar patterns can be grouped in one species type or in a leaf functional group that encompasses several species. This may not just save time for phenologists but also complement the phenological interpretation of the data collected.

Almeida et al. [7] have proposed the use of machine-learning methods to find similar patterns in the digital images and then check if those patterns correspond to similar species or functional groups. Their work was focused on the intraspecies analysis, i.e., on detecting different individuals of a same species. However, different species from a same functional group may exhibit a similar phenological pattern [4], confounding the classifiers. Hence, many questions arise when considering interspecies interactions, i.e., the recognition of individuals from different species but belonging to the same functional group [4].

[☆] This paper has been recommended for acceptance by Jenny Qian Du.

* Corresponding author at: Institute of Science and Technology, Federal University of São Paulo – UNIFESP, 12247-014 São José dos Campos, SP, Brazil. Tel.: +55 12 3309 9500; fax: +55 12 3921 8857.

E-mail addresses: ffaria@unifesp.br, juruna18@gmail.com (F.A. Faria), jurandy.almeida@unifesp.br (J. Almeida), bru.alberton@gmail.com (B. Alberton), pmorella@rc.unesp.br (L.P.C. Morellato), rtorres@ic.unicamp.br (R. da S. Torres).

In this work, we adopt the strategy of characterizing phenological patterns from time series and distinguishing species from the same leaf functional group in plant species recognition tasks. In fact, several time series representations have been proposed in the literature. Some successful approaches include data-adaptive (e.g., SAX [26] and APCA [23]) and non-data adaptive representations (e.g., wavelets [31]). A good survey upon this subject can be found in [40].

In this work, we present a novel approach for time series representation, which is based on a technique for nonlinear data analysis called recurrence plots (RP). Different from other time series representations, RP provides a visual mechanism for pattern identification, being suitable for combining with state-of-the-art computer vision description approaches. This work has also been motivated by the results of [24] and [37]. Both studies indicate that it is possible to perform classification tasks through the use of recurrence plots and texture feature extraction approaches. RP technique has been used successfully in different application domain, such as action recognition [24], identification of diabetes analysis of epilepsy [1], and detection of financial crisis [2].

In our experiments, we performed four rigorous comparative analysis to show the robustness of RP-based representations for plant recognition tasks. We begin with an effectiveness study evaluating the performance of RP-based classifier associated with time series of different hours of day. Then, we compared the proposed approach with another time series representation proposed by [6], called Visual Rhythm (VR). Next, we performed a correlation analysis to find out agreement/disagreement between all classifiers involved between RP and VR-based representations. Finally, we adopt a successful classifier fusion framework [16] to combine the most suitable classifiers. Experimental results show that the combination of RP- and VR-based representations yields better results than their use in isolation.

In summary, the main contributions of this work are: (i) a new representation of time series based on recurrence plots technique for plant recognition; (ii) effectiveness study of the recurrence plots approach in different hours of day; (iii) effectiveness comparative study between recurrence plots and visual rhythm approaches; (iv) correlation analysis between recurrence plots and visual rhythm approaches; (v) use of a classifier fusion framework to combine the most suitable classifiers using both approaches.

The remainder of this paper is organized as follows. Section 2 presents the phenological data acquisition process considered in our study. Section 3 presents the recurrence plot approach and how to use it for phenological time series representation. Section 4 describes the experimental protocol adopted to validate the proposed approach. Section 5 reports the results of our experiments and compares the proposed approach with another time series representation. Finally, we offer our conclusions and directions for future work in Section 6.

2. Phenological data acquisition

A digital hemispherical lens camera (Mobotix Q24) was set up in an 18 m-high tower in a Cerrado *sensu stricto*, a neotropical savanna vegetation located at Itirapina, São Paulo State, Brazil [4,32]. Fig. 1 shows all steps of the time series acquisition process used in our work.

Firstly, we set up the camera to take a daily sequence of five JPEG images (at 1280×960 pixels of resolution) per hour, from 6:00 to 18:00 h (UTC-3). The present study was based on the analysis of over 2700 images (Fig. 1a), recorded at the end of the dry season, between August 29th and October 3rd 2011, day of year 241 to 278, during the main leaf flushing season [4].

Next, the image analysis has been conducted by defining different regions of interest (ROI), as described in [33] and defined

by [4] for our target species. Then, we analyzed 22 ROIs (Fig. 1b) obtained from a random selection of six plant species identified manually by phenology experts in the hemispheric image [4]:

- (i) Three regions associated with *Aspidosperma tomentosum* (green areas).
- (ii) Four regions for *Caryocarp brasiliensis* (blue areas).
- (iii) Two regions for *Myrcia guianensis* (orange areas).
- (iv) Six regions for *Miconia rubiginosa* (magenta areas).
- (v) Two regions for *Pouteria ramiflora* (cyan areas), and
- (vi) Four regions for *Pouteria torta* (red areas).

We analyzed each ROI in terms of the contribution of the primary colors (R, G, and B), as proposed by [34] and described in [4]. Initially, we analyze each color channel and compute the average value of the pixel intensity (Fig. 1c). After that, we compute the normalized brightness of each color channel (RGB Chromatic coordinates) (Fig. 1d). The normalization of those values reduces the influence of the incident light, decreasing the color variability due to changes on illumination conditions [4,11,41]. Finally, by computing those values along the whole period (August 28th to October 3rd, 2011), we obtained time series to use as input data for our proposed framework (Fig. 1e).

According to the leaf exchange data from the on-the-ground field observations on leaf fall and leaf flush at our study site, those species were classified on three functional groups [4,27]: (i) deciduous, *A. tomentosum* and *C. brasiliensis*; (ii) evergreen, *M. guianensis* and *M. rubiginosa*; and (iii) semideciduous, *P. ramiflora* and *P. torta*.

3. Recurrence plots for plant species recognition

This section introduces our approach for phenological time series representation using recurrence plots. Section 3.1 describes how to compute recurrence plots (RP) from time series, while Section 3.2 presents how we use RP for representing phenology data.

3.1. Recurrence plots (RP)

Recurrence plots (RP), proposed by [15] in dynamical systems, is an advanced technique of nonlinear data analysis. RP technique has been used to visualize repeated events (the recurrence of states) of higher dimensional phase spaces through projection into the two or three dimensional sub-spaces. This technique is able to investigate recurrent behavior (periodicity) at time series (m -dimensional phase space) through a two-dimensional representation, such as a distance square matrix.

Recurrence Plot might be defined by:

$$R_{i,j} = \Theta(\epsilon_i - \|x_i - x_j\|), x_i \in \mathcal{R}^m, i, j = 1 \dots N \quad (1)$$

where N is the number of considered states (dots at the time series) x_i , ϵ_i threshold distance, $\|\cdot\|$ a norm between the states (e.g., Euclidean norm), m is the embedding dimension, and $\Theta(\cdot)$ the Heaviside function. This discontinuous function has value 0 for negative argument and 1 for positive argument.

Eq. (1) provides an $N \times N$ image, which shows us whether there are recurrent states or not, along the target trajectory. This created image might be binary or grayscale, depending on the choice of a threshold or not. Fig. 2 shows a real time series from the dataset and two examples of recurrence plot considering those real time series, unthresholded and thresholded.

The choice of an generalizable threshold to perform matching between two RP is a non-trivial task, but it can be possible with few heuristics [36]. As in this work we aim to make use of color and texture information, we have adopted the unthresholded approach using the Manhattan norm and $m = 1$. This unthresholded approach is defined in Eq. (2). However, we can not rule out that

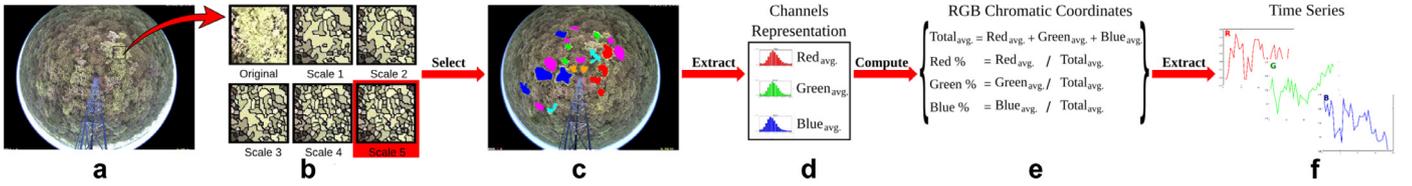


Fig. 1. The time series acquisition process pipeline. (a) Sample image of the cerrado savanna; (b) Different segmentation scales are computed and the coarse scale is selected; (c) Hemispherical image with the selected ROI's species; (d) Channel representation are extracted from ROI's; (e) RGB chromatic coordinates are computed; (f) Phenological time series extracted from digital images.

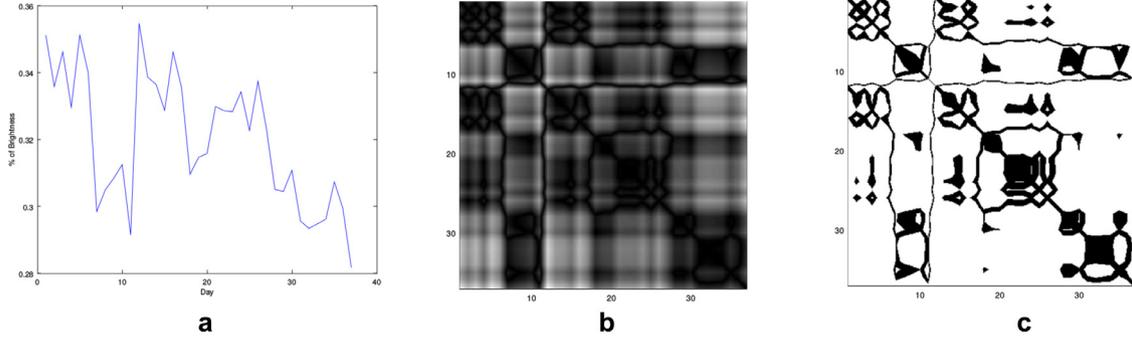


Fig. 2. In (a), example of time series, (b) is an example of unthresholded RP, and (c) is an example of thresholded RP with $\Theta(\cdot) < 20$, $N = 37$, and $m = 1$.

the extraction of shape information [13] is a research line to be explored in future work.

$$R_{i,j} = |x_i - x_j|, x_i \in \mathcal{R}^1, i, j = 1 \dots N \quad (2)$$

3.2. Recurrence plot for phenological time series representation

From phenology time series, it is possible to estimate changes on phenological events, such as leaf flushing when analyzing the green channel, or leaf color change and senescence using values from the red channel [3,33]. It requires the analysis of time series related to different color channels. The changing patterns along time are then validated with on-the-ground phenology observations.

However, color information in the RGB channels are highly correlated (i.e., changes in one channel may lead to variations in another), making harder to detect temporal changes in recurrent phenology events.

The novelty of this paper is to extend the notion of recurrence plots for the context of phenology. Here, we combine the time series into a single representation, making the phenological change analysis easier. The key contribution of our idea is the combination process we design to encode the time series into a single image.

Fig. 3 (a)–(c) illustrates the process of computing RP-based representations. In (a), given the obtained time series in Fig. 1, the recurrence plot algorithm is applied on these time series to create their two-dimensional representation (distance matrix). Even at this step, a normalization technique is applied to convert real values to integer values in the typical range of grayscale images ([0, 255]). Then, in (b), a merging process is performed to join the three distance matrices into a single color image. Finally, in (c), many image descriptors may be used to encode visual properties (e.g., color, texture, and shape) from this color image into feature vectors. The feature vectors created through the use of image descriptors are later used as input to a machine learning method (e.g., support vector machine and k -Nearest Neighbors).

Fig. 4 shows the recurrence plots computed using the defined ROIs. As RP provides a two-dimensional representation (distance square matrix), we can observe many different colors in the matrices, which are associated with the distance values between plots on the time series. The lowest distance values are closer to the

blue color and the highest distance values are closer to the red color. Notice that there are pattern differences among all of species used in this work.

4. Experimental setup

This section describes the baseline used, which is based on visual rhythms (Section 4.1); presents the classifier fusion framework adopted to combine VR-based and RP-based classifiers (Section 4.2), and, finally, introduces the experimental protocol (Section 4.3).

4.1. Baseline: visual rhythms (VR)

An effective way to analyze temporal properties from video data is by means of visual rhythms [28]. The objective is to create an abstraction of a video by coding the temporal change of pixel values along a specific sampling line [20].

In the context of phenology, a time series comprised of images taken by digital cameras at fixed time intervals can be viewed as a video of the vegetation, as proposed by [5]. However, instead of specific lines (e.g., diagonal, horizontal, and vertical), the interest here is to analyze unshapely regions related to plant species that are identified by phenology experts (see the initial step in Fig. 3).

Motivated by such limitations, [6] have generalized the notion of visual rhythms. From a generic point of view, a visual rhythm consists of temporal data samples grouped in an orderly manner. For that, they have designed a mapping function to convert a ROI into a vertical line. In the following, we briefly describe their strategy for extracting visual rhythms from phenology time series. For more details regarding their approach, refer to [9].

Let $\mathcal{V} = \{f_t\}, t \in [1, T]$ be a video, in domain $2D + t$, with T frames of dimensions $W_V \times H_V$; and \mathcal{I} be a binary image, with the same dimensions of \mathcal{V} , in which white pixels indicate a ROI. Initially, the image \mathcal{I} is converted into a list of Cartesian coordinates $\mathcal{L}_{xy} = \{(x, y) \mid \mathcal{I}(x, y) = 1\}$. Next, this list is used for computing the geometric center (x_c, y_c) of the ROI. Then, the Cartesian coordinate system of the elements in \mathcal{L}_{xy} is translated to have its origin at the point (x_c, y_c) . After that, the Cartesian coordinates \mathcal{L}_{xy} are converted to the polar coordinate system, creating a list of

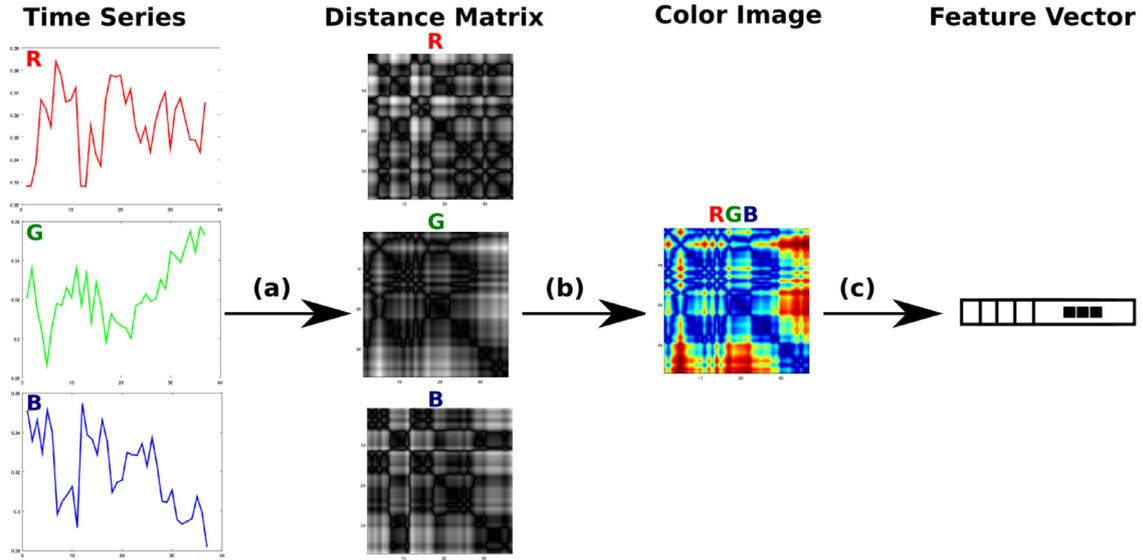


Fig. 3. The steps for feature extraction from the recurrence plot representation.

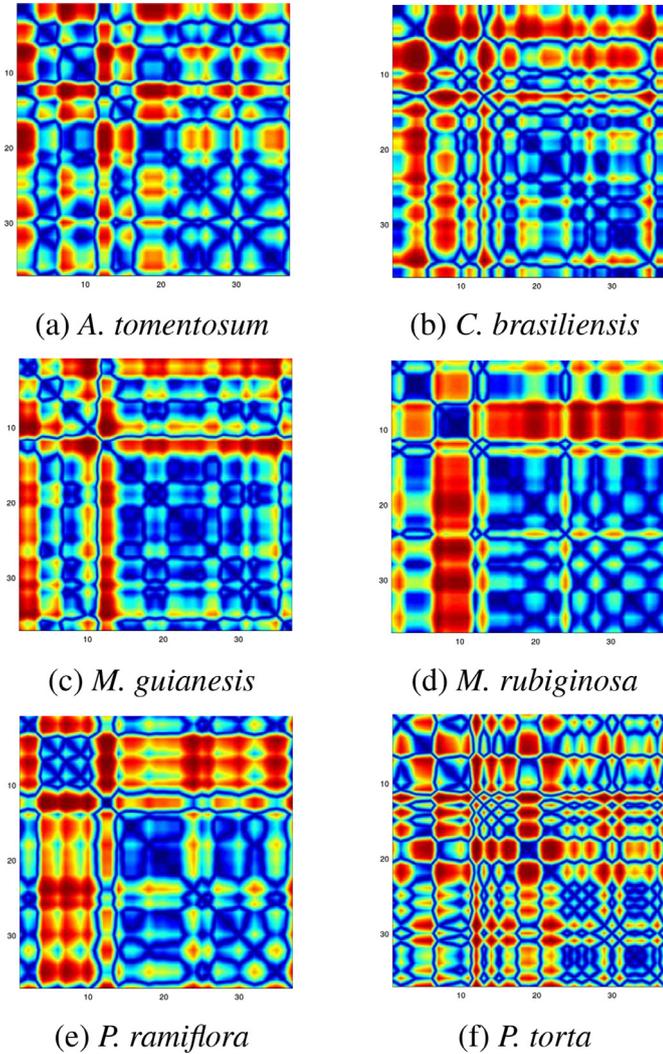


Fig. 4. Examples of recurrence plots for ROIs of each plant specie. These images were created using the *Octave* software [14] with one-dimensional interpolation method.

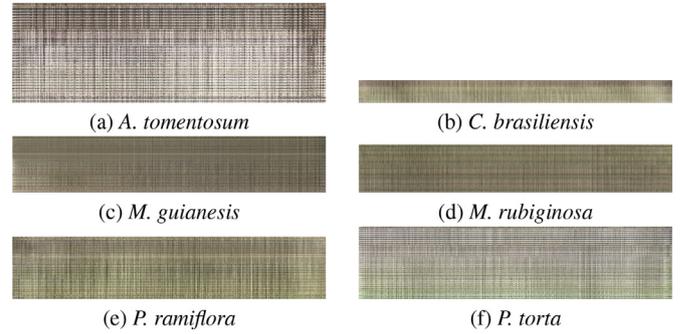


Fig. 5. Visual rhythms obtained for ROIs of some plant species. Text and figure have been extracted from [6].

polar coordinates $\mathcal{L}_{r\theta}$. Thereafter, an index $\mathcal{K} = \{k \mid \forall(r, \theta) \in \mathcal{L}_{r\theta}, k = 2\pi r + \theta\}$ is created to assign a unique value to each element in $\mathcal{L}_{r\theta}$. Finally, the keys in the index \mathcal{K} are sorted in an increasing order and then used to arrange the elements in \mathcal{L}_{xy} .

In this way, they defined a visual rhythm as a mapping of each frame f_t into a vertical line on an image \mathcal{R}^* , in domain $1D + t$, such that [6]:

$$\mathcal{R}^*(t, z) = f_t(\mathcal{L}_{xy}(z)), t \in [1, W_{\mathcal{R}^*}], z \in [1, H_{\mathcal{R}^*}],$$

where $W_{\mathcal{R}^*} = T$ and $H_{\mathcal{R}^*} = |\mathcal{L}_{xy}|$ are its width and height, respectively. Figure 5 presents the visual rhythms produced by their approach using the ROIs from Fig. 3(b).

4.2. The classifier fusion framework

This section presents a framework for classifier selection and fusion (FSVM), as devised in [16]. The objective of the fusion framework is to exploit the degree of agreement/disagreement among classifiers, concept known as diversity, with the objective of selecting the most suitable ones to be used in a combination scheme.

In the context of classifier fusion area and in this work, a classifier might be defined as a tuple containing a learning method (e.g., kNN) and a description technique (e.g., Color Histogram). Classifiers learn patterns from training instances and use learned models to assign unseen instances to appropriate classes. Once the training step is finished, a selection process is performed to define

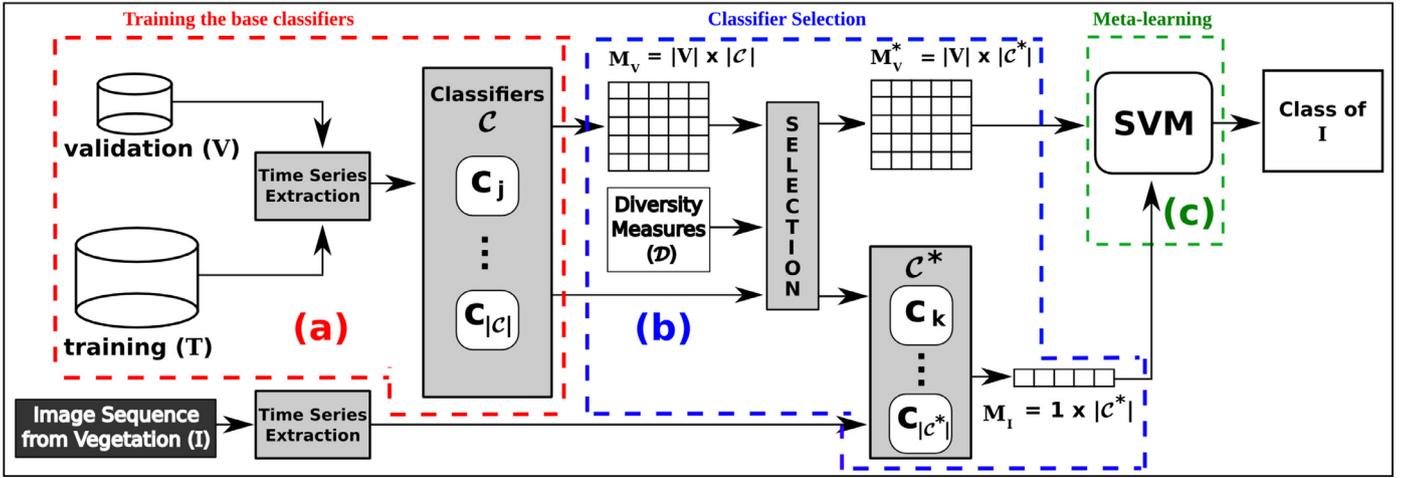


Fig. 6. Time series-based classifier selection and fusion framework adapted from [16]. In (a), given a classification problem with training examples, different classifiers are trained using data from training set T . In (b), the most discriminating classifiers are selected (C^*) by taking into account diversity measures (D). Finally, in (c), classifiers are combined in a meta-learning approach using a SVM technique.

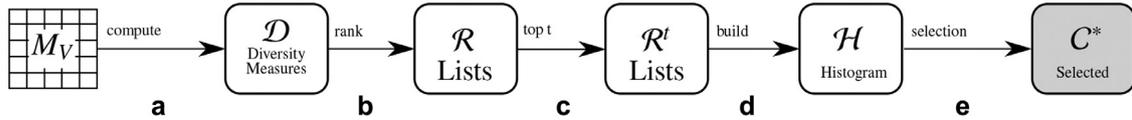


Fig. 7. The five steps for classifier selection are: (a) Compute diversity measures from the validation matrix M_V ; (b) \mathcal{R} lists sorted by diversity measures scores; (c) \mathcal{R}^t lists with top t ; (d) counts the number of occurrences of each classifier that satisfies a defined threshold; (e) Selected classifiers $|C^*|$. Text and figure have been extracted from [16].

classifiers whose combination, usually based on another learning method (meta-learning), is more promising. The objective is to determine the most discriminative methods, and, at the same time, boosting the classification performance at test time by selecting less, but more effective, classifiers.

Fig. 6 illustrates the framework FSVM for combining classifiers, while Fig. 7 illustrates the adopted five-step approach for selecting classifiers based on diversity measures, previously introduced by Faria et al. [16].

First, diversity measures (set \mathcal{D} in Fig. 7) are used to assess the degree of agreement among available classifiers in \mathcal{C} by taking into account the M_V matrix previously computed. That step is represented by arrow (a) in Fig. 7. Pairs of classifiers are then ranked according to their diversity score. Each diversity measure defines a different ranked list and, at the end of this step, a set \mathcal{R} of ranked lists is produced (arrow (b)). In the following, a novel set of ranked lists \mathcal{R}^t is computed by selecting the top t pairs of classifiers from each ranked list in \mathcal{R} (arrow (c)), and a histogram \mathcal{H} that counts the number of occurrences of a classifier in all ranked lists of \mathcal{R}^t is computed (arrow (d)). Finally, the most frequent classifiers in \mathcal{H} , whose accuracy is greater than a given threshold \mathcal{T} , are combined by a fusion approach (arrow (e)). \mathcal{T} is a threshold defined in terms of the average accuracy among all classifiers using the validation set V .

4.3. Experimental protocol

In this work, we adopted the evaluation method used in [7]. It relies on the classification of time series extracted from pixels associated with individuals of a same species. For that, we used the algorithm introduced by [19] to segment the hemispheric image into small polygons, obtaining 8, 849 segmented regions (SR). Then, we associated each SR with a single ROI aiming to label it. A labeled region is created if there is at least 80% of overlapped area between an SR and a ROI. Finally, we extracted a time series from

each labeled region using the approach described in Section 2. In this way, we built a dataset of 892 time series separated into six classes, one for each plant species: *A. tomentosum* (96), *C. brasiliensis* (346), *M. guianensis* (36), *M. rubiginosa* (195), *P. ramiflora* (50), and *P. torta* (169).

In the following, we present four experiments performed to validate the use of the RP representation in plant recognition tasks.

First, in Section 5.1, we performed an effectiveness study concerning the performance of classifiers that exploit RP-based features associated with time series of different hours of day. In this experiment, we used the k-Nearest Neighbors (kNN) learning method. We set $k = \{1\}$, which achieved the best results in [12]. For describing time series encoded into a RP representation, we used seven traditional and recently proposed image descriptors: ACC [21], BIC [38], CCV [30], and GCH [39], for encoding color information; GFD [42], GIST [29], and HWD [22], for analyzing texture properties.

In the second experiment, in Section 5.2, we compared the proposed RP-based approach with another time series representation proposed by [6], called Visual Rhythm (VR). Two different effectiveness experiments have been done:

- (1) Coarse-grained Analysis (Section 5.2.1) that evaluates the effectiveness results of the recurrence plot-based and visual rhythms-based classifiers; and
- (2) Fine-grained Analysis (Section 5.2.2) that evaluates the best effectiveness results achieved by both representations (RP and VR) using different image descriptors (GFD and BIC).

In the third experiment, in Section 5.3, we performed a correlation analysis to find out agreement/disagreement between all classifiers involved in the previous experiment using RP and VR representations.

In the fourth experiment, in Section 5.4, we adopt a successful classifier fusion framework [16] to combine the most suitable classifiers using both approaches (RP and VR) representations.

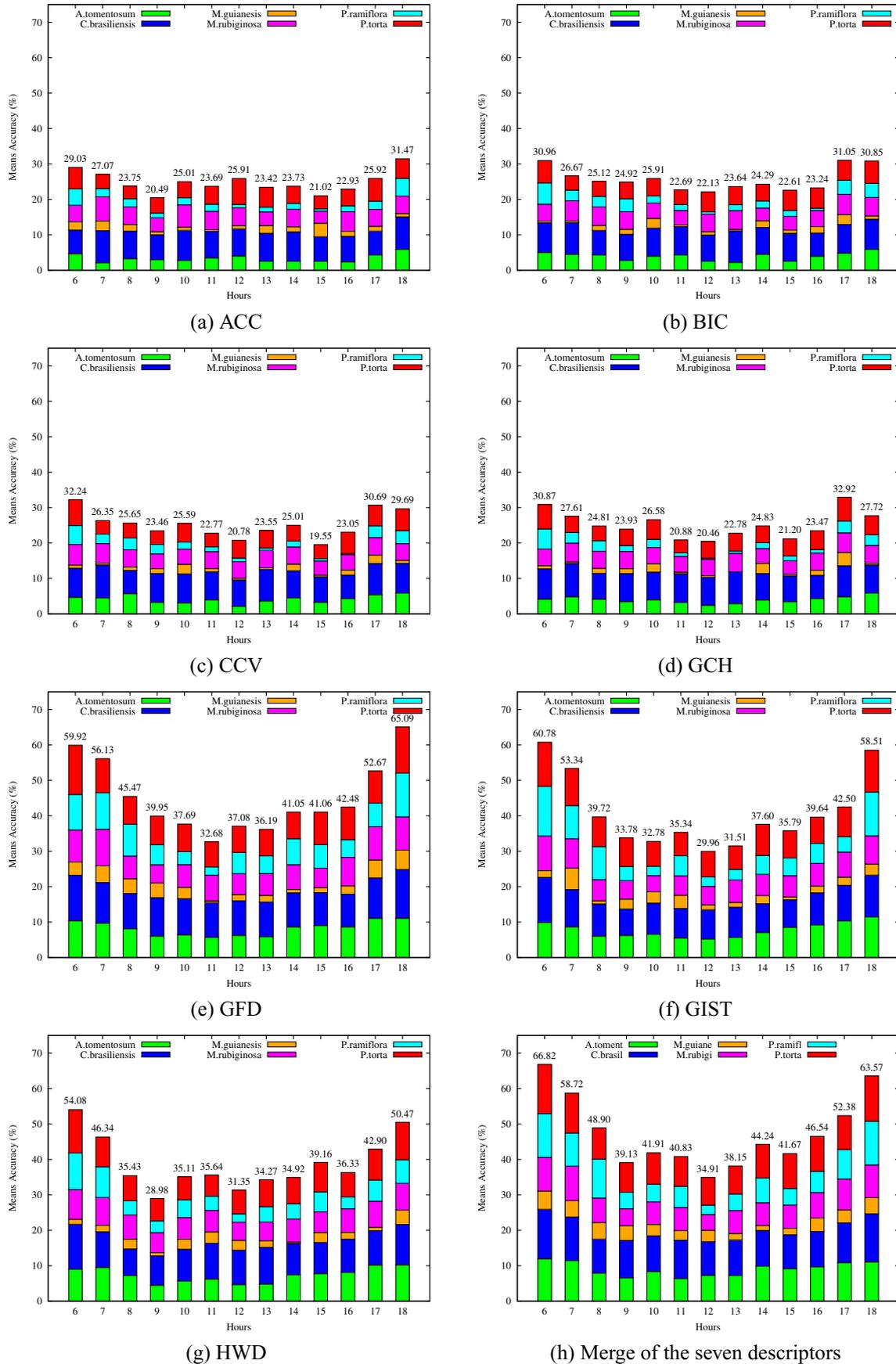


Fig. 8. In (a)–(g), effectiveness results of the kNN learning method using each of seven different image descriptors and in (f), results for all seven descriptors together..

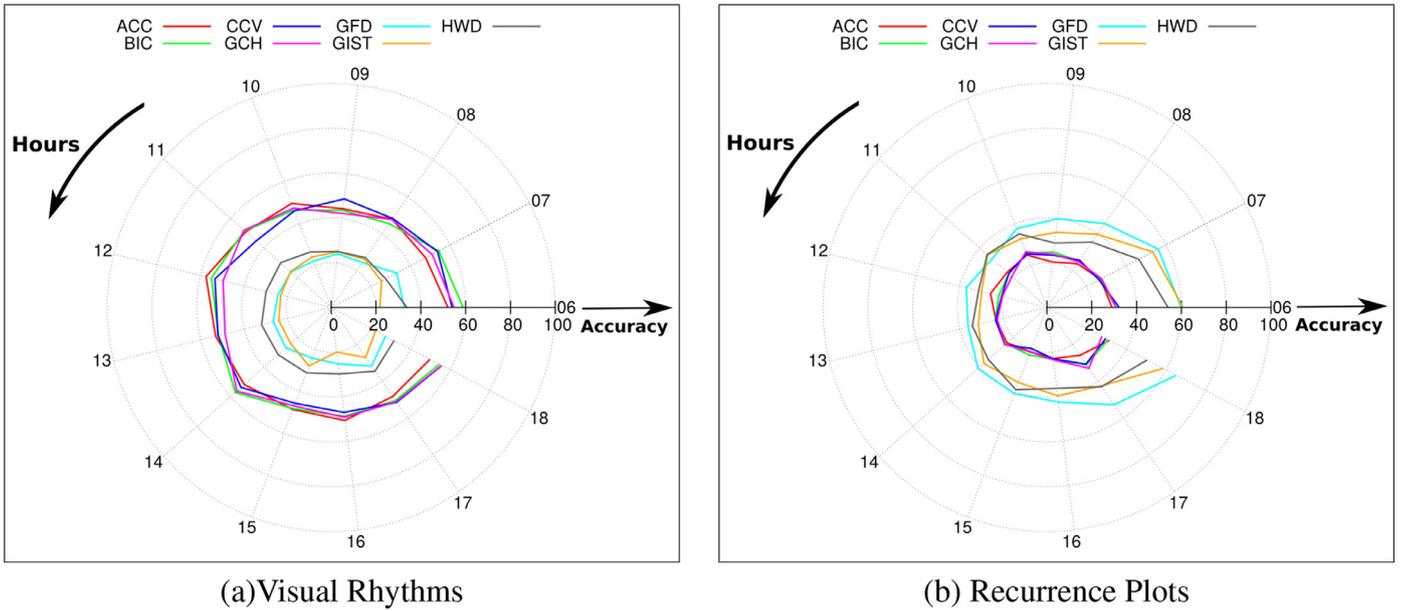


Fig. 9. Effectiveness results of the (a) visual rhythms-based and (b) recurrence plot-based kNN classifiers using seven image descriptors..

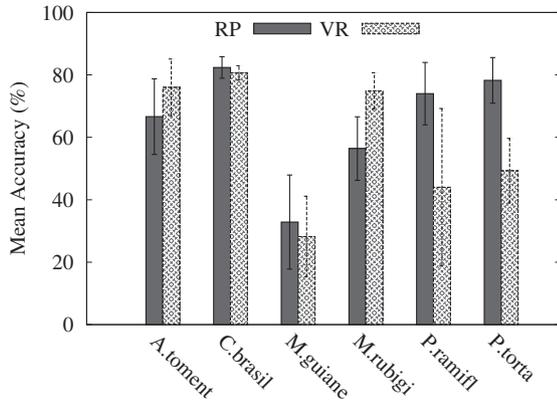


Fig. 10. The best effectiveness results for each representation approach. BIC color descriptor for Visual Rhythm approach and GFD texture descriptor for recurrence plots approach.

Finally, in Section 5.5, we conducted a behavioral study of the different approaches in small training scenarios.

5. Experimental results and discussion

5.1. Effectiveness results of RP-based classifiers

In these experiments, we analyze the behavior of the kNN learning method with each image descriptor (ACC, BIC, CCV, GCH, GFD, GIST, and HWD) for different hours of the day (from 6h to 18h).

We have adopted a 5-fold cross validation protocol, thus the evaluation measure used in this work is the arithmetic mean of the five accuracies (one accuracy per testing set). Furthermore, as the dataset is unbalanced (i.e., the number of samples in different classes is unequal), we computed each accuracy taking into account the size of classes in the testing set (weighted accuracy).

Fig. 8 (a)–(g) presents the effectiveness results for each of the image descriptors used in this work. Furthermore, in (h) we present the effectiveness results of all descriptors together. Notice that the x-axis refers to the hours of the day, while the y-axis refers to the used evaluation measure (Mean Accuracy).

As it can be observed, *C. brasiliensis* (in blue) and *P. torta* (in red) species may be considered the easiest species to be recognized among all the species, considering all hours of the day and all the image descriptors. However, *M. guianensis* (in orange) and *P. ramiflora* (in cyan) species are the hardest ones to be recognized considering all the six species. In almost all the experiments, the best results were achieved in the extreme hours of the day (6, 7, 17, and 18).

In relation to visual property (color and texture), the best effectiveness results achieved by the RP-based methods are those related to the use of texture descriptors (GFD, GIST, and HWD). Notice in Fig. 8(e) that our approach achieved the best result when consider a single descriptor (GFD) at 18h (65.09%). The worst result of a single descriptor was observed when the CCV color descriptor was used: 19.55% (Fig. 8(c) at 15h). However, the best effectiveness results for almost all hours of day have been achieved in (h), with the merge of all image descriptors.

5.2. Comparison of RP-based and VR-based classifiers

In this section, we compare the RP-based representations with the VR-based ones [6]. In the experiments with VR, we adopted the same experimental protocol (i.e., 5-fold cross validation and the same image descriptors and classifier) used with our RP-based approach.

5.2.1. Coarse-grained analysis

Fig. 9 shows two radar charts with effectiveness results of the recurrence plot-based and visual rhythms-based classifiers. These charts are composed of three parts, (1) radius, (2) slice of the perimeter, and (3) lines. The radius means the effectiveness measure values (mean balanced accuracy). The slices of the perimeter means each hours of day, from 6:00 to 18:00 h (UTC-3). Finally, the line means the performance of each kNN-1 learning method with one different image descriptor (ACC, BIC, CCV, GCH, GFD, GIST, and HWD).

Fig. 9 (a) are the seven image descriptors used in this work for visual rhythms-based classifiers and Fig. 9(b) are the same image descriptors for recurrence plots-based classifiers.

As it can be observed, in (a) the best results have been achieved by classifiers that use color image descriptors (ACC, BIC, CCV, and

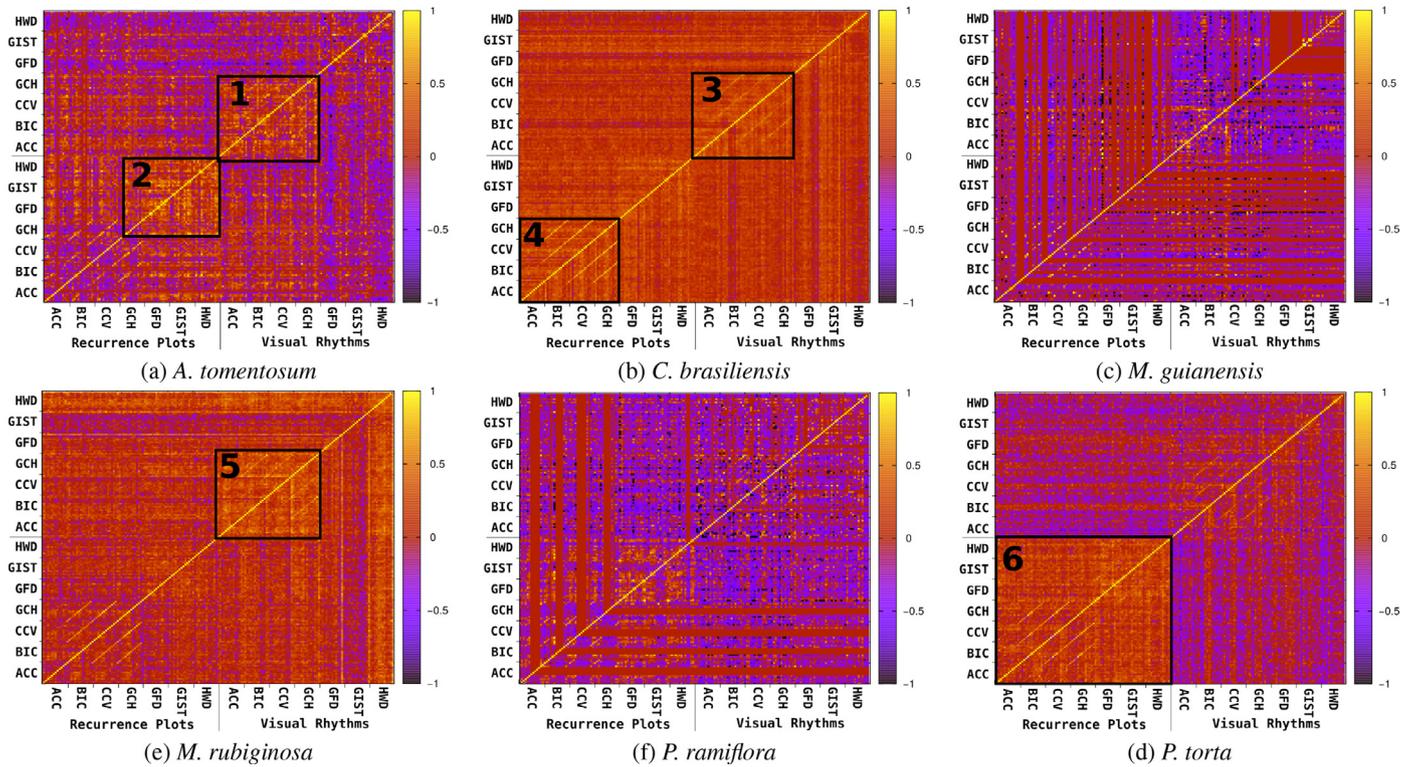


Fig. 11. Correlation analysis considering all 182 available classifiers (2 representation approaches \times 7 image descriptors \times 13 hours = 182 classifiers). The lowest correlation coefficients are closer to the purple color (-1) and the highest coefficients are closer to the yellow color (+1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

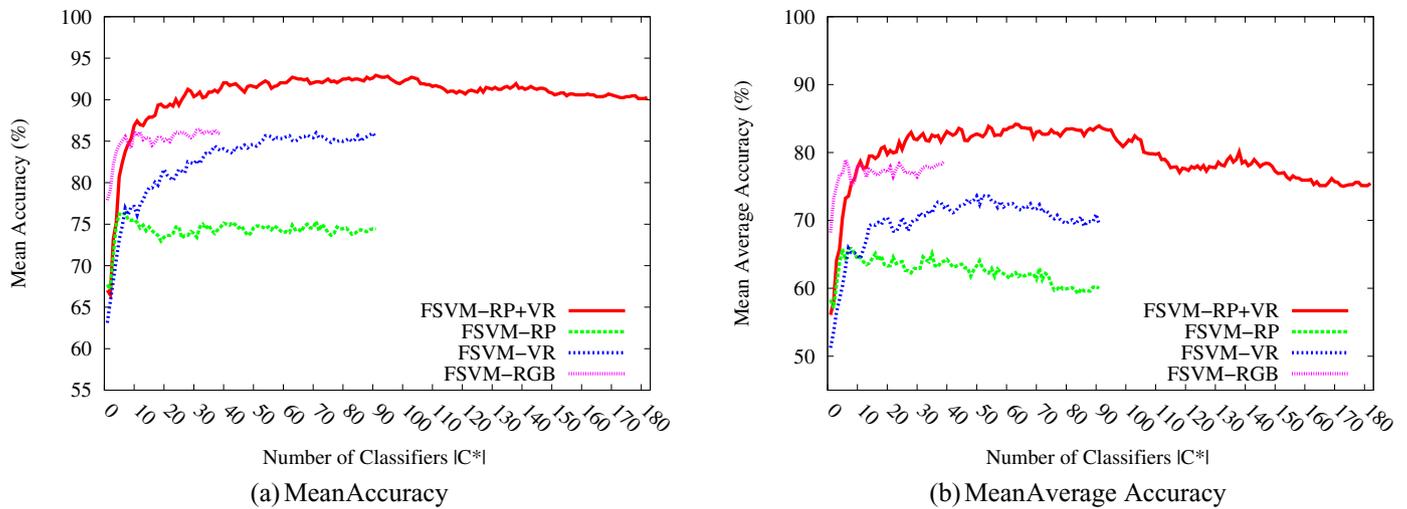


Fig. 12. Effectiveness performance considering (a) the mean accuracy and (b) the mean average accuracy when classifiers using RGB channels, RP, VR and both representations are combined.

GCH). However, in (b) the classifiers with texture image descriptors have achieved the best results (GFD, GIST, and HWD).

5.2.2. Fine-grained analysis

For a more detailed analysis, we performed a comparison between the RP and VR approaches considering their performance per class for the best achieved results. The best results of the RP and VR approaches were observed when used with GFD (texture) at 18h and BIC (color) at 6h, respectively. Fig. 10 shows this fine-grained analysis.

We can observe that the RP approach achieved better results in four (*C. brasiliensis*, *M. guianensis*, *P. ramiflora*, and *P. torta*) out

of the six species (see Fig. 10). The VR approach, in turn, yields better results for *A. tomentosum* and *M. rubiginosa*. Furthermore, in the experiments with *P. ramiflora*, and *P. torta* species, it is possible to notice a large difference between results achieved by the RP-based and VR-based approaches, around 30% of accuracy in both species.

5.3. Correlation analysis between RP-based and VR-based classifiers

This section presents the correlation analysis of each pair of classifiers for all 182 available classifiers aiming to identify

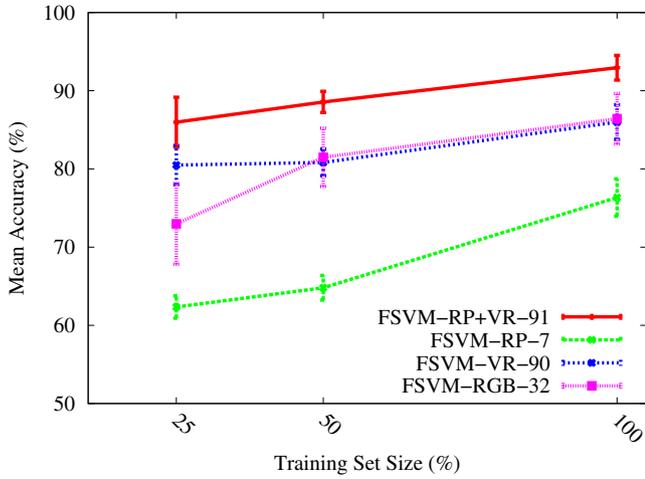


Fig. 13. Mean accuracy results of the best amount of classifiers for each approach using training sets with different sizes.

whether they might be combined by any fusion technique (e.g., Adaboost [18], Bagging [10], and meta-learning approach [16]).

The Correlation Coefficient ρ (COR) [25] has been used to assess the correlation of two classifiers c_i and c_j :

$$COR(c_i, c_j) = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}, \quad (3)$$

where a is the percentage of time series that both classifiers c_i and c_j classified correctly in a validation set. Value b is the percentage of time series that c_j hit and c_i missed, c is the opposite of b . The value d is the percentage of time series that both classifiers missed. The pairs of classifiers with lower COR values have greater degree of complementarity and are more likely to yield better results when combined. Range of COR is in $[-1, +1]$.

Fig. 11 presents the COR values for all possible combinations of pairs of classifiers considering the six classes. The lowest correlation coefficients are closer to the purple color (-1) and the highest coefficients are closer to the yellow color ($+1$). Furthermore, in this figure, there are six important regions that have been highlighted and they are explained below.

As we can observe in Fig. 11(a), in region (1), the RP-based classifiers with texture descriptors are more correlated among them. In region (2), the VR-based classifiers with color descriptors are more correlated among them. This fact means that different classifiers that use the same kind of visual properties (color or texture) in these approaches have predicted similar instances of the *A. tomentosum* species. However, outside these two regions, the same scenario can not be observed.

In general, although there are classifiers with high correlation (regions 3, 4, 5, and 6), the great majority of classifiers are non-correlated to each other (see purple color) for different species. Thus, there is a strong indication that the RP-based and VR-based classifiers used in this work might be combined to achieve better results for plant recognition task.

In this sense, we adopt a well successful classifier fusion framework [16] (Section 4.2) to address complementary information provided by RP-based and VR-based classifiers.

5.4. Fusion of RP-based and VR-based classifiers

In this section, we adopt the framework reported in Section 4.2 (FSVM), with the objective of demonstrating that is possible to combine different RP-based and VR-based classifiers and to improve the effectiveness results in the plant recognition

task. In addition, we consider the fusion of RGB channels as a baseline approach [17].

Fig. 12 shows four curves, FSVM-RGB, FSVM-RP, FSVM-VR, and FSVM-RP+VR, which refer to the fusion of the RGB channels, recurrence plot, visual rhythms, and combination of RP and VR, respectively. The suffix FSVM refers to the selection and fusion framework, which was implemented using the SVM classifier in the fusion step (meta-learning).

Two evaluation measures, Accuracy and Average Accuracy, are considered. In these experiments, the framework combines each approach for different numbers of classifiers $|C^*|$. In pink, the framework combines $|C^*| = \{1, \dots, 39\}$ RGB channels-based classifiers ($39 = 3 \text{ channels} \times 13 \text{ h}$). In green, the framework combines RP-based classifiers $|C^*| = \{2, \dots, 91\}$ ($91 = 7 \text{ descriptors} \times 13 \text{ h}$). In blue, it uses the VR-based classifiers $|C^*| = \{2, \dots, 91\}$. Finally, in red, the framework combines all available classifiers (91 RP-based plus 91 VR-based classifiers, leading to a total of 182 classifiers). As we can observe, the fusion framework using both representation (RP+VR) achieved a huge improvement when compared to the use of the two representations in isolation.

As it can be observed, the FSVM-RGB approach has achieved the better results on the range $\{1, \dots, 8\}$ classifiers. However, FSVM-RP+VR achieved the best results when more than 8 classifiers are considered ($\{9, \dots, 182\}$). Therefore, we could show that recurrence plot and visual rhythms representations address different and complementary information that might be combined to improve the effectiveness of plant identification systems.

5.5. Training set size impact

In this section, we have conducted a study considering three different sizes for the training set (25%, 50%, and 100%), which represents 15%, 30% and 60% of the entire datasets, respectively. These subsets have been selected from original training set.

Fig. 13 shows comparative analysis of the four different approaches FSVM-RGB-32, FSVM-RP-90, FSVM-VR-7, and FSVM-RP+VR-91. Notice that FSVM approaches show up with a numerical suffix N (e.g., $N = \{7, 32, 90, 91\}$), which means the number of classifiers each approach has used achieving the best effectiveness results. Furthermore, we used again the 5-fold cross-validation protocol.

Notice that the FSVM-RP+VR-91 approach, which combines RP-based and VR-based classifiers, has achieved better results than other approaches for any training set size.

6. Conclusion

In this work, we proposed the use of a technique of nonlinear data analysis for time series representation in plant species recognition task. This technique, called recurrence plot, allows us to represent repeated events (the recurrence of states) on time series into two-dimensional representation. We have extracted feature vector from this new representation through use of different kinds of image descriptor (e.g., color and texture). The created feature vector is then used as input to a learning method, in our case, the k -Nearest Neighbor method.

The experiments performed in this work showed that there are differences in terms of classification performance depending on the plant species considered, as well as, the low correlation that exists between almost all used classifiers. In these experiments, we observed that texture descriptors describe better the image provided from recurrence plots representation. However, the visual rhythms achieved better results when used in color descriptors.

A correlation analysis between all of classifiers have been performed and we could observe that recurrence plot-based and visual rhythms-based classifiers have low correlation coefficients, i.e.,

both approaches describe complementary information that might be combined by fusion techniques with the objective of producing better effectiveness results in plant recognition tasks.

Furthermore, we have adopted a successful classifier fusion framework [16] to combine RP-based and VR-based classifiers and then improve the effectiveness results. We could show in practice the huge complementarity degree between those time-series representations.

Future work includes the use of other image descriptors to extract different visual features (e.g., shape description approaches based on contour and regions [13]). In addition, the proposed framework can be augmented to consider learning-to-rank methods (e.g., genetic programming [8]) for combining different descriptors. Another point to be explored is the use of RP for multispectral images or multisensor systems, which have more than three time series. In these scenarios, remote sensing approaches for channel selection and combination can be used. Finally, we also plan to perform an extensive study on different strategies for feature selection and classifier fusion.

Acknowledgments

This research was supported by the São Paulo Research Foundation FAPESP and Microsoft Research Virtual Institute (grants #2010/52113-5, #2013/50169-1, and #2013/50155-0). BA received a master scholarship from CAPES and a doctoral fellowship from FAPESP (grant #2014/00215-0); LPCM and RST receive a Productivity Research Fellowship from CNPq (grants 310761/2014-0 and 306580/2012-8). Also, we have been benefited from funds of CAPES, CNPq, and FAPESP (grants #2009/18438-7 and #2010/51307-0).

References

- [1] U.R. Acharya, S. Vinitha Sree, G. Swapna, R.J. Martis, J.S. Suri, Automated eeg analysis of epilepsy: a review, *Knowl. Based Syst.* 45 (2013) 147–165.
- [2] P.M. Addo, M. Billio, D. Guegan, Nonlinear dynamics and recurrence plots for detecting financial crisis. Documents de travail du Centre d'Economie de la Sorbonne 13024, Universit Panthon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne, 2013.
- [3] H. Ahrends, S. Etzold, W. Kutsch, R. Stoeckli, R. Bruegger, F. Jeanneret, H. Wanner, N. Buchmann, W. Eugster, Tree phenology and carbon dioxide fluxes: Use of digital photography for process-based interpretation at the ecosystem scale, *Clim. Res.* 39 (2009) 261–274.
- [4] B. Alberton, J. Almeida, R. Henneken, R. da S. Torres, A. Menzel, L.P.C. Morellato, Using phenological cameras to track the green up in a cerrado savanna and its on-the-ground validation, *Ecol. Inform.* 19 (2014) 62–70.
- [5] J. Almeida, J.A. dos Santos, B. Alberton, L.P.C. Morellato, R. da S. Torres, Plant species identification with phenological visual rhythms, in: Proceedings of the IEEE International Conference on eScience (eScience'13), 2013a, pp. 148–154.
- [6] J. Almeida, J.A. dos Santos, B. Alberton, L.P.C. Morellato, R. da S. Torres, Visual rhythm-based time series analysis for phenology studies, in: Proceedings of the IEEE International Conference on Image Processing (ICIP'13), 2013, pp. 4412–4416.
- [7] J. Almeida, J.A. dos Santos, B. Alberton, R. da S. Torres, L.P.C. Morellato, Applying machine learning based on multiscale classifiers to detect remote phenology patterns in cerrado savanna trees, *Ecol. Inform.* 23 (2014) 49–61.
- [8] J. Almeida, J.A. dos Santos, W.O. Miranda, B. Alberton, L.P.C. Morellato, R. da S. Torres, Deriving vegetation indices for phenology analysis using genetic programming, *Ecol. Inform.* 26 (2015) 61–69.
- [9] J. Almeida, J.A. dos Santos, B. Alberton, L.P.C. Morellato, R. da S. Torres, Phenological visual rhythms: compact representations for fine-grained plant species identification, *Pattern Recogn. Lett.* (2016), doi:10.1016/j.patrec.2015.11.028.
- [10] L. Breiman, Bagging predictors, *Mach. Learn.* 24 (2) (1996) 123–140.
- [11] H.-D. Cheng, X. Jiang, Y. Sun, J. Wang, Color image segmentation: advances and prospects, *Pattern Recogn.* 34 (12) (2001) 2259–2281.
- [12] J.C. Conti, F.A. Faria, J. Almeida, B. Alberton, L.P.C. Morellato, J.L. C., R. da S. Torres, Evaluation of time series distance functions in the task of detecting remote phenology patterns, in: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR'14), 2014, pp. 3126–3131.
- [13] R. da S. Torres, M. Hasegawa, S. Tabbone, J. Almeida, J.A. dos Santos, B. Alberton, L.P.C. Morellato, Shape-based time series analysis for remote phenology studies, in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'13), 2013, pp. 3598–3601.
- [14] J.W. Eaton, D. Bateman, S. Hauberg, GNU Octave Manual Version 3, Network Theory Ltd., 2008.
- [15] J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems, *Europhys. Lett.* 4 (9) (1987) 973.
- [16] F.A. Faria, J.A. dos Santos, A. Rocha, R. da S. Torres, A framework for selection and fusion of pattern classifiers in multimedia recognition, *Pattern Recogn. Lett.* 39 (2014) 52–64.
- [17] F.A. Faria, J. Almeida, B. Alberton, L.P.C. Morellato, A. Rocha, R. da S. Torres, Time series-based classifier fusion for fine-grained plant species recognition, *Pattern Recogn. Lett.* (2016), doi:10.1016/j.patrec.2015.10.016.
- [18] Y. Freund, R.E. Schapire, Experiments with a new boosting algorithm, *Update* (1996) 148–156.
- [19] L. Guigues, J. Cocqueret, H. Le Men, Scale-sets image analysis, *Int. J. Comput. Vis.* 68 (2006) 289–317.
- [20] S.J.F. Guimarães, M. Couprie, A.A. Araújo, N.J. Leite, Video segmentation based on 2D image analysis, *Pattern Recogn. Lett.* 24 (7) (2003) 947–957.
- [21] J. Huang, R. Kumar, M. Mitra, W. Zhu, R. Zabih, Image indexing using color correlograms, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'97), 1997, pp. 762–768.
- [22] C.E. Jacobs, A. Finkelstein, D.H. Salesin, Fast multiresolution image querying, in: Proceedings of the International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'95), 1995, pp. 277–286.
- [23] E.J. Keogh, K. Chakrabarti, S. Mehrotra, M.J. Pazzani, Locally adaptive dimensionality reduction for indexing large time series databases, in: Proceedings of the ACM SIGMOD International Conference on Management of Data (ACM SIGMOD'01), 2001, pp. 151–162.
- [24] K. Kulkarni, P. Turaga, Recurrence textures for human activity recognition from compressive cameras, in: Proceedings of the IEEE International Conference on Image Processing (ICIP'12), 2012, pp. 1417–1420.
- [25] L.I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, Wiley-Interscience, 2004.
- [26] J. Lin, E.J. Keogh, L. Wei, S. Lonardi, Experiencing SAX: a novel symbolic representation of time series, *Data Min. Knowl. Discov.* 15 (2) (2007) 107–144.
- [27] L.P.C. Morellato, R.R. Rodrigues, H.F. Leitão Filho, C.A. Joly, Estudo comparativo da fenologia de espécies arbóreas de floresta de altitude e floresta mesófila semidecídua na serra do Iapí, Jundiá, São Paulo, Brazil. *J. Botany* 12 (1989) 85–98.
- [28] C.W. Ngo, T.C. Pong, R.T. Chin, Detection of gradual transitions through temporal slice analysis, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'99), 1999, pp. 1036–1041.
- [29] A. Oliva, A. Torralba, Modeling the shape of the scene: a holistic representation of the spatial envelope, *Int. J. Comput. Vis.* 42 (3) (2001) 145–175.
- [30] G. Pass, R. Zabih, J. Miller, Comparing images using color coherence vectors, in: Proceedings of the ACM International Conference on Multimedia (ACM-MM'96), 1996, pp. 65–73.
- [31] I. Popivanov, R.J. Miller, Similarity search over time-series data using wavelets, in: Proceedings of the IEEE International Conference on Data Engineering (ICDE'02), 2002, pp. 212–221.
- [32] P. Reys, M.G.G. Camargo, A.P. Teixeira, M.A. Assis, M.T. Grombone-Guaratini, L.P.C. Morellato, Estrutura e composição florística entre borda e interior de um cerrado sensu stricto e sua importância para propostas de recuperação, *Hoennea* 40 (3) (2013) 437–452.
- [33] A.D. Richardson, B.H. Braswell, D.Y. Hollinger, J.P. Jenkins, S.V. Ollinger, Near-surface remote sensing of spatial and temporal variation in canopy phenology, *Ecol. Appl.* 19 (2009) 1417–1428.
- [34] A.D. Richardson, J.P. Jenkins, B.H. Braswell, D.Y. Hollinger, S.V. Ollinger, M.L. Smith, Use of digital webcam images to track spring green-up in a deciduous broadleaf forest, *Oecologia* 152 (2007) 323–334.
- [35] M.D. Schwartz, Phenology: An Integrative Environmental Science, Springer, 2013.
- [36] D.F. Silva, V.M.A. de Souza, G.E.A.P.A. Batista, Time series classification using compression distance of recurrence plots, in: Proceedings of the IEEE International Conference on Data Mining (ICDM'13), 2013, pp. 687–696.
- [37] V.M.A. Souza, D.F. Silva, G.E.A.P.A. Batista, Extracting texture features for time series classification, in: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR'14), 2014, pp. 1425–1430.
- [38] R. Stehling, M. Nascimento, A. Falcao, A compact and efficient image retrieval approach based on border/interior pixel classification, in: Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM'02), 2002, pp. 102–109.
- [39] M. Swain, D. Ballard, Color indexing, *Int. J. Comput. Vis.* 7 (1) (1991) 11–32.
- [40] X. Wang, A. Mueen, H. Ding, G. Trajcevski, P. Scheuermann, E.J. Keogh, Experimental comparison of representation methods and distance measures for time series data, *Data Min. Knowl. Discov.* 26 (2) (2013) 275–309.
- [41] D.M. Woebecke, G.E. Meyer, K. Von Barga, D.A. Mortensen, Color indices for weed identification under various soil, residue, and lighting conditions, *Trans. ASAE* 38 (1995) 259–269.
- [42] D. Zhang, G. Lu, Shape-based image retrieval using generic fourier descriptor, *Signal Process.: Image Commun.* 17 (10) (2002) 825–848.