Ecological Informatics 36 (2016) 1–7

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

P-type based dimensionality reduction for open contours of Colombian Páramo plant species

Mary Lee Berdugo-Lattke [a], Fabio Gónzalez [b], J. Orlando Rangel-Ch [a], Francisco Gómez [c,*]

[a] Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá, Colombia
[b] Department of Computer Science, Universidad Nacional de Colombia, Bogotá, Colombia
[c] Department of Mathematics, Universidad Nacional de Colombia, Bogotá, Colombia

ARTICLE INFO

Article history:
Received 16 February 2016
Received in revised form 3 September 2016
Accepted 8 September 2016
Available online 09 September 2016

Keywords:
Colombian Páramo
Leaf morphology
Elliptic Fourier
P-type
Open contours

ABSTRACT

Páramos are a biome above the forest tree-line characterized by high native vegetation diversity. In recent years, several biota and environmental features of the Colombian Páramos have been described. Nevertheless, despite these advances, many relevant features such as leaf morphoecology of Páramo vegetation remain poorly understood. In particular, leaf morphology in Páramo plants has been sparsely described. Traditionally, leaf morphological description has been based on visual categorical systems, which can be highly biased by expert's opinion. In recent years, an objective method based on elliptic Fourier transforms and Principal Component Analysis (PCA) has been adopted by ecologists and phytologists to quantitatively assess leaf shape morphology. This approach is critically held on the idea that the leaf margin can be described as a closed contour, although, such assumption is not satisfied by several Páramo plant species because a considerable percentage of them are characterized by sessile leaves, i.e., leaves without petiole. Sessile leaves can be only described by using open contours. In this work, we propose a shape description for open contours based on p-type Fourier representations and PCA. The method was evaluated on 115 Páramo leaves from 12 species and compared with related approaches. The results show that the proposed method is able to capture high levels of variability and improve the mathematic representation of shape when compared to conventional methods based on simple geometrical quantities and elliptic Fourier transform descriptors.

1. Introduction

Páramos are a biome above the forest tree-line (at elevations between 3100 and 4700 m) located on northern Andes of South America and adjacent southern Central America. These areas are characterized by highly endemic vegetation mainly composed of giant rosette plants, scrubs, and grasses (Luteyn, 1999; Rangel-Ch, 2001). The Páramo region is recognized by a highly efficient water regulation system, which has allowed the establishment of diverse plant communities (Madriñán et al., 2013; Rangel-Ch, 2000, 2015). Because of these features, Páramo plays a fundamental ecological, economic and social role in the Andean region. In the recent years, Páramo areas have been overexposed to fire, grazing and harvesting, leading to loss of biodiversity, resulting in a reduction of water retention capacity (Buytaert et al., 2006). Páramo vegetation has been proved to be a critical regulator of the water cycle (Buytaert et al., 2006). Nevertheless, the morphoecology of the Páramo vegetation remains poorly understood. In particular, the morphological description of leaves in Páramo plants has been sparsely described in the literature (Rangel-Ch, 2006; Spehn et al., 2006).

Previous studies in Páramo leaf morphology have been mainly developed for paleoecology and ecology (Cuatrecasas, 1934; Mora-Osejo and Sturm, 1994; Rangel-Ch, 2010; Sturm and Rangel-Ch, 1985); these studies mainly describe diversity, structure, physiology and anatomical adaptations of vegetation. These studies have linked several morphological features of the plants concerning the protection mechanisms developed to support extreme environmental conditions, which characterize the Páramo region. Some of these features include the presence of pubescence on the leaves, necromass, a parabolic geometry of leaves related to the caulinar axis, and mucilages at the leaves base (Baruch, 1979; Cuatrecasas, 1934; Goldstein and Meinzer, 1983; Larcher, 1975; Meinzer et al., 1994; Miller, 1994; Mora-Osejo and Sturm, 1994; Rada et al., 1985; Smith and Young, 1987; Sturm and Rangel-Ch, 1985). These studies have provided important insights about the diverse shapes and adaptations exhibited by leaves. However, in most of the cases, these studies lack quantitative analyses to provide more precise biological interpretations (Hartendorp et al., 2012).

Plant shape description is commonly based on visual categories contained in botanical dictionaries or manuals (Ball et al., 1962;

* Corresponding author at: Av. Carrera 30 N. 45-03, Edificio 405 - Oficina 336, Bogotá, Colombia.
E-mail addresses: mlberdugol@unal.edu.co (M.L. Berdugo-Lattke), fagonzalezo@unal.edu.co (F. Gónzalez), jorangelc@unal.edu.co (J.O. Rangel-Ch), fagomezj@unal.edu.co (F. Gómez).

http://dx.doi.org/10.1016/j.ecoinf.2016.09.001
1574-9541/

Beentje and Cheek, 2003; Ellis et al., 2009; Harris and Woolf, 2001). These categorization systems are defined in terms of visual guides devised to capture the different aspects of plant parts, including, leaf, leaflets, flowers, leaf venation and roots, among others. Visual categorical systems are proposed by different authors based on her/his expertise and the data available in their area of work (Beentje and Cheek, 2003; Ellis et al., 2009). Visual shape assessment based on these systems can be a highly specialized and several times, a difficult task (Hartendorp et al., 2012; Wilder et al., 2011). Visual descriptions may constitute a good approach to the human interpretation of leaf shape; however, this strategy can be biased by the expert opinion. Such bias can be attributed, for instance, to the high degree of subjectivity implicit in the recognition of biological objects (DiCarlo et al., 2012; Hartendorp et al., 2012; Wilder et al., 2011).

An alternative approach to performing this characterization could be made by using objective geometrical features, such as length, width, and area, which describe general shape properties of the leaf (Bréda, 2003; Wright et al., 2004). Total leaf length, maximum leaf width and length-width proportion have been studied in *Espeletia pycnophylla* Cuatrec. (Frailejon) to investigate leaf morphology at Páramo vegetation (Benavides et al., 2010). Similar features, including leaf area, were also used for the description of other Páramo species (Halloy and Mark, 1996). These approaches have provided valuable insights into the underlying mechanisms of plant morphology. However, these results may be limited because they do not evaluate all different leaf shapes exhibit by Páramo plants (Ginter et al., 2012). This last aspect would be critical in the study of this ecosystem (Roy and Foote, 1997), because Páramo vegetation has been shown to be highly diverse (Rangel-Ch, 2000).

In the recent years, the phytologists and ecologists have explored a set of objective methods for the quantitative assessment of leaf shape (Chitwood et al., 2014). These methods are based, for instance, on digital morphology and dimensionality reduction techniques (Iwata and Ukai, 2002). One of the most popular approaches is based on Elliptic Fourier (EF) transforms and Principal Component Analysis (PCA) (Iwata and Ukai, 2002). In this method, a shape signature is firstly computed by using information of the leaf boundaries. Then, a real-valued Discrete Fourier Transform (DFT) is applied on this signature to capture shape variations in a set of harmonics (Kuhl and Giardina, 1982). Nevertheless, because of the periodic conditions imposed by the DFT, this representation strategy requires the description of leaf border as a closed contour. Unfortunately, this requirement may not be satisfied by several Páramo leaves because a considerable part of them are sessile, i.e., their shape is characterized by an absence of petiole. This shape can only be properly described by using open contours, as illustrates in Fig. 1.

In this paper, we propose a novel method for representation, visualization and analysis of open leaf contours from sessile leaves. The method is based on p-type Fourier representations (Uesaka, 1984) and PCA. Our main contribution is the use of p-type descriptors for the representation of the contour of Páramo leaves. In addition, we provide an objective evaluation of the dimensionality reduction quality of the proposed method compared to other related geometrical representations, in particular, simple geometrical quantities and EF.

2. Materials and methods

2.1. Leaves data

Leaves data were obtained from plants growing at the Chingaza Páramo (Chingaza National Park, Colombia, 3100–4000 m) and Pantano de Martos (Guatavita, Cundinamarca 2900–3600 m). A number between five and 10 leaves per individual were collected for 12 species [*Acaena cylindristachya* Ruiz & Pav., *Calamagrostis effusa* (Kunth) Steud., *Diplostephium phylicoides* Cuatrec., *Hypericum juniperinum* (Kunth) Wedd., *Oreobolus goeppingeri* Suess., *Paepalanthus alpinus* Körn., *Paepalanthus karstenii* Ruhland., *Puya goudotiana* Mez., *Puya santosii* Cuatrec., *Valeriana pilosa* Ruiz & Pav., *Carex jamesonii* Boott. and *Rhynchospora macrochaeta* Steud. ex Boeckeler] for a total of 115 leaves. Foliar perimeter was manually traced and rasterized into a digital format at 300 dpi using a desktop scanner (HP Scanjet G2710) (O'Neal et al., 2002). Adjustments for tonal curves, contrast, and contour closing were applied to the images; Fig. 1 shows one leaf sample per species.

The species were selected based on their ecological dominance (coverage percentage) in plant communities and geographical distribution (Rangel-Ch, 2000). Samples include dominant species of wide distribution for several terrestrial vegetation types, such as *C. effusa*, *H. juniperinum*, *R. macrochaeta*, *A. cylindristachya* and *V. pilosa*, some dominants species of the Páramo region around Bogotá highland as *D. phylicoides*, *P. karstenii*, *P. alpinus* and also dominant and frequent species in marsh areas such as *O. goeppingeri*, *P. goudotiana*, *P. santosii* and *C. jamesonii* species. It is worthy to recall that this study only considers the shape variation in open contours, i.e., leaves and/or leaflets without petiole (sessile leaves).

2.2. Step by step proceeding of algorithm

Fig. 2 shows the proposed algorithm. First, the leaf samples were stored as binary images. The leaf contour for each sample was represented as a series of complex numbers. A normalization process was

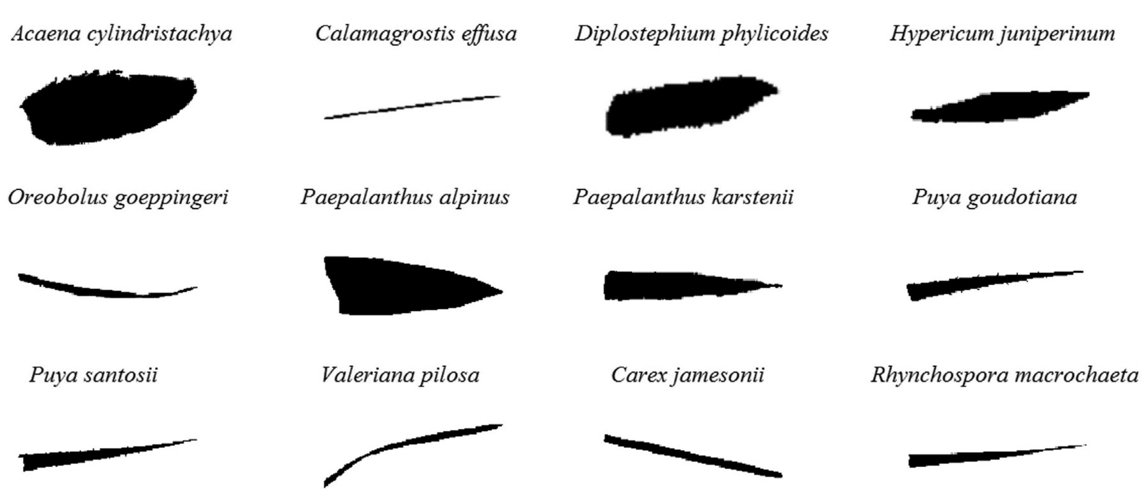

Fig. 1. Digital leaves samples for 12 Colombian Páramo species. These leaves lack petiole and should be characterized by open contours. All species have 10 samples except by *P. santosii* in which the number of samples was five.