

## Pattern recognition tool based on complex network-based approach

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2013 J. Phys.: Conf. Ser. 410 012048

(<http://iopscience.iop.org/1742-6596/410/1/012048>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 143.107.180.158

The article was downloaded on 14/02/2013 at 12:35

Please note that [terms and conditions apply](#).

# Pattern recognition tool based on complex network-based approach

Dalcimar Casanova<sup>1</sup>, André Ricardo Backes<sup>2</sup> and Odemir Martinez Bruno<sup>1</sup>

<sup>1</sup> Instituto de Física de São Carlos (IFSC), Universidade de São Paulo, Av. Trabalhador São Carlense, 400, 13560-970 - São Carlos, SP, Brasil

<sup>2</sup> Faculdade de Computação, Universidade Federal de Uberlândia, Av. João Naves de Ávila, 2121, 38408-100, Uberlândia, MG, Brasil

E-mail: arbackes@yahoo.com.br, dalcimar@gmail.com and bruno@ifsc.usp.br

**Abstract.** This work proposed a generalization of the method proposed by the authors: 'A complex network-based approach for boundary shape analysis'. Instead of modelling a contour into a graph and use complex networks rules to characterize it, here, we generalize the technique. This way, the work proposes a mathematical tool for characterization signals, curves and set of points. To evaluate the pattern description power of the proposal, an experiment of plant identification based on leaf veins image are conducted. Leaf vein is a taxon characteristic used to plant identification proposes, and one of its characteristics is that these structures are complex, and difficult to be represented as a signal or curves and this way to be analyzed in a classical pattern recognition approach. Here, we model the veins as a set of points and model as graphs. As features, we use the degree and joint degree measurements in a dynamic evolution. The results demonstrates that the technique has a good power of discrimination and can be used for plant identification, as well as other complex pattern recognition tasks

## 1. Introduction

In recent years, many real-world problems and phenomenas are being studied with the use of complex network theory [6]. Between these problems the computer vision is one field that has benefited of these new developments. In this way, the work of [1] uses a complex network to measure features of generic shapes and use it in a classification problem.

In fact, this method have several interesting properties as (1) rotation and scale invariance: shapes with different sizes or positions have the same properties; (2) noise tolerance: the network model allows to deal with different levels of noise; (3) robustness: the network model do not need space or sequential information about the shape points. This properties allows the method to be extended for a widely applications. As example [14] use the same idea to measure the similarity of graphs on a problem of 3D objects classification.

In same way, here we generalize the technique to apply the method in a plant identification problem based on venation of their leaves. We explore the fact of the method does not require sequential informations, to model the leaf veins as a graph and measure the degree features. These features will be used in classical pattern recognition problem.

On the other hand, we propose a method to solve the problem of automated plant taxonomy [4], more specifically identify plants by analysis of their leaves.

After their shape, the next most studied aspect of leaves is the vein structure, also referred to as the venation [10]. Leaf venation corresponds to the blood vessel of organisms and provide the transport mechanism for water, minerals, sugars and other substances. The venation structure can be used to identify a plant. Although the fine detail may vary, the overall pattern of veins is conserved within many species [4].

So, in this work we use the complex network method to characterize the leaf venation. Few attempts are made to analyze or compare the venation in literature [11, 13, 12], the majority of studies is devoted to the process of extracting venation [4].

## 2. Complex network model and measurements of the vein leaf network

We can describe the research in complex networks as the intersection between graph theory and statistical mechanics. It integrates computer sciences, mathematics and physics [5], so that, this area presents a truly multidisciplinary nature.

In the last few years, many researches have focused on applying the developed concepts of the complex networks to many real data and situations. This is motivated because of the great flexibility and generality of the complex networks to model virtually any problem.

Here we propose to apply the complex network theory in a classification problem. Let  $P$  a set of  $N$  points in the  $R^n$  space. Each point is a typical vector in the form of  $p_i = [x_i, \dots, x_n]$ . To apply the Complex Networks Theory we must build a representation of this set  $P$  as a network  $G(V, E)$ . To accomplish this, we consider each point in the set  $P$  as a vertice in the network (i.e.,  $P = V$ ). Then, we add an edge binding each pair of vertices,  $e_{i,j} \in E : V \times V$ , where its weight  $w(e_{i,j})$  is the Euclidean distance between them:

$$w(e_{i,j}) = |p_i - p_j| \quad (1)$$

To ensure scale invariance of the network, we normalize the weight into the interval  $[0, 1]$ ,

$$w(e_{i,j}) = \frac{w(e_{i,j})}{\max_{w(e_{i,j}) \in E}} \quad (2)$$

Initially, all vertices in the network are connected to each other, so that, all vertices present the same number of connections. This is a regular network and it does not present properties relevant for classification. Thus, in order to extract relevant properties of this network, we propose to apply a dynamic evolution process to this network [5]. To achieve this, we apply a threshold  $t$  over the edges  $E$  to achieve a new set of edges  $E^*$ , where

$$E^* = \delta_t(E) = \{e_{i,j} \in E | w(e_{i,j}) \leq t\}. \quad (3)$$

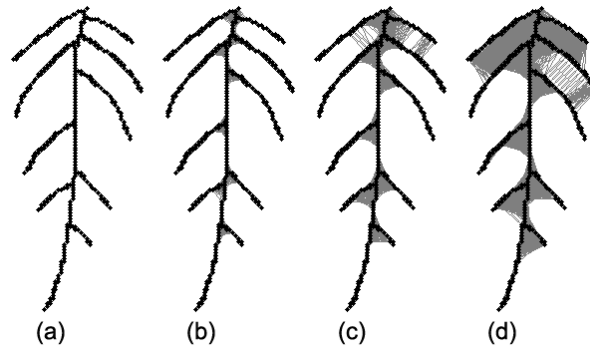
From this new set of edges  $E^*$ ,  $E^* \subseteq E$ , arises a new network  $G^* = (V, E^*)$  which can be interpreted as a intermediate step in the evolution of the network  $G$  and it owns relevant properties of the network topology. In our approach, we characterize each network  $G^*$  by using two degree measurements: *maximum degree*

$$k_\kappa = \max_i k_i \quad (4)$$

and *average degree*

$$k_\mu = \frac{1}{N} \sum_{i=1}^N k_i \quad (5)$$

where  $k_i$  is the degree of the vertice  $v_i$  in the network  $G^*$ . By considering different threshold values  $t$ , it is possible to build a feature vector that characterizes the evolution of the network  $G$  and, consequently, the original set of points  $P$ :



**Figure 1.** (a) leaf nerv; (b)-(d) complex network computed for different  $t$  values,  $t = 0.050, 0.075, 0.100$ .

$$[k_\mu(t_1), k_\kappa(t_1), k_\mu(t_2), k_\kappa(t_2), \dots, k_\mu(t_m), k_\kappa(t_m)]. \quad (6)$$

The complete procedure is showed in Figure 1.

### 3. Experiments and Results

Since [10], in his pioneer work on classification of Dicotyledonous leaves, the venation structure is used to help the plant identification and taxonomy. In fact, after their shape, the next most studied aspect of leaves is the vein structure.

Although the fine detail may vary, the overall pattern of veins is conserved within many species. Several studies also show properties superintendent this taxonomic character. According [2], for example, the angles between the veins are well defined and are directly related to the segment size. These properties are then used to build a mathematical model based on fractals to model the leaf venation.

The big challenge with automated venation analysis, however, is the extraction of the vein networks (segmentation), where we have several works dedicated to solve this problem [13, 8, 9, 3, 12].

Here this complex task is carried out in five steps, as in [13]: (i) high pass filtering; (ii) segmentation; (iii) morphological filters; (iv) thinning; (v) and outline subtraction.

In the first stage, high frequency signals were intensified through a high pass filtering of the Fourier transform. The second stage segmentation was done using the morphological Laplacean method. An morphological filters applied, removing the isolated pixel groups and leaving only the veins and the leaf outlines. A thinning algorithm was adopted for reducing the vein map to 1 pixel width and finally the last step is the subtraction of the leaf outline. This process results in a database with 44 leaf veins grouped into 11 classes.

In the experiments, we computed an initial network  $G$  for each leaf vein. Then, we generate different manifestations of  $G$ ,  $G^*$ , by applying the  $\delta$  transformation for  $t = 0.050, 0.175, 0.300, 0.425, 0.550$ . In the sequence, we computed the maximum and average degree from each network  $G^*$ . This results in a feature vector containing 10 features for each sample. We used these feature vectors to describe each leaf vein.

We evaluated the method by applying the Linear Discriminant Analysis (LDA), a supervised statistical classification method, over the feature vectors [7]. We also used the leave-one-out cross-validation scheme to evaluate the samples. The method was able to correctly classify 41 of 44 samples in their respective classes, which gives us a success rate of 93.18%.

#### 4. Conclusions

The results show that is possible generalize the complex network method in order to characterize the leaf venation. Due the properties of rotation and scale invariance, noise tolerance and robustness is possible build a graph model from the leaf venation and evaluate their similarity using measurements of their graph model. We believe that the same approach can be used on others similar problems with good results too.

#### References

- [1] A. R. Backes, D. Casanova, and O. M. Bruno. A complex network-based approach for boundary shape analysis. *Pattern Recogn.*, 42(1):54–67, 2009.
- [2] S. Bohn, B. Andreotti, S. Douady, J. Munzinger, and Y. Couder. Constitutive property of the local organization of leaf venation networks. *Phys. Rev. E.*, 65(6):061914, 2002.
- [3] O. M. Bruno, R. de Oliveira Plotze, M. Falvo, and M. de Castro. Fractal dimension applied to plant identification. *Inform. Sci.*, 178(12):2722–2733, 2008.
- [4] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin. Plant species identification using digital morphometrics: A review. *Expert Syst. Appl.*, 39(8):7562–7573, 2012.
- [5] L. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas. Characterization of complex networks: A survey of measurements. *Adv. Phys.*, 56(1):167–242, 2007.
- [6] L. F. Costa, O. N. Oliveira Jr., G. Travieso, F. A. Rodrigues, P. R. Villas Boas, L. Antiqueira, M. P. Viana, and L. E. C. Rocha. Analyzing and modeling real-world phenomena with complex networks: A survey of applications. *Adv. Phys.*, 60(3):329–412, 2008.
- [7] B. S. Everitt and G. Dunn. *Applied Multivariate Analysis*. Arnold, 2nd edition edition, 2001.
- [8] H. Fu and Z. Chi. Combined thresholding and neural network approach for vein pattern extraction from leaf images. *IEE Proc.-Vis. Image Signal Process.*, 153(6):881–892, 2006.
- [9] F. Gouveia, V. Filipe, M. Reis, C. Couto, and J. Bulas-Cruz. Biometry: the characterization of chestnut-tree leaves using computer vision. In *ISIE*, 1997.
- [10] L. J. Hickey. Classification of the architecture of dicotyledons leaves. *Amer. J. Bot.*, 60: 17–33, 1973.
- [11] Y. Nam, E. Hwang, and D. Kim. Clover: A mobile content-based leaf image retrieval system. In *ICADL*, pages 139–148, 2005.
- [12] R. O. Plotze and O. M. Bruno. Automatic leaf structure biometry: Computer vision techniques and their applications in plant taxonomy. *Int. J. Pattern Recogn.*, 23(2):247–262, 2009.
- [13] R. O. Plotze, M. Falvo, J. G. Pádua, L. C. Bernacci, M. L. C. Vieira, G. C. X. Oliveira, and O. M. Bruno. Leaf shape analysis using the multiscale minkowski fractal dimension, a new morphometric method: a study with passiflora (passifloraceae). *Can. J. Bot.*, 83(3): 287–301, 2005.
- [14] J. Tang, B. Jianga, C.-C. Chang, and B. Luo. Graph structure analysis based on complex network. *Digit. Signal Process.*, 22(5):713–725, 2012.