Sensors and Actuators B 249 (2017) 278–295

Contents lists available at ScienceDirect

Sensors and Actuators B: Chemical

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

Comparison of decision tree based classification strategies to detect external chemical stimuli from raw and filtered plant electrical response

Shre Kumar Chatterjee[a], Saptarshi Das[a,*], Koushik Maharatna[a], Elisa Masi[b], Luisa Santopolo[b], Ilaria Colzi[b], Stefano Mancuso[b], Andrea Vitaletti[c]

[a] School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
[b] Department of Agri-Food Production and Environmental Science, University of Florence, Florence, Italy
[c] DIAG, Sapienza University of Rome, via Ariosto 25, 00185 Rome, Italy

ARTICLE INFO

Article history:
Received 17 March 2016
Received in revised form 8 March 2017
Accepted 12 April 2017
Available online 14 April 2017

Keywords:
Decision tree
Multiclass classification
Discriminant analysis
Mahalanobis distance classifier
Statistical features
Plant electrical signal processing
Time series analysis

ABSTRACT

Plants monitor their surrounding environment and control their physiological functions by producing an electrical response. We recorded electrical signals from different plants by exposing them to Sodium Chloride (NaCl), Ozone (O_3) and Sulfuric Acid (H_2SO_4) under laboratory conditions. After applying pre-processing techniques such as filtering and drift removal, we extracted few statistical features from the acquired plant electrical signals. Using these features, combined with different classification algorithms, we used a decision tree based multi-class classification strategy to identify the three different external chemical stimuli. We here present our exploration to obtain the optimum set of ranked feature and classifier combination that can separate a particular chemical stimulus from the incoming stream of plant electrical signals. The paper also reports an exhaustive comparison of similar feature based classification using the filtered and the raw plant signals, containing the high frequency stochastic part and also the low frequency trends present in it, as two different cases for feature extraction. The work, presented in this paper opens up new possibilities for using plant electrical signals to monitor and detect other environmental stimuli apart from NaCl, O_3 and H_2SO_4 in future.

1. Introduction

Plants, such as *Mimosa pudica* (Touch-me-not) and *Helianthus annuus* (Sunflower), show some form of physical changes due to external stimuli in the form of touch and sunlight respectively [1]. The wilting of general plants due to dry environmental conditions is also commonly found. For many years, researchers have tried to establish the relationship between these reactions of the plants and the surrounding environmental conditions [1]. It has been found that the underlying phenomenon behind this is the plant electrophysiological mechanism which may be traced in the electrical response of the plant to the external stimulus [1]. These electrical signals, which control various physiological functions in the plants, hold useful information about the external stimulus (which causes the electrical signal in the plant) contained within its deterministic and stochastic parts to different extents. Analysis using low frequency (trend) part of the plant electrical signal to study the external chemical or light stimulus has been reported in Chatterjee et al. [2,3]. Also, other studies on plant electrical signal processing have been reported in Refs. [4–14], in particular use of classification techniques to find out the applied external stimuli, through various statistical features computed from the recorded plant electrical signal, was reported first in Ref. [2]. Since the statistical features in Ref. [2] were extracted from raw plant signals (with low frequency trends or drifts), a background (pre-stimulus) information subtraction method was adopted in the classification process to focus only on the incremental values in each feature due to the application of the stimulus.

In this paper, we initially focus on the information contained in the stochastic part of the plant electrical signals by applying a high pass filtering on the raw signals to remove the inconsistent trends or drifts. We also used raw signals with the trends (using the background information subtraction method as reported in Ref. [2])

* Corresponding author.
E-mail addresses: skc105@ecs.soton.ac.uk (S.K. Chatterjee), s.das@soton.ac.uk, sd2a11@ecs.soton.ac.uk (S. Das), km3@ecs.soton.ac.uk (K. Maharatna), elisa.masi@unifi.it (E. Masi), luisa.santopolo@unifi.it (L. Santopolo), ilaria.colzi@unifi.it (I. Colzi), stefano.mancuso@unifi.it (S. Mancuso), vitaletti@dis.uniroma1.it (A. Vitaletti).

http://dx.doi.org/10.1016/j.snb.2017.04.071
0925-4005/

to show a comparative analysis between the classification performance of the filtered and raw plant signals. Thus, we here explore, if there is any improvement in the classification process while using only the detrended random part rather than the raw signal containing small local fluctuations superimposed on relatively larger change in the trends.

In order to develop a classification strategy for detecting the external chemical stimuli, here we used 15 features out of which 11 features have been reported in Ref. [2]. In addition to these 11 features, four additional statistical features have been explored along with independent testing of the classifiers with a much larger dataset. In this paper, we report the use of *discriminant analysis* and *Mahalanobis distance* based classification algorithms to establish a decision tree based classification system using both *One-Versus-One* (OVO) and *One-Versus-Rest* (OVR) strategy [15,16]. We also report the validation scheme of the classifiers in two different ways – (1) Leave One Out Cross Validation (LOOCV) on ~73% of the available data (*retrospective study*) and (2) independent testing of the remaining ~27% of the data (*prospective study*).

Datasets from experiments on plants using three different stimuli – NaCl, H_2SO_4 and O_3 have been used in this exploration and can be found on the EU FP7 funded project PLants Employed As SEnsor Devices (PLEASED) website. The work presented in this paper may help in taking a step closer to the concept of a plant electrical signal based external stimuli sensing platform which the PLEASED project aims to develop. Such a device, if successful, will aid in monitoring a large geographical area for multiple environmental stimuli or pollutants of interest. In order to proceed towards the realization of such a stimulus classification scheme using plant electrical signals, the steps shown in Fig. 1 were followed.

The work presented in this paper presents the following salient contributions over previous approaches:

- The present work uses filtered (containing only stochastic part of the signals) as well as raw plant signal (containing both deterministic and stochastic parts of the signals) for classifying the external stimuli for a comparative analysis between two pre-processing techniques affecting the final classification results.
- An exhaustive set of experimental data with 28,070 data blocks were used for training the classifiers, which is ~7.4 times higher than the data used in our previous work [2].
- Apart from Tomato and Cucumber plants which were used for the experiments reported in Ref. [2], Cabbage was included in the present work for extracting the experimental data. This helped to analyse the electrical response in a pool of three different species due to chemical stimulus.
- Features from NaCl of different concentration (5 ml and 10 ml) were combined to be labelled as NaCl in the present classification work, whereas in the previous work [2], these two stimuli were considered as separate classes.
- In the previous study [2], the classification results were obtained using individual (univariate analysis) or feature pairs (bivariate analysis), whereas multivariate feature analysis have been carried out here.
- Apart from the previously explored 11 statistical features in Ref. [2], here we explore four additional features of the plant electrical signal for classification of the external stimuli viz. higher order central moments (hyperskewness and hyperflatness), fano factor, and correlation dimension.
- The classification results reported in the previous study [2] was average results for six binary stimuli combinations that in a way averages the detectability of one class with the others. Whereas a systematic decision tree has been developed in the present study that helps answering the question which chemical stimulus can be easily detected from the plant signals and which stimuli are hard to differentiate from the rest.
- Along with reporting the retrospective classification accuracy employing LOOCV, a separate held out dataset was used for prospective validation in this work, whereas only retrospective LOOCV results were reported in the previous work [2], on a much smaller dataset.

2. Recording electrical signal from plants

Raw electrical signals from different experiments involving O_3, NaCl and H_2SO_4 as external stimuli, under laboratory conditions, were acquired from different plants. Each experiment was conducted on a new plant, thereby eliminating the risk of any residual effect of the previous experiments infiltrating the current electrophysiological condition of the plants. The experiments were conducted inside a plastic transparent box placed in a Faraday cage kept in a dark room, as shown in Fig. 2, to minimize any exter-

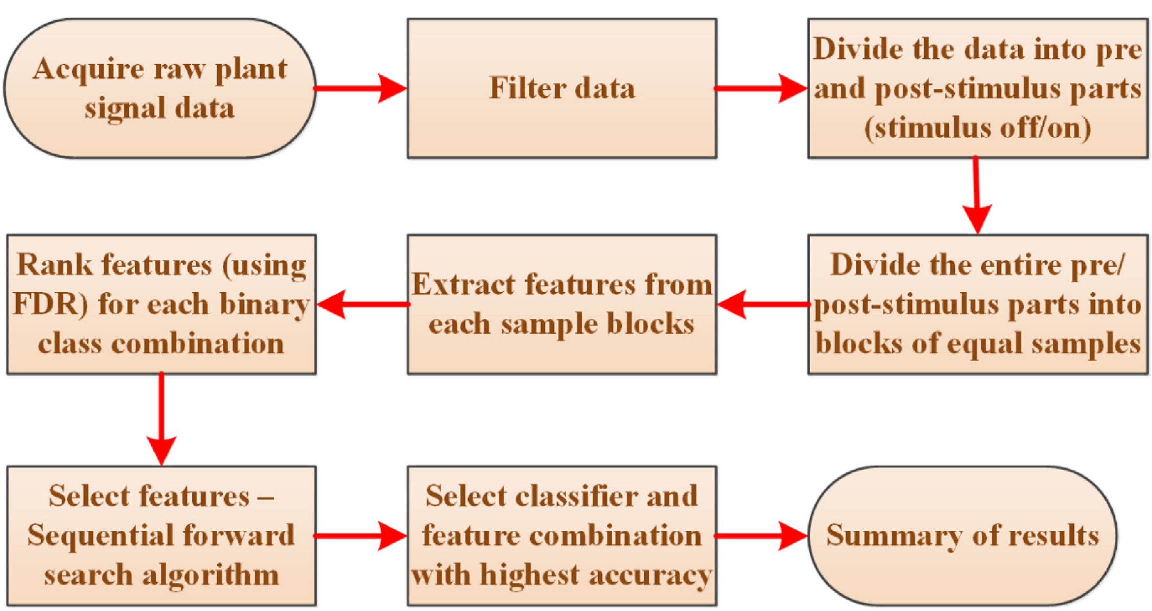

Fig. 1. Steps for classification of environmental stimuli using plant electrical signal.