Journal of Food Engineering 203 (2017) 25–31

Contents lists available at ScienceDirect

Journal of Food Engineering

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

An optimization of the MOS electronic nose sensor array for the detection of Chinese pecan quality

Keming Xu, Jun Wang*, Zhenbo Wei, Fanfei Deng, Yongwei Wang, Shaoming Cheng

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

ARTICLE INFO

Article history:
Received 2 July 2016
Received in revised form
20 January 2017
Accepted 23 January 2017
Available online 25 January 2017

Keywords:
Array optimization
e-nose
Feature matrix
PCA
PLSR
BPNN

ABSTRACT

In this research, an embedded metal oxide semiconductor (MOS) electronic nose (e-nose) was designed to detect Chinese pecan quality. To improve the performance of e-nose, three types of features were extracted to form initial feature matrix, including mean-differential coefficient value, stable value, and response area value. Furthermore, followed by the non-search feature selection strategy, optimized feature matrix was obtained through the procedure of mean analysis, variation coefficient analysis, cluster analysis and correlation analysis. It was observed that pecans were better classified after the optimization of initial feature matrix, shown by principal component analysis (PCA) score plot. And also the regression models of optimized feature matrix established by partial least squares regression (PLSR) ($R^2 = 0.9377$) and back propagation neural networks (BPNN) ($R^2 = 0.9787$) presented a better prediction capacity than these of initial one (PLSR: $R^2 = 0.8887$; BPNN: $R^2 = 0.9093$). In conclusion, the optimization method not only reduced data dimensionality but also improved electronic nose performance.

1. Introduction

Chinese pecan is a popular nut for its high nutritional value and unique flavor. As the country with a massive production of pecan, China yielded approximately 1.42 million tons of pecans in 2013, nearly half of the world production in total. However, pecan kernel is prone to rancidity during storage because of environmental factors such as light, oxygen and humidity (Jin et al., 2011), which will lead to over-high peroxide value as well as acid value and also do harm to people's health. Hence, pecan quality detection has practical significance.

Pecan quality detection involves some conventional methods that are time-consuming or money-consuming, such as sensory evaluation, microbial detection and physicochemical indexes' assessment. Furthermore, they are subjective and need complex sample preparation, which restrict their applications. Some nondestructive detection methods including near-infrared spectroscopy analysis (Nakariyakul, 2014; Moscetti et al., 2014) and machine vision analysis (Pablo et al., 2016) are increasingly applied to the field of pecan quality detection. Nevertheless, their applications have been held back due to the dependence on complicated detection devices and the protection of pecan's shell leading to the decrease of detection performance. Different from aforementioned methods, electronic nose (e-nose) is a bionic device designed to obtain internal information intelligently via the "fingerprint figure" of sample's volatiles. During the storage, specific volatiles are released through pores of pecan's shell and their components change with storing time. As a consequence, e-nose can be employed to detect pecan quality.

In addition, literature have been published about pecan quality detection based on e-nose (Zhang et al., 2008; Wei et al., 2015; Marion et al., 2011; Tian et al., 2013; Jiang and Wang, 2016). However, e-noses applied in researches mentioned above all are commercial e-nose, and their sensor arrays are not designed based on specific samples, which may have influence on the accuracy of detection. Meanwhile, volatiles of pecans are complicated and just have small difference at different storage periods, it's therefore vital to optimize the sensor array to improve the performance of e-nose in detecting pecans. Recent decades have witnessed the development of many kinds of e-noses, among which metal oxide semiconductor (MOS) e-nose holds the widest application for its rapid response, high sensitivity and most importantly low cost during the process of development. Hence, a self-designed MOS e-nose with the optimized sensor array was applied to detect pecan quality in this paper.

Optimization of sensor array is usually based on feature

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

http://dx.doi.org/10.1016/j.jfoodeng.2017.01.023
0260-8774/

selection, which means that the sensors corresponding to the selected features were chosen to form the optimized sensor array. Generally, there are two different methods for feature selection, searching method and non-searching method. The searching type, such as genetic algorithms (Shi et al., 2013; Gardner et al., 2005), simulated annealing algorithms (Llobet et al., 2007), random search algorithms (Bertolazzi et al., 2016) and sequential forward algorithms (Jeong et al., 2014), is a method mainly based on algorithms selecting the best feature subset from all feature combinations. However, their applications are limited by complicated calculation process and easily converging to local minima, and the process of them is just like a black box (i.e., the reason why these features are selected is unknown). While non-searching type employed in this paper refers to the combination of multivariate statistical methods such as ANOVA, coefficient of variation analysis and correlation analysis, which leads to efficient computing and visualizes the optimization process (i.e., which features and the reason why these features are excluded of each step can be inferred).

For the non-searching feature selection method, some researches have been done to get the optimized sensor array. However, in these studies, just one feature was extracted representing one sensor to generate the initial feature matrix, then non-searching selection method was employed to obtain the optimized feature matrix step by step, and finally the corresponding optimized sensor array was gained (Zhang et al., 2007; Fei et al., 2012; Cetó et al., 2014; Wang et al., 2015). During the optimization above, the elimination of one feature means the exclusion of corresponding sensor from the sensor array (i.e., optimization based on sensors), which lead to the ignorance of some useful information of sensor's response signals. As a result, this method may fail to get the best performed sensor array. In this paper, three kinds of features (mean-differential coefficient value (MDCV) (Zhao et al., 2007), stable value (SV) (Hui et al., 2015) and response area value (RAV) (Wei et al., 2013)) of each sensor were extracted to form the initial feature matrix and optimized it based on the non-searching feature selection method step by step, which includes mean analysis, variation coefficient analysis, cluster analysis and correlation coefficient analysis. This procedure could be considered as optimization based on features, which means that one sensor is excluded only if all its corresponding features are eliminated, and the optimized feature matrix obtained will show better performance in pattern recognition. To verify the validity of optimization, the classification performances based on initial feature matrix and optimized one were compared through principal component analysis (PCA). Furthermore, partial least squares regression (PLSR) and back propagation neural networks (BPNN) were applied and their prediction abilities were compared.

2. Materials and methods

2.1. Materials

A batch of fresh Chinese pecans was harvested from Longgang Town, Lin'an City in Zhejiang Province. Pecans with basically the similar size and color were selected and randomly divided into 4 groups (30 sample sets per group and 20 pecans per set). And then accelerated aging process was implemented to shorten the time consumed in the experiment. Pecans stored in a 4 °C cold storehouse for 1 year and 2 years can be simulated by keeping them in an incubator (STIK (Shanghai) CO., China) at 35 °C and 30% relative humidity (RH) for 10 days and 20 days respectively (Wang et al., 2006; Ling et al., 2013). Therefore, pecans were placed in the incubator for artificial aging. The first day when samples were placed into the incubator was defined as day0. Then, each group of pecans was taken out from the incubator every 5 days (defined as day5, day10, day15 respectively).

2.2. Electronic nose (e-nose)

An embedded MOS e-nose designed by the agricultural equipment and intelligent detection (AE&ID) team of Zhejiang University was applied in the experiment, which consisted of a gas sensor array, tube-shaped chamber, signal conditioning circuit, digital signal processor (DSP) control system, programmable intelligent touch screen, SD card and Bluetooth module, etc.

Based on existed gas chromatography-mass spectrometer (GC-MS) results for pecan volatiles (Zhou et al., 2012), 13 MOS sensors were preliminarily selected according to the main components in pecan volatiles such as aldehydes, ene alkanes, alcohols, acids. Table 1 lists the 13 MOS sensors and describes their main attributes. The working voltage of each sensor was 3.3 V, and the heating voltage was 5 V. All procedures were performed within the environmental temperature at range of 25 °C ± 2 °C.

To reach the standard working temperature (above 200 °C), the e-nose was preheated for 120 min before detection. In this experiment, each group of pecans was randomly divided into 30 sets, with 20 pecans in each set (about 70 g), and each set was placed in one breaker (500 mL) sealed with plastic wrap for 60 min, making sure the volatile to fill the breaker and to get equilibrium. E-nose was cleaned with zero gas (clean air) for 90s to get its reference value before each experiment. Subsequently, the odors from those pecan samples were pumped into the e-nose through a needle for the detection, each detection lasted 60 s, and one signal was recorded per second. Each measurement was followed by the recalibration of the e-nose.

2.3. Optimization of the sensor array

Each step of non-searching method selects features according to its different characteristics, so each selection process based on different statistical method needs a corresponding selection criteria. Hence, the procedure of non-searching method can be considered as selection based on multi-criteria.

Properties of features, such as divergence, replicability, and correlation, were treated as evaluation indexes for the optimization of sensor array in this paper. And the classification result of different feature matrixes, based on different values of evaluation indexes, were compared through PCA to determine the best selection criteria. First, the features with large divergence (i.e., features show well ability to discriminate different samples) were singled out through the mean analysis, because the larger divergence means the better classification performance. After the initial

Table 1
Primary sensors and their performance.

Sensor	Number	Main attribute
TGS2600	S1	Cigarette gas, Lampblack, Hydrogen, Alcohol, Methane,
TGS2602	S2	Organic vapors, Hydrogen sulfide, Formaldehyde
TGS822	S3	Alcohol, Acetone, Benzene, Hexane
TGS825	S4	Hydrogen sulfide
TGS2444	S5	Ammonia
TGS2611	S6	Methane
MQ138	S7	Aldehydes, Alcohols, Ketones
TGS2620	S8	Alcohol, Methane, Isobutane
WSP2110	S9	Benzene, Formaldehyde, Aromatic hydrocarbons
TGS826	S10	Ammonia
TGS2442	S11	Carbon monoxide
TGS813	S12	Methane, Propane, Butane
TGS816	S13	Flammable gases (Methane)