Plant Leaf Recognition Using Zernike Moments and Histogram of Oriented Gradients

Dimitris G. Tsolakidis, Dimitrios I. Kosmopoulos, and George Papadourakis

Department of Informatics Engineering,
Technological Educational Institute of Crete,
GR-71004, Heraklion, Greece
dimitri_ts0@hotmail.com,
dkosmo@ie.teicrete.gr, papadour@cs.teicrete.gr

Abstract. A method using Zernike Moments and Histogram of Oriented Gradients for classification of plant leaf images is proposed in this paper. After preprocessing, we compute the shape features of a leaf using Zernike Moments and texture features using Histogram of Oriented Gradients and then the Support Vector Machine classifier is used for plant leaf image classification and recognition. Experimental results show that using both Zernike Moments and Histogram of Oriented Gradients to classify and recognize plant leaf image yields accuracy that is comparable or better than the state of the art. The method has been validated on the *Flavia* and the *Swedish Leaves* datasets as well as on a combined dataset.

Keywords: Leaf recognition, Zernike moments, Histogram of oriented gradients, support vector machine.

1 Introduction

Plants play a vital role in the environment. There is huge volume of plant species worldwide and their classification has become an active area of research [1]. A plant database is of obvious importance for archiving, protection and education purposes. Moreover, recognition of plants has also become essential for exploiting their medicinal properties and for using them as sources of alternative energy sources like bio-fuel. The classification of plant leaves is a useful mechanism in botany and agriculture [3]. Additionally, the morphological features of leaves can be employed in the early diagnosis of certain plant diseases [5].

Plant recognition is a challenging task due to the huge variety of plants and due to the many different features that need to be considered. There are various ways to recognize a plant, like flower, root, leaf, fruit etc. Recently, computer vision and pattern recognition techniques have been applied towards automated process of plant recognition [2]. Leaf recognition plays an important role in plant classification and its key issue lies in whether the chosen features have good capability to discriminate various kinds of leaves. Computer aided plant recognition is still challenging due to improper models and inefficient representation approaches.

A. Likas, K. Blekas, and D. Kalles (Eds.): SETN 2014, LNAI 8445, pp. 406–417, 2014.

A lot of work has been focused on the shape description of the leaves. In the past decade, research on contour-based shape recognition [18-19] was more active than that on region-based [17]. In [15] they introduced a multiscale shape-based approach for leaf image retrieval. The leaf represented by local descriptors associated with margin sample points. Within this local description, they studied four multiscale triangle representations: the well-known triangle area representation (TAR), the triangle side lengths representation (TSL) and two other representations, triangle oriented angles (TOA) and triangle side lengths and angle representation (TSLA). In this research they used 1-NN as classifier. In [16] they proposed a contour-based shape descriptor, named Multiscale Distance Matrix (MDM), to capture shape geometry while being invariant to translation, rotation, scaling, and bilateral symmetry and to classify the plants they used 1-NN as classifier. The color information was incorporated in the plant identification in [11] and [12] and RBPNN was used as classifier. However most researchers avoid using color, mainly due to its dependency on the illumination.

Some other researchers focused on green leaves and ignored color information on the leaf. In [10] they used PNN to classify 32 species of plants. All the plants they used in their research had green leaves. Also in [24] Zulkifli used General Regression Neural Networks (GRNN) and invariant moment to classify 10 kinds of plants. They did not include color features to the classifier. Furthermore, in [14] they used K-SVM to classify 32 species of plants and they also did not use any color features.

This paper differs from the previous approaches due to the fact that we propose a method for recognizing leafs using as shape descriptor the Zernike Moments (ZM) and as a descriptor for the interior of the leaf the Histogram of Oriented Gradients (HOG). Support Vector Machine (SVM) has been used as a classifier, which is among the best methods for discriminative models.

Experimental results on *Flavia* dataset [21] indicates that the proposed method yields an accuracy rate of 97.18%, on *Swedish Leaves* dataset [22] 98.13%. When we combine both *Flavia* and *Swedish Leaves* the obtained accuracy is 97.65%. To our knowledge these results are similar or better than the state of the art, and it is the first time someone combines these two popular databases.

An overview of the method is given in Fig. 1. More specifically we perform a preprocessing step, then we extract a feature vector per image and finally we do the classification of the image.

The rest of the paper is organized as follows. In the next section we describe the preprocessing steps. In section 3 we outline the feature extraction method and in section 4 we outline the classification method. In section 5 we give the experimental results of our method and section 6 concludes this paper.