Computers and Electronics in Agriculture 143 (2017) 293–301

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Discrimination among tea plants either with different invasive severities or different invasive times using MOS electronic nose combined with a new feature extraction method

Yubing Sun, Jun Wang*, Shaoming Cheng

Department of Biosystems Engineering, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, PR China

ARTICLE INFO

Keywords:
Tea plant
Damage
Electronic noses
Feature extraction

ABSTRACT

Damage of tea plant causes a lot of loss in tea production, but there is not an appropriate method to detect tea plants with pest damage. In this work, electronic nose (E-nose) and Gas Chromatography-Mass Spectrometer (GC-MS), as an auxiliary technique, were employed to detect tea plants with pest damage in two aspects, including tea plants with different invasive severities and with different invasive times, for giving a comprehensive results. A new feature extraction method based on a piecewise function was proposed and its performance was compared with those of the other three commonly employed models-polynomial functions, exponential functions, and Gaussian functions. Feature selection based on principal component analysis (PCA) and multi-layered perceptron (MLP) were employed for further feature reduction and classification, respectively. The results showed that feature extraction based on piecewise function was the best. The combination of feature extraction based on piecewise function, feature selection based on PCA and MLP was the best method and good enough for the classification in tea plants damage area. The results proved that E-nose was able to detect tea plants either with different invasive severities or different invasive times.

1. Introduction

Tea with great flavor and high content of beneficial substances (Suzuki et al., 2016), is the most widely consumed beverage aside from water (Vernarelli and Lambert, 2013). Especially in China, tea has been used as a daily beverage and crude medicine for thousands of years (Yen and Chen, 1995). Tea plant (*Camellia sinensis*) is grown in about 60 countries (Ma et al., 2016) and an important crop in many countries, such as China, Japan, Kenya and India (Dong et al., 2011; Ramya et al., 2013), etc. The tender leaves tea plant are the raw materials of tea (Saravanakumar et al., 2007). The health of tea plant is therefore a crucial factor for producing high yield and high quality of tea. However, many pests attack tea leaves during tea plant growth, causing 5–55% losses in tea production and 0.5 billion–1 billion dollars in economic losses (Hazarika et al., 2009). Tea geometrid (*Ectropis obliqua*, a chewing insect) is one of the most common pests of tea plant in China and causes severe damage by feeding on tea leaves (Ma et al., 2016; Wang et al., 2016). In this work, the damage caused by *Ectropis obliqua* was studied. Invasive severity and invasive time are two important parameters for pest damage detection. Hence, in this study, two aspects, including tea plants with different invasive severities and with different invasive times, were discussed for giving a comprehensive information of pest damage.

A lot of researches about the reaction of plant to pests have been published and the results showed that volatile organic compounds (VOCs) emitted by plant change when the plant is attacked by pests (Henderson et al., 2010; Holopainen and Gershenzon, 2010; Snoeren et al., 2007; Yu et al., 2008b). Some researches also showed response changes of tea plant when exposed to multiple stresses (Fu et al., 2015; Mei et al., 2016), and VOCs vary both quantitatively and qualitatively with infestation duration and herbivore density (Cai et al., 2014). Electronic nose (E-nose) (Kiani et al., 2016; Markom et al., 2009), which is a nondestructive technique that mimics the human olfactory system detecting based on volatiles emitted by samples, is therefore possible to detect pest damage (Lampson et al., 2014). Besides, it is an instrument embedded with a chemical sensor array of partial specificity cooperating with appropriate pattern recognition algorithms used for detection of chemicals (Gardner and Bartlett, 1999). Gas Chromatography-Mass Spectrometer (GC-MS), which is a precise detection technique and able to determine the constituents of volatiles and their contents (Ivanova et al., 2013), was also employed and combined with E-nose for giving a more reliable results.

* Corresponding author.
E-mail address: jwang@zju.edu.cn (J. Wang).

https://doi.org/10.1016/j.compag.2017.11.007
Received 20 December 2016; Received in revised form 12 October 2017; Accepted 7 November 2017

Metal oxide semiconductor (MOS) gas sensors, which have the advantages of cross-sensitivity, broad spectrum response and low-cost, have been widely used in E-nose application (Zhang et al., 2014). The electrical signal of E-nose based on MOS sensors consists of a few hundred measured values, which are too many to be signals as input of analytical methods and exist data redundancy. Feature extraction, as a process of finding a small set of values that largely represented the whole signals, is constantly required (Carmel et al., 2003). Furthermore, feature extraction method is one of the key points of performance improvement of E-nose systems because feature extraction is the first step of the sensor signal processing (Gardner et al., 2005). The feature extraction methods from original response curves, such as maximum value, average value, maximum slope, maximum variance, standard deviation and root mean square (Distante et al., 2002; Yu et al., 2008a), are commonly applied for E-nose because they are visual, intuitive and fast to compute. In particular, maximum value and average value are frequently-used features (Dai et al., 2015).

However, the feature extraction methods from original response curves leave some potential information carried in the signals unutilized. Although these methods have been used for many simple applications successfully, it is widely accepted that more sophisticated features are required when turning to more demanding tasks (Hong et al., 2014). Feature extraction methods based on curve fitting extract features depending on the whole dataset of E-nose and largely overcome this shortcoming. In curve fitting, the critical and complex problem is how to select and envision the specific form of the unknown function (Yan et al., 2015). The most commonly used models (Yan et al., 2015), such as polynomial functions, exponential functions, and Gaussian functions, are all nonlinear, which makes the fitting process complicated and long. A feature extraction method based on a new curve fitting model, a piecewise function, is introduced in this study. This model transforms the nonlinear data into linear first, and then linear equation is employed to fit the linear data and calculate the unknown parameters as features. Hence, this feature extraction method is simple and quick to compute.

In this work, E-nose was employed, and GC-MS was used as an auxiliary technique for proving the potential of E-nose in detecting tea plants either with different invasive severities or different invasive times. A new feature extraction method based on a piecewise function for MOS sensor response was proposed for detecting tea plants either with different invasive severities or different invasive times and its performance was compared with those of feature extraction methods based on the three commonly applied models-polynomial function, exponential function, and Gaussian function. Curve fitting figure and root mean squared error (RMSE) were introduced to evaluate the fitting performance of four curve fitting models, while correct classification rate obtained by multi-layered perceptron (MLP) was calculated for assessing classification performance as well as proving the advantage of E-nose in tea plant detection. The main objectives of this research are: (1) to prove the ability of E-nose in detecting tea plants either with different invasive severities or with different invasive times, (2) to propose a new feature extraction method for E-nose data analysis, (3) to explore if the new proposed method would outperform the feature extraction method based on conventional widely applied models.

2. Experimental

2.1. Experimental sample and design

This study was carried out using the tea plant cultivar "clone Longjing43", which was 20–30 cm high and had 10 leaves roughly for each plant, and pest (*Ectropis obliqua*) at the 3th larval stage was provided by Tea Research Institute, Chinese Academy of Agricultural Sciences. Each tea plant was cultivated in tea plantation (at field scale) in Hangzhou, China, and watered weekly. They were taken care of carefully making sure that each tea plant was healthy and undamaged. And the chosen tea plants were transplanted to laboratory one week before experiment and all free from insects and mites, kept under controlled conditions (24 ± 2 °C, 75–85% of relative humidity). The pests were reared on fresh tea leaves (cultivar "clone Longjing43") for at least two generations and maintained in a growth chamber under controlled conditions (24 ± 2 °C, 75–85% of relative humidity, and L16:D8 photoperiod), as described by Cai et al. (2014). The experiment was also carried out under controlled conditions (24 ± 2 °C and 75–85% of relative humidity). The experiment began on 10th September 2015, and two independent GC-MS and E-nose experiments were taken.

Experiment 1 (invasive severity) consisted of four groups of tea plant samples, including untreated tea plants group, group of tea plants with low invasive severity (tea plants invaded by 5 pests), group of tea plants with medial invasive severity (tea plants invaded by 10 pests), and group of tea plants with high invasive severity (tea plants invaded by 15 pests), labeled as groups Untreated, 5Pests, 10Pests and 15Pests, respectively. The time point putting pest on tea plant was 8AM. The VOCs' collection process for GC-MS was from 8AM to 4PM, while the time point for E-nose detection was 4PM. Each group had 3 replications for GC-MS detection and 20 replications for E-nose detection. One tea plant was applied for each replication.

Experiment 2 (invasive time) consisted of four groups of tea plant samples, including untreated tea plants group, group of tea plants with short invasive time (tea plants with 8 h invaded), group of tea plants with medial invasive time (tea plants with 16 h invaded), and group of tea plants with long invasive time (tea plants with 24 h invaded), labeled as groups Untreated, 8 h, 16 h and 24 h, respectively. Each group was invaded by 10 pests and the time point putting pest on tea plant was 12Midnight. The VOCs' collection process for GC-MS was from 12Midnight to 8 AM, 8 AM to 4 PM and 4 PM to 12Midnight for groups 8 h, 16 h and 24 h, respectively. While the time point for E-nose detection was 8 AM, 4 PM and 12Midnight, respectively. For group Untreated, the VOCs' collection process for GC-MS was from 8 AM to 4 PM, while the time point for E-nose detection was 4PM. Each group had 3 replications for GC-MS detection and 20 replications for E-nose detection. One tea plant was applied for each replication.

2.2. GC-MS measurements

Purge-and-trap coupled with GC-MS was employed to determine VOCs emitted by different samples (Foreman et al., 2015; Andrews et al., 2015). Purge-and-trap system, which was applied to collect volatiles, consisted of eight parts (Fig. 1), including pump, charcoal filter, flowmeter, split Teflon plate, trestle table, glass cover, cylindrical glass tubes and Porapak-Q adsorbent trap. The pump pushed air going through charcoal filter that was used to purify the air and then in sequence through the whole system from the left side of figure to the right side. During this process, the flow rate, which was controlled by flowmeter, was keep at 0.4 L/min. When the purified air passed through the glass cover including the intact tea plant, the VOCs released by tea plant was extruded and went through Porapak-Q adsorbent trap, which contained Porapak-Q adsorbent that was able to adsorb VOCs passing through it.

In this experiment, 8 h was used for collecting VOCs. Then, the collected VOCs were eluted with 600 μL of dichloromethane. Ethyl caprate was taken as internal standard, and 3 μL of Ethyl caprate dichloromethane (50 μL/L) was injected into the eluted solution by a syringe with a range of 10 μL. The solution made above was analyzed by GC-MS.

In this study, VOCs were analyzed in an HP 6890 series gas chromatograph equipped with a flame ionization detector, and coupled to an HP 5973 mass spectrometer selective detector (Agilent Technologies, Palo Alto, CA, USA). An HP-5 methyl siloxane chromatographic column (30 m, 0.25 mm internal diameter, and 0.25 m film thickness; Alltech, Deerfield, IL, USA) was used for separation. Helium (24 mL/min) was