# An Efficient Electronic Nose System for Odour Analysis and Assessment

**Wentian Zhang**

Faculty of Engineering and Information Technology
University of Technology Sydney

This dissertation is submitted for the degree of
*Doctor of Philosophy*

March 2020

# Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: 2019.12.19

# Acknowledgements

I would like to express my great appreciation to my supervisor A/Prof. Steven W. Su for his continual support, guidance, help and encouragement during my PhD study. A/Prof. Steven has brought me into the topics of electronic nose, and provided an excellent research platform and brilliant insights into my research works. It is my honor to have a supervisor who always inspires me to achieve higher targets. His conscientious and meticulous attitude on research has had a significant influence on my work.

I would also like to demonstrate my sincere gratitude to my co-supervisors Prof. Shari L. Forbes (Université du Québec à Trois-Rivières, Canada), and Dr. Maiken Ueland (University of Technology Sydney, Australia), for their solid support regarding the odour analysis in terms of the chemical compound analysis field, which enables me to have a deeper understanding of my research works.

Many thanks to the other co-supervisors of mine: Dr. Rosalind X. Wang (Commonwealth Scientific and Industrial Research Organisation, Australia), and Prof. Hung Nguyen (Swinburne University of Technology), for their precious comments and suggestions on my research.

Special thanks to my colleagues in A/Prof. Steven W. Su' research group, in particular, Dr. Tao Zhang, Dr. Lin Ye, Taoping Liu, Kairui Guo, Wenhui Chen, Yao Huang and Hairong Yu, for their selfless help. Working together with them will be a good memory. I am also grateful to my friends, Dr. Ye Shi, Dr. Zhichao Sheng, Haimin Zhang, Zhiyuan Shi, Dr. Daniel Roxby, Cheng Yong, Wei Tian, for their warm help.

I also wish to express my appreciation to the staff members in the school of biomedical engineering, University of Technology Sydney.

My deepest gratitude goes to my family for their immeasurable support and encouragement.

This project was supported by the Wildlife Crime Tech Challenge (WCTC) and funded by the United States Agency for International Development (USAID).

Thanks to DATA61 for providing PhD top-up scholarship.

Thanks to UTS FEIT Data Arena Research Exhibit Grant for supporting this research.

# Abstract

An electronic nose (e-nose) is capable of identifying chemical compounds through sensing and analysing odour molecules. As a type of machine olfaction, e-nose plays a significant role in the odour analysis area and has received considerable attention from researchers all over the world. The e-nose system comprises a set of active gas sensors that detect the odour and transduce the chemical vapours into electrical signals. The odour "fingerprint" captured by the gas sensors can then be analysed and identified with pattern analysis methods, e.g., Principal Component Analysis (PCA), Cluster Analysis (CA), Support Vector Machine (SVM), and Artificial Neural Networks (ANNs). E-nose has been extensively applied in the areas of agriculture, medical diagnosis, environmental monitoring and protection, food safety, the military, cosmetics and pharmaceuticals.

In order to meet the growing demand from the global odour analysis market, a novel e-nose system, which has a high-efficiency and low-cost odour analysis, was designed and built in this dissertation through collaboration with different research areas. Firstly, inspired by the knowledge of the human olfactory system, an automated fault monitoring and alarming electronic nose (e-nose) system, named “NOS.E”, for odour detection and identification has been designed. This design is based on reliable hardware and software designs as well as an airflow intake system design which is the most significant part of NOS.E. Just as the air inhalations are important and necessary activities for the olfactory perception by controlling the airflow in the human olfactory system, the airflow control design is a crucial and essential element to guarantee the precise odour analysis procedure in the e-nose system. Different parts of the NOS.E are built together under a particular control logic, which was designed to improve the e-nose test efficiency by saving operation time. In addition, the fault detection and

alarming design generates a high-reliability performance for the e-nose by constantly monitoring the working status of the air intake system, to make sure all the actuators are working under the guidance of the proposed control logic.

A novel e-nose data pre-processing method, based on a recently developed non-parametric kernel-based modelling (KBM) approach is presented. The experimental results show that when extracting the derivative-related features from signals collected by the NOS.E, the proposed non-parametric KBM odour data pre-processing method achieves more reliable and stable pre-processing results compared with other pre-processing methods such as wavelet package correlation filter (WPCF), mean filter (MF), polynomial curve fitting (PCF) and locally weighted regression (LWR). Moreover, this dissertation also proposes a novel e-nose pattern analysis algorithm, which is a hybrid of genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM). GA was used to select the informative features and the optimal model parameters of FSVM. FSVM was used as a fitness evaluation criterion and the sequent odour classifier, which would reduce the outlier effects to provide a robust classifier which has a steady classification accuracy.

In addition, several studies were conducted with the NOS.E system. The first was to evaluate the performance of NOS.E based on data collected from different types of alcohols. A comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC-TOFMS) was used to provide the standard comparison for the evaluation in this study. The second study focused on the effectiveness of KBM data pre-processing method and FSVM odour pattern analysis method. The third study explores the potential to implement NOS.E in the biomedical engineering area, while the fourth study applied NOS.E in the wildlife protection area by rapidly identifying legal from illegal wildlife parts. As a proof-of-concept test, water buffalo horn and rhinoceros horn samples were selected as the test targets in this study.

The study results indicated the reliability and effectiveness of the developed NOS.E system. The NOS.E system is able to be applied to various applications based on the user-friendly and rapid odour analysis tests. Moreover, the NOS.E has the potential to be a universal odour analysis platform implemented in different applications.

# Publications

The contents of this thesis are based on the following papers that have been published, accepted, or submitted to peer-reviewed journals and conferences.

**Journal Papers:**

1. Zhang, Wentian, Taoping Liu, Lin Ye, Maiken Ueland, Shari L. Forbes, and Steven W. Su. "A novel data pre-processing method for odour detection and identification system." Sensors and Actuators A: Physical, 287, 113-120, 2019.

2. Zhang, Wentian, Taoping Liu, Maiken Ueland, Shari L. Forbes, Rosalind X. Wang, and Steven W. Su. "Design of an efficient electronic nose system for odour analysis and assessment." Measurement (Under review).

3. Zhang, Wentian, Taoping Liu, Steven W. Su, Shari L. Forbes, and Maiken Ueland. "Development of the Electronic Nose for Wildlife Products Identification." Forensic science international (Submitted for publication).

4. Liu, Taoping, Wentian Zhang, Peter McLean, Maiken Ueland, Shari L. Forbes, and Steven W. Su. "Electronic Nose-Based Odor Classification using Genetic Algorithms and Fuzzy Support Vector Machines." International Journal of Fuzzy Systems 20, no. 4: 1309-1320, 2018.

5. Liu, Taoping, Wentian Zhang, Lin Ye, Maiken Ueland, Shari L. Forbes, and Steven W. Su. "A novel multi-odour identification by electronic nose using non-parametric modelling-based feature extraction and time-series classification." Sensors and Actuators B: Chemical, 298, 126690,2019.

**Conference Papers:**

1. Zhang, Wentian, Taoping Liu, Miao Zhang, Yi Zhang, Huiqi Li, Maiken Ueland, Shari L. Forbes, X. Rosalind Wang, and Steven W. Su. "NOS. E: A New Fast Response Electronic Nose Health Monitoring System." In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4977-4980. IEEE, 2018.

# Table of contents

# List of figures

# List of tables

# Chapter 1

# Introduction

This chapter starts with the motivation and scope of this thesis, then introduces some related research topics, and finally shows the outline of this thesis.

## 1.1 Problem Statement

The electronic nose (e-nose) is a device that imitates the mammalian olfactory system and is designed to detect and classify different aroma mixtures. The e-nose system comprises several active gas sensors that detect the odour and transduce the chemical vapours into electrical signals [1, 2]. The odour "fingerprint" captured by the gas sensors can then be analysed and identified with pattern analysis methods, e.g., Principal Components Analysis (PCA), Cluster Analysis (CA), Support Vector Machine (SVM), and Artificial neural networks (ANNs). As a type of machine olfaction, e-nose plays a significant role in the odour analysis area and has received considerable attention from researchers all over the world [3–6]. E-nose devices have been extensively applied in the areas of agriculture, medical diagnosis, environmental monitoring and protection, food safety, the military, cosmetics and pharmaceuticals [3, 6–13].

Existing research has made significant contributions to e-nose technologies [14–21]. The studies on e-nose mainly focus on two different parts: 1. The design of a hardware system (such as sensor design and main control system design) [11, 13, 22–27]; 2. The algorithms for e-nose, such as data pre-processing methods and odour pattern

analysis methods [11, 16, 28–34]. Moreover, some researchers develop their e-nose research based on the commercially available e-nose products (such as the *fox* e-nose (*Alpha MOS, France*), the portable Cyranose 320 (*Cyrano Science, USA*), Airsense *PEN2* and *PEN3* (*Airsense Analytics GmbH, Germany*) [3, 35–41]. These studies on e-nose have made considerably advances in this area.

Generally speaking, most of the existing works have been laboratory based rather than field based [16, 18, 19]. Furthermore, the costs of current commercial e-nose products are too high for the majority of consumers [20, 37]. In addition, sensor drift, sensitivity and system reliability are the main issues which the industry are particularly concerned about in regards to implementing e-nose systems in industrial contexts[5, 42]. Considering its high efficiency and cost-effectiveness compared with other odour analysis techniques (e.g. gas chromatography and mass spectroscopy), e-nose technology has the greatest potential to be applied in different fields [43–45].

## 1.2 Motivations and Aims

Many studies are devoted to improving the sensor drift and sensitivity by using different algorithms and sensor materials [46–50]. This dissertation focuses on the application of specific instruments with designed airflow and control system to improve the reliability of the e-nose systems [5, 44]. Based on the innovative engineering technologies such as the automated air intake system and control logic, an efficient e-nose system (named "NOS.E") was developed to bring lab-based research into a low cost portable device for the use of a variety of commercial applications (e.g. meat freshness monitoring, environmental protection, health monitoring, etc.). The automated air intake control system in the NOS.E does not need the users to configure the test manually for each test. Therefore, it can improve the test efficiency by saving operation time when performing the odour analysis test multiple times. The automated test mode eliminates the incorrect manipulations caused by the manual test, thus improving the data quality and test performance [51–53]. Moreover, the additional design of the airflow control fault detection and monitoring can improve the reliability of the NOS.E system and guarantee a valid test dataset by terminating the system when the faults

occur. Therefore, NOS.E has the potential to be a universal odour analysis platform implemented in diverse applications [44].

In order to produce reliable pattern analysis results for e-nose applications, data pre-processing methods are used to improve the stability of the feature extracted from the pre-processed odour data. These pre-processing techniques mainly include wavelet package correlation filter (WPCF), mean filter (MF), polynomial curve fitting (PCF) and locally weighted regression (LWR). [54, 55]. Even though these methods are relatively mature and efficient, at times they are unable to obtain reliable results due to individual variations in the test system and unexpected responses caused by the gas interference or fluctuations of environmental parameters [51–53]. Moreover, these unexpected responses treated as noise will potentially reduce the stability and reliability of certain features. Experimental results will be significantly influenced especially for the derivative-related features which are sensitive to noise.

To seek a data pre-processing method which can overcome the drawbacks of current data pre-processing techniques for e-nose systems, this dissertation proposes a novel non-parametric kernel-based modelling (KBM) data pre-processing method. Furthermore, this method is tested by the proposed NOS.E odour detection and identification system. The NOS.E system mainly comprises an efficient power system, an automated air intake system, an interchangeable metal-oxide (MOX) gas sensor array board, and a fast data acquisition module. The target odour is drawn into the mixing chamber by the gas sampling pump, before going into the gas chamber, where the sensor array senses the odour stimulus.

The performance of the pattern analysis algorithm in the various e-nose platforms and applications are different, and there is no single pattern analysis technique that could be applied in all the e-nose applications to achieve the best performance [32]. Therefore, based on the NOS.E system, a hybrid of genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM) odour pattern analysis methods is proposed for perfumes classification. The accuracy of classification for the proposed method is evaluated with the other popular pattern analysis algorithms.

The aim of this dissertation is to present an automated e-nose system equipped with a rapid data acquisition and control system, unique air intake system, novel data

pre-processing and pattern analysis methods. The proposed e-nose system is designed as a reliable and efficient odour analysis platform which could be implemented in different applications.

As such, the research in this project has the following aims:

1. Develop an odour analysis system through the combination of the commercial gas sensors, reliable power supply system, rapid data acquisition and control system, and an unique air intake system.

2. Implement the special air flow control logic in the proposed odour analysis system for automated control purpose.

3. Design an effective test protocol which contains the configuration parameters for different user scenarios.

4. Perform the entire system integration to produce an automated odour analysis system.

5. Apply the novel kernel-based modelling data pre-processing method to provide a reliable key feature for the odour analyser.

6. Propose a hybrid of genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM) algorithm as the odour pattern analysis method for different perfumes.

7. Complete experimental subjects with different studies, collect and process data to assess the performance, reliability, ease-of-use, and effectiveness of the developed odour analysis prototype.

8. Discuss the plan of further developments and improvements to advance the prototype designs toward future commercialisation target.

## 1.3 Research Topics in the Proposed Odour Analysis System

In this dissertation, to fulfill the future commercialisation plan, several research topics have been studied to design the proposed automated odour analysis system. The topics being considered are separately discussed below.

### 1.3.1 Comprehensive electronic nose equipment design

Contrasting with the majority e-nose systems which had been studied in the research lab, this dissertation tries to design a universal odour analysis platform which contains both specific instruments, novel data processing algorithm and general test protocol designs. Therefore, except the internal high performance sensor technology, some external factors such as environmental parameters (humidity, ambient temperature and atmosphere), the presence of other interfering odours during standard test procedure, the quality of the power system, data processing circuit, accuracy in analogue to the digital data conversion process, etc. must be given special consideration to improve the performance of the odour analysis system [20, 37, 43–45].

#### Hardware design

Since this dissertation will not focus on the innovation of the specific sensor techniques, metal oxide gas sensors are chosen as the primary sensor array solution considering their high efficacy and popularity in the global market [5, 42]. Signal processing circuits are designed to remove noise from the outputs of these gas sensors [2, 3]. A reliable power system will also improve the signal quality by providing the stable and clean power rails for the whole system. Moreover, an external high performance analogue to the digital (A/D) unit will help to improve the accuracy of A/D conversion tasks, and transmit the signals rapidly to the reliable digital signal processor (DSP) for further processing. In addition, some electrical/signal isolation design will also help to improve the reliability of the e-nose system.

#### Air intake system design

Based on the mechanical and electronics designs, the performance of the gas sensors is improved by using the special air intake system which is designed to isolate the interfering odours and improve the sensitivity of gas sensors. Moreover, the feasibility and effectiveness of different reference gas sources are also take into account in the comprehensive e-nose equipment design topic [56–60].

#### Automated airflow control method

Currently, there are different test platforms that contribute to the e-nose research area, while some of them are commercial equipment, the other parts are self design and develop system. With the exception of some commercial odour analysis equipments (e.g. GC×GC, GCMS, *fox* e-nose, etc.), most of the odour analysis test platforms follow various test protocols which make it difficult to make a comparisons between test results. Furthermore, commercial odour analysis equipment is considered expensive and time consuming for the majority of researchers [20, 37]. Hence, based on the hardware and air intake system design, the unique automated airflow control method and general test protocol are studied to improve the repeatability and efficiency of odour analysis tests by providing an automated and standard test protocol to allow inter-study data comparisons.

### 1.3.2 Data Analysis Methods

While the other odour analysis techniques (e.g. GC×GC, GCMS) focus on the analysis of all chemical compounds in the target items, the e-nose based odour analysis platform is mainly designed to sense the mixture of VOCs/odour profile of the test items without identifying the specific chemical compounds [2, 3]. In contrast, other data analysis methods (e.g. pre-processing, feature selection, pattern analysis, etc.) are used to detect different items by recognising the "odour-prints" which is the unique signal matrix generated by gas sensor arrays according to different odour sources [1, 2, 32]. Considering the differences of the e-nose platforms (e.g. sensor type, dimensions of the gas chamber, design of the air intake system, etc.), in order to find the best data

analysis method based on the e-nose platform which the user tested, multi-methods need to be evaluated [2, 32]. The block diagram of the typical e-nose data analysis flow is shown in Figure 1.1. The data pre-processing methods are used to remove the noise from the e-nose signals, extract and select the useful key features. Based on these key features, the pattern analysis methods are introduced to make the machine understand the meaning of the odour signals.

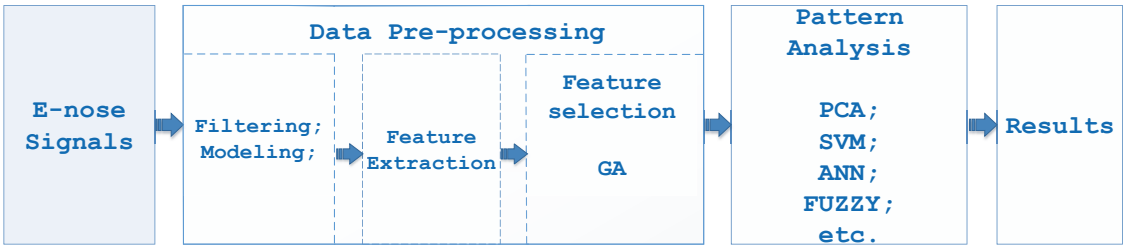


**Fig. 1.1** The block diagram of the e-nose data analysis flow.

### Data Pre-processing

In order to get reliable pattern analysis results for e-nose applications, data pre-processing methods were used to improve the stability of the feature extracted from the pre-processed odour data. These pre-processing techniques mainly include wavelet package correlation filter (WPCF), mean filter (MF), polynomial curve fitting (PCF) and locally weighted regression (LWR), etc. [54, 55]. Even though these methods are relatively mature and efficient, they are sometimes unable to obtain reliable results due to individual variations in the test system and unexpected responses caused by the gas interference or fluctuations of environmental parameters [51–53]. Moreover, these unexpected responses treated as noise will potentially reduce the stability and reliability of features. Experimental results will be significantly influenced especially for the derivative-related features which are sensitive to noise. Therefore, a new non-parametric KBM method is exploited to model the gas sensor response [61–63] and guarantee the smoothness of the filtered signals. Unlike most of the previous data pre-processing methods, in which the filtering techniques are used to remove the noise, our approach reduces the noise based on non-parametric modelling to improve the

classification accuracy; nevertheless, the identification of the model is based on the raw data.

In the odour pattern analysis stage, the features are often extracted from the data recorded until the steady-state response of gas sensors is reached. Feature extraction methods generally fall into two categories: the human-supervised extraction based on expert knowledge and the automatic feature extraction methods that are completely data-driven. In this dissertation, we have extracted the features by using a human-supervised method as the proof-of-concept process. Considering that the derivative-based features are sensitive to noise, to demonstrate the effectiveness of our proposed method in dealing with noisy signals, we chose six derivative-based features commonly used by other works [64, 65]: the response of the maximum $1^{st}$ derivative ($D_{res}$), the response of the maximum $2^{nd}$ derivative ($D_{resx}$), the response of the minimum $2^{nd}$ derivative ($D_{resn}$), time interval between gas-in and maximum $1^{st}$ derivative of response ($t_{Dres}$), the time interval between gas-in and maximum $2^{nd}$ derivative of response ($t_{Dresx}$), and the time interval between gas-in and minimum $2^{nd}$ derivative of response ($t_{Dresn}$). The diagram of these features is shown in Fig. 1.2.

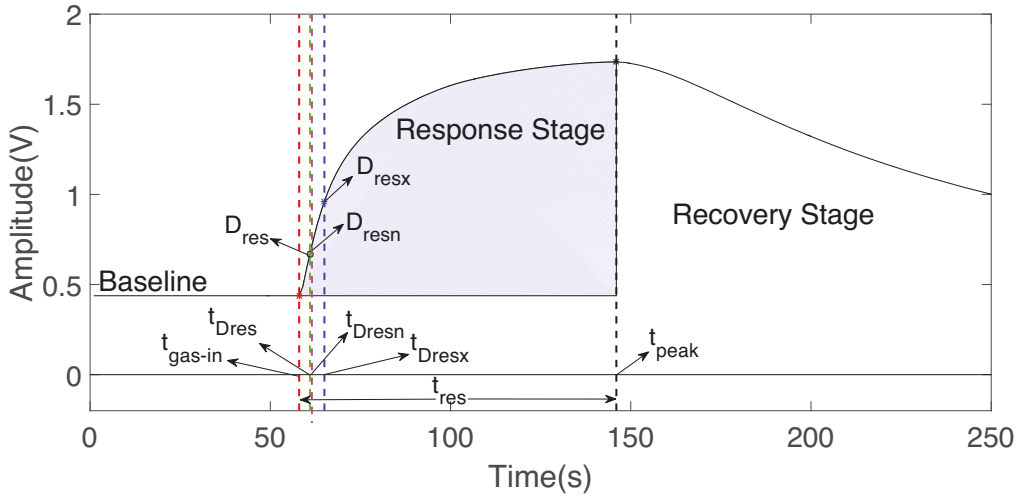


**Fig. 1.2** Features from the signal used for pattern analysis. Red line represents the time of gas-in, $t_{gas-in}$. The green line is the time point which has maximum $1^{st}$ derivative of response, $t_{Dres}$. The pink line is the time point which has minimum $2^{nd}$ derivative of response, $t_{Dresn}$. The blue line is the time point which has maximum $2^{nd}$ derivative of response, $t_{Dresx}$. The black line is the time point which has maximum value of response, $t_{peak}$.

### Pattern analysis

Usually, classification problems need to process a massive amount of data, and it is not easy to find useful information from these chaotic data unless some key information in these data is grabbed. That is why feature extraction is the most important phase in pattern analysis tasks. The pattern analysis method is used to make the machine understand the meaning of the odour datasets. Currently, there are several different pattern analysis techniques available [2], and the various analysis results are achieved based on these techniques [32]. Genetic algorithm (GA)-based techniques have an advantage over statistical methods because they are distribution-free, i.e., no prior knowledge is needed about the statistical distribution of the data. GA also automatically discovers the discriminant features for a class. Considering the outlier effects and noise issues, some fuzzy based classification algorithms such as fuzzy c-mean, fuzzy art, fuzzy NN, etc. are also implemented to the e-nose field. Fuzzy neural network has a considerable improvement in e-nose classification performance compared to a common back-propagation network [66]. Fuzzy ARTMAP has been applied to discriminate three different e-nose datasets and the performance is better than back-propagation trained multilayer perceptron (MLP) [67]. The performance of fuzzy clustering c-mean (FCM) combined with SVM e-nose classifier is better than the other well-known machine learning algorithms [68]. Fuzzy-wavelet neural network model has also been implemented to the e-nose field for food quality class evaluation and prediction [69].

In this dissertation, in addition to the application of the PCA and SVM in our e-nose studies, we also explored a novel e-nose pattern analysis method which is the combination of self-supervised GA and supervised fuzzy support vector machine (FSVM). The genetic algorithm was used to improve the classifier accuracy by optimising the feature matrix and the optimal model parameters of FSVM. FSVM was used as the odour classifier, which would reduce the outlier effects. A robust classifier is provided, which has a steady classification accuracy by applying a fuzzy membership to each input point and reformulated the SVMs. Different types of fuzzy membership can be set to solve classification problems, and it can handle class imbalance datasets problem in the presence of outliers and noise [70–72].

### 1.3.3 Industrial need

Based on researchers' achievements in the e-nose applications, the advantages of applying e-noses in the industrial field are obvious [20, 37]. Most studies have shown association with industrial partners regarding the availability of the samples and the user scenarios [5]. Hence, it is critical and essential for the e-nose researchers to pay more attention to the industrial need in future research and development activities.

It is likely that only experiments showing stable use will be a convincing factor for the industry when considering the uptake of such a device [5]. Therefore, reliability is the first factor that needs to provide high-level consideration within the design of the NOS.E odour analysis system. The reliability factor takes into account repeatability, meaning that the operation of the NOS.E should be consistently safe and result able to be repeated within a similar test environment. The reliability also requires reasonable adaptability in the various working environment.

Durability must also be considered during the entire phase of the development, because the market will abandon a product which needs frequent repairs when performing a scheduled odour analysis task. Durable design is required mainly for the purpose of maintenance, which will consist of electrical, mechanical and PCB design, and when considering the installation of parts and sensors in the sensor chamber.

Another important concept that needs to be considered is ease of use for the end users. If the system is too complicated to use, then it would be difficult to operate for the majority of potential users. Therefore, a user-friendly interface will be designed to complete all the operations related to the NOS.E odour analysis task

When people are considering commercialisation, the cost is an important factor. Balancing the cost and performance (or service) will enhance the competitive power in the global market [5, 73, 74]. In the design of the NOS.E, we start considering this problem when selecting appropriate gas sensors. A Metal Oxide Semiconductor (MOS) gas sensor was chosen for its reliable performance and its ability to sense more sensitive odours compared with other types of gas sensors. The MOS gas sensors are quite popular in the global market and thus more readily available, which will be preferable when further maintenance is required. It also means the cost of the MOS gas sensor

is lower than the other types of gas sensors. In addition to the sensor selection, the cost-performance balance concept will be embedded in the design of the other parts such as power system design, air intake system design, mechanical design, etc. [5, 74].

## 1.4 Dissertation Contribution

This dissertation presents the research and development of a novel automated odour detection and identification system with applications in health monitoring and wildlife products identification. The primary contributions of this dissertation are summarised as follows:

- Firstly, this dissertation proposed a new automated e-nose system which has the potential to be a universal odour analysis platform to allow inter-study data comparisons.

  Inspired by the knowledge of the human olfactory system, an automated fault monitoring and alarming electronic nose (e-nose) system, named ”NOS.E”, for odour detection and identification has been designed. This design is based on the reliable hardware and software designs as well as airflow intake system design which is the most significant part of the NOS.E. The hardware design contains an isolated power system; an automated air intake system (which consists of air input and exhaust ports, filters, solenoid valves, pumps, and a sensor chamber); a fault detection and alarming based airflow control feedback monitoring system; an interchangeable gas sensor array; as well as a fast data acquisition and control system. The software design is mainly based on the control logic and a remote user interface (named as "NOS.E Analyzer") which is used to control NOS.E equipment, record, store and analyse the data. Based on these designs, the NOS.E was developed to bring lab-based research into a low cost portable device for the use of a variety of commercial applications (e.g., meat freshness monitoring, environmental protection, health monitoring, etc.). The automated air intake control system can improve the test efficiency by saving operation times when performing the odour analysis test multiple times. The automated test mode can eliminate the incorrect manipulations caused by the manual test, and thus

improve the data quality and test performance. Moreover, the additional design of the airflow control fault detection and monitoring can improve the reliability of the NOS.E system and guarantee a valid test dataset by terminating the system when the faults occur. Therefore, NOS.E is able to become a universal odour analysis platform to allow inter-study data comparisons.

- Secondly, a novel non-parametric KBM odour data pre-processing method is introduced in the e-nose application as a reliable and effective data pre-processing method. This data pre-processing method can provide more reliable and stable pre-processing results compared with the other pre-processing methods.

  Unlike most of the previous data pre-processing methods, in which the filtering techniques are used to remove the noise, our approach reduces the noise based on non-parametric modelling to improve the classification accuracy; nevertheless, the identification of the model is based on the raw data. Generally, researchers use filtering based methods to process the raw data of the MOS gas sensor response. However, sometimes, the filter based pre-processing methods cannot obtain the desired effects due to the individual variation of the test system and the unexpected responses caused by the interference gas. Especially, as the filter based methods cannot always guarantee the smoothness of the filtered signals, in feature extraction stage, unacceptable outliers might be generated. In order to obtain a better result and facilitate the automated feature extraction, we adopted a non-parametric kernel-based modelling method (which applied the finite impulse response to describe the system's characteristics) as the novel e-nose data pre-processing method. Furthermore, this method is tested by three standard signals (linear signal, logarithmic signal and sigmoid signal) and the recently developed NOS.E odour detection and identification system. Then the goodness-of-fit and the Coefficient of Variation (CoV) method is employed to evaluate the performance of the non-parametric KBM odour data pre-processing method. According to the test results, when extracting derivative-related features, the proposed non-parametric KBM method provides more reliable and stable pre-processing results compared with the other pre-processing methods such as wavelet package correlation filter (WPCF), mean filter (MF),polynomial curve fitting (PCF) and locally weighted regression (LWR).

- Thirdly, a hybrid of a genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM) pattern analysis method is explored in this dissertation. This method is used to solve the uncertainty of high dimensional singularity in the feature set and handle the imbalance of datasets in the present outliers and noise problems in the e-nose area.

  In order to increase the veracity and reliability of an e-nose system, multi-sensor is widely used in odour pattern analysis. The feature vectors extracted from all chemical sensors compose a high-dimensional feature set, and this dimensionality will lower its discrimination. This problem can be addressed by using GA based feature selection algorithm, which aims to choose only the informative features. The proposed method is one kind of wrapper technique, which means that accuracy of a classifier matters in this situation. The feature set extracted from gas sensors contains noise, and many training datasets have unbalanced data. To solve these problem, the accuracy of FSVM can be regarded as an evaluation criterion. In addition to selected informative features, model parameters of FSVM can also be optimized simultaneously in genetic irritation.

  In this dissertation, we present a novel e-nose pattern analysis method which is the combination of self-supervised genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM). Genetic algorithm was used to improve the classifier accuracy by optimizing the feature matrix and the optimal model parameters of FSVM. As a robust odour classifier, FSVM is used to reduce the outlier effects by applying different types of fuzzy membership to each input point and reformulated the SVMs, and it can handle class imbalance datasets problem in the presence of outliers and noise.

## 1.5 Dissertation Outline

The outline of the dissertation is as follows:

**Chapter 1**

This chapter presents the problem statements, motivation and aims, the research topics and the outline of the dissertation.

**Chapter 2**

A brief review of the olfaction system and odour analysis platforms is firstly presented in this chapter. Then, an overview of data pre-processing methods for e-nose is introduced. Finally, an over of some popular odour pattern recognition algorithms for e-nose is provided.

**Chapter 3**

To obtain reliable experimental results, an automated odour detection and identification system is developed. We first design a proof-of-concept prototype which is based on an isolated power system, an automated air intake system (which consists of air input and exhaust ports, filters, solenoid valves, pumps, and a sensor chamber), a fault detection and alarming based airflow control feedback monitoring system, an interchangeable gas sensor array, and the fast data acquisition and control system. Through the optimisation of this design (which mainly includes the optimisation of mechanical, electrical and PCB design), we develop a new prototype which has the potential to be a universal odour analysis platform that could be applied in different user scenarios.

The work in this chapter has been submitted for publication:

- Wentian Zhang, Taoping Liu, Maiken Ueland, Shari L. Forbes, X Rosalind Wang and Steven W. Su, " Design of an efficient electronic nose system for odour analysis and assessment." Measurement. (Under Review)

**Chapter 4**

The design of data processing and analysis methods for e-nose in NOS.E systems is investigated in this chapter. The traditional data techniques mainly include wavelet package correlation filter (WPCF), mean filter (MF), polynomial curve fitting (PCF) and locally weighted regression (LWR), etc. [54, 55]. Some unexpected responses (caused by the gas interference or fluctuations in environmental parameters) which are treated as noise will potentially reduce the stability and reliability of features. Experimental results will be significantly influenced especially for the derivative-related features which are sensitive to noise. The novel non-parametric kernel-based modelling (KBM) data pre-processing method can overcome this drawback of current data pre-

processing techniques and provide a relatively smooth estimated signal response for e-nose system.

The work in this chapter has been published in:

- Zhang, Wentian, Taoping Liu, Lin Ye, Maiken Ueland, Shari L. Forbes, and Steven W. Su. "A novel data pre-processing method for odour detection and identification system." Sensors and Actuators A: Physical, 287, 113-120, 2019.

**Chapter 5**

A novel e-nose pattern analysis algorithm which is a hybrid of genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM) is studied in this chapter. GA was used to select the informative features and the optimal model parameters of FSVM. FSVM was used as fitness evaluation criterion and the sequent odour classifier, which would reduce the outlier effects to provide a robust classifier which has a steady classification accuracy. This proposed algorithm has been compared with some commonly used learning algorithms, such as support vector machine, the k-nearest neighbors and other combination algorithms. The experiment results show that the proposed odour classification algorithm can significantly improve the classification accuracy by selecting high-quality features and reach to 92.05% classification accuracy.

The work in this chapter has been published in:

- Liu, Taoping, Wentian Zhang, Peter McLean, Maiken Ueland, Shari L. Forbes, and Steven W. Su. "Electronic Nose-Based Odor Classification using Genetic Algorithms and Fuzzy Support Vector Machines." International Journal of Fuzzy Systems 20, no. 4: 1309-1320, 2018.

**Chapter 6**

The study of exhaled breath for health diagnosis and monitoring is becoming an increasingly popular area of research. Unlike most traditional health monitoring and diagnosis through the use of bodily fluids (e.g. through blood, urine), breath collection is non-invasive and convenient. Furthermore, studies have shown there are many metabolomic compounds in human breath that could be used for health monitoring

and disease diagnosis. Metabolomics are influenced by an individual's lifestyle, diet, and the environment. These factors make breath analysis an attractive option for personalised health care.

To explore the potential biomedical applications, we apply NOS.E for the rapid detection and identification of human health conditions in this chapter. By detecting the changes in the composition of an individual's respiratory gases, which have been shown to be linked to changes in metabolism, e-nose systems can be used to characterize the physical health condition. We demonstrated our system's viability with a simple dataset consisting of breath collected under three different scenarios from one volunteer. Our preliminary results show the popular classifier SVM can discriminate NOS.E's responses under the three scenarios with high performance. In future work, we will aim to gather a more varied dataset to test NOS.E's abilities.

The work in this chapter has been published in:

- Wentian Zhang, Taoping Liu, Miao Zhang, Yi Zhang, Huiqi Li, Maiken Ueland, Shari L. Forbes, X. Rosalind Wang, and Steven W. Su. "NOS. E: A New Fast Response Electronic Nose Health Monitoring System." In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4977-4980. IEEE, 2018.

**Chapter 7**

The utilisation of NOS.E in wildlife products identification is considered in this chapter.

This chapter proposed the e-nose prototype (NOS.E II) which used an efficient and reliable method to identify illegal wildlife parts. The novel mechanical and airflow designs as well as kernel based data preprocessing methods were implemented in the NOS.E II to improve its sensitivity and portability. As a proof of concept test, water buffalo horn and rhinoceros horn samples were selected as the test targets to identify legal from illegal wildlife parts. The NOS.E prototype also has potential to be used globally to differentiate a diverse range of trafficked species including large cats (e.g. tigers, leopards), elephants (including ivory), pangolins, bears, sea turtles, sharks, and a range of exotic birds and reptiles, all of which have distinct odour signatures.

The work in this chapter has been submitted for publication:

- Wentian Zhang, Taoping Liu, Steven W. Su, Shari L. Forbes, and Maiken Ueland. “Development of the Electronic Nose for Wildlife Products Identification." Forensic science international.

**Chapter 8**

This chapter summarizes the works of this PhD dissertation and presents the future research developments.

# Chapter 2

# Background

In this chapter, we first provide a brief introduction and understanding of mammalian olfactory systems (using human olfactory as an example), then we describe the structure and principle of machine olfactory systems (electronic nose). Finally, the data processing theories, as the mathematical foundation for solving the odour identification problems in this dissertation, are introduced.

## 2.1 Olfactory System in Humans and Machines

### 2.1.1 Human olfactory system

The sensitivity and range of human olfactory systems are remarkable, enabling people to detect and identify between thousands of odours [75]. Over the past decades, based on the solid understanding of the olfactory system which people achieved through the studies of molecular, physiological, genetic, and developmental biology studies. Growing interest in the biological modelling of the olfactory system has attracted many researchers to reveal two fundamental and significant topics: what is olfaction and how do human scent [75–78]?

Generally speaking, olfaction is a chemoreception that forms the sense of smell. It occurs when odorant molecules are detected by odour receptors (ORs) and are

transmitted to the brain through the olfactory bulb for further recognition. However, the actual mechanism is more complicated than we understand [75, 76, 78–82].

In air-breathing animals including humans, odorant molecules are volatile, hydrophobic, with low molecular mass (less than 300 daltons) organic compounds [2, 75, 79]. These odorant molecules are uncharged and vary widely in structure to include many chemical molecular classes, including organic acids, alcohols, aldehydes, amides, amines, aromatics, esters, ethers, halogenated hydrocarbons, hydrocarbons, ketones, nitriles, other nitrogen-containing compounds, phenols, and sulfur-containing compounds [2, 75, 79].

The functional organization of the olfactory system is shown in Figure 2.1. Odorant molecules reach the ORs via orthonasal transport through the nares when sniffing or from the oral cavity through the pharynx (back of the throat) when eating [2, 81]. Odorant molecules are captured by specific ORs located in the olfactory epithelium at the top of the nasal cavity. The olfactory information is transported through the olfactory nerves across the tiny holes in the cribriform plate of the ethmoid bone to the olfactory bulb, where the olfactory nerves make their first synapses with second-order neurons (mitral cell) in intricate spherical masses of neuropil called glomeruli. The olfactory tract is responsible for transmitting the information from the olfactory bulb to the anterior olfactory nucleus, the olfactory tubercle, the prepyriform cortex, and the amygdala, and eventually to higher brain centres that process the olfactory signals [2, 83].

Three different cells constitute (or make up) the olfactory epithelium: the supporting cell which is a type of glia cell; receptor cell which is treated as the bipolar olfactory sensory neurons with dendritic cilia projecting from their terminal ends in a thin mucus layer (10–100 $\mu M$ thick); and basal cells (like stem cells) which make new olfactory receptor cells [2, 83].

### 2.1.2 Machine olfactory system

While gas chromatography mass spectrometry instruments (GCMS) have been identified as a gold standard for odour analysis and the evolution of the more sophisticated

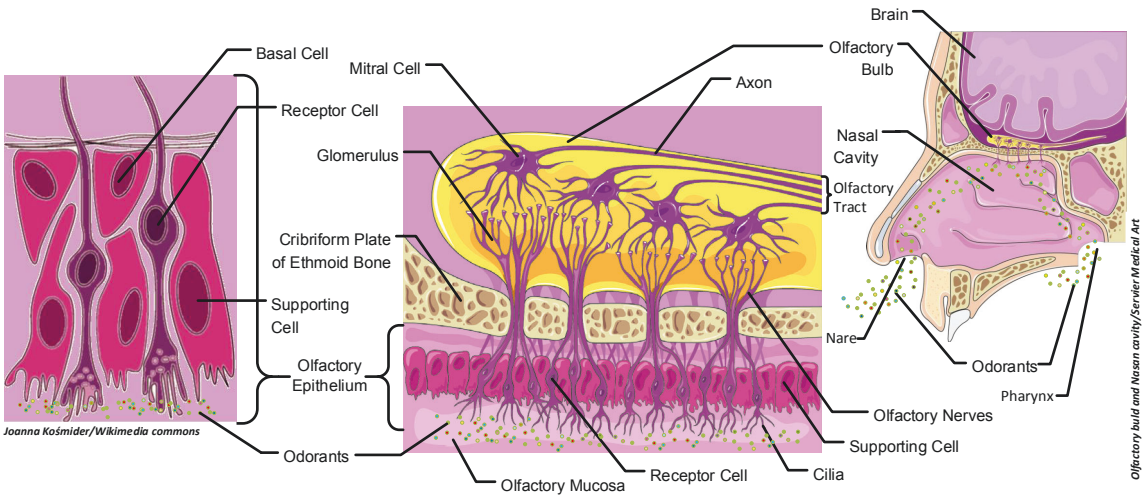


**Fig. 2.1** Structure of Olfactory System.

two dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC-TOFMS) have both become popular in the analysis of scent in recent machine olfactory system studies [84–87]. Electronic nose (e-nose) technology still draws enough attention on the global market and has the great potential to be applied in different fields for its high efficiency and cost-effectiveness compared with other machine olfactory techniques [43–45].The odour detection and identification system presented in this dissertation is mainly based on the e-nose technique.

### E-nose technique

E-nose is capable of identifying chemical compounds through sensing and analysing odour molecules. As a kind of machine olfaction, e-nose technology has the great potential to be applied in different fields [1, 3–6].

Based on the human olfactory system (Figure 2.1), and the similarities between human olfactory system and e-nose system (Figure 2.2), a sensor array is used to mimic the olfactory receptor (OR) to transform the odorants information to the electrical signals. These signals can be processed by the "olfactory bulb" of the e-nose which is a micro-controller. Moreover, to recognise the odorants information, a computer is used to represent the brain, and the pattern recognition algorithms are the "content of the brain" in this computer.

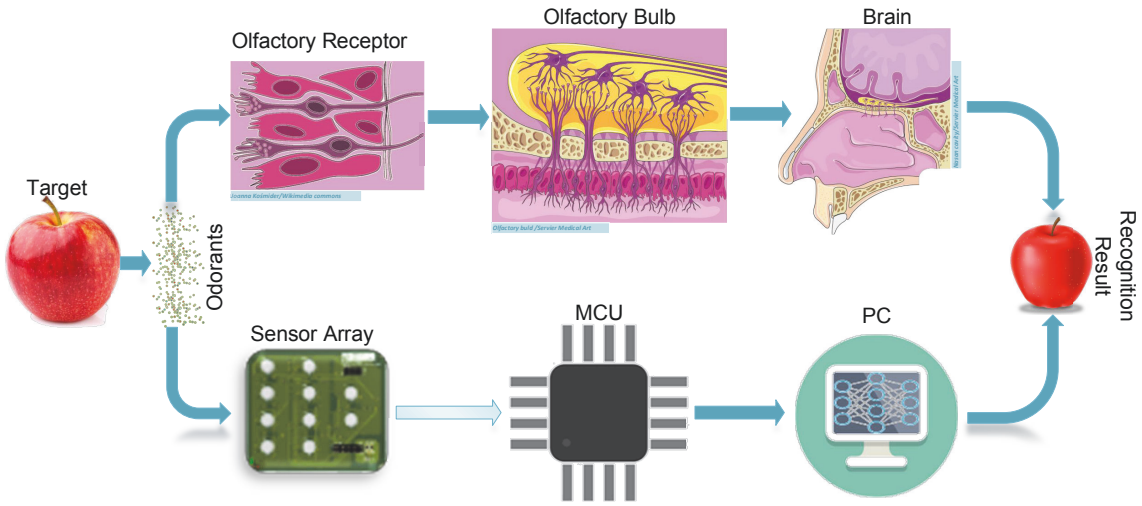


**Fig. 2.2** The similarities between human olfactory system and e-nose system.

The block diagram of a typical e-nose system is shown in Figure 2.3. According to the different applications and project requirements, we choose a particular number of gas sensors to build a sensor array. The odorant molecules of target items will be extracted into the e-nose by an air intake system, which consists of different actuators, a gas sensor array and data acquisition and control modules. The sensor array will generate specific electric signals once they are contacted by the target gas appropriately. The data acquisition system is used to collect these data and data preprocessing methods are used to process them before extracting the key features for the odour classifier. The odour classifier will take responsibility to identify and recolonize the target using appropriate pattern analysis algorithms and will send the results to the users [1–4, 14].

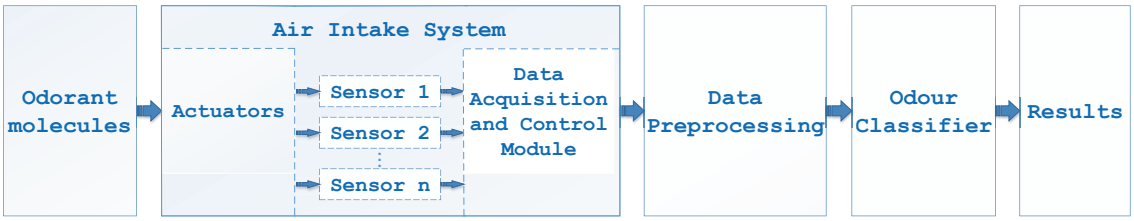


**Fig. 2.3** The block diagram of a typical e-nose system

Based on the above typical e-nose system structure, extensive studies have been carried out as highlighted below:

- Zhang, Lei, et al. [11] introduced a rapid scheme (GAT–RWLS method) which was used in the e-nose system to solve the discreteness and improve the reproducibility of MOS gas sensors. Additionally, this paper also demonstrates that one sensor's discreteness is steady and keeps unchanged when the sensor is exposed to different kinds of gas components. The performance of the e-nose system used in this paper would be better if the author redesigned the air intake system in their design (e.g. smaller gas chamber and isolated sensor array).

- Herrero, José Luis, et al. [13] designed a wireless portable e-nose prototype with a web-based application to analyse the pollutants in water. Different types of Neural Networks (Feedforward using Backpropagation learning algorithm and Radial-Basis Functions based neural network) were used in this paper and achieved a 91% and 94% success of classification based on 12 water samples with different pollutants. Although the design of the prototype can be improved (such as mechanical design), the e-nose framework proposed in this paper is quite valuable for the future e-nose works.

- Zhao, Zhenzhen, et al. [16] applied a feature-based spectrum analysis technique algorithm in optical e-nose signal processing area. This algorithm includes the primary feature extraction mechanism, and the enhanced technique with fault tolerance. It shows a valuable solution for optical e-nose by demonstrating the proposed feature extraction mechanism can extract the features from the sensing responsive spectra of the odours and can uniquely represent the odour. However, it seems that the optical e-nose system used in this paper is more complex than the general e-nose system. Perhaps this e-nose system is low portability and sophisticated to operate for the users.

- Sun, Hao, et al. [22] proposed an e-nose system which was used to detect the common bacteria in wound infection. Based on some sensor array optimisation methods (e.g. Wilks' $\Lambda$ statistic, Mahalanobis distance, principal component analysis (PCA), linear discriminant analysis (LDA), and genetic algorithm), they achieved considerable accuracy of classification (up to 96.15%) for eight different kinds of samples detected, such as culture medium, Escherichia coli, Staphylococcus aureus, and Pseudomonas aeruginosa, etc.. However, using 34

gas sensors to construct the original sensor array will raise the cost and power consumption for the e-nose system. Perhaps it would be more efficient by the surveys regarding the response characteristics of the sensors and the chemical compounds of the target items.

- Westenbrink, E., et al. [23] developed an e-nose instrument equipped with an array of 13 electrochemical and optical sensors for the detection of colorectal cancer. Based on the 92 urine samples dataset, their e-nose system can identify colorectal cancer against irritable bowel syndrome with a sensitivity of 78% and a specificity of 79% by using an (n-1) K-nearest-neighbour algorithm. However, for a diagnostically related paper, how to avoid the cross infection issues is not mentioned in this paper.

- Li, Dong, et al. [26] reported a novel headspace integrated e-nose used to discriminate Chinese medical herbs. Principal components analysis and Sammon mapping were used as the feature extraction method before applying Fisher discriminant analysis as the recognition solution. The correct classification accuracy of the headspace integrated e-nose for discriminating thirteen species of herbs is 100% which is an exciting result for e-nose research. Although temperature modulation technology was widely used to enhance the sensor selectivity, for MOS type gas sensors, perhaps this technique would affect the sensor performance due to the variable heating temperature inside these sensors.

- Zhang, Lei, et al. [28] proposed an odour recognition model which is named a cross-domain discriminative subspace learning (CDSL) for multiple electronic noses (E-noses). The authors considered the drift of the gas sensor as a cross-domain classification problem; therefore they can apply a CDSL model which is a cross-domain learning framework to solve this problem. The proposed model can guarantee that the well-trained classification method can be easily applied to different e-nose systems. Accordingly, in the future, there is a possibility that the model studied in this paper could provide some support to the commercialisation procedure of the e-nose system.

- Jing, Ya-Qi, et al. [32] developed a bioinspired neural network data processing method by applying the mechanism of the olfactory system in their e-nose

research. The primary objective of their method is to improve the efficiency of data processing in the e-nose area. Despite having only one parameter ($\mu$ which controls the speed of the slow variable current) in their neuron model, the performance of their method still shows great potential for a bioinspired approaches.

Other than the studies listed above which are based on the self-designed e-nose system, many researchers also proposed their work by using the popular commercial products which can be summarised as follows:

- Cui, Shaoqing, et al. [35] identified American ginseng and Asian ginseng based on the FOX 4000 (Alpha MOS, Toulouse, France) e-nose system. The e-nose test result is compared with GCMS test result and achieved a similar recognition result; therefore, they claim that e-nose could be used as a rapid and nondestructive method to identify the species of the ginseng.

- de Swart, Jara, et al. [36] evaluated the effect of smoking on faecal VOC composition by using Cyranose 320 (Cyrano Science, USA) e-nose equipment. The research outputs in this paper demonstrated that for the future faecal VOC profile study, the researchers should consider the smoking status of the patient.

- Yang, Wenjian, et al. [38] studied the effect of hot air drying on volatile compounds of Flammulina velutipes by using FOX 3000 (Alpha MOS, Toulouse, France) and GCMS apparatus (7890A/5975C, Agilent Technologies, Santa Clara, CA, USA). The results in this paper may provide a theoretical basis for the formation mechanism of flavour substances in dried F. velutipes.

- Xu, Lirong, et al. [40] analysed the degree of the oxidation in edible oil by using the PEN3 portable e-nose (Win Muster Airsense Analytics Inc., Schwerin, Germany). Combined with CA, PCA, and LDA pattern recognition methods, this paper claimed that the e-nose technique could be used as an alternative to the American Oil Chemists' Society (AOCS) Official Method for rapid detection of edible oil oxidation.

### Metal oxide semiconductor gas sensor

There are many different types of gas sensors which are already implemented in the e-nose area [2, 5, 3, 14, 88–95]. A brief summary is listed below:

- Metal Oxide Semiconductor (MOS) type gas sensor: [10, 96–102], etc.
- Electrochemical type gas sensor: [103–108], etc.
- Conducting polymers and non-conducting polymers type gas sensor: [109–114], etc.
- Optical type gas sensor: [115–119], etc.
- Quartz crystal microbalance type gas sensor: [120–125], etc.
- Surface Acoustic Wave (SAW) type gas sensor: [126–131], etc.
- Biology type gas sensor: [132–136], etc.

Among these sensors, the MOS type gas sensor is the most widely used and easily achieved in the market gas sensor [3]. Although the MOS type gas sensor has some drawbacks (such as high power consumption, sensitive to humidity, poor precision, etc.), it still attracts considerable attention from the e-nose researchers and manufactures for its high sensitivity, stable performance, limited sensing range, rapid response and recovery times for low mol. wt. compounds (not high), etc. [3, 14, 31, 137]. In this dissertation, considering the stable performance, low operational expenses and maintenance costs as well as easy to use characteristics, the MOS based gas sensors are adapted in the proposed odour detection and identification system.

The MOX sensor array of an e-nose is designed either to detect gases or vapors. These sensors are not tuned to a single chemical, but to detect families of chemicals. All of the sensors start out at a measured resistance, their baseline resistance. If there has been no change in the composition of the air, the outputs stay at the baseline resistance and the percent change is zero, as shown in Figure 2.4 (b). Each sensor changes its resistance, by a different amount, making a pattern of the change once the gas compound is changed as shown in Figure 2.4 (a). If a different compound had

caused the air to change, the pattern of the sensors' change would be different which is shown in Figure 2.4 (c). This creates a pattern of sensor responses, whereby the machine can be trained to recognize the target items [137, 138].

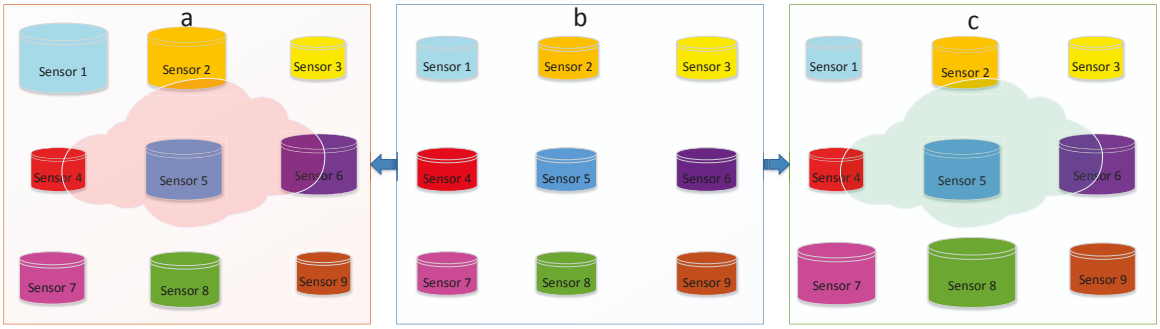


**Fig. 2.4** Gas Sensor Array operating principle

### 2.1.3 The contrast between human olfactory and Machine olfactory system

In the human olfactory system, sniffs are derived in the olfactory cortex and are rapidly modulated in an odorant-dependent fashion by a dedicated olfactometer system. Sniffs are not merely a stimulus carrier but are rather a central component of the olfactory perception [80]. Just as deviations in nasal airflow can distort olfactory perception [80, 139], so do deviations in e-nose airflow influence the odour analysis results. Like when the olfactory scene disappears when no air flows in the nasal cavity [80, 140], so does the stimulus for the sensor array disappear when the airflow system is closed in the e-nose system. Therefore, the innovative engineering technologies such as the automated air intake system and control logic could be used to improve the performance of the e-nose system. Moreover, the functional organization of the machine olfactory system is similar to the human olfactory system. Thus, to find the similarities and differences between human olfactory and machine olfactory system is highly necessary to the design of any e-nose system [2].

Firstly, olfactory perception has been treated as a "holistic" sense. That is, a mixture that contains many individual odorants can be immediately identified, holistically, as one specific entity, without having to first identify each component odorant within the mixture. For example, there are more than 100 odorants in coffee, but it could

be recognised immediately without particularising the odour of any constituent. Even roses have more than 270 odorants, yet still could be identified as rose instantly. A similar holistic approach is used to identify unique perfumes, people using some specific adjective descriptors (e.g. sweaty, herbal, spicy, floral, etc.) to classify different odours [2]. Indeed, most "odours" represent mixtures of hundreds of odorants that blend together in a recognisable fusion. This holistic approach to olfaction is believed to be a critical survival feature that was optimised during evolution to instantly distinguish complex mixtures from each other, without time-consuming preliminary "chemical analyses" of each component in the mixture [2, 79]. Similarly, the e-nose odour analysis technique also completes the recognition task based on the signal matrix, which contains the sensor responses to the stimulus from all the odourants from the test sample. Therefore, e-nose is used to rapidly identify the different odour as a "holistic" fashion, rather than precise individual chemical analyses [10].

Secondly, the relationship among the nasal airflow, the number of odorant molecules, and the olfactory response is complex. It has been reported that three primary variables are used to determine the response of the olfactory system: the number of molecules ($N$), the duration of the air inhalations ($T$), and the volume of the air inhalations($V$). These primary variables, in turn, define the three additional derived variables of concentration ($C = N/V$), delivery rate ($D = N/T$), and airflow velocity ($F = V/T$). Together, these six variables characterise the nature of the response in the olfactory system. They are not independent variables, such that an increase in $N$ will also increase $C$ and $D$, if $T$ and $V$ remain unchanged. So, if an increase in $N$ results in an increase in the olfactory response, it may also be attributable to the changes in $C$ and $D$ [79].

Since the functional organization of the e-nose system is similar to the human olfactory system, the response of the e-nose system also follows the definition discussed above. Therefore, the sensitivity of the e-nose system can be improved by reducing the dimensions of the sensor chamber. The airflow velocity and the duration of the test can also be modified according to different applications. While a lower air flow speed and longer test time was used for a low concentration target odour, a higher air flow speed and shorter test time will be applied for a high concentration target odour respectively [2, 79, 80].

In addition, while the olfactory system uses a combinatorial receptor coding scheme to encode odour identities, the e-nose system is designed based on the similar method which uses the combination of odour signal "finger-prints" generated by the gas sensor array to analyse the odour of target item [2, 3, 141]. Just like one OR recognises multiple odorants, one odorant is recognised by multiple ORs. Moreover, different odorants are recognised by different combinations of ORs. This is similar for the e-nose system,whereby one gas sensor recognises multiple odorants, one odorant is recognised by multiple gas sensors, and different combinations of gas sensors recognise different odorants. That is the basic working principle of the sensor array in an e-nose system [2, 3, 14, 141].

Moreover, there are some drawbacks for the odour classification of the human olfactory system. First, there are natural interindividual variations in the emotional properties of odours. Tags such as happy, charming, offensive, and sickening are common associations with odours, and these subjective evaluations can affect the odour analysis results [83]. Second, there are individual deviations in the olfactory perception based on genetic differentiation [2]. Third, there are individual variations in the use of odour descriptors even among qualified experts. Fourth, there are limited descriptors used to represent a large number of odour sensations [2]. For this last reason, measures of similarity rather than adjective descriptors have been used to quantify odour quality by arranging odour sensations in multidimensional spaces.

In the e-nose system, emotions will have no influence for the odour classification results. The individual differences caused by the deviations in different equipment (e.g. machining errors, data acquisition errors, the value of electronic components errors, etc.) can be controlled within the reasonable and acceptable range based on the rich engineering experiences, sophisticated mathematics methods and high-quality components [5, 42, 44]. As for the odour description, unlike the human olfactory system, based on the specific mathematical theories, the odour analysis results in the e-nose system will be calculated and then presented use some graphic patterns (e.g. radar figure, parallel coordinator graph, 3D ) or some index parameters (e.g. air quality index, mead freshness index) [2]. Therefore, the odour analysis results are much more objective compared with the descriptors used in the human olfactory system.

## 2.2 Data Pre-processing Methods for E-nose

Data pre-processing is an important technique in an e-nose system; it can improve the recognition result by guaranteeing the reliable information extracted from the pre-processed sensor responses. [1–4, 54, 142, 143]. Signal preprocessing techniques include noise reduction, baseline manipulation, normalization, feature extraction, the reduction of array inputs, and the scaling of individual sensor signals [2, 52, 142–151].

This dissertation applied a new non-parametric kernel-based modelling data pre-processing solution to improve the reliability and stability of the e-nose classifier. In addition, to optimizing the feature matrix and the optimal model parameters of fuzzy support vector machine (FSVM) based e-nose classifier, genetic algorithm (GA) was also used as a pre-processing method in this dissertation.

In this section, we first review non-parametric kernel-based modelling approach genetic algorithm method. Then describe the application of GA in the e-nose area.

### 2.2.1 Kernel based modelling method

Recently, some key mathematical tools and concepts as well as the computational aspects from machine learning techniques were introduced to the control community, e.g. reproducing kernel Hilbert spaces, kernel methods and regularisation networks, the representer theorem and the connection with the theory of Gaussian processes [62]. In particular, a method which focuses on kernel-based regularisation and its connections with reproducing kernel Hilbert spaces (RKHS) and Bayesian estimation of Gaussian processes has great potential to be applied in different fields [62].

For time-invariant linear dynamical systems the output is obtained as a convolution between the input and the system's impulse response (IR). This means that system identification is an example of an inverse problem: indeed, finding the IR from observed data is a deconvolution problem. Such problems are quite ubiquitous and appear in different applications e.g. in medicine, geophysics, and image restoration [62]. In this dissertation, this method is regarded as a novel data pre-processing method which is used to reduce the noise based on non-parametric modelling to improve the classification

accuracy. The main idea is to model the gas sensors' response based on the IRs which are found from the gas sensors' raw signal. The rebuilt response is smoother than the original sensors' response, and can therefore provide more useful information for the odour classifier.

Although the KBM method performs like a "filter", it's more accurate to be considered it as a kind of curve fitting method. Compared with classical curve fitting, dynamical modelling takes both the current input and the previous behaviour of the system into consideration, and therefore, the modelling-based approach can provide more curve structure information than curve fitting-based methods. But, for traditional dynamical modelling approaches, these models are usually parametric models (e.g. ARX model). A specified order of the system should be pre-set before estimating the parameters, that is, the complexity of the parametric model is usually bounded. In fact, it is difficult to identify the exact order when the input is a step signal and the observation is polluted with noise. For this reason, a non-parametric modelling approach, IR model, is introduced. Unlike the estimation of a parametric model, the KBM method is more flexible since the estimation of impulse response does not need to consider the order of the system. However, finding the impulse responses from the noised observation is an ill-posed inverse problem, which is insufficient for the stimulation of the system to estimate a large number of parameters of the IR model. To solve this problem, a newly developed non-parametric modelling approach [61, 152, 62] was introduced, which adds a regularisation term with a well-selected kernel function. The kernel-based regularisation helps prevent overfitting caused by the noise-polluted observations, and obtain a more accurate and robust estimation [153, 154]. In addition, this kernel-based method projects the parameters of IR into a RKHS which can reduce high-frequency components in the IR model.

Furthermore, for e-nose data, considering the efficiency factor, only the raising part of the signal response (like a step response) is commonly used for odour pattern analysis tasks; the recovery of signal response is ignored. The proposed KBM data pre-processing method is specifically designed for the step response signals, especially like e-nose data.

When $t$ with sampling time $T$ is selected as the time index. The relationship between the input ($u$) odour intake, which can be approximately treated as a step stimulation compared with the reference gas (baseline), and the output ($y$) gas sensor response can be described by a single input single output (SISO) dynamic system. Hence, the discrete time output $y$ can be calculated by the impulse response (IR) of this system by the following equation [63, 152]:

$$y(t) = \sum_{\tau=1}^{N} u(t-\tau)g(\tau) + \epsilon(t), \quad t = 1, 2, 3 \cdots, N \tag{2.1}$$

where $g(\cdot)$ represents the parameters of the finite impulse response (FIR). $q$ represents the shift operator, i.e. $qu(t) = u(t+1)$, $\epsilon(t)$ is the Gaussian white noise and N is the total sampling number. Traditionally, the parameter $g(\cdot)$ is estimated by minimize the following cost function derived from Eq. 2.1:

$$y(t) = \sum_{t=1}^{N}(y(t) - \sum_{\tau=0}^{\infty} u(t-\tau)g(\tau))^2. \tag{2.2}$$

Assuming that function $g \in \mathbb{R}^m$, then function $g$ in the regularisation term can be projected into a reproducing kernel Hilbert space (RKHS).

The IR model can be identified by minimising the cost function [63, 155]:

$$\min_{g} \left( \sum_{t=1}^{N}(y(t) - \sum_{\tau=0}^{\infty} u(t-\tau)g(\tau)[g])^2 + \gamma \|g\|_{\mathcal{H}}^2 \right). \tag{2.3}$$

### 2.2.2 Genetic algorithm

As a global search optimisation technique, Genetic Algorithm (GA) is a population-based evolutionary computation algorithm. It aims to find an "individual" with the best fitness from the searching space composed of many feasible solutions. The main idea of GA is the survival of the fittest. While applying GA to solving an optimisation problem, specific variables should be determined in advance: a population set, a fitness function, a stopping criterion and control parameters like population size, maximum iterations as well as crossover and mutation rate [156, 157].

A standard GA follows these three main steps:

- 1. Initialization: generate an initial population of individuals composed of zeros and ones.
- 2. Perform the following sub-steps iteratively until the current generation meet the stopping criterion:
  - 2.1. Evaluation: calculate the fitness of each individual in the current generation based on a predefined fitness function;
  - 2.2. Selection: a portion of existing individuals were selected to breed a new generation relying on a fitness based process, where fitter individuals are more likely to be selected:
  - 2.3. Genetic operation: breed a new population by applying the following operators:

    2.3.1. Reproduction: copy the selected individual as parent to the new population;

    Let us assume $N_{pj}$ represents the genotypic length of the $j-th$ parameter which need to be optimised, then $N_{pj}$ can be calculated by:

$$N_{pj} = round\left[log_2\left(\frac{\theta_{pj,upper} - \theta_{pj,lower} + \Delta}{\Delta_{pj}}\right)\right] + 1, \tag{2.4}$$

    where $\theta_{pj,upper}$ and $\theta_{pj,lower}$ represent the upper and lower bound of searching domain respectively. $\Delta_{pj}$ is a pre-set parameter and denotes the required precision. The genotype $g_{pj}$ for parameters $j$ should be decoded into phenotype $\theta_{pj}$ by:

$$\theta_{pj} = \theta_{pj,lower} + (\theta_{pj,upper} - \theta_{pj,lower})\left(\frac{\sum_{i=1}^{N_{pj}}(g_{pj}^{(i)})^{N_{pj}-1}}{2^{N_{pj}}}\right), \tag{2.5}$$

    where $g_{pj}^{(i)}$ denotes the value of the $i-th$ position of $g_{pj}$.

    2.3.2. Crossover: bred new offspring by recombining the chosen nodes from two parent vectors randomly;

Parts of genes between two genotypic individuals are exchanged, and the number of crossover points is decided by the number of optimisation objects.

2.3.3. Mutation: create a new offspring by randomly mutating a randomly chosen node of one selected individual.

One gene may be altered randomly. A 0-valued gene will be changed to 1 or vice versa in the followed mutation step.

- 3. Return the parameters of the tree to obtain the highest fitness as the best approximate solution.

As an effective search algorithm, GA is used to solve the optimisation problems [158, 159]. In the e-nose area, GA has also been applied for feature selection or sensor array optimisation problems [5, 22, 160]. In order to improve e-nose performance by optimizing feature selection techniques, Gardner, JW et al. introduced a novel search method procedure, V-integer genes genetic algorithms (GA), and compared with other search methods such as sequential forward or backward searches (SFS or SBS) and X-binary genes GAs. Their test results show that the V-integer genes GA approach is an accurate, and importantly, a very fast search method when compared to some other feature selection techniques [161]. After analyzing the average fingerprint spectrum and its principal component scores, Shi, Bolin et al. applied GA to select and optimize the effective sensors which can significantly contribute to identifying Xihu-Longjing tea from three producing areas and two tree species [162].

## 2.3 Pattern Analysis Techniques for E-nose

Similar to the function of the brain in the human olfactory system, the pattern analysis method is used to make the machine understand the meaning of the odour datasets collected by the e-nose. Currently, there are many different pattern analysis techniques available [2]. Some popular papers that focused on e-nose pattern analysis can be summarised as:

- Principal Components Analysis (PCA): [40, 58, 163–172], etc.

- Support Vector Machine (SVM): [10, 172–176, 59, 177–179], etc.
- Linear Discriminant Analysis (LDA),: [164, 165, 170, 171, 180–185], etc.
- Cluster Analysis (CA): [40, 186–191], etc.
- Artificial Neural Networks (ANNS): [164, 185, 187, 192–198], etc.

Except these classical e-nose pattern analysis methods, some novel odour analysis solutions are listed below:

- Zhang, Lei, et al. [28] proposed an odour recognition model which is named as cross-domain discriminative subspace learning (CDSL) for multiple electronic noses (E-noses).
- Jing, Ya-Qi, et al. [32] proposed a bioinspired neural network e-nose classification method. This neural network mimics the main structures of the mammalian olfactory system. It contains olfactory sensing neurons, mitral cells, and granule cells.
- Jiang, Xue, et al. [33] developed an efficient active learning (AL) e-nose classification algorithm which is based on improved query by committee (QBC) for RBFNN (name as EQBC-RBFNN).
- Esme, E., et al. [68] studied a hybrid e-nose classifier which is the combination of unsupervised fuzzy clustering c-mean (FCM) and supervised support vector machine (SVM).
- Luo, Hanxiao, et al. [199] presented a quantum-behaved particle swarm optimisation-based restricted Boltzmann machine (QPSO-RBM) method as a novel e-nose classifier.
- Zhang, Lei, et al. [200] applied an evolutionary cost-sensitive extreme learning machine (ECSELM) method in e-nose as a new odour classification method.
- Peng, Chao, et al. [201] presented a quantum-behaved particle swarm optimisation-based kernel extreme learning machine (QPSO-KELM) for e-nose classification.

- Zhang, Lei, et al. [202] proposed a novel multi-feature kernel semi-supervised joint learning model (MFKS), which is a semi-supervised learning approach that can be used as the classification framework for e-nose.
- He, Aixiang, et al. [203] presented a novel dictionary learning approach to improve the classification performance of the e-nose.
- Li, Ming, et al. [204] proposed a spiking neural network (SNN)-based odour recognition method for e-nose.

In this section, we first describe principal components analysis (PCA) and support vector machine (SVM) which are most popular and classical e-nose pattern analysis methods. We then briefly explore the implementation of the fuzzy support vector machine (FSVM) in e-nose research.

### 2.3.1 Principal components analysis

PCA is a conventional unsupervised pattern analysis method used as a visual tool to show groupings and separations within datasets. The mechanism of this method is to transform the odour information from the high-dimensional feature space to a low-dimensional visual space (e.g. 2D or 3D) by finding new coordinates that maximize variances of the test samples [2, 205]. The unsupervised pattern analysis feature indicates that this method is closer to the way that the human olfactory system works using intuitive associations with no, or little, prior knowledge. The visualised characteristic provides the user a high-efficiency method to view the odour data from its most informative viewpoint [2, 206].

Given $\mathbf{G}$=$(G_1, G_2, \cdots, G_m)$ as a $m$-dimension variable, $m$ is the channel numbers of the sensor array. $G_a$=$[F_1, F_2, \cdots, F_n]$, $a$=$1, \cdots, m$ represents the feature vector where $n$ is the numbers of the feature. Therefore, the covariance matrix of these features can be described as:

$$\mathbf{F} = E[GG^T] = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} \end{bmatrix}. \tag{2.6}$$

The portion of data variance accounted for the components can be expressed as [207]:

$$\kappa_{1:j} = \frac{\sum_{i-1}^{j} \lambda_i}{\sum_{i-1}^{m} \lambda_i}, \tag{2.7}$$

where $j$ is the order of components, $\lambda_i$ is the eigenvalue associated with the $i-th$ component.

Let us select $\alpha \in (0, 1)$ as the key features percent ratio with respect to the whole features, select $s$ as the first $s$ components. Then the first $s$ components key features percent ratio can be expressed as:

$$P = \kappa_1 + \kappa_2 + \cdots + \kappa_s, \tag{2.8}$$

when $s < j$, and $P \geq \alpha$, then we can use the first $s$ components to represent the original information.

### 2.3.2 Support vector machine

SVM is a practical supervised pattern analysis algorithm which is theoretically well founded and with an excellent generalisation ability [173]. As a classical pattern analysis algorithm, SVM already proved to be successful in a number of practical applications and has been used regularly in e-nose research [10, 172–176, 59, 177–179]. An SVM learns a discriminant function that separates positive and negative examples with the maximum margin. The equation of hyperplane is in the form of:

$$\omega^T x + b = 0, \tag{2.9}$$

where $\omega$ and $b$ represent the weight vector and bias respectively.

The margin $\gamma$ can be write as:

$$\gamma = \frac{2}{\|\omega\|}. \tag{2.10}$$

The objective of SVM is to find an optimal hyperplane to separate two different classes of samples. Nonnegative slack variables $\xi = (\xi_1, \xi_2, ..., \xi_M)$ are introduced to measure the misclassification degree of the training samples. Therefore, the mathematical formula of optimisation problem can be deduced by solving:

$$\min_{\mathrm{w},b}\{\tfrac{1}{2}\|\mathrm{w}\|^2 + C\sum_{i=1}^{M}\xi_i\}$$
$$s.t.\begin{cases} y_n(\omega^T x^{(i)} + b) \geq 1 - \xi_i \\ \qquad\qquad\qquad\qquad , i = 1, 2, ...M. \\ \xi_i \geq 0 \end{cases} \tag{2.11}$$

Where the parameter $C$ is an error penalty term, which determines the influence of the misclassification on the objective function. Increasing $C$ will give more importance to the errors on the training set in determining the optimal hyperplane.

In addition, kernel function $K : (x_1, x_2) \rightarrow K(x_1, x_2)$ could be applied in SVM to solve the nonlinear classification problems by transforming the original low-dimensional space into a higher-dimensional space [173, 208, 209, 155]. In this dissertation the *RBF* kernel is implemented in the SVM classifier.

### 2.3.3 Fuzzy support vector machine

SVM is an effective algorithm in dealing with classification problems, but there are still some limitations of this tool especially in classifying real-world data. Because some training samples are more meaningful than other data points. These important samples must be classified perfectly even if some noise or outliers are neglected [70]. Fuzzy SVM (FSVM) applies a fuzzy membership function to every training data $p_n$, hence the input data are transferred to fuzzy training samples, expressed as [210]:

$$\{(p_n, y_n, s_n | p_n \in \mathbb{R}^n), \sigma < s_n \leq 1\}, n = 1, ..., N, \tag{2.12}$$

where each training point $p_n$ is given a label $y_n \in \{1, -1\}$, $\sigma$ is a sufficiently small positive number. And the fuzzy membership $s_n$ is a function of time $t_n$:

$$s_n = f(t_n) = (1-\sigma)\left(\frac{t_n - t_1}{t_n = t_1}\right)^2 + \sigma, \tag{2.13}$$

where $s_n$denotes the weight of the corresponding training point towards one class and $(1-\ s_n)$ is the weight of noise or less important points. Therefore, the hyperplane optimisation problem can be defined as [70, 210]:

$$\min_{\mathrm{w},b,\xi} \frac{1}{2}\|\mathrm{w}\|^2 + cs^T\xi, \tag{2.14}$$

restrictions on condition to:

$$\begin{cases} y_n(\mathrm{w}^T p_n - b) \geq 1 - \xi_n \\ \quad \xi_n \geq 0, n = 1, ...N. \end{cases} \tag{2.15}$$

A smaller $s_n$ decreases the influence of the parameter $\xi_n$, such that the corresponding sample $p_n$ is regarded less substantial. In a similar way as SVM, the Lagrangian multiplier function can be constructed as [70]:

$$\begin{aligned} \min_{\mathrm{w},b,\xi}\max_{\alpha,\beta}\{\tfrac{1}{2}\|\mathrm{w}\|^2 + cs^T\xi - \textstyle\sum_{n=1}^{N}\alpha_n[y_n(\mathrm{w}^T x_n - b) \\ +\xi_n - 1] - \textstyle\sum_{n=1}^{N}\xi_n\beta_n\}. \end{aligned} \tag{2.16}$$

Then the optimal problem can be transferred to [70]:

$$\begin{aligned} &\max_{a}\{\textstyle\sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}(a_m a_n y_m t_n p_m^T p_n)\} \\ &s.t. \begin{cases} 0 \leq a_n \leq s_n C \\ \qquad\qquad\quad , n = 1, ...N. \\ \quad \sum_{n=1}^{N} a_n y_n \end{cases} \end{aligned} \tag{2.17}$$

The parameter $a_i$ can be solved by the Sequential Minimal Optimisation (SMO) technique, which was proposed by Haykin [211].

# Chapter 3

# The Design of NOS.E Equipment

This chapter presents an efficient electronic nose (e-nose) system, named "NOS.E", for odour analysis and assessment. In addition to the reliable hardware and software designs, an airflow intake system is implemented to ensure the precise odour analysis procedure in the NOS.E system. Additionally, a particular control logic was introduced to improve the test efficiency of the NOS.E by reducing operation time. Furthermore, the fault detection and alarming design can generate a high-reliability performance by constantly monitoring its working status. To evaluate the performance of the NOS.E, three types of alcohols were tested by the NOS.E and compared to data collected by comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC-TOFMS). The results indicate that the NOS.E can successfully distinguish three different alcohols with high efficiency and low cost and has the potential to be a universal odour analysis platform implemented in various applications.

## 3.1 Introduction

The electronic nose (e-nose) is a device that imitates the mammalian olfactory system and is designed to detect and classify different aroma mixtures. The e-nose system comprises several active gas sensors that detect the odour and transduce the chemical vapours into electrical signals [2]. Some classical pattern analysis algorithms (e.g.

Principal Components Analysis (PCA), Support Vector Machine (SVM), and Artificial neural networks (ANNs), etc.) can be used to analyse these e-nose electrical signals.

Existing research has made great contributions to e-nose technologies [14–21, 212–215]. However, most of the existing works have been laboratory based rather than field based [16, 18, 19, 214]. Additionally, the cost of current commercial e-nose products are too high for the majority of consumers [20, 37]. In addition, sensor drift, sensitivity and system reliability are the main issues which the industry are particularly concerned about in regards to implementing e-nose systems in industrial contexts [5, 42]. Considering its high efficiency and cost-effectiveness compared with other odour analysis techniques (e.g. gas chromatography and mass spectroscopy), e-nose technology has the great potential to be applied in different fields [43–45].

Many studies have been devoted to improving the sensor drift and sensitivity by using certain algorithms and sensor materials [46–50]. Inspired by the knowledge of olfactory systems [2, 80], we present an e-nose instrument with the well-designed airflow and control systems which can improve the reliability of the e-nose system [5, 44]. In the human olfactory system, air inhalations are derived in the olfactory cortex and are rapidly adjusted and controlled by a specialised olfactometer system in an odorant-dependent mode. These inhalations are not only stimulus carriers but also essential elements of the olfactory perception [80]. For instance, deviations in nasal airflow can distort olfactory perception [80, 139], and the olfactory scene disappears when no air flows in the nasal cavity [80, 140]. Likewise, in an e-nose system, different test targets will correspond to the different air flow parameters (such as duration and speed). Moreover, deviations in airflow can distort odour analysis results, and stimulus for the sensor array disappear when the airflow system is closed. Therefore, by using the innovative engineering technologies such as the automated air intake system and control logic, an automated e-nose system (named "NOS.E") was developed to bring lab-based research into a low cost portable device for the use of a variety of commercial applications (e.g. meat freshness monitoring, environmental protection, health monitoring, etc.).

The automated air intake control system in NOS.E does not need the users to configure the test manually for each test. Therefore, it can improve the test efficiency

by saving operation time when performing the odour analysis test multiple times. The automated test mode eliminates the incorrect manipulations caused by the manual test, thus improving the data quality and test performance [51–53]. Moreover, the additional design of the airflow control fault detection and monitoring can improve the reliability of the NOS.E system and guarantee a valid test dataset by terminating the system when the faults occur. Therefore, NOS.E is able to become a universal odour analysis platform to allow inter-study data comparisons [44].

In order to assess the performance of the proposed NOS.E system, we collected headspace samples from three different alcohols: Johnnie Walker Green Label Whisky (JW), Grey Goose Vodka (GG), and Jack Daniels Tennessee Whiskey (JD). The samples were analysed using comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC-TOFMS) and the NOS.E system. GC×GC-TOFMS has been identified as a gold standard for odour analysis and became the preferred method in recent studies [84, 85, 216, 217]. Using data collected and processed with the recently developed non-parametric kernel based modelling (KBM) algorithm [218] and principal component analysis (PCA), a comparison between the GC×GC-TOFMS and NOS.E for alcohol odour analysis showed that the NOS.E system could successfully identify the three different alcohols with high efficiency and low cost. Moreover, by using the popular classifier, support vector machine (SVM), we can achieve 93.33% average classification accuracy to identify these three types of alcohols.

This chapter is structured as follows. Section 3.2 is devoted to the equipment design. Reliable automated airflow control designs are developed in section 3.3. Section 3.4 provides experimental material and methods. The experimental results are analysed in section 3.5. The conclusions are drawn in section 3.6.

## 3.2 NOS.E Equipment Design

### 3.2.1 Hardware design

Two NOS.E hardware platforms (Fig. 3.1 and Fig. 3.1) were built for different applications. Both these platforms consist of the power supply module, the sensor module, the driver module, the communication module, and the microcontroller (MCU). The power module (Fig. 3.3) provides different isolated power rails for the other parts of the equipment to enable the NOS.E system to work with the driver circuits (Fig. 3.4) to satisfy the proposed automated airflow control logic. The data processing in this chapter is based on the NOS.E II prototype. The NOS.E I related works will be demonstrated in Chapter 4, Chapter 5, and Chapter 6.

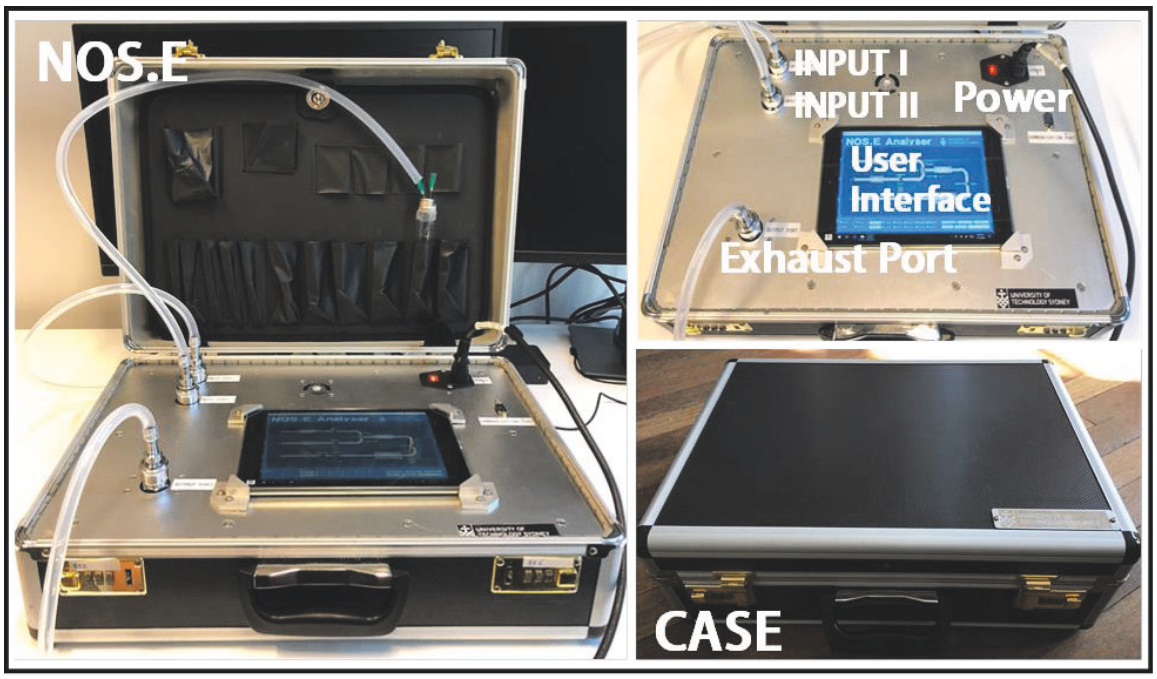


**Fig. 3.1** NOS.E I Equipment.

The design of the NOS.E power system is shown in Fig. 3.3. The $100-240VAC$ power source is converted to $20VDC$ which will be converted as isolated $VDD\_5V$ and $VCC\_12V$ power rails. The $VDD\_5V$ provides different power rails ($VDD\_3V3$ and $VDD\_1V9$) for the MCU. The $VAA\_5V$ (transferred from $VDD\_5V$) is used to provide the power for sensors as well as to generate the reference voltage for the

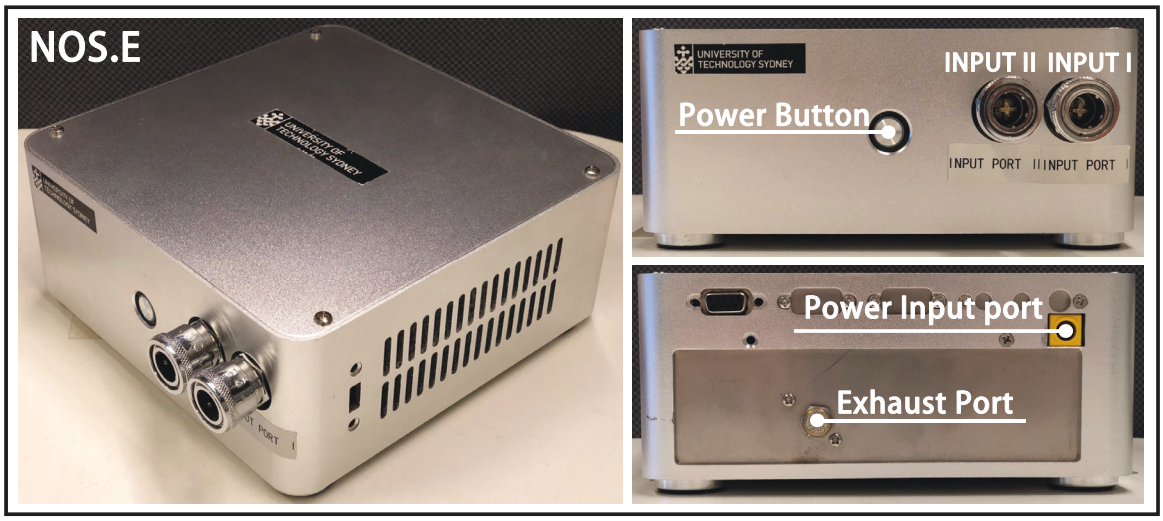


**Fig. 3.2** NOS.E II Equipment.

analogue to digital unit. The power for the actuators related circuits is provided by $VCC\_12V$.

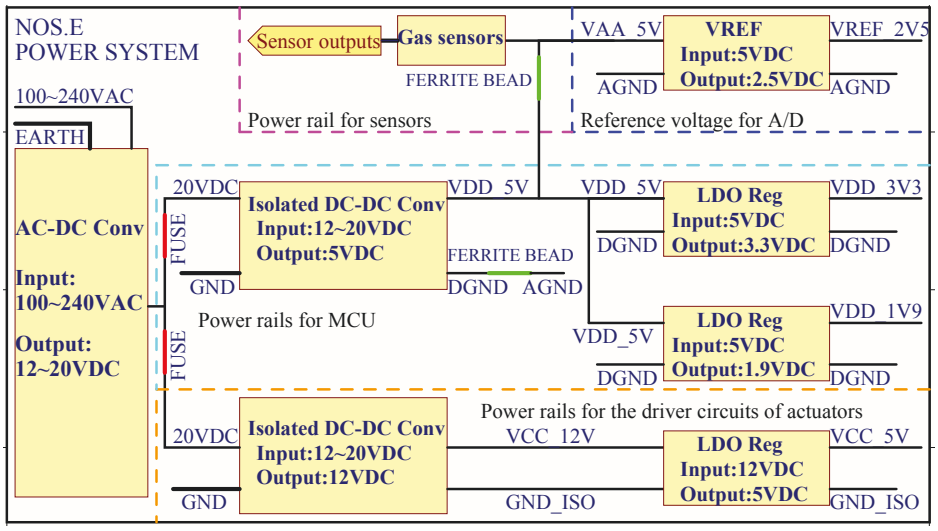


**Fig. 3.3** NOS.E Power System Diagram.

Fig. 3.4 shows the circuit used to control the different actuators. When this circuit receives a high logic level control signal from MCU, the base current is switched on $Q1$ and the forward current $I_F$ is flowed through the emitting diode of $U1$, $R_3$ and the collector-emitter of $Q1$. Inside the $U1$, an integrated photodetector detects the light generated by the emitting diode. According to the current transfer ratio ($CTR$) of $U1$ (Equation 3.1), the output current flows through $R_4$ and generates the Gate-Source

voltage ($V_{GS}$) to turn on $Q2$ to let $VCC\_12V$ provide the power to the actuator. Therefore, a high logic level signal is achieved through the Status port. If the Control port sends a low logic level signal to this circuit, likewise, a low logic level signal is collected through the Status port.

$$I_O = I_F * CTR, \tag{3.1}$$

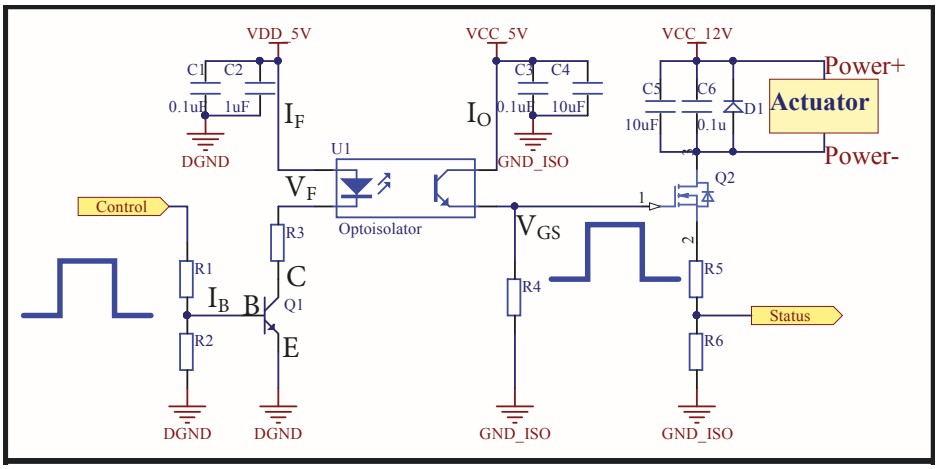


**Fig. 3.4** NOS.E Schematic of Driver Circuits. $R$ represents the resistor; $C$ represents capacitor; $Q1$ is NPN switching transistors; $U1$ is high speed optocoupler; $Q2$ is N-Channel MOSFET.

The working statuses of different actuators are monitored by the circuit shown in Fig. 3.5. While the high logic level Status signal represents that the actuator is opened, the low logic level Status signal means that the actuator is closed. The monitoring signal "Status" is isolated by the high-speed optocoupler, then the isolated signal (actuator working status) is sent to a buffer before being collected by the MCU.

The interchangeable sensor array is built with eight commercially available metal oxide gas sensors: *TGS 2620, TGS 2602, TGS 2600,TGS 2600, TGS 2603, TGS 2610D, TGS 2611E*, and *TGS 2612* (FIGARO ENGINEERING INC, Mino, Osaka JAPAN). The target input gases stimulate the gas sensors to generate specific voltage outputs. An analogue to digital converter is used to convert these voltage outputs to digital signals which are sent to the MCU for data processing. The MCU completes the tasks such as data acquisition, actuator control, data pre-processing and communicating with the host computer.

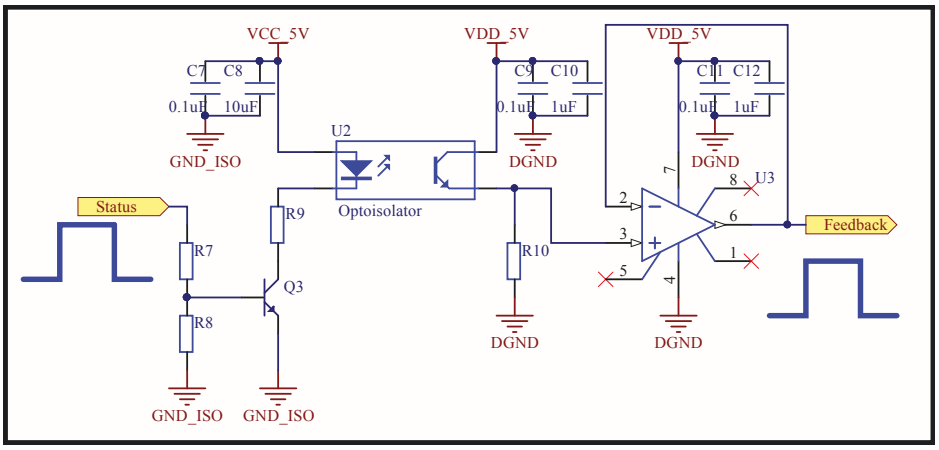


**Fig. 3.5** NOS.E Schematic of Automated Control Monitoring Circuits. *R* represents the resistor; *C* represents capacitor; *U*2 is high speed optocoupler; *U*3 is Operational Amplifier.

### 3.2.2 Software design

The software architecture (Fig. 3.6) of NOS.E is based on two main parts: The first part is the host computer software which is designed as the user interface, and the second part is the slave computer software, which is located inside the NOS.E instrument.

The host computer is designed as the user interface (see Fig. 3.7) which is used to configure the NOS.E equipment. This interface can send the commands to control the different actuators and show the real-time working status of these actuators by receiving the fault detection and alarming feedbacks from the slave computer. The user interface is also used to record, store, and analyse the data by using related widgets and commands. The output wave-forms of gas sensors is displayed to the user in real time. The collected odour data is stored on the local disk in the text format. Key features of the odour data are extracted and shown to the user in graph.

The slave computer can collect the odour data through an analogue to digital module and the data acquisition program embedded in the MCU, and transmit the odour data to the host computer through WIFI module. The WIFI module uses the universal asynchronous receiver transmitter (UART) interface to communicate with the MCU. The slave computer can also communicate with the host computer to manage all the actuators by using control commands and work status feedbacks.

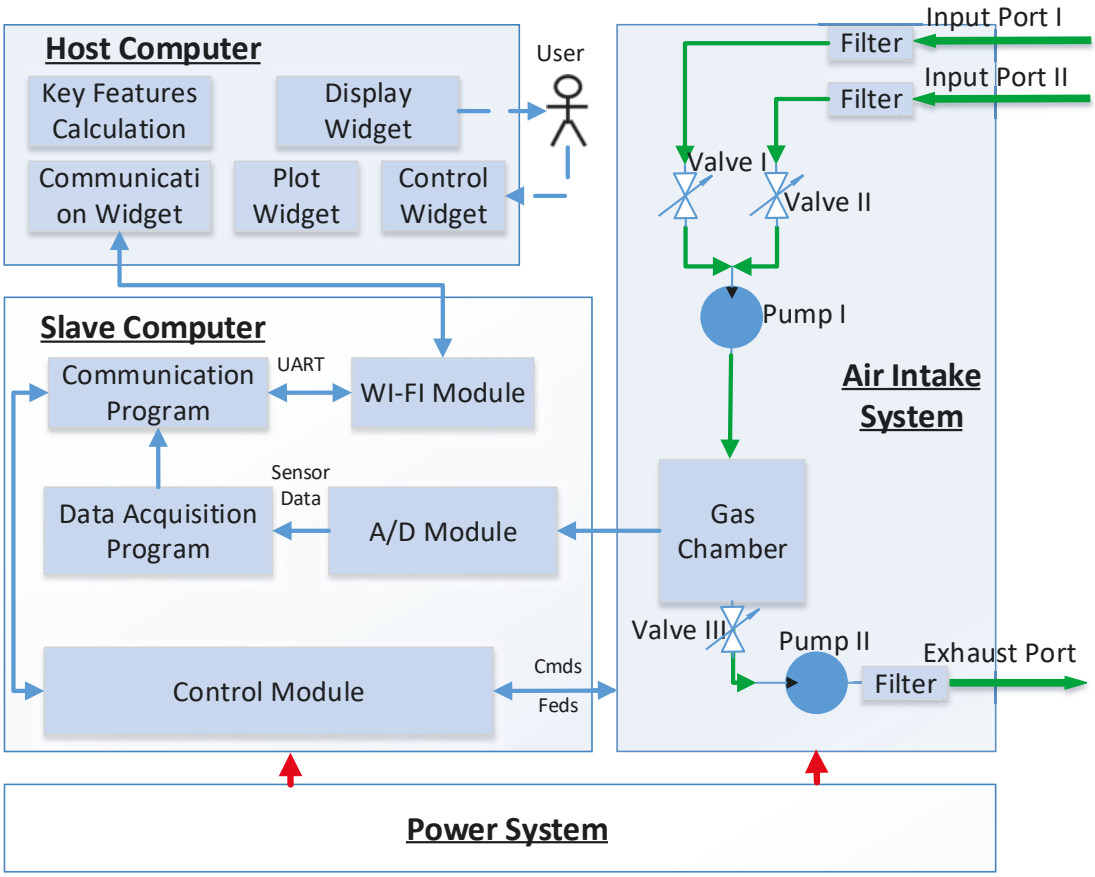


**Fig. 3.6** Block Diagram of NOS.E Software Architecture. Dashed blue line represents the actions related with users; Solid blue line represents the signal chain; Solid green line represents the airflow chain; Solid red line represents the power chain.

## 3.3 Reliable Automated Airflow Control Design for the NOS.E

The airflow control design is a critical element in the e-nose system, analogous to the important and necessary activities of air inhalations that control the airflow in the human olfactory system [80]. An e-nose's air intake system contains the mixing chamber, the gas chamber, the driver circuits, and the actuators (the sampling pumps and the solenoid valves). The primary function of the air intake system in the NOS.E is to ensure the odour tests are executed routinely and automatically by controlling the actuators and work with the MCU to implement the automated control logic.

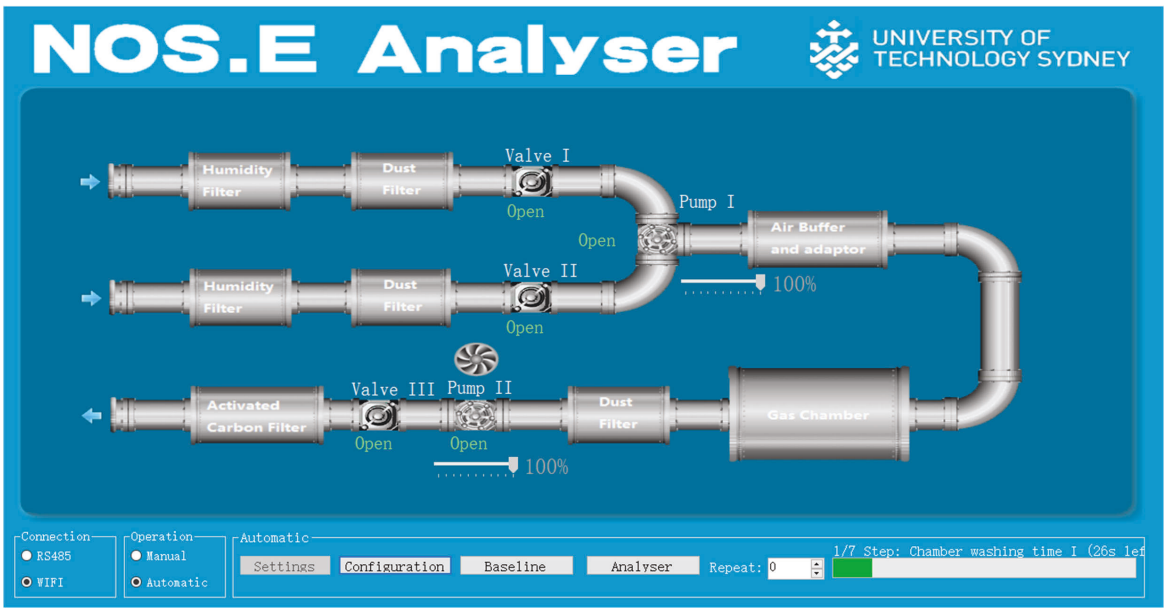


**Fig. 3.7** Control and Configuration Panel.

### 3.3.1 Automated airflow control phase

NOS.E's airflow control logic reduces the test errors caused by manual initialisation and configuration of the test system. Eight phases have been designed in the NOS.E control logic. Each working phase has different airflow parameters, which are fixed based on the manual training mode, then preprogrammed into the user interface. Moreover, these airflow parameters are also varied for different applications. Based on the target application selected by the users, the controller will send different commands to control the working status of the actuators. Table 3.1 shows the specific status of the different actuators which are related to the proposed automated control logic. In this table, $V_1, V_2, V_3, P_1$, and $P_2$ represent the control signal of valve I, valve II, valve III, pump I and pump II. $T_1, T_2$ and $T_3$ are the control phase ID signals sent by the controller. In addition, $H$ indicates the high logic level signal, and $L$ indicates the low logic level signal.

#### Phase I: Equipment warming up and chamber washing

To warm up the equipment and wash the chamber, the first step involves opening the solenoid valve I, solenoid valve III, pump I and pump II; and closing valve II. These

**Table 3.1** The Automated Control Logic for NOS.E

| Phase | $V_1$ | $V_2$ | $V_3$ | $P_1$ | $P_1$ | $T_1$ | $T_2$ | $T_3$ |
|---|---|---|---|---|---|---|---|---|
| I | H | L | H | H | H | L | L | L |
| II | L | L | H | L | H | L | L | H |
| PAUSE | L | L | L | L | L | L | H | L |
| III | H | L | H | H | H | L | H | H |
| IV | L | L | H | L | H | H | L | L |
| V | L | H | H | H | H | H | L | H |
| VI | H | L | H | H | H | H | H | L |
| VII | H | L | H | H | H | H | H | H |

**Note**: H: High logic level signal; L: Low logic level signal.

actuators will stay at this working status for 300 seconds before closing Valve I and Pump I. Then the instrument automatically switches to the next phase. The logic function of phase I is shown in Equation 3.2.

$$Y_I = V_1\bar{V}_2V_3P_1P_2\bar{T}_1\bar{T}_2\bar{T}_3. \tag{3.2}$$

### Phase II: Vacuuming I

To reduce the response time of the airflow system, a negative pressure environment is created by closing solenoid valve I and pump I, and opening the solenoid valve III, and pump II for 10 seconds. This negative pressure will enable the input gas contact with the gas sensor immediately once the valve I or valve II, valve III and pump I are opened. The logic function of phase II is shown in Equation 3.3.

$$Y_{II} = \bar{V}_1\bar{V}_2V_3\bar{P}_1P_2\bar{T}_1\bar{T}_2T_3. \tag{3.3}$$

### Phase PAUSE

After the vacuuming I phase, all the solenoid valves and pumps are closed for 10 seconds to let the user know the equipment is ready to collect data. At this point, NOS.E also reminds the user to connect the test sample to the sample input port. The logic function of the pause phase is shown in Equation 3.4.

$$Y_P = \bar{V}_1\bar{V}_2\bar{V}_3\bar{P}_1\bar{P}_2\bar{T}_1T_2\bar{T}_3. \tag{3.4}$$

**Phase III: Baseline setup**

To setup the baseline for the sensor array, NOS.E opens solenoid valve I, solenoid valve III, pump I, and pump II for 20 seconds. The logic function of phase III is shown in Equation 3.5.

$$Y_{III} = V_1\bar{V}_2V_3P_1P_2\bar{T}_1T_2T_3. \tag{3.5}$$

**Phase IV: Vacuuming II**

Before the testing phase, a negative pressure environment is created to reduce the response time of the airflow system. This step involves opening the solenoid valve III, and pump II; and closing the solenoid valve II and pump I. The status of these actuators will keep for 10 seconds. The logic function of phase IV is shown in Equation 3.6.

$$Y_{IV} = \bar{V}_1\bar{V}_2V_3\bar{P}_1P_2T_1\bar{T}_2\bar{T}_3. \tag{3.6}$$

**Phase V: Testing**

During the testing phase, NOS.E opens all actuators except the solenoid valve I. The targeted gas is quickly extracted into the airflow system, and starts to collect data for another 90 seconds. The logic function of phase V is shown in Equation 3.7.

$$Y_V = \bar{V}_1V_2V_3P_1P_2T_1\bar{T}_2T_3. \tag{3.7}$$

**Phase VI: Baseline Recovery**

To exhaust the target gas after the testing phase, NOS.E opens all actuators except the solenoid valve II for 90 seconds. The fresh air moves through the airflow system, as the concentration of the targeted gas decreases, the sensor array response returns toward the baseline. The odour data is saved in the local folder after this phase. The logic function of phase VI is shown in Equation 3.8.

$$Y_{VI} = V_1\bar{V}_2V_3P_1P_2T_1T_2\bar{T}_3. \tag{3.8}$$

**Phase VII: Chamber washing**

To clean the air intake system, NOS.E opens all actuators except the solenoid valve I. The fresh air is moves through the system for 300 seconds, then NOS.E sends notice to the user that the system will start the next round test automatically within 5 seconds. The logic function of phase VII is shown in Equation 3.9.

$$Y_{VII} = V_1\bar{V}_2V_3P_1P_2T_1T_2T_3. \tag{3.9}$$

### 3.3.2 Fault detection for airflow control

An incorrect airflow control logic will lead to the inappropriate working status of the actuators as well as the incorrect test data. There are certain risks that the test will fail when the users are performing the odour analysis tests without monitoring the equipment working status. Therefore, we designed an automated airflow control and fault detection system (shown in Figure 3.8) to improve the reliability of the NOS.E system. The actuators are controlled by the commands ($V_1, V_2, V_3, P_1$, and $P_2$) which come from the controller. According to the working status of actuators, some feedback signals ($V_1', V_2', V_3', P_1'$, and $P_2'$) are collected by the monitoring circuits (Figure 3.5). The control phase ID signals ($T_1, T_2$, and $T_3$) are combined with the control commands and feedback signals by using the shifting register before monitoring the feedback results. The feedback output is sent to the controller, and if the actuators' working status and phase ID are correct, the controller will send the next round of control commands; otherwise, if any unexpected errors occur, the controller will terminate the system then sends the specific fault and an alarm will inform the users [219–221].

The error of fault detection for automated airflow control system $f_N(k)$ can be calculated by the following equation:

$$f_N(k) = \sum_{i=1}^{n}[Y_N(k-i) - Y_N'(k-i)]^2, \tag{3.10}$$

where $N \in \{I, II, P, III, IV, \cdots VII\}$ is the current test phase, $n$ is the duration of the current working phase, $i$ is the length of the monitoring window, $Y_N(k-i)$ represent actuators' control signals sent by the controller, $Y_N'(k-i)$ indicates working status

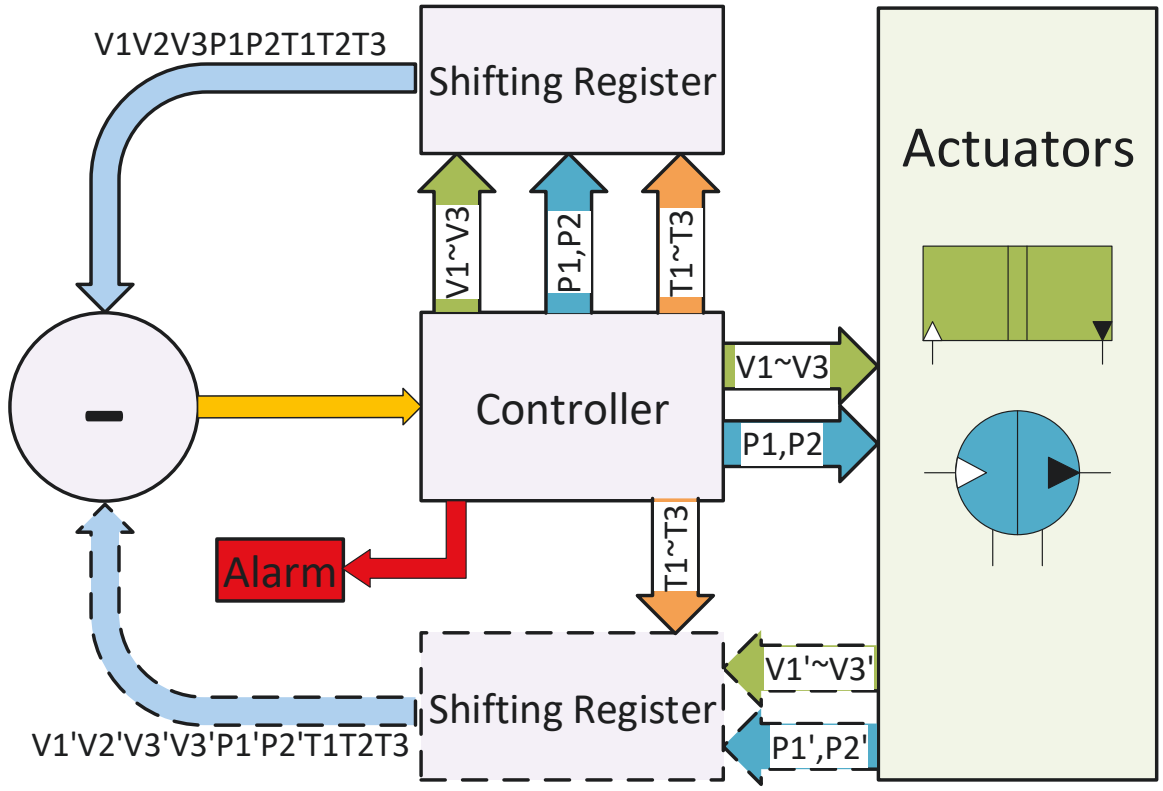


**Fig. 3.8** Diagram of NOS.E Automated AirFlow Control and Fault Detection Design. $V_1, V_2, V_3, P_1$, and $P_2$ are the valve and pump control signals sent by controller; $T_1, T_2$ and $T_3$ are the control phase ID signals sent by controller; $V_1', V_2', V_3', P_1'$, and $P_2'$ are the working status feedback signals collected from the solenoid valves and pumps.

feedback signals collected from the actuators.

Since there are eight working phase in the NOS.E automated control system, $\boldsymbol{f}_N(k)$ can also be written as:

$$\boldsymbol{f}(k) = [f_I(k), f_{II}(k), f_P(k), f_{III}(k), \cdots, f_{VII}(k)]^T. \tag{3.11}$$

$\boldsymbol{Y}(k)$ is a vector representation of the control signals in different working phases:

$$\boldsymbol{Y}(k) = [Y_I(k), Y_{II}(k), Y_P(k), Y_{III}(k), \cdots, Y_{VII}(k)]^T, \tag{3.12}$$

As in the above, $\boldsymbol{Y}'(k)$ is a vector representation of the working status feedback signals collected from the actuators in each working phase:

$$\boldsymbol{Y}'(k) = [Y'_I(k), Y'_{II}(k), Y'_P(k), Y'_{III}(k), \cdots, Y'_{VII}(k)]^T. \tag{3.13}$$

As found in Table 3.1, all the control and phase ID signals should match with the designed order and logic level. The NOS.E system cannot work appropriately if just one of the signals' logic level is mis-matched with the proposed control logic. The fault detection threshold (Equation 3.10) is set at 1, i.e. $f_N(k) = 1$. Once the calculated error is equal to 1, the controller will terminate the NOS.E system and send fault information and alarm the user.

## 3.4 Material and Methods for the NOS.E Reliability Tests

### 3.4.1 Automated airflow control and fault detection system

We first test the stability and reliability of NOS.E before using it on alcohol samples. This is done over a period of three days when the complete sequence of eight working phases was automatically run 90 times (30 times per day). The NOS.E user interface (see Fig. 3.7) was used to monitor the status of the different actuators throughout this period. The presence of green text "open" under an actuator means it is working, whereas the red text “closed” below an actuator indicates this actuator has stopped working and has closed. In addition, faults simulation tests were also performed by manually turning off/on the power of different actuators when the NOS.E system was running. The results were monitored by the NOS.E user interface.

### 3.4.2 Alcohol samples

In this dissertation, 24 samples were collected from three different alcohols bought from BWS, (Haymarket, Sydney, Australia) ($N_{JW} = 8$, $N_{GG} = 8$, $N_{JD} = 8$). Each sample (5mL) was put into 20 mL vials, which were sealed airtight with a screw cap containing a 1.3 mm thick polytetrafluoroethylene/silicone septum (Sigma-Aldrich, Castle Hill,

NSW, Australia). These samples were separated into two groups. Fifteen samples ($N_{JW} = 5$, $N_{GG} = 5$, $N_{JD} = 5$) were analysed by NOS.E in two consecutive days, the first three samples were analysed on day 1, and the other two samples were analysed on day 2. Another identical nine samples were tested by the GC×GC-TOFMS within one day.

### 3.4.3 GC×GC-TOFMS

Sample collections of the different alcohols were completed via headspace sampling via headspace sampling using a 50/30 $\mu m$ divinylbenzene/ carboxen/ polydimethylsiloxane (DVB/CAR/PDMS) 24 Ga Stableflex solid-phase microextraction (SPME) fibre and manual fibre holder (Supelco, Bellefonte, PA, USA). According to the manufacturer's recommendations, the fibre was initially conditioned for 60 min at 270 ℃ before its first use. The DVB/CAR/PDMS fibre was exposed to the sample headspace and the VOCs were allowed to adsorb onto the fibre for 60 min. The fibre was thermally desorbed for 5 min at 250 ℃ into the inlet of the GC×GC-TOFMS and carried through the column via a helium carrier gas for separation and analysis. The parameters of the oven were configured as follows; initial temperature of 35 ℃ held for 5 min, then increased to 240 ℃ at a rate of 5 ℃ per min and held for another 5 min. The secondary oven temperature offset was 10 ℃. The mass range observed was 29 to 450 amu at an acquisition rate of 100 spectra per second for 1000 seconds with an additional acquisition delay of 120 seconds. The total run time of GC×GC-TOFMS for each sample is 51 min. Fibre blanks were run in-between samples to ensure no cross contamination occurred between samples.

### 3.4.4 NOS.E

The NOS.E equipment was warmed up for 30 min by powering and opening all the actuators. The NOS.E Analyser was used to complete the configuration of the test time for each operation phase. (Phase I: 300 seconds; Phase II: 10 seconds; PAUSE: 10 seconds; Phase III: 20 seconds; Phase IV: 10 seconds; Phase V: 90 seconds; Phase VI: 90 seconds; Phase VII: 300 seconds). Repeat time, which is applied as 5 in this

dissertation, is used to setup the test rounds to let the NOS.E complete the data collection automatically. The reference gas (fresh air) was connected to the input port I of the NOS.E equipment. Two sampling needles were inserted in the vial's septum, while one needle was connected to the reference gas, another needle was used to connect the sample vial to the input port II, to allow the headspace samples to be analysed. After preparation, the automated NOS.E odour analysis tests were run through the NOS.E Analyser. The NOS.E system saves the test datasets and stops running automatically once three tests are completed. It took 830 seconds to run each NOS.E test.

### 3.4.5 Data processing

Principal component analysis (PCA) is a conventional unsupervised pattern analysis method used as a visual tool to show groupings and separations within datasets. The mechanism of this method is to transform the odour information from the high-dimensional feature space to a low-dimensional visual space (e.g. 2D or 3D) by finding new coordinates that maximize variances of the test samples [2, 205]. The visualised characteristic provides users with a highly efficient method to view the odour data from its most informative viewpoint [2, 206]. Therefore, the GC×GC-TOFMS chromatographic datasets and NOS.E key features datasets (which were extracted from the voltage signals collected from three different alcohols) were tested and compared using PCA techniques to evaluate the performance of the NOS.E system. Additionally, the most popular classification algorithm, SVM, was applied to verify the performance of the NOS.E system.

## 3.5 Results and Discussion

### 3.5.1 Automated airflow control and fault detection system

During the automated airflow control test where the full sequence of eight control phases were run, no fault alarms were triggered (results not shown), indicating the system ran exactly as configured. The automated operation mode was able to improve

the test efficiency by saving the user considerable time. Moreover, the automated test mode can eliminate the incorrect manipulations caused by manual testing, and thus improve the data quality and test performance. In addition, the faults simulation test results listed in Table. 3.2 showed that this design can timely terminate the NOS.E system and show the fault information. In this table, row $N/A$ represents the correct control commands and monitoring status for all the actuators in each phase. Row $V_1'$ to $P_2'$ indicates the abnormal working status (which is bold and underlined in each row) monitored by the MCU. This fault detection and monitoring design can improve the reliability of the NOS.E system and guarantee a valid test dataset by terminating the system when the faults occur.

**Table 3.2** The Simulation Test Results of NOS.E Fault Detection Design

| Phase | Fault Status | $V_1$ | $V_2$ | $V_3$ | $P_1$ | $P_1$ | NOS.E |
|---|---|---|---|---|---|---|---|
| I | $N/A$ | H | L | H | H | H | Running |
| | $V_1'$ | **L** | L | H | H | H | Terminated |
| | $V_2'$ | H | **H** | H | H | H | Terminated |
| | $V_3'$ | H | L | **L** | H | H | Terminated |
| | $P_1'$ | H | L | H | **L** | H | Terminated |
| | $P_2'$ | H | L | H | H | **L** | Terminated |
| II | $N/A$ | L | L | H | L | H | Running |
| | $V_1'$ | **H** | L | H | L | H | Terminated |
| | $V_2'$ | L | **H** | H | L | H | Terminated |
| | $V_3'$ | L | L | **L** | L | H | Terminated |
| | $P_1'$ | L | L | H | **H** | H | Terminated |
| | $P_2'$ | L | L | H | L | **L** | Terminated |
| P | $N/A$ | L | L | L | L | L | Running |
| | $V_1'$ | **H** | L | L | L | L | Terminated |
| | $V_2'$ | L | **H** | L | L | L | Terminated |
| | $V_3'$ | L | L | **H** | L | L | Terminated |
| | $P_1'$ | L | L | L | **H** | H | Terminated |
| | $P_2'$ | L | L | L | L | **H** | Terminated |
| III | $N/A$ | H | L | H | H | H | Running |
| | $V_1'$ | **L** | L | H | H | H | Terminated |
| | $V_2'$ | H | **H** | H | H | H | Terminated |
| | $V_3'$ | H | L | **L** | H | H | Terminated |
| | $P_1'$ | H | L | H | **L** | H | Terminated |
| | $P_2'$ | H | L | H | H | **L** | Terminated |
| IV | $N/A$ | L | L | H | L | H | Running |
| | $V_1'$ | **H** | L | H | L | H | Terminated |
| | $V_2'$ | L | **H** | H | L | H | Terminated |
| | $V_3'$ | L | L | **L** | L | H | Terminated |
| | $P_1'$ | L | L | H | **H** | H | Terminated |
| | $P_2'$ | L | L | H | L | **L** | Terminated |
| V | $N/A$ | L | H | H | H | H | Running |
| | $V_1'$ | **H** | H | H | H | H | Terminated |
| | $V_2'$ | L | **L** | H | H | H | Terminated |
| | $V_3'$ | L | H | **L** | H | H | Terminated |
| | $P_1'$ | L | H | H | **L** | H | Terminated |
| | $P_2'$ | L | H | H | H | **L** | Terminated |
| VI | $N/A$ | H | L | H | H | H | Running |
| | $V_1'$ | **L** | L | H | H | H | Terminated |
| | $V_2'$ | H | **H** | H | H | H | Terminated |
| | $V_3'$ | H | L | **L** | H | H | Terminated |
| | $P_1'$ | H | L | H | **L** | H | Terminated |
| | $P_2'$ | H | L | H | H | **L** | Terminated |
| VII | $N/A$ | H | L | H | H | H | Running |
| | $V_1'$ | **L** | L | H | H | H | Terminated |
| | $V_2'$ | H | **H** | H | H | H | Terminated |
| | $V_3'$ | H | L | **L** | H | H | Terminated |
| | $P_1'$ | H | L | H | **L** | H | Terminated |
| | $P_2'$ | H | L | H | H | **L** | Terminated |

**Note**: H: High logic level signal; L: Low logic level signal.

### 3.5.2 Alcohol samples analysis

The alcohol samples were first analysed using GC×GC-TOFMS and presented as total ion chromatograms (TIC) (Figure 3.9). Although the two whisky samples have similar chemical compounds, these alcohol samples still showed a distinct pattern based on the total ion chromatogram (TIC). The whisky samples (Figure 3.9a and Figure 3.9c) demonstrated a considerable variation in the TICs. Grey Goose Vodka samples showed a much purer compound pattern as evident in Figure 3.9b demonstrating a much less complex odour profile compared to the other two alcohols.

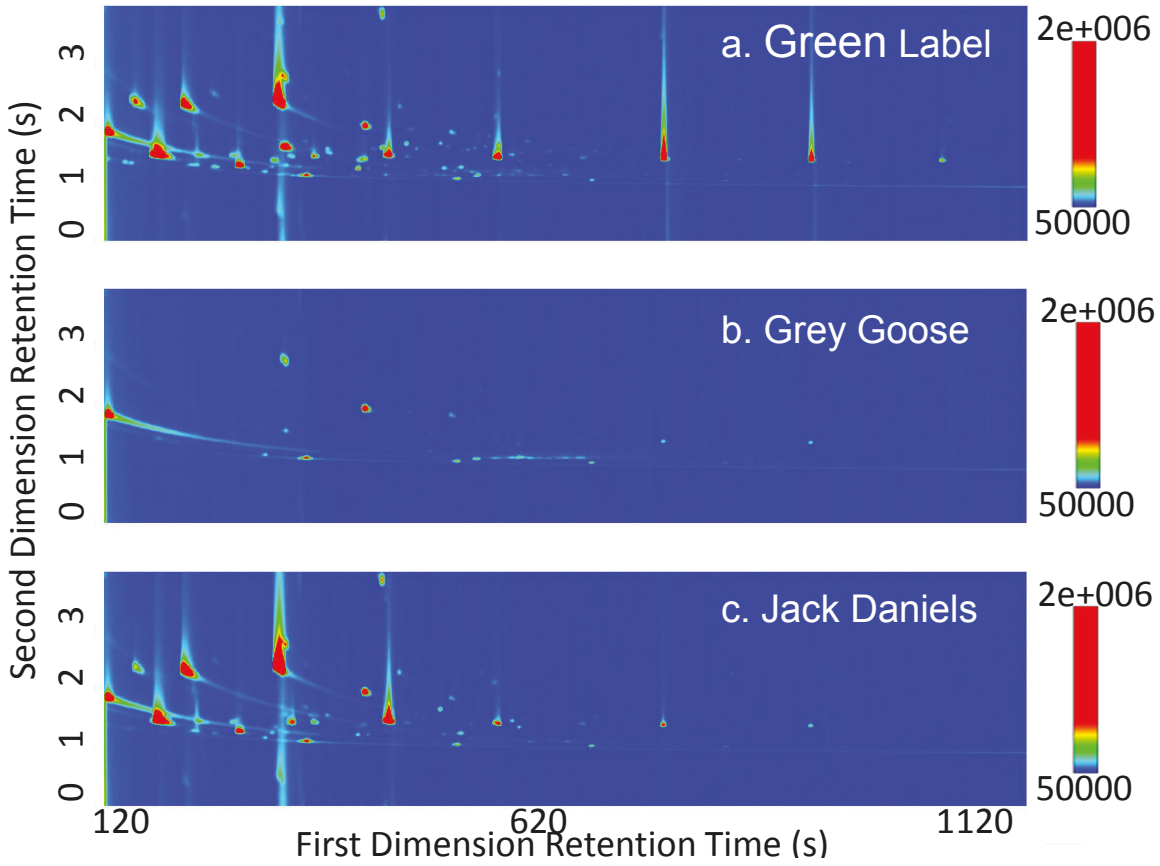


**Fig. 3.9** GC×GC-TOFMS TIC contour plots representative of (a) Green Label Whiskey samples, (b) Grey Goose Vodka samples, and (c) Jack Daniels Tennessee Whiskey samples.

Figure 3.11 and Figure 3.12 show the PCA results of three different alcohols by using GC×GC-TOFMS and the NOS.E platform. In Figure 3.11, all of the replicate alcohol samples tested were tightly clustered and illustrated similar VOC profiles. In addition, all three alcohol types were clearly distinguished from each other. According to the NOS.E PCA results, although the alcohol samples from the same origin were

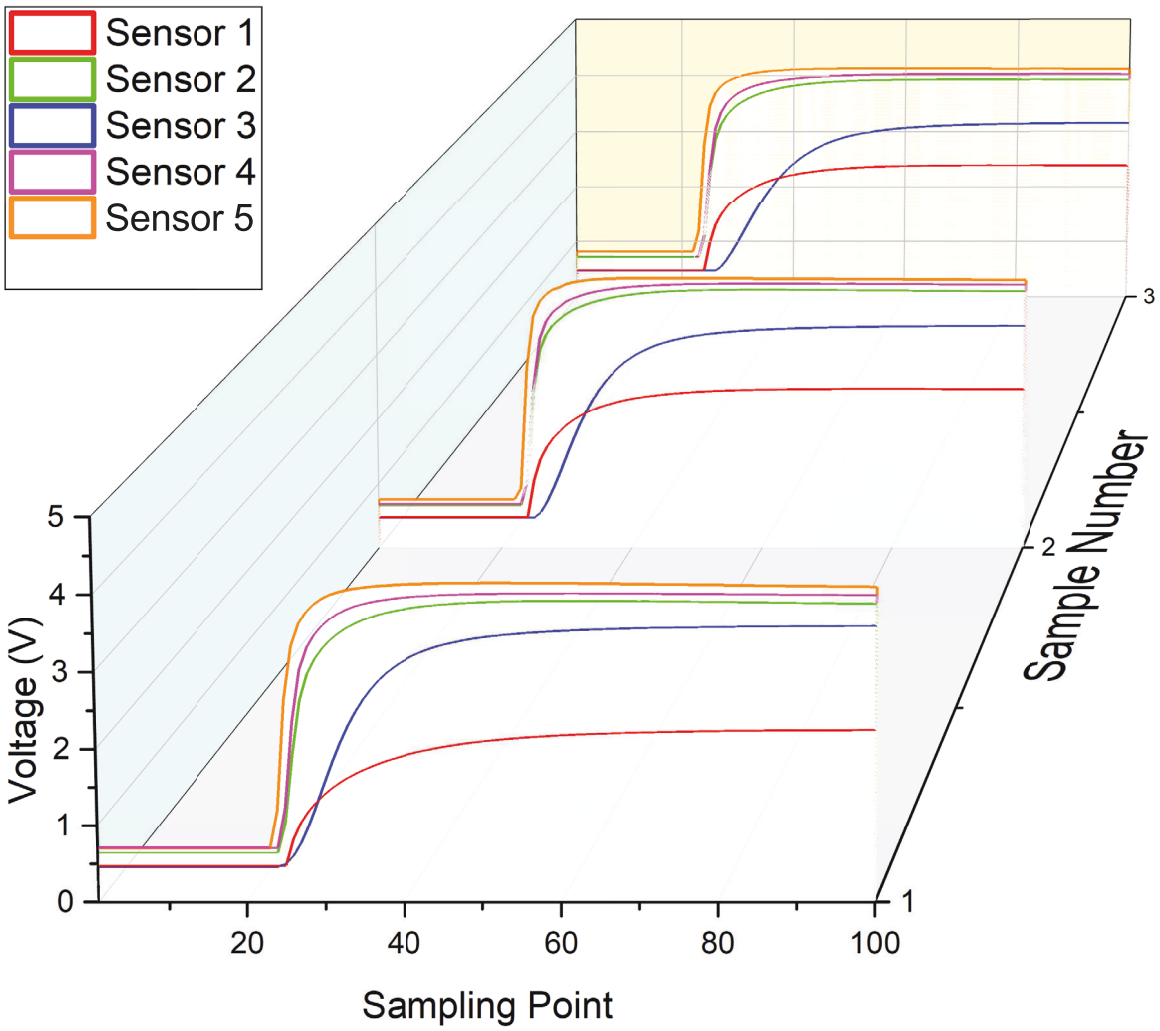


**Fig. 3.10** NOS.E Responses for Three Alcohol Samples.

not tightly clustered, these three alcohol types were still distinctly separated from each other and could be easily identified. Considering the low cost and fast analysis time, the performance of the NOS.E is acceptable. The classification results are shown in Table 3.3. We achieved 100% average accuracy of SVM classification when identifying GG with JW and JD samples. The average accuracy of SVM classification for the JW and JD samples is 97.8%. For the multi-classification of these three alcohol samples, we can recognise them with 93.33% average accuracy of SVM classification. Overall, it was determined that the NOS.E can provide an efficient method to identify different items based on their odour.

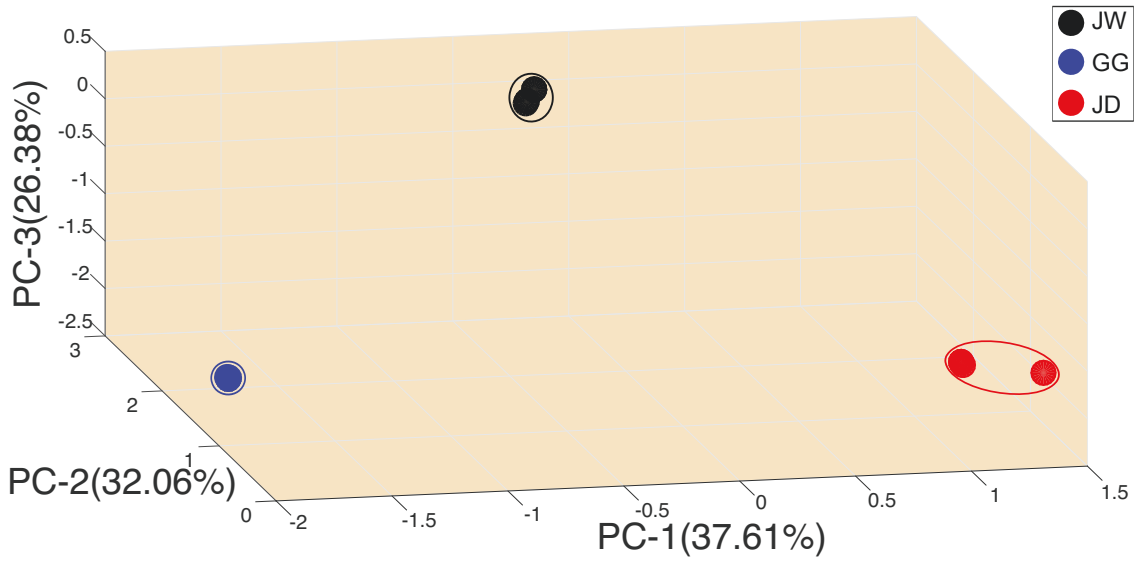


**Fig. 3.11** GC×GC-TOFMS PCA Analysis Results

**Table 3.3** Classification Results of three Alcohols

| Alcohols | Accuracy of SVM Classification |
|---|---|
| JW *VS* GG | 100% |
| JW *VS* JD | 97.8% |
| GG *VS* JD | 100% |
| JW *VS* GG *VS* JD | 93.33% |

## 3.6 Conclusions

This chapter proposed a new e-nose instrument, NOS.E, which was used for standardisation of odour detection and identification purposes. Based on the specific control logic, an automated air intake design (motivated by the human olfactory system) and the related fault detection and alarming design were equipped in the proposed NOS.E system. To validate the design of the automated airflow control and fault detection design, the NOS.E system was run 90 times over three days. The status of the airflow system was monitored by the NOS.E user interface. According to the test results, the NOS.E system can successfully execute the designed control logic. Moreover, the fault detection and monitoring system is able to provide the insurance to avoid wasting resources by timely terminating the system when a fault status of actuators is detected. The performance of the designed NOS.E system has been tested by three different alcohols and compared with the results of GC×GC-TOFMS by using PCA pattern recognition techniques. In summary, compared with the time consuming and expensive

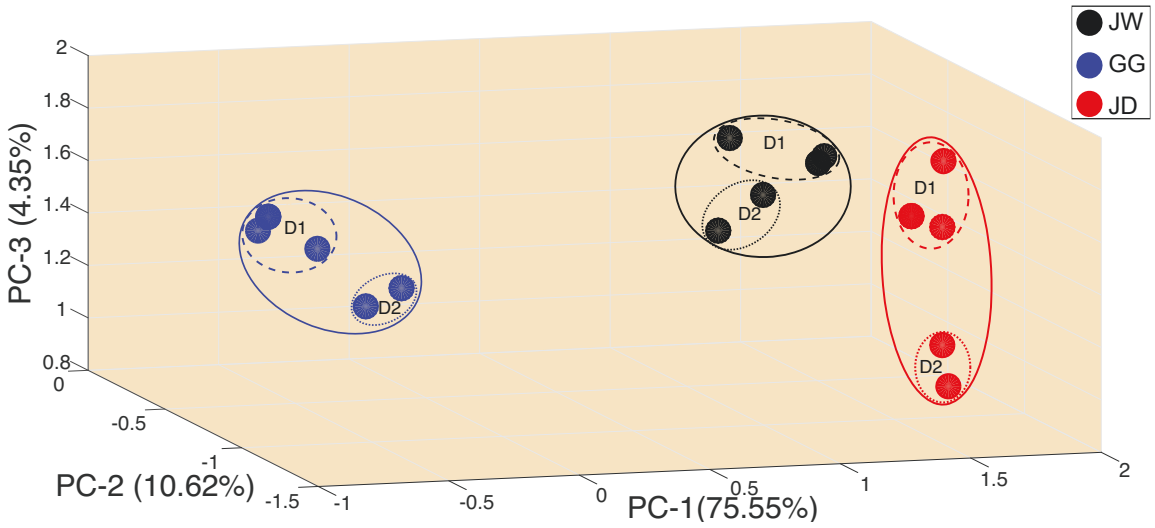


**Fig. 3.12** NOS.E PCA Analysis Results

GC×GC-TOFMS odour analysis method, the proposed e-nose system is a suitable odour analysis platform in terms of its low cost, high efficiency and reliability features. Future works will continue to develop the platform and improve the performance of the airflow control in the NOS.E. The intent is to build more NOS.E prototypes to validate its repeatability before implementing NOS.E in different applications (e.g. food quality assessment, illicit drug detection, wildlife products identification, etc.). Finally, a remote e-nose system based on the cloud and Narrow band Internet of Things (NB-IoT) technologies will be built.

# Chapter 4

# Development of the Data pre-processing and classification methods

## 4.1 Introduction

Electronic nose (e-nose) is capable of identifying chemical compounds through sensing and analysing odour molecules. As a kind of machine olfaction, E-nose plays a significant role in the odour analysis area and has received considerable attention from researchers all over the world [3–6]. The e-nose system comprises a set of active gas sensors that detect the odour and transduce the chemical vapours into electrical signals [3, 4]. The odour "fingerprint" captured by the gas sensors can then be analysed and identified with pattern classification methods, e.g., Principal Components Analysis (PCA), Cluster Analysis (CA), Support Vector Machine (SVM), and Artificial neural networks (ANNs). E-nose has been extensively applied in the areas of agriculture, medical diagnosis, environmental monitoring and protection, food safety, the military, cosmetics and pharmaceuticals [3, 6–13].

Currently, by using different e-nose platforms, the studies on e-nose mainly focus on two different parts: 1. The design of hardware systems (such as sensor design and main control system design) [11, 13, 22–27]; 2. The algorithms for e-nose, such as data

pre-processing methods and odour classification methods [11, 16, 28–34]. Moreover, some researchers develop their e-nose research based on the famous commercial e-nose products (such as the *fox* e-nose (*Alpha MOS, France*), the portable Cyranose 320 (*Cyrano Science, USA*), Airsense *PEN2* and *PEN3* (*Airsense Analytics GmbH, Germany*) [3, 35–41]. These studies on e-nose have made great progress in this area.

In order to get reliable classification results for e-nose applications, data pre-processing methods are used to improve the stability of the feature extracted from the pre-processed odour data. These pre-processing techniques mainly include wavelet package correlation filter (WPCF), mean filter (MF), polynomial curve fitting (PCF) and locally weighted regression (LWR), etc. [54, 55]. Even though these methods are quite mature and efficient, sometimes they are unable to obtain reliable results due to individual variations in the test system and unexpected responses caused by the gas interference or fluctuations of environmental parameters [51–53]. Moreover, these unexpected responses treated as noise will potentially reduce the stability and reliability of features. Experimental results will be significantly influenced especially for the derivative-related features which are sensitive to noise.

To seek a data pre-processing method which can overcome the drawbacks of current data pre-processing techniques for e-nose system, this dissertation proposes a novel non-parametric kernel-based modelling (KBM) data pre-processing method. Furthermore, this method is tested by recently developed NOS.E odour detection and identification system. The NOS.E system (shown in Fig. 4.1) mainly comprises an efficient power system, an automated air intake system, an interchangeable metal-oxide (MOX) gas sensor array board, and a fast data acquisition module. The target odour is drawn into the mixing chamber by the gas sampling pump, before going into the gas chamber, where the sensor array senses the odour stimulus.

Before applying the non-parametric KBM approach [61] [62] in the area of e-nose data pre-processing, three standard signals (linear signal, logarithmic signal and sigmoid signal) are used to test the performance of five different pre-processing methods (WPCF, MF, PCF and LWR). The results (presented in *Section* 4.3) show that the proposed method provides more reliable pre-processing results compared with other methods. Moreover, these data pre-processing methods are applied on real odour signals collected

by the NOS.E system and then the Coefficient of Variation (CoV) method is employed to evaluate the stability of derivative-related features [222]. The CoV analysis results (listed in Table 4.2) for different perfume features indicate that the proposed method is much better than other data pre-processing methods.

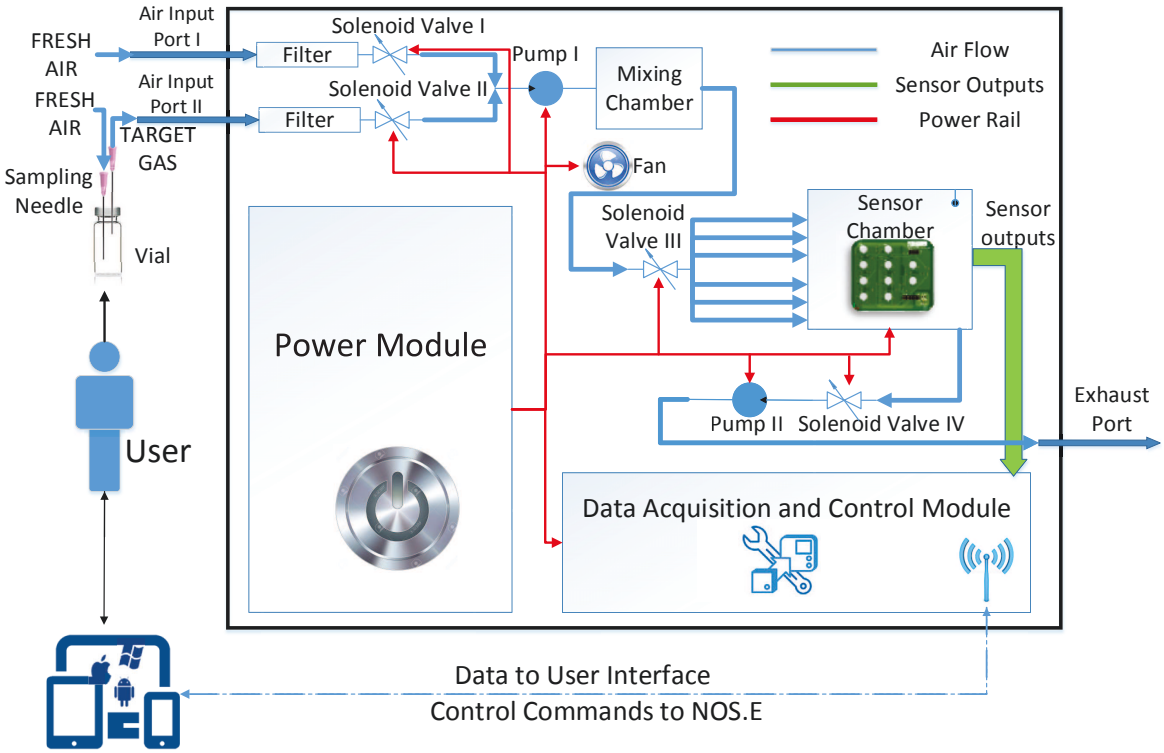


**Fig. 4.1** Block Diagram of NOS.E System.

## 4.2 Experimental Setup

### 4.2.1 Data pre-processing simulation setup

In this study, three different standard simulated test signals (as shown in Fig 4.2): linear signal (Lin) $y = x$; logarithmic signal (Log) $y = log(x)$ and sigmoid signal (sig) $y = \frac{1}{(1+exp(x))}$ are used to test the performance of five different pre-processing methods (KBM, WPCF, MF, LWR and PCF). Firstly, these signals are polluted by Gaussian noise with the Signal-to-Noise Ratio (SNR) ranging from 10dB to 50dB. Then the polluted signals are processed by the five different data pre-processing methods. Finally, the Normalized Root Mean Square Error (NRMSE) is used as a criterion to determine

the goodness-of-fit between the pre-processed signals and the original signal, where the NRMSE costs vary between -Inf (bad fit) to 1 (perfect fit).

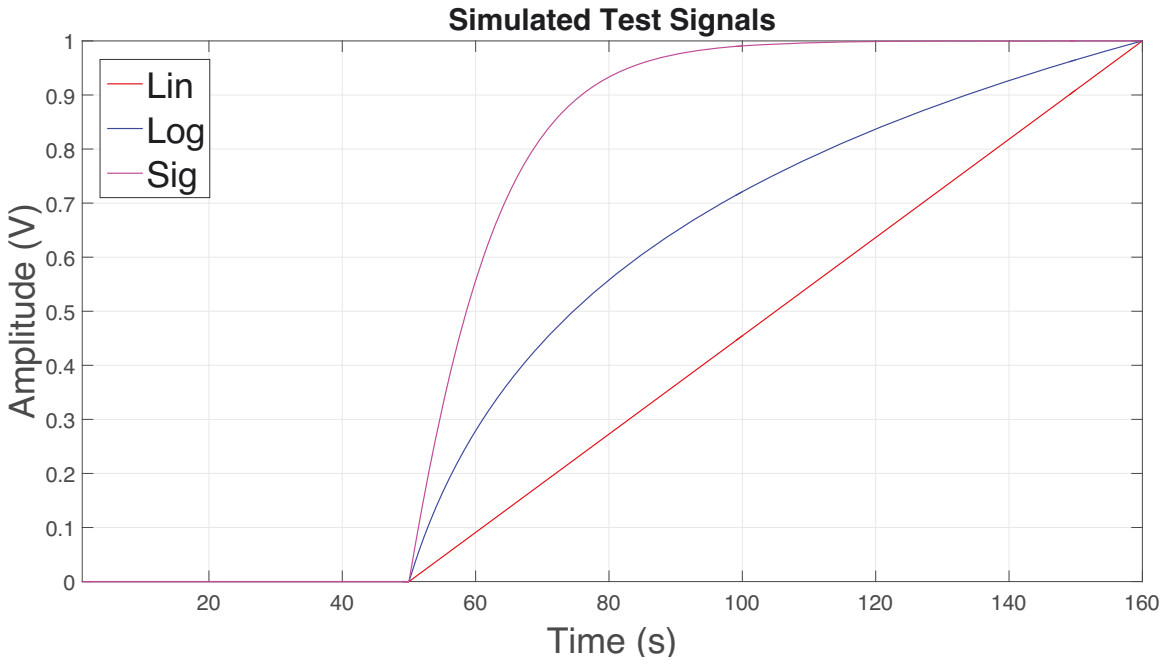


**Fig. 4.2** Simulated Test Singals.

### 4.2.2 Perfume test experimental setup

In this study, the NOS.E system is used to detect two different kinds of perfume samples (bought from David Jones, Bondi Junction, Sydney, Australia) to verify the performance of the proposed data pre-processing method. The diagram of NOS.E perfume test experimental setup is shown in Fig. 4.3. Users can operate the NOS.E system via a touchscreen Pad. The integral components of this system are assembled in a carrying case. The sensor array in NOS.E equipment is composed of ten commercially available metal oxide gas sensors: *TGS 2611E, TGS 2612, TGS 2610D, TGS 2611C,TGS 2610C, TGS 2602, TGS 2600, TGS 2620, TGS 2603*, and *TGS 2602*. The input gases are sensed by the sensor array, converted to digital signals and then sent to the host computer for data processing.

In this dissertation, perfume I is *CHANEL Chance*, and perfume II is *CHANEL Gabrielle*. The test protocol of the NOS.E system are listed in the following steps (using perfume I as an example).

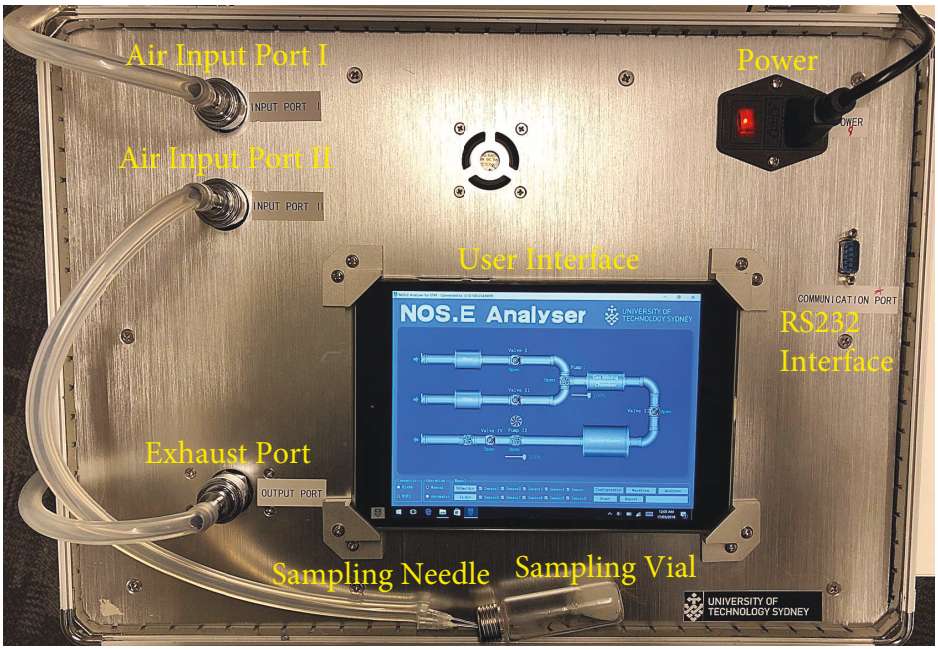


**Fig. 4.3** Diagram of NOS.E Perfume Test Experimental Setup.

- 1. Prepare the test sample;

  Prepare $1mL$ perfume using a $10mL$ headspace vial.

- 2. Power on NOS.E equipment and open all the actuators to warm up the NOS.E equipment for 60 minutes (to make sure the equipment works at the ideal working status);

- 3. Set up the test time for each phase using the NOS.E Analyser;

  Chamber washing time for 300 seconds; Vacuum I for 10 seconds; Baseline setup for 20 seconds; Vacuum II for 10 seconds; Test time for 90 seconds; Baseline recovery for 90 seconds and Chamber washing II for 300 seconds.

- 4. Connect the test sample I;

  As shown in Fig. 4.1, the headspace vial has two sampling needles, one for the fresh air input and the other for the target gas output to NOS.E equipment.

- 5. Configuration;

  Click Configuration button on the user interface to send all the operating parameters and instructions to the slave computer.

- 6. Baseline set up;

  Click Baseline Set up button, and the NOS.E system will start to collect the data until the baseline recovery phase is complete.

- 7. Save the data;

  Once the test is completed, the data will be saved in the local folder automatically.

- 8. Review the test results;

  Click Analyser button to review the sensor responses and key features.

- 9. Repeat the same sample followed from steps 5 to 8 for ten times, and these repetitions are counted as ten acquisitions.

Followed by the NOS.E odour test protocol, 60 perfume samples (30 perfume I samples, 30 perfume II samples) were collected by the NOS.E system. Each sample was collected 10 times, hence, 600 odour datasets (300 perfume I datasets, 300 perfume II datasets) were analysed. The odour datasets used in this dissertation were obtained under 25 ℃ to 27 ℃ ambient temperature and 50% RH ambient humidity.

## 4.3 Methodology

In this section, a new non-parametric KBM method is exploited to model the gas sensor response [61–63]. Unlike most of the previous data pre-processing methods, in which the denoising is based on the filtering techniques, our approach reduces the noise based on non-parametric modelling to improve the classification accuracy; nevertheless, the identification of the model is based on the raw data.

Generally, the researchers use filtering based methods to process the raw data of the MOX gas sensor response. However, sometimes, the filter based pre-processing methods cannot obtain the desired effects due to the individual variation of the test system and the unexpected responses caused by the interference gas [51–53]. Especially, as the filter based methods cannot always guarantee the smoothness of the filtered signals, in feature extraction stage, unacceptable outliers might be generated. In order

to obtain a better result and facilitate the automated feature extraction, we adopted a non-parametric modelling method which applied the finite impulse response to describe the system's characteristics.

In this dissertation, $t$ with sampling time $T$ is selected as the time index. The relationship between the gas input ($u$), which can be approximately treated as a step stimulation, and the gas sensor response ($y$) can be described by a single input single output (SISO) dynamic system. Hence, the discrete time output $y$ can be calculated by the impulse response (IR) of this system by the following equation:

$$y(t) = \sum_{k=1}^{\infty} g_k^0 q^{-k} u(t) + \epsilon(t), \quad k = 1, 2, 3 \cdots, \infty, \quad t = 1, 2, 3 \cdots, N \tag{4.1}$$

where $g_k^0$ represents the coefficient of the impulse response. $q$ represents the shift operator, i.e. $qu(t) = u(t+1)$, $\epsilon(t)$ is the Gaussian white noise.

Considering the impulse response decays exponentially for linear stable systems, we here express the system by using the $m$th finite impulse response (FIR) as:

$$G(q, \boldsymbol{c}) = \sum_{k=1}^{m} g_k q^{-k}, \quad \boldsymbol{c} = [g_1, g_2, \cdots, g_m]^T. \tag{4.2}$$

Hence, the model in Eq.(4.1) is able to be transferred as:

$$y(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{c} + \epsilon(t), \tag{4.3}$$

where $\boldsymbol{\varphi}(t)$ contains the input information of the system:

$$\boldsymbol{\varphi}(t) = [u(t-1), u(t-2), \cdots, u(t-m)]^T. \tag{4.4}$$

Then, the FIR model can be written as:

$$\boldsymbol{Y}_N = \boldsymbol{\phi}_N \boldsymbol{c} + \boldsymbol{\varepsilon}_N, \tag{4.5}$$

where $N = M - m$, and $M$ is the number of data point that we collected. $\boldsymbol{Y}_N$ is a vector representation of sensor's responses:

$$\boldsymbol{Y}_N = [y(1), y(2), \cdots, y(t), \cdots, y(N)]^T, \tag{4.6}$$

where $y(t)$ denotes the $t$-th element of $\boldsymbol{Y}_N$.
$\boldsymbol{\epsilon}_N$ is a vector representation of Gaussian white noise:

$$\boldsymbol{\epsilon}_N = [\epsilon(1), \epsilon(2), \cdots, \epsilon(t), \cdots, \epsilon(N)]^T, \tag{4.7}$$

where $\epsilon(t)$ denotes the $t$-th element of $\boldsymbol{\epsilon}_N$.
The $i$-th row of $\boldsymbol{\phi}_N \in \mathbb{R}^{N \times m}$ is $[u(m+i-1), u(m+i-2), \cdots, u(i)]$.

Assuming that function $g \in \mathbb{R}^m$, then function $g$ in the regularisation term can be projected into a reproducing kernel Hilbert space (RKHS).

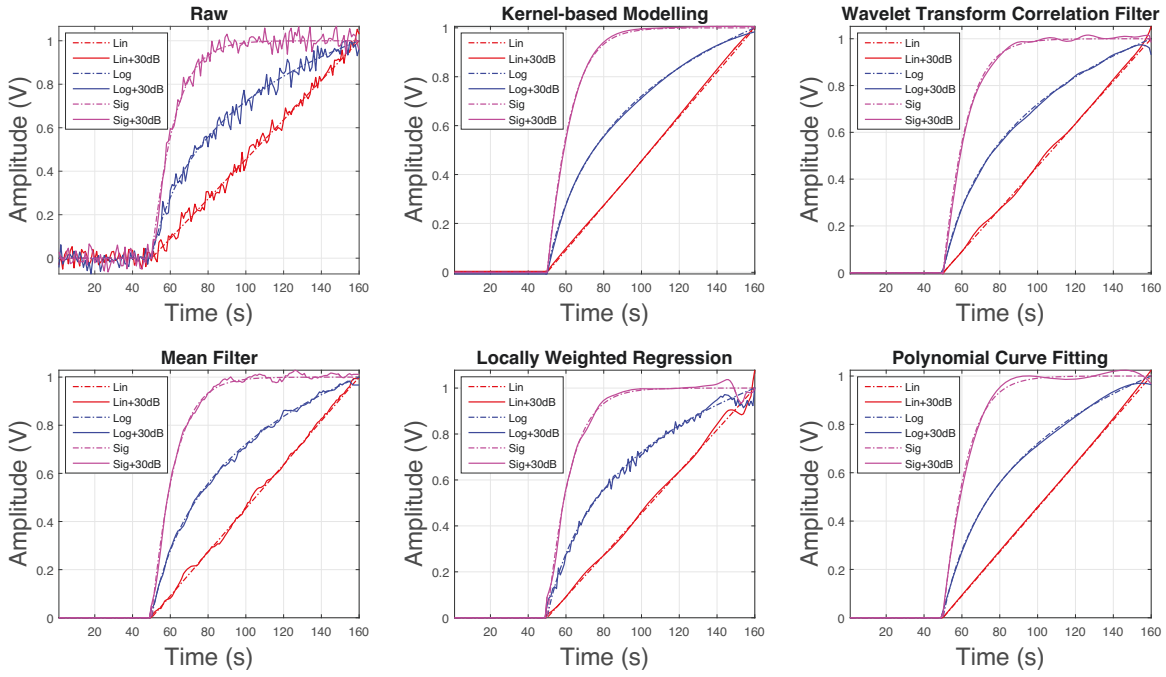


**Fig. 4.4** Simulated Response Comparison for Different Data Pre-processing Methods.

The IR model can be identified by minimising the cost function:

$$\hat{\boldsymbol{c}} = arg \min_{\boldsymbol{c} \in \mathbb{R}^m} ||\boldsymbol{Y}_N - \boldsymbol{\phi}_N \boldsymbol{c}||_2^2 + \gamma \boldsymbol{c}^T \boldsymbol{P}^{-1} \boldsymbol{c}, \tag{4.8}$$

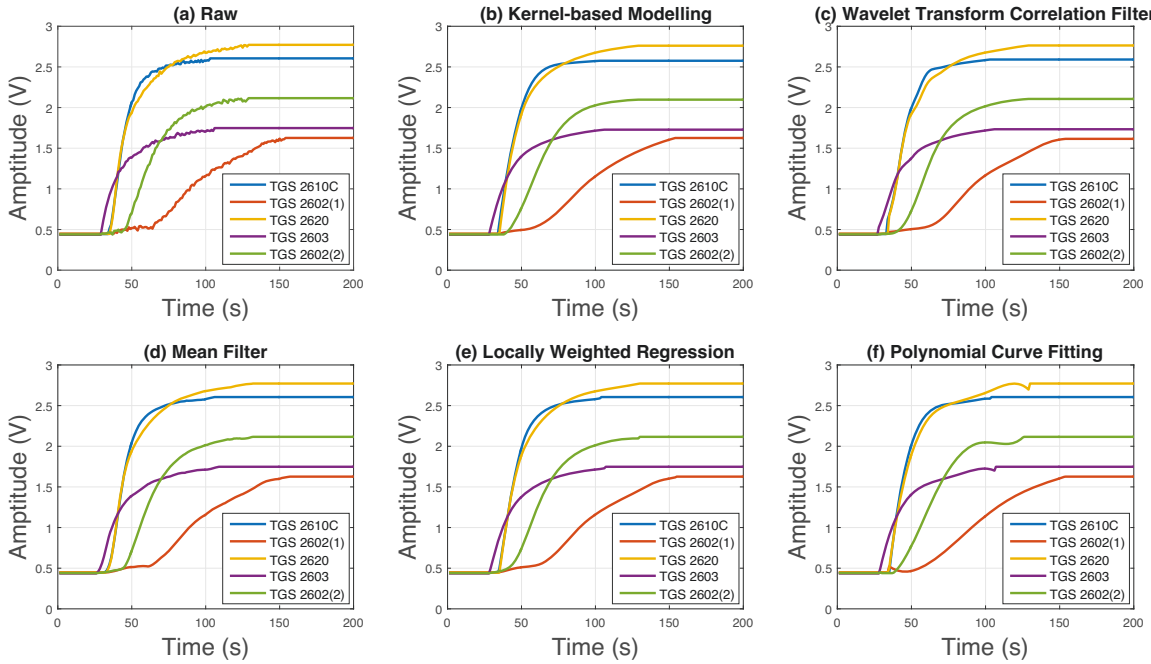


**Fig. 4.5** Data pre-processing Results for Perfume I.

where $\boldsymbol{P}$ represents the kernel matrix which is defined as:

$$p(i,j) = e^{-\rho\|i-j\|^2}, \rho > 0. \tag{4.9}$$

The estimated IR model from Eq.(4.8) can provide better and smoother results compared to LASSO (Least Absolute Shrinkage and Selection Operator) or Ridge regression using the prior information in kernel matrix $\boldsymbol{P}^{-1}$ [62]. Furthermore, as the system is stable, after a while, the impulse response will be close to zero. Hence, when $m$ is too big, we expect that the last several parameters of the estimated FIR approach to zero. Therefore, an extra $\mathcal{L}_1$ regularisation was added to sparsify the transfer function identified, and the cost function can be rewritten as:

$$\hat{\boldsymbol{c}} = arg \min_{\boldsymbol{c}\in\mathbb{R}^m} ||\boldsymbol{Y}_N - \boldsymbol{\phi}_N\boldsymbol{c}||_2^2 + \gamma\boldsymbol{c}^T\boldsymbol{P}^{-1}\boldsymbol{c} + \alpha||\boldsymbol{c}||_1. \tag{4.10}$$

where $\alpha$ is a positive coefficient to control the trade off between $\mathcal{L}_1$ regulariser and kernel regulariser $\gamma\boldsymbol{c}^T\boldsymbol{P}^{-1}\boldsymbol{c}$.

**Table 4.1** goodness-of-fit Result

| Method | Test Signal | goodness-of-fit Result | | | | |
|---|---|---|---|---|---|---|
| | | Signal-to-Noise Ratio | | | | |
| | | 10dB | 20dB | 30dB | 40dB | 50dB |
| N/A | Linear | 0.0567 | 0.7016 | 0.9056 | 0.9702 | 0.9906 |
| N/A | Logarithmic | 0.1713 | 0.7379 | 0.9171 | 0.9738 | 0.9917 |
| N/A | Sigmoid | 0.2970 | 0.7777 | 0.9297 | 0.9778 | 0.9930 |
| KBM | Linear | **0.8411** | 0.9492 | 0.9818 | 0.9900 | 0.9912 |
| KBM | Logarithmic | **0.8605** | **0.9553** | **0.9843** | **0.9916** | **0.9927** |
| KBM | Sigmoid | **0.8815** | **0.9620** | **0.9859** | **0.9914** | **0.9922** |
| WCTF | Linear | 0.6740 | 0.8950 | 0.9605 | 0.9661 | 0.9635 |
| WCTF | Logarithmic | 0.7139 | 0.9063 | 0.9568 | 0.9670 | 0.9696 |
| WCTF | Sigmoid | 0.7557 | 0.9137 | 0.9573 | 0.9633 | 0.9638 |
| MF | Linear | 0.7057 | 0.9066 | 0.9694 | 0.9875 | 0.9912 |
| MF | Logarithmic | 0.7414 | 0.9179 | 0.9730 | 0.9887 | 0.9919 |
| MF | Sigmoid | 0.7802 | 0.9296 | 0.9750 | 0.9860 | 0.9876 |
| LWR | Linear | 0.7579 | 0.9256 | 0.9766 | 0.9923 | 0.9969 |
| LWR | Logarithmic | 0.7936 | 0.9349 | 0.9784 | 0.9905 | 0.9927 |
| LWR | Sigmoid | 0.8253 | 0.9432 | 0.9772 | 0.9841 | 0.9850 |
| PCF | Linear | 0.8409 | **0.9498** | **0.9841** | **0.9950** | **0.9984** |
| PCF | Logarithmic | 0.8598 | 0.9542 | 0.9815 | 0.9873 | 0.9881 |
| PCF | Sigmoid | 0.8783 | 0.9538 | 0.9711 | 0.9735 | 0.9738 |

**Note**: N/A: Not Applicable; KBM: Kernel-based modelling; WTCF: Wavelet Transform Correlation Filter; MF: Mean Filter;
LWR: Locally Weighted Regression; PCF: Polynomial Curve Fitting.

The above equation can be considered as a special case of elastic net [223] which the $\mathcal{L}_2$ norm regularisation is weighted by kernel matrix $\boldsymbol{P}^{-1}$. We here rearrange Eq.(4.10) and define two new parameters as:

$$\boldsymbol{\phi}_N^* = \frac{1}{\sqrt{1+\gamma}} \begin{bmatrix} \boldsymbol{\phi}_N \\ \sqrt{\gamma}\boldsymbol{B} \end{bmatrix}, \tag{4.11}$$

where $\boldsymbol{B}$ is the upper triangular matrix from Cholesky factorisation of kernel matrix $\boldsymbol{P}^{-1}$ ($\boldsymbol{P}$ is symmetric) and

$$\boldsymbol{Y}_N^* = \begin{bmatrix} \boldsymbol{Y}_N \\ \boldsymbol{0} \end{bmatrix}. \tag{4.12}$$

Then, the cost function Eq.(4.10) can be written as:

$$\hat{\boldsymbol{c}}^* = arg \min_{\boldsymbol{c}^* \in \mathbb{R}^m} ||\boldsymbol{Y}_N^* - \boldsymbol{\phi}_N^* \boldsymbol{c}^*||_2^2 + \frac{\alpha}{\sqrt{1+\gamma}} ||\boldsymbol{c}^*||_1, \tag{4.13}$$

where $\boldsymbol{c}^*$ is defined as:

$$\boldsymbol{c}^* = \sqrt{1+\gamma}\boldsymbol{c}. \tag{4.14}$$

Due to the limitation of the input signal, the input matrix $\boldsymbol{\phi}_N^T \boldsymbol{\phi}_N$ is not orthogonal.

As efficient algorithms for solving wide classes of convex optimisation problems, interior-point methods are always efficient in terms of computation time and resource consumptions [224]. The Eq. (4.13) is convex but not differentiable. The $\mathcal{L}_1$ regularisation LSP (Least Squares Problems) can be transformed to a convex quadratic problem, with linear inequality constraints. Therefore, we adopt an interior-point method (primal-dual interior-point method) [225] for this $\mathcal{L}_1$ norm regularisation for $\hat{\boldsymbol{c}}^*$. As the result, $\hat{\boldsymbol{c}}^*$ can be restored to $\hat{\boldsymbol{c}}$ as:

$$\hat{\boldsymbol{c}} = \frac{1}{\sqrt{1+\gamma}}\hat{\boldsymbol{c}}^*. \tag{4.15}$$

## 4.4 Experimental Results and Discussion

### 4.4.1 Data pre-processing simulation results

The datasets listed in Table. 4.1 come from the average of 100,000 times simulations, which show the goodness-of-fit test results for three test signals. According to this table, although PCF achieves higher goodness-of-fit results for linear signal (SNR from 20dB to 50dB), the proposed non-parametric KBM method can also achieve quite close results. As for the logarithmic signal and sigmoid signal, the proposed method achieves higher goodness-of-fit results than the other methods, which indicates the signals processed by the proposed method are more close to the original signals. Therefore, the proposed non-parametric KBM method provides more reliable pre-processing results. Moreover, since the original test signals (linear signal, logarithmic signal and sigmoid signal) are smooth and noiseless, when using these signals to evaluate the different data pre-processing methods, the method achieves higher goodness-of-fit is considered as a more smooth method. In addition, the simulated response results (see Fig. 4.4) for different data pre-processing methods indicate that the proposed method can provide more smooth results compared with other pre-processing methods.

### 4.4.2 Perfume test results

Based on the different data pre-processing methods, perfume I is used as an example to show the waveforms for the sensors which have responsed to the test samples, and results are plotted in Fig. 4.5. From this figure, it can be more directly seen that the proposed non-parametric KBM method can provide more smooth response waveforms compared with the other methods. Therefore, the proposed method can achieve much more reliable derivative-related features.

In odour classification stage, the features are often extracted from the data recorded until the steady-state response of gas sensors are reached. Feature extraction methods generally fall into two categories: the human-supervised extraction based on expert knowledge and the automatic feature extraction methods that are completely data-driven. In this study, we have extracted the features by using a human-supervised method as the proof-of-concept of the presented method. Considering the derivative-based features are sensitive to noise, to demonstrate the effectiveness of our proposed method in dealing with noised signals, we in particular chose six derivative-based features commonly used by other works [64, 65]: the response of the maximum $1^{st}$ derivative ($D_{res}$), the response of the maximum $2^{nd}$ derivative ($D_{resx}$), the response of the minimum $2^{nd}$ derivative ($D_{resn}$), time interval between gas-in and maximum $1^{st}$ derivative of response ($t_{Dres}$), the time interval between gas-in and maximum $2^{nd}$ derivative of response ($t_{Dresx}$), and the time interval between gas-in and minimum $2^{nd}$ derivative of response ($t_{Dresn}$). The diagram of these features is shown in Fig. 1.2.

All the derivative-related features listed in Fig. 1.2 are extracted based on the six different pre-processed datasets. Then the Coefficient of Variation (CoV: $Var_{coef}(x)$) assessment method was used to evaluate the stability of these features. According to Eq. 4.16, a smaller CoV value will indicate the better feature stability:

$$Var_{coef}(x) = \frac{\sigma(x)}{\bar{x}} \times 100\%, \tag{4.16}$$

where $\sigma(x)$ is the standard deviation of the feature $x$, and $\bar{x}$ is the mean value of the feature $x$.

Based on the CoV analysis results (listed in Table 4.2) for features of different types of perfumes, the proposed non-parametric KBM method has a smaller CoV value compared with the other methods, which means this method could provide more stable feature datasets in terms of improving the performance of the classification task.

**Table 4.2** The Odour Features' Coefficient of Variation for Perfume I and Perfume II

| Features | Perfume I | | | | | | Perfume II | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | KBM | WTCF | MF | LWR | PCF | N/A | KBM | WTCF | MF | LWR | PCF | N/A |
| TGS 2610$C$ $D_{res}$ | 1.64 | 19.85 | 1.94 | **1.22** | 17.45 | 41.18 | 1.40 | 29.94 | 2.27 | **1.13** | 19.65 | 47.25 |
| TGS 2610$C$ $D_{resx}$ | **0.71** | 11.16 | 4.10 | 1.49 | 28.24 | 12.31 | **0.72** | 7.05 | 4.32 | 1.49 | 16.13 | 14.42 |
| TGS 2610$C$ $D_{resn}$ | 0.90 | 56.94 | 1.82 | **0.84** | 57.13 | 34.82 | **0.61** | 66.92 | 1.87 | 0.72 | 43.51 | 70.85 |
| TGS 2610$C$ $t_{Dres}$ | **0.59** | 6.32 | 2.28 | 0.80 | 33.66 | 31.17 | **0.66** | 3.42 | 2.59 | 0.97 | 33.87 | 31.48 |
| TGS 2610$C$ $t_{Dresx}$ | **2.22** | 53.52 | 12.31 | 14.02 | 26.18 | 48.91 | **2.58** | 82.42 | 14.96 | 15.59 | 30.55 | 50.81 |
| TGS 2610$C$ $t_{Dresn}$ | 11.07 | 18.50 | 17.10 | 57.13 | 29.37 | **15.93** | 18.69 | 21.88 | 18.14 | 56.91 | **16.13** | 17.70 |
| TGS 2602$_I$ $D_{res}$ | **1.91** | 112.57 | 10.99 | 23.57 | 83.36 | 37.07 | **2.44** | 90.16 | 11.35 | 17.94 | 54.53 | 81.65 |
| TGS 2602$_I$ $D_{resx}$ | **1.68** | 66.50 | 10.27 | 10.77 | 17.66 | 22.97 | **1.85** | 52.55 | 11.12 | 11.20 | 19.49 | 39.82 |
| TGS 2602$_I$ $D_{resn}$ | **1.58** | 43.97 | 16.55 | 48.21 | 60.17 | 31.33 | **1.79** | 33.04 | 17.94 | 51.49 | 67.42 | 33.99 |
| TGS 2602$_I$ $t_{Dres}$ | **10.90** | 30.30 | 17.52 | 120.75 | 37.24 | 16.510 | **13.35** | 33.49 | 19.31 | 95.21 | 15.38 | 19.25 |
| TGS 2602$_I$ $t_{Dresx}$ | **3.43** | 103.43 | 14.50 | 34.42 | 109.80 | 43.46 | **2.54** | 73.34 | 13.09 | 20.40 | 64.10 | 80.94 |
| TGS 2602$_I$ $t_{Dresn}$ | **2.64** | 85.17 | 16.38 | 8.09 | 18.11 | 22.66 | 3.67 | 87.15 | 17.98 | **3.03** | 15.61 | 41.02 |
| TGS 2620 $D_{res}$ | **4.00** | 47.97 | 18.04 | 62.07 | 60.17 | 30.60 | **4.92** | 39.21 | 18.67 | 63.21 | 66.40 | 31.17 |
| TGS 2620 $D_{resx}$ | 3.70 | 6.46 | **3.50** | 3.78 | 6.32 | 6.54 | 2.56 | 3.63 | **2.39** | 2.55 | 3.23 | 3.67 |
| TGS 2620 $D_{resn}$ | 2.94 | 3.66 | **2.79** | 3.04 | 3.69 | 3.69 | 3.00 | 4.02 | **2.85** | 3.05 | 4.28 | 4.06 |
| TGS 2620 $t_{Dres}$ | 3.94 | 4.36 | **3.69** | 3.89 | 4.33 | 4.41 | 4.60 | 5.17 | **4.29** | 4.56 | 5.12 | 5.23 |
| TGS 2620 $t_{Dresx}$ | 3.34 | 4.18 | **3.13** | 3.52 | 5.15 | 4.23 | 2.93 | 4.00 | **2.74** | 3.15 | 4.49 | 4.04 |
| TGS 2620 $t_{Dresn}$ | 1.12 | 6.78 | 2.05 | **0.95** | 5.63 | 6.04 | 1.19 | 6.57 | 2.25 | **0.73** | 4.60 | 9.89 |
| TGS 2603 $D_{res}$ | **0.52** | 27.88 | 4.40 | 2.61 | 71.53 | 26.75 | **0.40** | 21.38 | 4.35 | 4.00 | 5.49 | 23.89 |
| TGS 2603 $D_{resx}$ | **0** | 7.65 | 1.63 | 0.56 | 21.44 | 7.93 | 0.40 | 5.11 | 1.54 | **0.22** | 8.90 | 10.12 |
| TGS 2603 $D_{resn}$ | **0** | 8.21 | 2.53 | 1.49 | 9.30 | 15.89 | **0** | 4.48 | 2.69 | 1.64 | 6.92 | 12.92 |
| TGS 2603 $t_{Dres}$ | **0** | 6.76 | 3.66 | 0.76 | 5.93 | 43.86 | **0** | 5.35 | 3.98 | 1.02 | 5.24 | 40.47 |
| TGS 2603 $t_{Dresx}$ | **4.66** | 5.65 | 36.88 | 62.61 | 69.01 | 44.10 | 8.71 | **5.49** | 37.96 | 65.71 | 5.49 | 42.65 |
| TGS 2603 $t_{Dresn}$ | **3.03** | 36.98 | 5.91 | 58.37 | 79.67 | 53.22 | **0** | 14.10 | 3.16 | 37.05 | 22.34 | 46.67 |
| TGS 2602$_{II}$ $D_{res}$ | **0** | 10.71 | 9.56 | 0.53 | 36.24 | 49.88 | **0** | 8.50 | 4.41 | 0.70 | 38.62 | 53.05 |
| TGS 2602$_{II}$ $D_{resx}$ | **0.10** | 12.20 | 7.82 | 52.17 | 15.26 | 41.60 | **0.06** | 5.53 | 8.93 | 55.90 | 14.76 | 39.37 |
| TGS 2602$_{II}$ $D_{resn}$ | **0.66** | 20.53 | 12.96 | 9.22 | 60.15 | 38.52 | **0.43** | 20.77 | 9.78 | 16.62 | 53.35 | 39.57 |
| TGS 2602$_{II}$ $t_{Dres}$ | **1.33** | 43.22 | 17.30 | 42.88 | 65.11 | 48.52 | **3.46** | 17.41 | 13.45 | 44.96 | 34.93 | 40.79 |
| TGS 2602$_{II}$ $t_{Dresx}$ | **2.52** | 17.24 | 33.01 | 5.96 | 9.05 | 40.38 | 3.82 | 18.06 | 26.64 | **2.11** | 6.67 | 44.34 |
| TGS 2602$_{II}$ $t_{Dresn}$ | **0** | 9.60 | 23.55 | 38.84 | 76.17 | 39.82 | **0** | 5.75 | 26.79 | 40.32 | 75.87 | 40.30 |

### 4.4.3 Classification results

Based on different data pre-processing methods (non-parametric KBM, WPCF, MF, PCF and LWR), we extracted six derivative-related key features for the classification of the two types of perfumes by using the SVM classifier with RBF kernel. We used the five-fold cross-validation method for classification. The full dataset is partitioned into training and test sets by randomly selecting 20% of the data from each sample to form the test set and the remaining data to form the training set. This process was performed five times: where each data sample appears exactly once in a test set. The five-fold cross-validation method is run ten times and the averaged accuracy is used to assess the perfume classification accuracy.

The classification results are listed in Table 4.3. Without any pre-processing methods, the SVM classifier achieved a 62.57% average accuracy of classification. When the odour dataset was processed by WPCF, we achieved 94.12% average classification

accuracy. The MF pre-processing method for SVM classifier has a 68.57% average classification accuracy. The PCF and LWR pre-processing methods attained 87.80% and 71.50% average classification accuracy, respectively. According to these classification results, the proposed non-parametric KBM method achieved a much higher average classification accuracy (96.23%) than other methods. The confusion matrix for the classification accuracy based on different data pre-processing methods is shown in Fig 4.6. The classification accuracy clearly indicates that the proposed data pre-processing method can provide more stable features to improve the performance of the odour classifier.

**Table 4.3** Classification Results

| Pre-processing Method | Parameters of Pre-processing Method | Average Accuracy of SVM Classification (%) |
|---|---|---|
| N/A | $N/A$ | 62.57 |
| Kernel-based modelling method | $Kernel = RBF$, $\rho = 0.0005$, $m = 200$ | 96.23 |
| Wavelet Package Correlation Filter | $Wavelet = Daubechies - 3$, $Level = 5$ | 94.12 |
| Mean Filter | $Window\ width = 7$ | 68.57 |
| Polynomial Curve Fitting | $Order = 3$ | 87.80 |
| Locally Weighted Regression | $\tau = 5$ | 71.50 |

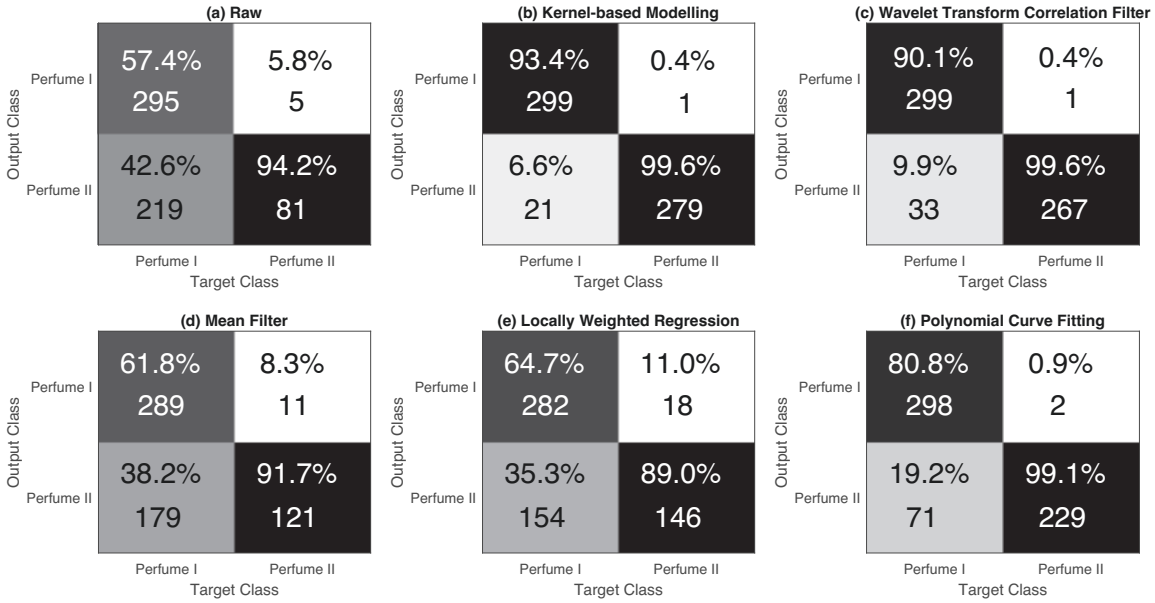


**Fig. 4.6** The confusion matrix for the classification accuracy for different data-preprocessing methods.

## 4.5 Conclusion

A novel non-parametric KBM odour data pre-processing method has been presented and its effectiveness has been tested on the recently developed NOS.E system. According

to the test results, when extracting derivative-related features, the proposed non-parametric KBM method provides more reliable and stable pre-processing results compared with the other pre-processing methods. Based on these derivative-related features, NOS.E system can detect and identify two different perfumes with a 96.23% classification accuracy by using popular SVM classifier.

Further study will focus on the areas below: 1. Collecting different samples to validate the implementation of the NOS.E system across various odour identification and classification applications; 2. Improving the efficiency and functionality of the NOS.E system by further research outputs with the proposed non-parametric KBM data pre-processing method and NOS.E hardware design optimisation; 3. Considering discriminative models for feature extraction [226]); 4. Applying some cost-sensitive classification algorithms to improve the performance of NOS.E system [200].

# Chapter 5

# Electronic Nose based Odour Classification using Genetic Algorithm and Fuzzy Support Vector Machine

## 5.1 Introduction

The electronic nose (e-nose) is a device intended to detect and identify odours or flavors based on the analysis of the responses of an array of gas sensors by using machine learning algorithms. Over the last decades, in order to improve its performance, the algorithms that have been employed to analyze electronic nose data have undergone several important developments being driven by the fast growth of machine learning technologies [3, 4, 82].

Benedetti, Simona, et al. [173] used Principal Component Analysis (PCA) and Artificial Neural Network (ANN) to evaluate seventy samples of honey of different geographical and botanical origin data and showed a 83.5% accuracy rate. By utilizing support vector machine (SVM) for the calibration of the e-nose arrangement, Brudzewski, Kazimierz, et al. presented an interesting case study for milk recognition [173]. The results of numerical experiments in this paper showed that for the

e-nose, SVM algorithm has good generalization ability for the reasonably small size of training datasets. In order to improve e-nose performance by optimizing feature selection techniques, Gardner, JW et al. introduced a novel search method procedure, V-integer genes genetic algorithms (GA), and compared it with other search methods such as sequential forward or backward searches (SFS or SBS) and X-binary genes GAs in paper [67]. Their test results showed that the V-integer genes GA approach is an accurate and a fast search method when compared to some other feature selection techniques.

Based on the dataset collected by PEN2, Yu, Huichun et al. [164] achieved an 80% to 100% correct rate when classifying Longjing green-tea quality grade using the ANN algorithm. To correlate the e-nose measurements with the tea taster's scores in paper [227], Bhattacharyya, Nabarun, et al. obtained 81% to 85% classification rate for Back-Propagation-Multilayer Perceptron (BP-MLP), 86% to 91% classification rate for Radial Basis Function (RBF) network, and 91% to 94% classification rate for probabilistic neural network (PNN) with unknown tea samples. In paper [228], Cynkar, Wies, et al. predicted the geographical origin of Tempranillo wines produced in Australia and Spain by using PCA, partial least squares discriminant analysis (PLS-DA), and stepwise linear discriminant analysis (SLDA) with full cross validation (leave-one-out method) to analyse the data collected by a mass spectrometry based electronic nose (MS-EN). The PLS-DA showed 85% accuracy rate while SLDA showed 86% accuracy rate. After analyzing the average fingerprint spectrum and its principal component scores in paper [162], Shi, Bolin et al. applied GA to select and optimize the effective sensors which can significantly contribute to identifying Xihu-Longjing tea from three producing areas and two tree species. Based on the e-nose data collected by Wedge et al. at 2009, Gromski, Piotr S et al. compared the classification accuracy of four pattern recognition algorithms, which include linear discriminant analysis (LDA), PLS-DA, random forests (RF) and SVM. They recommended that SVM with a polynomial kernel should be favoured as a classification method over the other statistical models that they assessed [229].

Usually, classification problems need to process a massive amount of data, and it's not easy to find useful information from these chaos data unless we try to grab the key information. That's why feature extraction is the most important phase in classification

problems. GA-based techniques have an advantage over statistical methods because they are distribution-free, i.e., no prior knowledge is needed about the statistical distribution of the data. GA also automatically discovers the discriminant features for a class. Considering the outlier effects and noise issues, some fuzzy based classification algorithms such as fuzzy c-mean, fuzzy art, fuzzy NN, etc. are also implemented in the e-nose field. Fuzzy neural network has a considerable improvement in e-nose classification performance compared to a common back-propagation network [66]. Paper [67] indicated that Fuzzy ARTMAP can be applied to discriminate three different e-nose datasets and the performance is better than back-propagation trained multilayer perceptron (MLP). Paper [68] enunciated the performance of fuzzy clustering c-mean (FCM) combined with SVM e-nose classifier is better than the other well-known machine learning algorithms. Moreover, [69] implemented fuzzy-wavelet neural network model to the e-nose field for food quality class evaluation and prediction.

In this dissertation, we present a novel e-nose classification method which is the combination of self-supervised genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM). GA was used to improve the classifier accuracy by optimizing the feature set and the optimal model parameters of FSVM. FSVM was adopted as the evolution criterion and the sequent odour classifier, which can reduce the outlier effects to provide robust and accurate classification.

## 5.2 Methodology

### 5.2.1 Genetic algorithm

Genetic Algorithm (GA) is a population-based global searching optimisation technique. According to Darwin's principle of natural selection, GA aims to find an individual with the best fitness value from the searching space composed of many feasible solutions. These solutions are represented by genotypes, and the genotypes with better fitness value may have a higher probability to be selected as parents for the upcoming generation. Crossover and mutation operators help the algorithm converge efficiently and hence the algorithm has less chance to get local optima than other methods. Also,

GA does not require any knowledge of the domain and any information about the structure of the problem to be optimized. It also works well even if the objective function is not smooth where the derivative methods cannot be applied. Therefore, GA can be regarded as a useful tool in feature selection. Fig. 5.1 depicts the basic evolutionary cycle of GA.

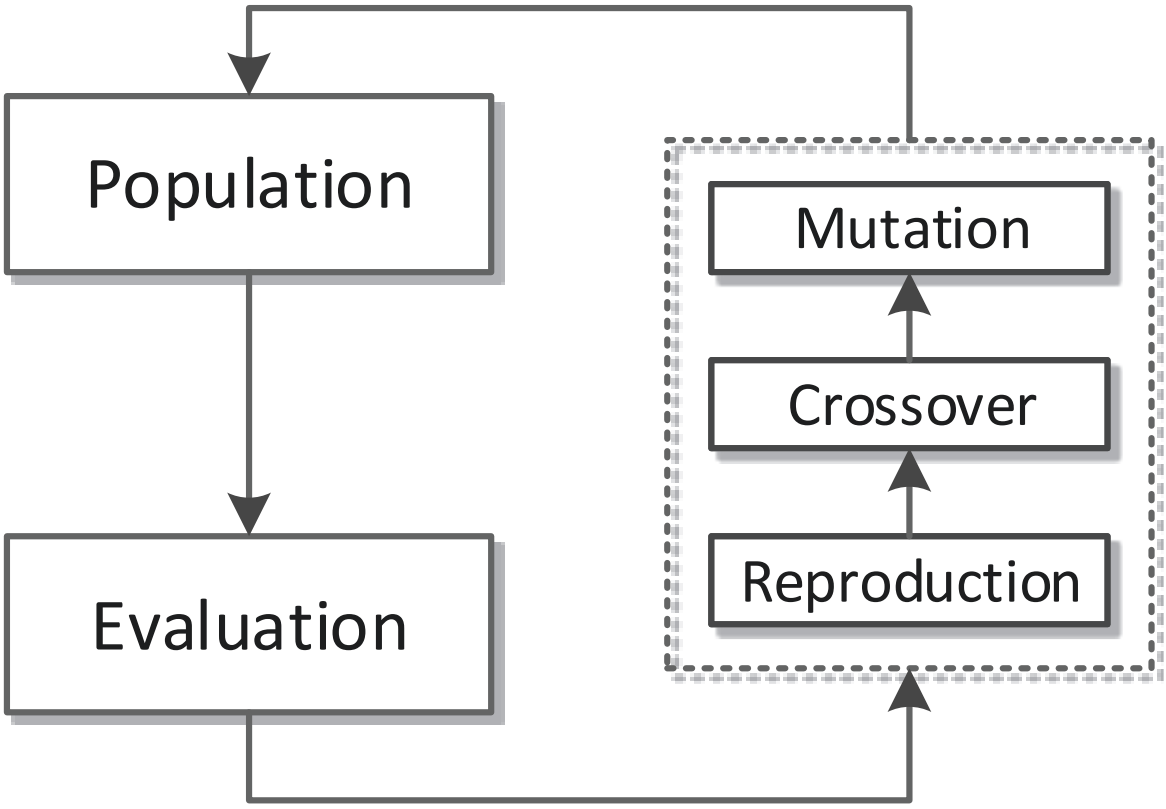


**Fig. 5.1** Basic Evolutionary Cycle of GA.

### 5.2.2 Fuzzy support vector machine

SVM is an effective algorithm in dealing with classification problems, but there are still some limitations of this tool especially in classifying real-world data. Some training samples are more meaningful than other data points. These important samples must be classified perfectly even if some noise or outliers are neglected. Fuzzy SVM (FSVM) applies a fuzzy membership function to every training data $p_n$, hence the input data

are transferred to fuzzy training samples, which are expressed as [70, 210, 230]:

$$\{(p_n, y_n, s_n | p_n \in \mathbb{R}^n), \sigma < s_n \leq 1\}, n = 1, ..., N, \tag{5.1}$$

where each training point $p_n$ is given a label $y_n \in \{1, -1\}$, $\sigma$ is a sufficiently small positive number. And the fuzzy membership $s_n$ is a function of time $t_n$:

$$s_n = f(t_n) = (1 - \sigma) \left( \frac{t_n - t_1}{t_n = t_1} \right)^2 + \sigma, \tag{5.2}$$

where $s_n$denotes the weight of the corresponding training point towards one class and $(1- s_n)$ is the weight of noise or less important points. Therefore, the hyperplane optimisation problem can be defined as [70]:

$$\min_{\mathrm{w},b,\xi} \frac{1}{2} \|\mathrm{w}\|^2 + cs^T \xi. \tag{5.3}$$

restrictions on condition to:

$$\begin{cases} y_n(\mathrm{w}^T p_n - b) \geq 1 - \xi_n \\ \qquad\qquad\qquad\qquad , n = 1, ...N. \\ \qquad\qquad \xi_n \geq 0, \end{cases} \tag{5.4}$$

A smaller $s_n$ decreases the influence of the parameter $\xi_n$, such that the corresponding sample $p_n$ is regarded less substantial. In a similar way as SVM, the Lagrangian multiplier function can be constructed as [70]:

$$\begin{aligned} \min_{\mathrm{w},b,\xi} \max_{\alpha,\beta} \{ \tfrac{1}{2} \|\mathrm{w}\|^2 + cs^T \xi - \textstyle\sum_{n=1}^{N} \alpha_n [y_n(\mathrm{w}^T x_n - b) \\ + \xi_n - 1] - \textstyle\sum_{n=1}^{N} \xi_n \beta_n \}. \end{aligned} \tag{5.5}$$

Then the optimal problem can be transferred to [70]:

$$\max_{a}\{\sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}(a_m a_n y_m t_n p_m^T p_n)\}$$
$$s.t. \begin{cases} 0 \leq a_n \leq s_n C \\ \qquad\qquad\qquad , n = 1, ...N. \\ \sum_{n=1}^{N} a_n y_n \end{cases} \tag{5.6}$$

The parameter $a_i$ can be solved by the Sequential Minimal Optimisation (SMO) technique, which was proposed by Haykin [211].

### 5.2.3 Proposed method

SVM is a powerful classifier, but there are still some limitations of this algorithm in the application of e-nose. In SVM, each training sample belongs to either one class or the other exactly, and all training samples are treated uniformly for each class. However, features extracted from all chemical sensors are affected by the noise and sensor drift, which leads to the overlapping of features with each other, as shown in the parallel coordinate plot (Fig. 5.2). In this case, each training sample no more belongs to one of the two classes exactly. It may belong to one class with a large probability and be meaningless with a low probability. It is required that the meaningful training samples should be classified correctly without caring about whether the meaningless samples are misclassified or not. Therefore, fuzzy membership associated with each training sample is introduced to solve this problem. Besides, some less discriminable features are existed in the feature set. This high-dimensionality may lower the classification accuracy. To solve the above mentioned problems, the proposed method applied the accuracy of FSVM as an evaluation criterion to select the most informative features. Besides feature selection, model parameters of FSVM can also be optimized simultaneously in genetic irritation. The detailed coding and decoding system, crossover and mutation operation, fitness evaluation, and system architecture for the proposed method are described as follows.

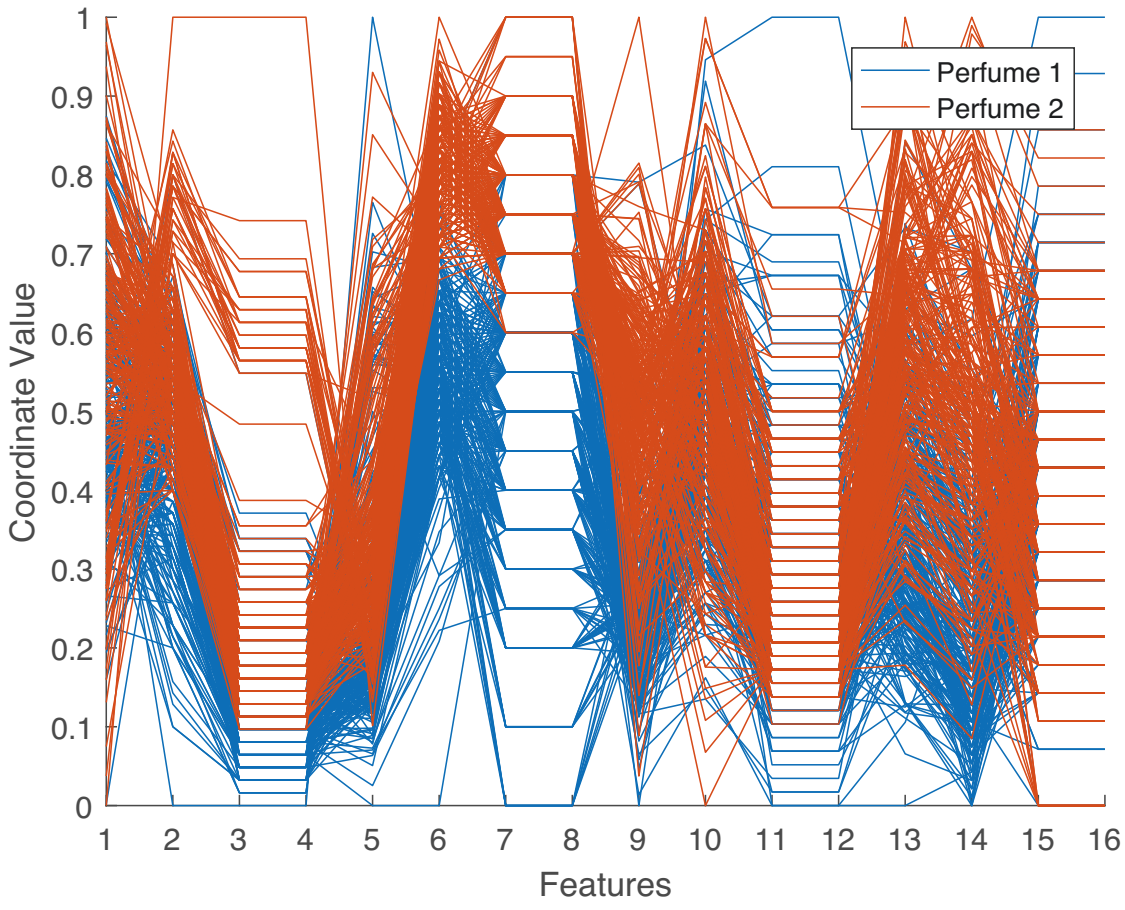


**Fig. 5.2** Parallel Coordinate Plot of Perfumes Feature Set.

### Coding and Decoding

In order to implement the proposed method, a coding and decoding system is required. A common method is to use a vector constituted only by zeros and ones as a genotypic representation of actual problems. The proposed method is able to select features and optimize classifier simultaneously. Therefore, the genotype consists of two parts: feature mask and FSVM parameters. Fig. 5.3 shows an example of the genotype of one individual.

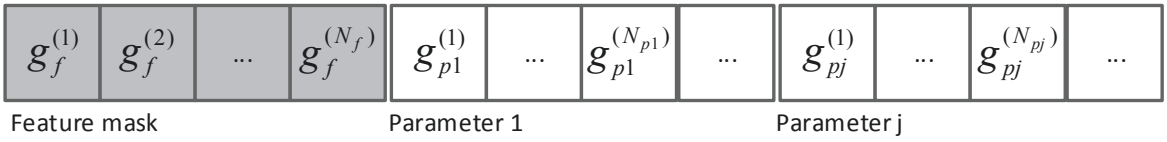


**Fig. 5.3** Genetic Constitution of an Individual.

Feature mask $g_f$ is a binary vector with a length of $N_f$, where a 0 or 1 at the $i-th$ position represents the absence or presence of the $i-th$ feature. When initializing the feature mask, the value of the $i-th$ position is decided by:

$$g_i = \begin{cases} 1, P(g_f^i = 1) = N_{selected}/N \\ \\ 0, P(g_f^i = 0) = 1 - N_{selected}/N \end{cases}, i = 1, 2...N_f. \quad (5.7)$$

Where $N_{selected}$ is a pre-set parameter and denotes the number of selected features. This strategy helps reducing features in genetic evolution.
The relationship between feature mask $g_f$ and its phenotype representation is:

$$\tilde{X} = X \cdot diag(g_f). \quad (5.8)$$

Where $X$ and $\tilde{X}$ denote the original dataset and the dataset after feature selection respectively.

$N_{pj}$ represents the genotypic length of the $j-th$ parameter of FSVM. For the RBF kernel function based FSVM, there are two parameters. The parameter C determines the complexity of the SVM model. The value of $\gamma$ affects the shape of RBF function. $N_{pj}$ is calculated by:

$$N_{pj} = round\left[log_2\left(\frac{\theta_{pj,upper} - \theta_{pj,lower} + \Delta}{\Delta_{pj}}\right)\right] + 1. \quad (5.9)$$

Where $\theta_{pj,upper}$ and $\theta_{pj,lower}$ represent the upper and lower bound of searching domain respectively. $\Delta_{pj}$ is a pre-set parameter and denotes the required precision.

The genotype $g_{pj}$ for parameters $j$ should be decoded into phenotype $\theta_{pj}$ by:

$$\theta_{pj} = \theta_{pj,lower} + (\theta_{pj,upper} - \theta_{pj,lower}) \left( \frac{\sum_{i=1}^{N_{pj}} (g_{pj}^{(i)})^{N_{pj}-1}}{2^{N_{pj}}} \right). \tag{5.10}$$

Where $g_{pj}^{(i)}$ denotes the value of the $i-th$ position of $g_{pj}$.

### Crossover and Mutation Operation

Homologue crossover operator is adopted in this research, which is a random mechanism for exchanging parts of genes between two genotypic individuals. The number of crossover points is decided by the number of optimisation objects. In mutation operation, the proposed method adopts two different strategies for FSVM parameters and feature mask. In terms of parameters optimisation, one gene may be altered randomly. A 0-valued gene will be changed to 1 or vice versa. In terms of feature selection, two genes may occasionally be selected and then swap their value, which may keep $N_{select}$ from change. Fig. 5.4 illustrates the crossover and mutation operation of the proposed method.

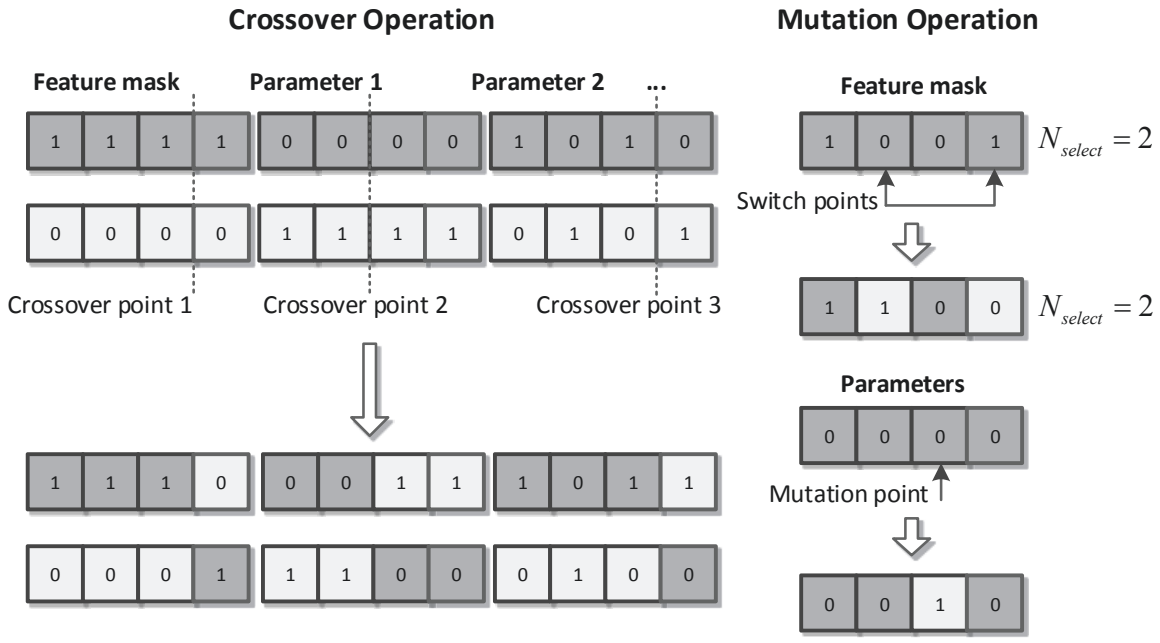


**Fig. 5.4** Crossover and Mutation Operation of Proposed Method.

**Fitness Evaluation**

The accuracy of FSVM is applied as an evaluation criterion. Considering about the unbalanced training set, the balanced accuracy is applied to evaluate the fitness of each GA individual:

$$fitness = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right). \tag{5.11}$$

Where True Positive ($TP$), True Negative ($TN$), False Negative($FN$) and False Positive($FP$) stand for the numbers of true positive, true negative, false negative and false positive respectively. For example, in the proposed two perfumes classification problem, for a given odour sample, $TP$ increase 1 when this sample was classified to perfume 1 and the true class of this sample is also perfume 1. $FP$ increase 1 when the sample was predicted as perfume 1 but the true class label is perfume 2. The weight is set to 1/2 because the two classes are equally important.

In order to evaluate the fitness of the current generation more rationally, the proposed method splits the training set into 10 equal size subsets and performs 10-fold cross validation (10-CV). Of the 10 subsets, one single subset, defined as subset B, is retained as the validation data, and the remaining 9 subsets consist of subset A as training data. The 10 statistical results (TP, TN, FN and FP) can then be averaged to calculate the fitness. Fig. 5.5 reveals the process of performing FSVM based fitness evaluation with 10-CV.

**System Architecture**

The architecture of the proposed method is presented in Fig. 5.6. This diagram is drawn mainly relying on the data flow between each component. Starting from the top-left corner of the diagram, the information of gas molecules was recorded by chemical sensor array and stored in a dataset. Then pre-processing techniques were adopted to ensure the high quality of the extracted features. In order to test the performance of the methodology, the dataset composed by features was hence divided into two parts, namely, training set and testing set. The GA-FSVM algorithm had only access to the

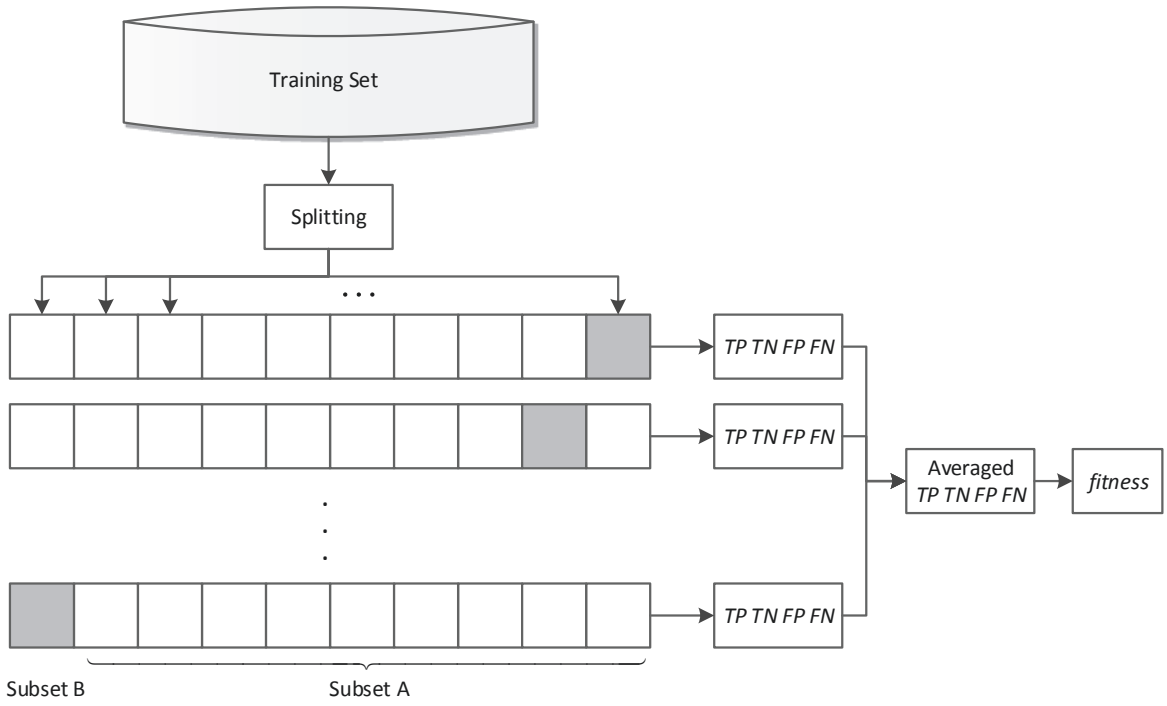


**Fig. 5.5** Fitness Evaluation using 10CV.

training set. After the evolution of several generations, the optimal feature selection rules and the best FSVM classifier were obtained as the outputs. The rules and the classifier were applied to the testing set. The classification accuracy of the testing set can be used to verify the performance of the algorithm. In the GA-FSVM module, the number of selected features should be set in advance, which would influence the initial population, as described in the previous section. The training set is split into subset A and B. These two sets perform like training and validation procedure in testing the classifier. Subset A is used to train a FSVM model with rough parameters and features, while subset B is used to tune the variables based on the fitness function.

## 5.3 Materials and Experiment

### 5.3.1 Experiment set up

500 test samples from two different perfumes, were prepared for testing. 250 samples for the perfume I, Hollister cologne; the other 250 samples were prepared for the perfume

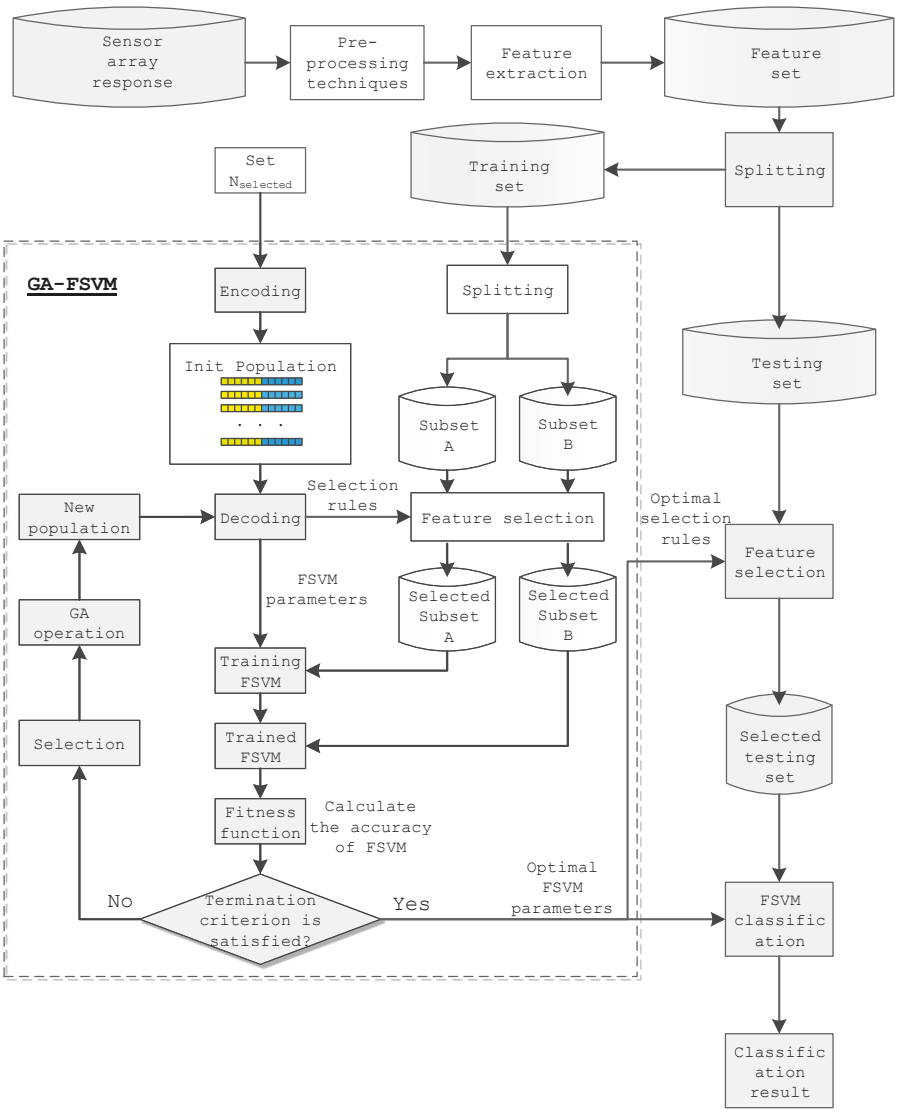


**Fig. 5.6** Algorithm Flow Chart.

II, Vera Wang Princess Perfume. For each sample, we used the perfume bottle to spray six times towards the 10mL volume headspace crimp top vial, then closed it. As Fig. 5.7 shows, the headspace was sampled and injected into the NOS.E gas input port by the sampling needle.

The basic test system diagram of our e-nose was shown in Fig 5.8, the power supply, NOS.E equipment, ventilation system, WI-FI communication port and NOS.E analyzer, a friendly user interface used to collect and analyze odour data for NOS.E. The headspace of the sample vial was sampled through the sampling needle to NOS.E

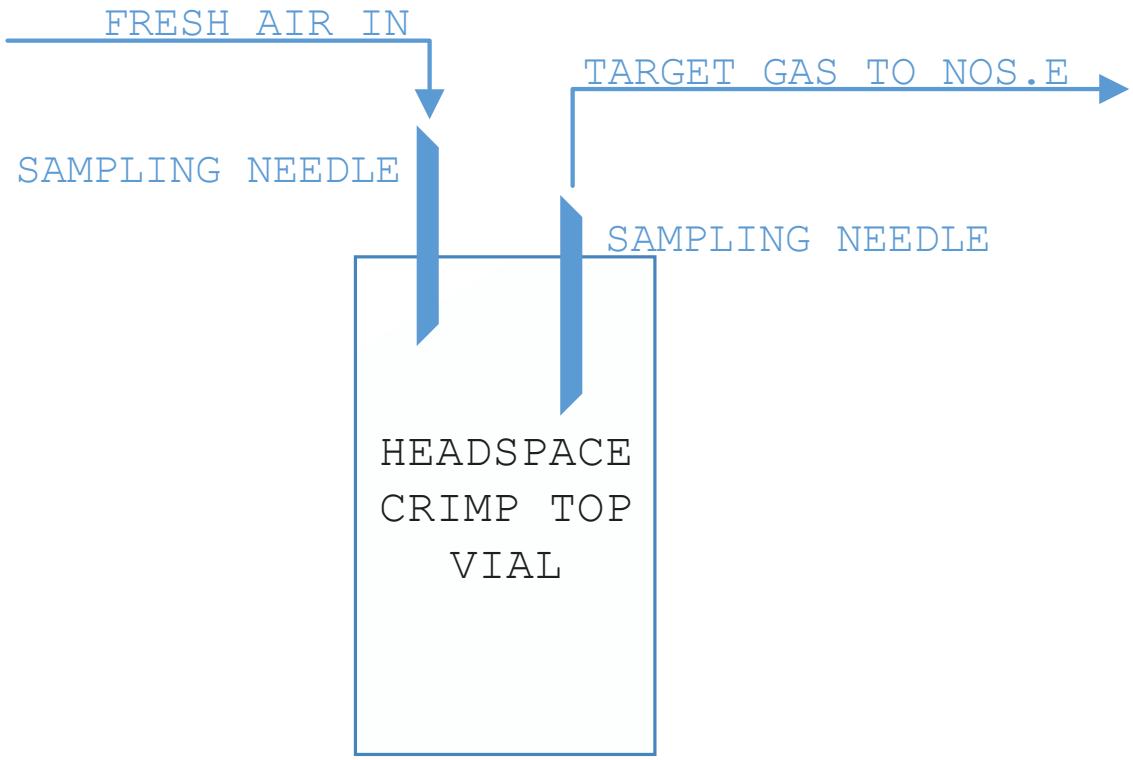


**Fig. 5.7** Samping Preparation.

**Table 5.1** Feature Types and Descriptions Used in Perfume Classification

| Type | Symbol | Description |
|---|---|---|
| Response | $x_m$ | The maximum response value |
| Time Parameter | $t_{dres}$ | Time interval form time of gas-in to time of maximum $1^{st}$ derivative in response stage |
| Time Parameter | $t_{dresx}$ | Time interval from time of gas-in to time of maximum $2^{nd}$ derivative in response stage |
| Integral | $intP_{res}$ | Integral from time of gas-in to time of peak |

gas input port. While the different perfume odours go into our equipment, 4 metal oxide gas sensors will produce the different responses. Each test lasted for 1220 seconds which included five phases, equipment preparation phase, baseline setup phase, test phase, baseline recovery phase and equipment rest phase. All the test data was stored as TXT format in the personal computer for the algorithms training.

The odour datasets used in this chapter were collected at 25 ℃ to 27 ℃ (ambient temperature) and 50% RH ambient humidity.

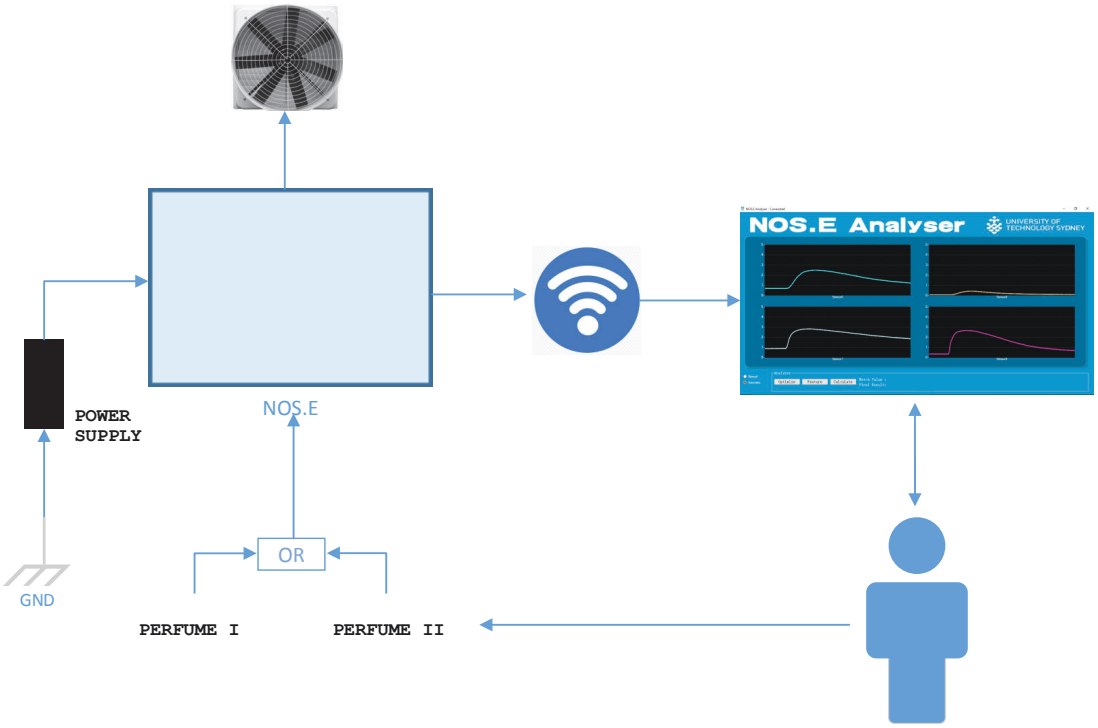


**Fig. 5.8** Test System Diagram.

**Table 5.2** Comparison of Different Feature Reduction Methods

| Method | Number of selected features | Number of total features | Maximum accuracy (%) |
|---|---|---|---|
| Proposed Method | 6 | 16 | 92.05 |
| PCA+FSVM | 11 | 16 | 79.31 |
| SAE+FSVM | 10 | 16 | 80.37 |

### 5.3.2 Pre-processing

The raw response obtained directly from the chemical sensor array cannot be used in the e-nose system due to its low signal-to-noise ratio (SNR) and the sensor drift. A pre-processing step is therefore needed.

Wavelet transform domain filter [231] is one of the most popular data pre-processing method in increasing SNR of e-nose system, it is used in this chapter to show the effectiveness of the proposed classification method. The used wavelet basis was 'daubechies 3', and decomposition level was set to 6. Fig. 5.9 shows an example of raw response curve and filtered result of the first sensor (MQ2) for Vera Wang Princess Perfume in one test.

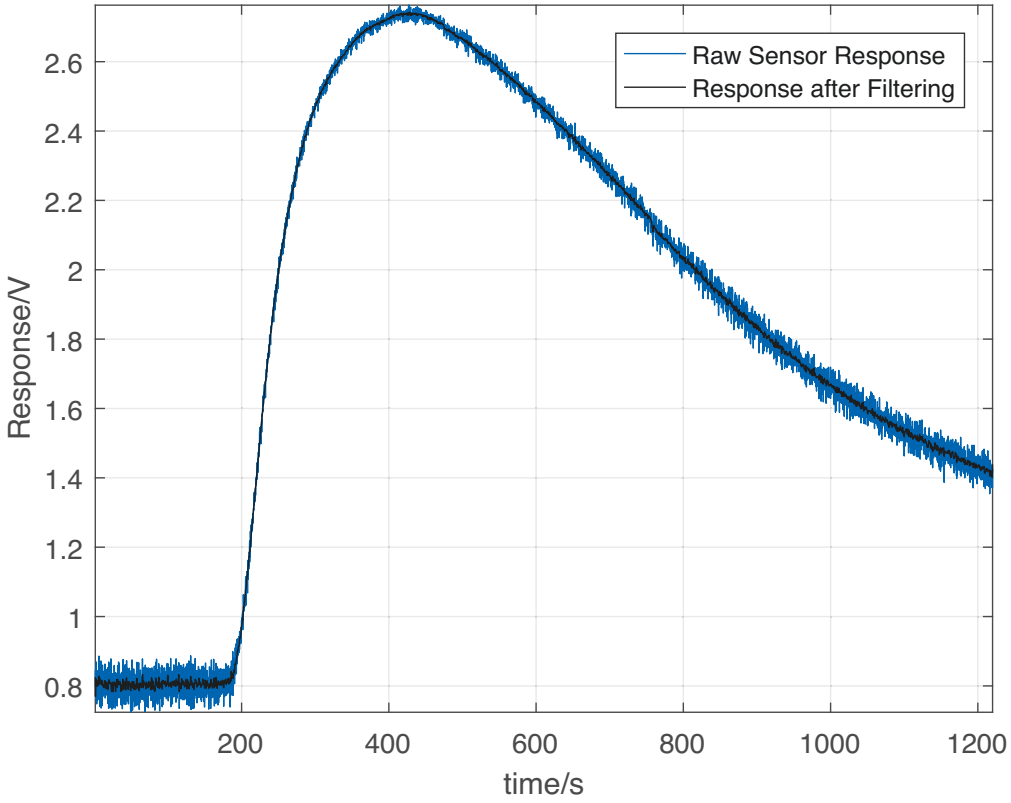


**Fig. 5.9** Raw Response and Filtered Curve of MQ2 Gas Sensor for Vera Wang Princess Perfume.

In order to reduce the influence of sensor drift, normalization is adopted in this chapter. This could eliminate the influence of concentration. The relative response is calculated by:

$$R(t) = \frac{(Y(t) - Baseline)}{Baseline}. \tag{5.12}$$

Where $Y(t)$ denotes the raw response of sensor.

The response of the gas sensors after filtering and normalizing are presented in Fig. 5.10. It is obvious that the baseline of each sensor is in the same level. In addition, these curves are smoother than the raw data.

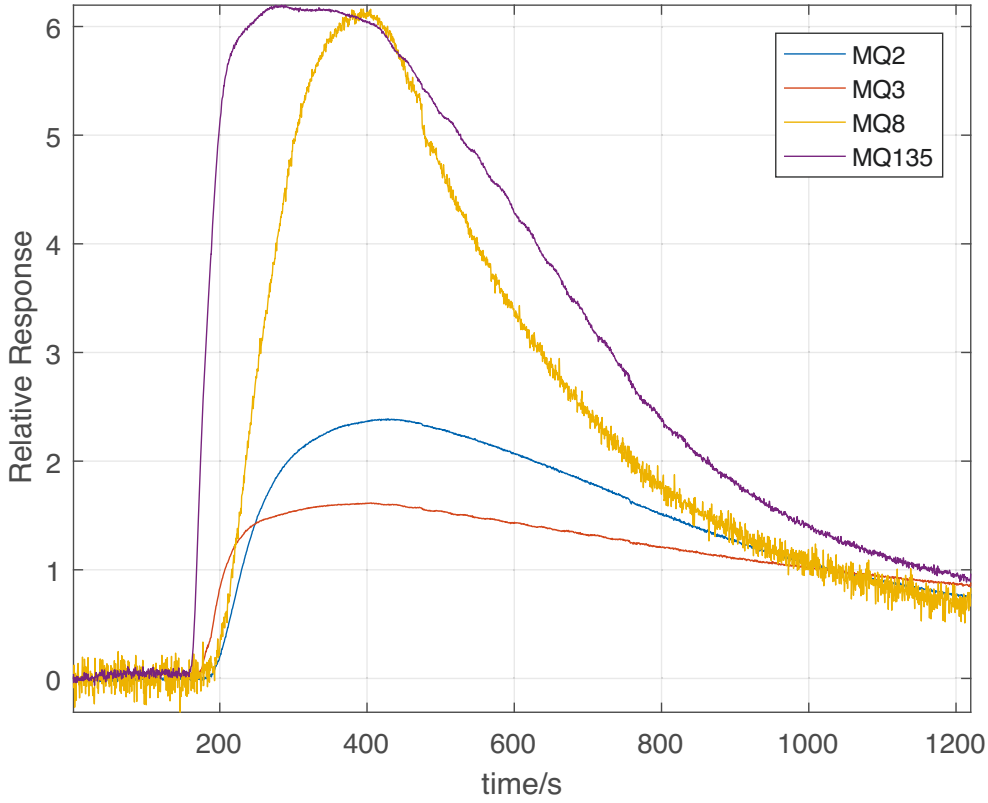


**Fig. 5.10** Sensor Responses after Pre-processing.

### 5.3.3 Feature extraction

Features extracted from gas sensor's steady state response are widely used in odour classification. In this study, traditional features according to the characteristics of the response are chosen, as shown in Table 5.1. Four features were extracted from the response of one sensor, hence in this experiment, the number of the feature set is 16.

## 5.4 Comparison Results

### 5.4.1 Feature reduction

Two feature reduction techniques were chosen to compare with the proposed hybrid approach by averaging the result of 500 tests.

Principal Component Analysis (PCA) with SVM is a classical combination in the field of odour classification. PCA reduces the interrelated variables and retains the most informative principal components as selected features. The state-of-the-art feature reduction method is Sparse Auto-Encoder (SAE) Neural Network. SAE compresses the input features to the comprehensive interlayer neurons.

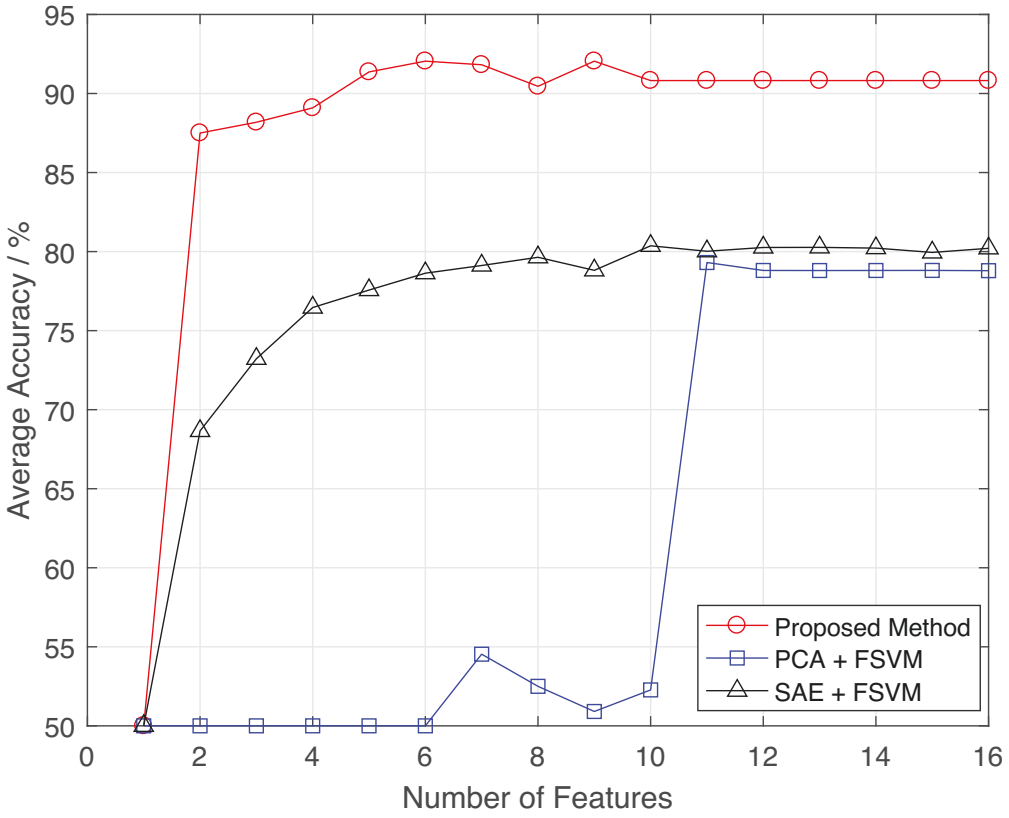


**Fig. 5.11** Relation Curves of Mean Accuracy with Number of Selected Features in 500 Tests.

**Table 5.3** Details of Selected Features

| Selected features | Sensor | Description |
|---|---|---|
| Feature 5 | MQ3 | Integral from time of gas-in to time of peak |
| Feature 6 | MQ3 | The maximum response value |
| Feature 7 | MQ3 | Gas-in to time of maximum $1^{st}$ derivative in response stage |
| Feature 8 | MQ3 | Time interval from time of gas-in to time of maximum $2^{nd}$ derivative in response stage |
| Feature 9 | MQ8 | Integral from time of gas-in to time of peak |
| Feature 14 | MQ135 | Integral from time of gas-in to time of peak |

The same FSVM classifier is used in this section. The proposed method is able to select features and optimize the classifier at the same time. Therefore, in the case of

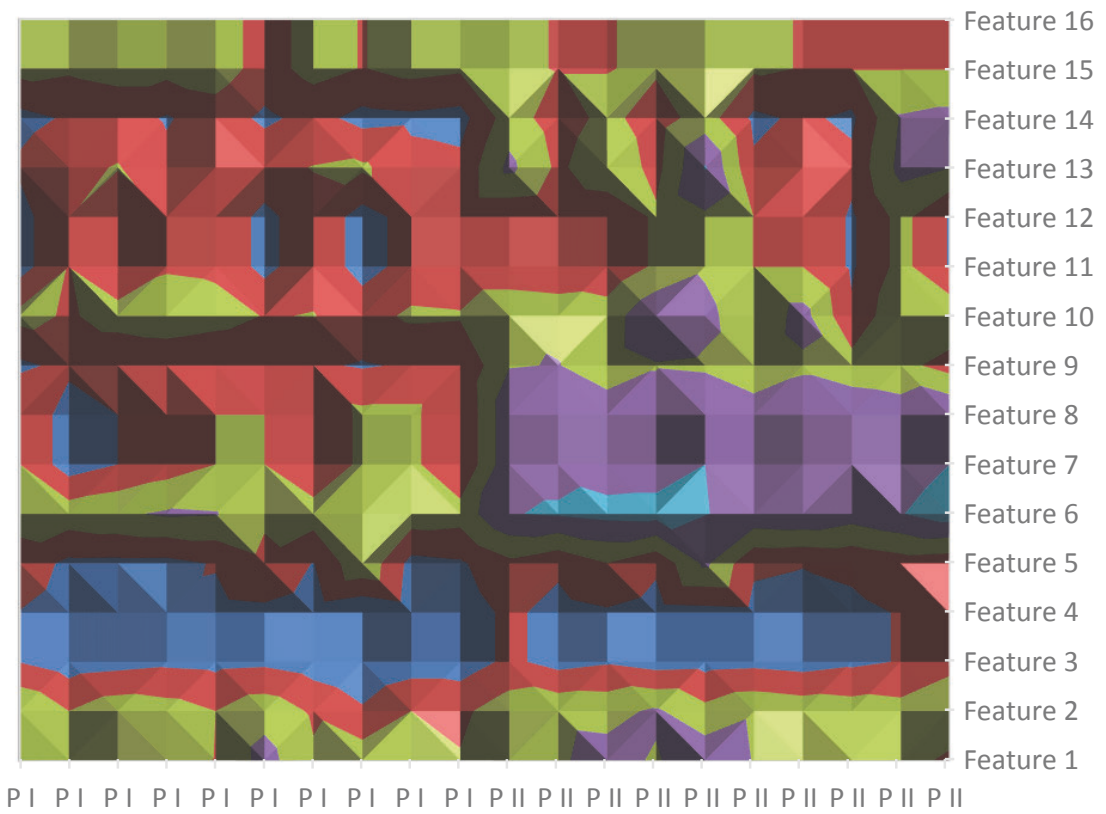


**Fig. 5.12** Features for Two Perfumes.

PCA and SAE, another GA is applied to find the best parameters for each classifier respectively.

The comparison of different feature reduction methods is shown in Table 5.2. Fig. 5.11 shows the relation curve of mean accuracy with the number of features. The accuracy curve of the proposed method reached a peak at 92.05%, where the number of selected features is 6. Then the accuracy curve gradually declined and stabilized at 90.81%. The SAE based algorithm worked well while PCA performed worst in this perfume classification task. The specific information of selected features is listed in Table 5.3. Most features are extracted from sensor 2 (MQ3), while sensor 1 (MQ2) have no contribution to the accuracy. In addition, Integral is the most effective feature in odour classification.

In order to verify the correctness of the proposed method in selecting informative features, we used a visualization technique to reveal the change of feature space, whose

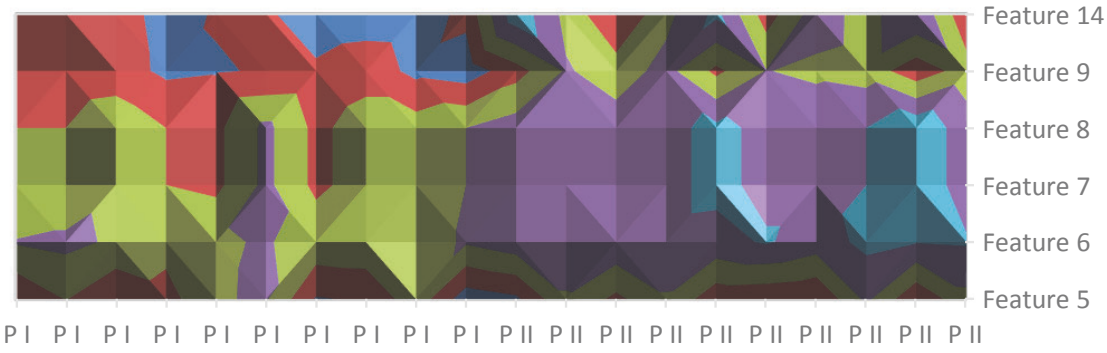


**Fig. 5.13** Features Selected for Two Perfumes.

**Table 5.4** Comparison of Different Algorithms for Perfume Classification Problem

| Method | Training samples | | Test samples | | Mean Accuracy In 500 Tests (%) |
|---|---|---|---|---|---|
| | Perfume I | Perfume II | Perfume I | Perfume II | |
| KNN[232] | 125 | 125 | 125 | 125 | 51.37 |
| SVM[233] | 125 | 125 | 125 | 125 | 86.04 |
| FSVM | 125 | 125 | 125 | 125 | 90.81 |
| PCA+SVM | 125 | 125 | 125 | 125 | 63.47 |
| PCA+FSVM | 125 | 125 | 125 | 125 | 79.31 |
| GA + SVM[234] | 125 | 125 | 125 | 125 | 89.82 |
| SAE + FSVM | 125 | 125 | 125 | 125 | 80.37 |
| Proposed Method | 125 | 125 | 125 | 125 | 92.05 |

results are shown in Fig. 5.12 and Fig. 5.13. Fig. 5.12 presents all 16 features of randomly chosen 10 samples of each perfume respectively, and Fig. 5.13 shows the feature selection result. Feature values have been normalized and quantized to different colours. It is obvious that the selected feature space is more separable than the original one.

### 5.4.2 Classification

The proposed hybrid approach is compared with other classification methods by averaging the results of 500 tests. The comparison result is listed in Table 5.4.

The proposed method has the highest average classification accuracy compared with other algorithms. This accuracy derives from two aspects: (a) when $N_{selected}$ was set to 16, the accuracy represents the performance of FSVM classifier, according to Table 5.4, FSVM is more precise than SVM in our experiment; (b) GA selected the most informative features and provided the best model parameters to FSVM.

## 5.5 Conclusions

Odour classification is a challenging machine learning problem with a high dimensional singularity in the feature set and imbalance of datasets in the present outliers and noise problem. We have shown that the proposed odour classification algorithm is able to handle the feature selection and datasets imbalance issues and practically even for small datasets. Implementing the proposed odour classification method to other e-nose applications is currently under consideration.

# Chapter 6

# A New Fast Response Electronic Nose Health Monitoring System

## 6.1 Introduction

The study of exhaled breath for health diagnosis and monitoring is becoming an increasingly popular area of research. Unlike most traditional health monitoring and diagnosis through the use of bodily fluids (e.g. through blood, urine), breath collection is non-invasive and convenient. Furthermore, studies have shown there are many metabolomic compounds in human breath that could be used for health monitoring and disease diagnosis [37, 235–239]. Metabolomics is influenced by an individual's lifestyle, diet, and the environment. These factors make breath analysis an attractive option for personalised health care.

Previous works have shown human breath can be used in monitoring and diagnosing conditions such as lung cancer [37, 235, 240], diabetes [237, 241], urinary track infections [242], asthma [168], malaria [243], etc. Many of these conditions are becoming more prevalent in the modern society due to aging populations and increased pollution [244, 245]. For example, researchers from the U.S. Centers for Disease Control and Prevention (CDC) have shown that about 1 in 7 older U.S. adults have lung diseases such as asthma or chronic obstructive pulmonary disease (COPD). Lung obstruction is the third leading cause of death in the United States [246].

Many current breath analysis studies are based on gas chromatography mass spectroscopy instruments (GCMS), which provide a vast amount of information to researchers and clinicians [243, 247–250]. However, GCMS is bulky, expensive and requires expert knowledge for operation; making it unsuitable for personalized health care. Electronic nose (e-nose) instruments are an attractive alternative to GCMS because they have smaller dimensions and are lighter weight. E-nose systems mimic natural olfactory systems by using a combination of gas sensors with overlapping and distinct responses to different input odours. These signals from the sensors are collected by a data acquisition system, and then pre-processed. Data preprocessing can take many different forms such as baseline correction, normalization, feature extraction, etc. Modern machine leaning techniques can then be employed to identify the input gases. Fig. 6.1 shows a block diagram of a typical e-nose system.

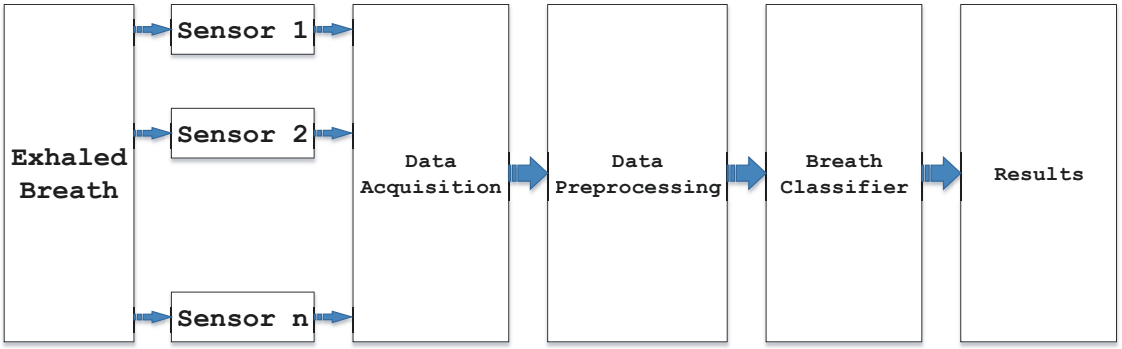


**Fig. 6.1** Basic Block Diagram of Electronic Nose.

We present here a new e-nose system that has a fast response time of under three minutes, making it a suitable system for rapid diagnosis. As a proof of concept we analyzed breath samples from one healthy adult in three different conditions: normal, smoking and alcohol consumption, and used a well-known classifier to discriminate between the classes.

## 6.2 Electronic Nose System

We designed and developed a new e-nose system called “NOS.E” for monitoring exhaled breath (Fig. 6.2). This system contains a PAD-based control and display panel, an

air-intake system, power system, a gas-sensor array, and a controller unit all housed inside a handheld case. Fig. 6.3 shows the handhold part which was used as the breath input port. The sensor array is composed of four commercially available metal oxide gas sensors: *TGS 2600, TGS 2602, TGS 2603*, and *TGS 2620*. Table 6.1 lists the sensitivity characteristics of these sensors. The input gases are analyzed by the sensor array and then converted to digital data, and transmitted to the controller unit for data processing.

**Fig. 6.2** UTS e-nose System "NOS.E".

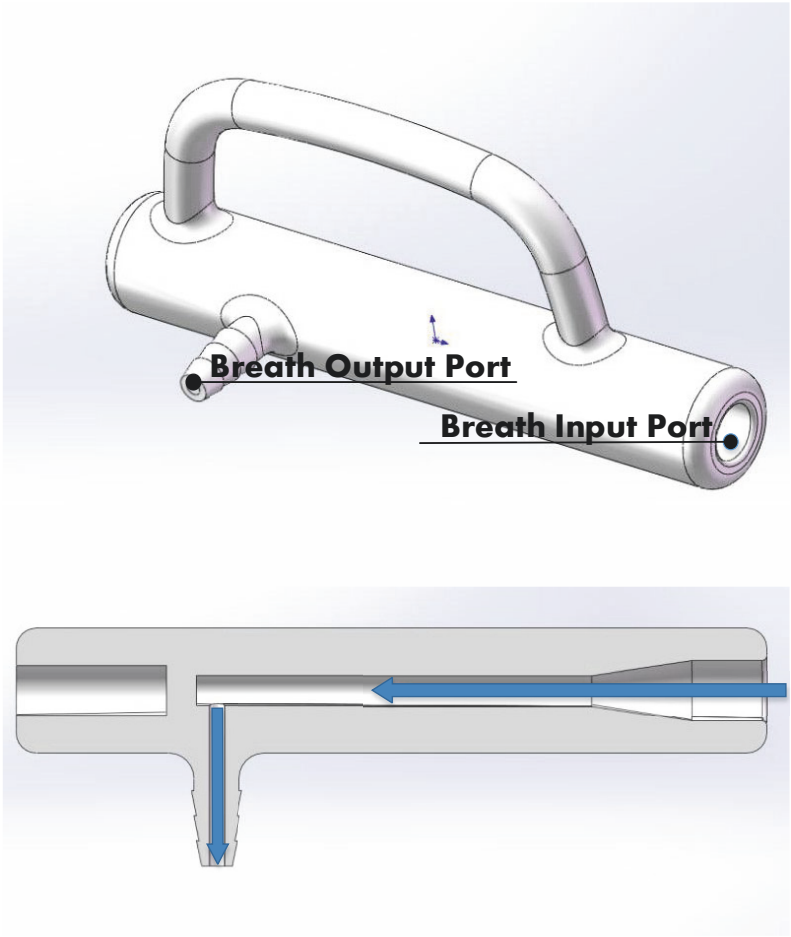


**Fig. 6.3** Breath Input Port of the NOS.E.

Table 6.1 Gas Sensors Sensitivity Characteristics

| Sensor Name | Target Gases |
|---|---|
| Sensor 1: TGS 2620 | Ethanol, Hydrogen, Iso-butane, CO, Methane, etc. |
| Sensor 2: TGS 2602 | VOCs, Ammonia, $H_2S$, etc. |
| Sensor 3: TGS 2600 | Hydrogen, Ethanol, Iso-butane, etc. |
| Sensor 4: TGS 2603 | Trimethylamine, Methyl mercaptan, etc. |

## 6.3 Data Processing

### 6.3.1 Feature extraction

Feature extraction of gas sensor steady-state responses are widely used in odour classification. Applications of feature extraction methods generally fall into two categories: manual extraction using expert knowledge and automatic methods that are data focused. In this study, we chose to extract the features manually as a proof-of-concept for the instrument. We chose four features common to many other works [64, 65]: maximum response value ($x_m$), time interval between gas-in and maximum $1^{st}$ derivative of response ($t_{dres}$), time interval between gas-in and maximum $2^{nd}$ derivative of response ($t_{dresx}$), and integral between time of gas-in and time of peak ($intP_{res}$).

### 6.3.2 SVM

To verify the performance of our NOS.E system, we chose one of the most popular classification algorithms, *SVM*, to identify the three different exhaled breath conditions. *SVM* is an effective algorithm in dealing with classification problems. An *SVM* learns a discriminant function that separates positive and negative examples with the maximum margin. The objective of *SVM* is to find an optimal hyperplane to separate two different classes of samples. The equation of hyperplane is in the form of:

$$\omega^T x + b = 0, \tag{6.1}$$

where $\omega$ and $b$ represent the weight vector and bias, respectively.

Nonnegative slack variables $\xi = (\xi_1, \xi_2, ..., \xi_M)$ are introduced to measure the misclassification degree of the training samples. Therefore, the mathematical formula

of optimisation problem can be deduced by solving:

$$\min_{\mathrm{w},b}\{\tfrac{1}{2}\|\mathrm{w}\|^2 + C\sum_{i=1}^{M}\xi_i\}$$
$$s.t.\begin{cases} y_n(\omega^T x^{(i)} + b) \geq 1 - \xi_i \\ \qquad\qquad\qquad\qquad , i = 1, 2, ...M. \\ \xi_i \geq 0 \end{cases} \tag{6.2}$$

Where the parameter $C$ is an error penalty term, which determines the influence of the misclassification on the objective function. Increasing $C$ will give more importance to the errors on the training set in determining the optimal hyperplane. In this study, by using genetic algorithms ($GAs$) we set the value of $C$ as 901.

Instead of calculating distances between the transformed patterns in the new feature space, distances can still be measured in the original space with the introduction of the so-called kernel functions $K : (x_1, x_2) \rightarrow K(x_1, x_2)$. Normally, we defined the kernel instead of getting it from transformations, because it is much easier to work with kernel over the original low dimensional data space; it can be proven that the final classification function depends on the kernels only. In this study we use the Gaussian ($RBF$) kernel:

$$K(x_1, x_2) = exp\left(\frac{(x_1 - x_2)^2}{2\sigma^2}\right). \tag{6.3}$$

We apply GAs to select $\sigma$=6.01 as the optimal value for the parameter of RBF kernel.

The final classification function, effectively implementing the maximal margin hyperplane in the feature space, is:

$$F(x) = sign\left(\sum_i y_i \alpha_i K(x, x_i) - b\right), \tag{6.4}$$

where the $\alpha_i$ are the solutions of the dual problem and are non-zero only for a subset of vectors $x_i$ called support vectors [251].

## 6.4 Experimental Setup and Results

### 6.4.1 Experimental setup

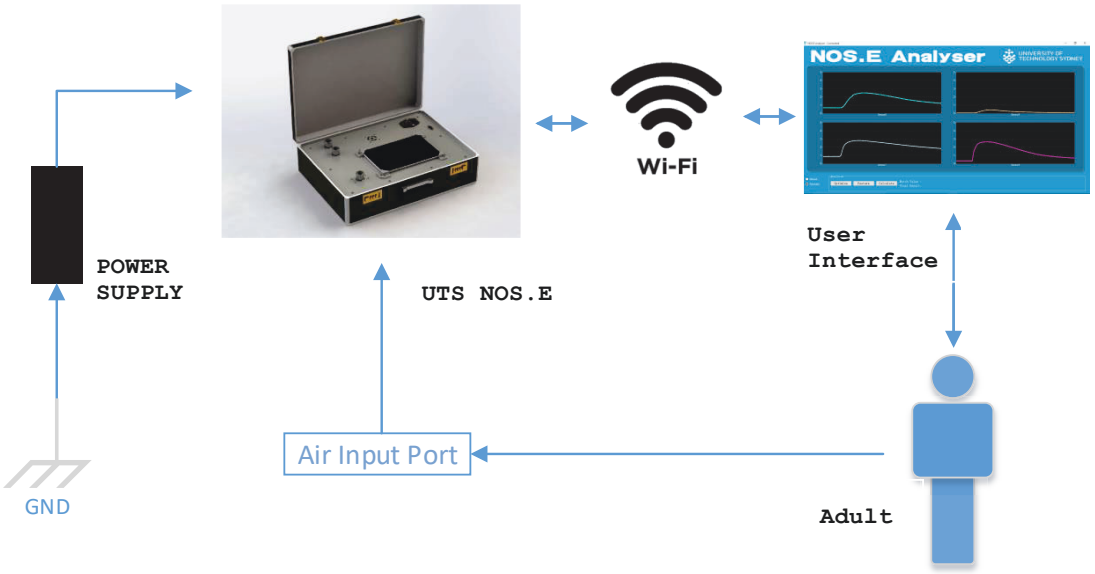


**Fig. 6.4** Test System Diagram of UTS NOS.E.

Sixty total exhaled breath samples (three different scenarios) were analyzed by the NOS.E system: 20 normal breath; 20 exhaled breath after smoking; and 20 exhaled breath after alcohol consumption. One volunteer was used for consistency and was instructed to breathe through an air input port to the NOS.E equipment, which then analysed the sample. The volunteer was asked to breath out using a "ha" expiration 12 times for each sample.

All samples were collected within seven hours. The set of samples from each scenario took a total of 60 minutes to collect. Normal breath samples were collected first, followed by smoking, and then alcohol consumption. Between normal and smoking samples, the equipment was allowed to run a blank for 115 minutes in order to ensure no sample was left in the air system. The volunteer spent 5 minutes smoking one cigarette just before smoking samples were taken. Between the smoking and alcohol consumption samples, the equipment was cleaned for 110 min by allowing room air to

flow through the system. The volunteer drank $100mL$ of red wine in 10 minutes before the last set of breath tests were conducted. Our experimental setup here means the samples taken for each scenario should be very similar, we emphasise here that the aim of this preliminary experiment is to show initial proof-of-concept for our e-nose system, more rigorous trials will be scheduled in the next phase.

Fig. 6.4 shows the basic test diagram which comprised of the power supply, NOS.E equipment, Wi-Fi communication unit and NOS.E analyzer (a user friendly interface used to collect and analyze odour data for NOS.E.). All the data were stored and later analyzed for classification.

Example responses of the gas sensor array under the three proposed scenarios are shown in Fig. 6.5. The odour datasets used in this chapter were collected under 25 ℃ to 27 ℃ ambient temperature and 60% RH ambient humidity.

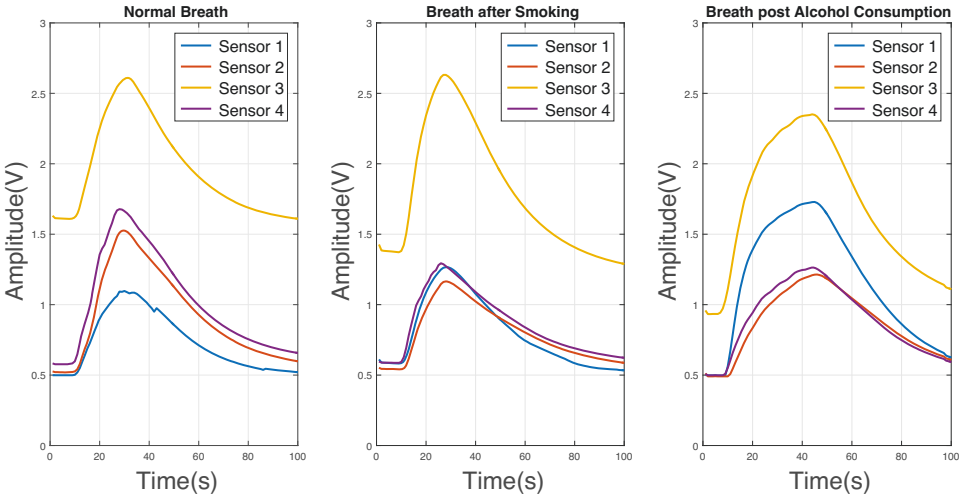


**Fig. 6.5** Example responses from *NOS.E* for the three scenarios we tested.

## 6.4.2 Cross validation

We used a five fold balanced cross validation method for classification. The full dataset is partitioned into training and test sets by randomly selecting 20% of the data from each scenario to form the test set and the remaining for the training set. This process was performed five times: where each data sample appears exactly once in a test set.

SVM was run as a binary classifier as well as a multi-class classifier. In the former case, two scenarios were tested against each other, thus yielding three classification problems: normal vs. smoking, normal vs. alcohol, and smoking vs. alcohol. The RBF kernel ($C = 901, \sigma = 6.01$)is implemented in the binary SVM calssifier. In the multi-class classification, one-versus-one (*1V1*) strategy is applied. It evaluates all possible pairwise classifiers and thus induces $k(k-1)/2$ (where $k$ is the number of classes) individual binary classifiers. Applying each classifier to a test example would give one vote to the winning class. A test example is labeled to the class with the most votes [252].

### 6.4.3 Results

Table 6.2 shows the results of the four classifications performed. From the table, it was evident that when separating breath samples post alcohol consumption with normal and smoking breath samples, 100% and 95% classification accuracy were achieved, respectively. The accuracy of identification for the normal breath and smoking breath samples lies on 97.5%. Moreover, for the multi-classification task, we can identify these three breath types with 98.334% classification accuracy.

According to these preliminary test results, the NOS.E system can identify different human breath under three different common conditions, which means the proposed solution has certain possibilities to implement to monitor the human health conditions with well-trained data processing algorithms. In the next step, we will schedule more test scenarios to improve the NOS.E system design and the performance of the human breath classifier.

Table 6.2 Classification Accuracy of three Breath Scenarios

| Breath Scenarios | SVM Classification Accuracy |
|---|---|
| Normal *VS* Cigarette | 97.5% |
| Normal *VS* Wine | 100% |
| Wine *VS* Cigarette | 95% |
| Normal *VS* Cigarette *VS* Wine | 98.334% |

## 6.5 Conclusion and Future Works

We present here a new fast response electronic nose system, NOS.E, which can analyse a sample in under three minutes. We tested our system with some initial human breath sample from one healthy volunteer in one day under three different metabolomic scenarios. Using simple features from the responses and support vector machine classifier, we were able to easily distinguish the breaths from each scenario. We achieved high classification accuracy of over 95% in all comparisons. This initial study showed that our e-nose system can detect changes to odour components in exhaled breath and could be in the future applied in the human health monitoring area.

For future work, we will recruit a small group of healthy volunteers and collect samples under a variety of scenarios and at different times, which will allow us to test NOS.E's capability. Ultimately, we aim to build and test an e-nose system that can be used to monitor human health through breath.

We envisage the NOS.E system can be deployed as part of a cloud based health monitoring system (shown in Fig. 6.6). The hardware e-nose equipment will be installed on site to monitor the exhaled breath of patients. The data collected can be uploaded to a remote server, which will monitor or diagnose the health conditions. Users can check these data and control the e-nose equipment through a cell phone or personal computer. Furthermore, health professionals can also monitor the same data and provide their feedback to the users. This cloud based comprehensive system would allow continuous and efficient monitoring of users' health. By introducing doctors and the remote server in this new e-nose health monitoring system, its application could be expanded to medical areas such as general health screening and disease diagnosis. Moreover, this new system can also bring a lot of benefits for the users by saving a lot of time and improve the diagnosis efficiency.

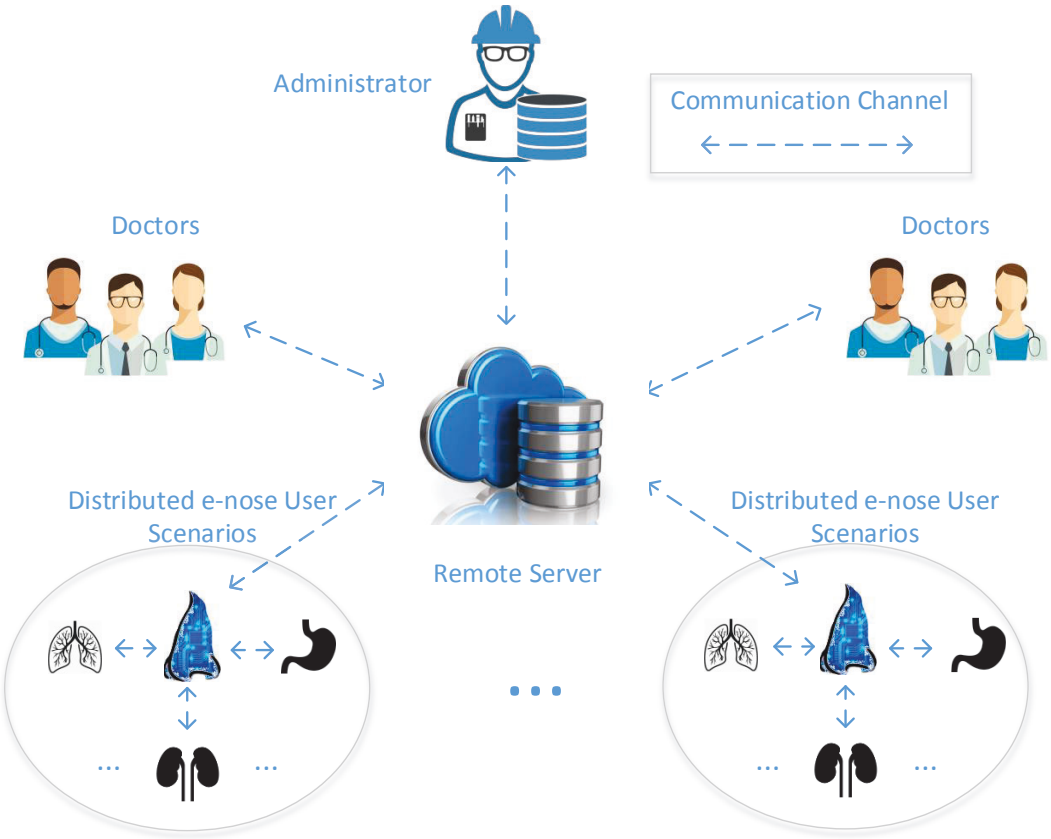


**Fig. 6.6** Remote Server based Electronic Nose System.

# Chapter 7

# Development of the New Electronic Nose for Wildlife Products Identification

The scale of illegal wildlife trade has rapidly increased in recent decades due to the huge profit margin. At the same time, this trend is also threatening biodiversity and accelerating the extinction for many species (such as rhinoceros, elephants, sea turtles etc.). In order to combat the illegal wildlife trade, law enforcement has collaborated with researchers to apply several strategies for inspecting and identifying legal from illegal wildlife parts. Generally, the suspect wildlife parts are analysed in the laboratory by using isotopic fingerprinting or DNA analysis. However, these tests are time-consuming, complex and cannot be conducted on-site. Although wildlife detection dogs are used to detect illegal wildlife parts on-site in some countries, there are many factors such as behavioural and individual deviations that can affect these results. Based on the electronic nose (e-nose) prototypes developed by the NOS.E team (University of Technology Sydney), this chapter proposes a new e-nose prototype (NOS.E II) which uses an efficient and reliable method to identify illegal wildlife parts. Novel mechanical and airflow designs as well as kernel based data preprocessing methods were implemented in the NOS.E II to improve its sensitivity and portability. As a

proof-of-concept test, water buffalo horn and rhinoceros horn samples were selected as the test targets to identify legal from illegal wildlife parts.

## 7.1 Introduction

Illegal wildlife trade has rapidly developed into an international crisis due to the enormous amount of contraband involved in this area. The value of the international organised environmental trade in flora and fauna (excluding fisheries and timber) is estimated to be between 7 and 23 billion USD annually [84, 253–255]. These illegal activities are endangering biodiversity and driving many species towards extinction [256–259]. The main problem in the enforcement of the illegal wildlife trade is the lack of an efficient and reliable solution to identify legal from illegal wildlife parts, such as water buffalo horn from rhinoceros horn [253].

Generally, laboratory analysis such as isotopic fingerprinting or DNA analysis is required to confirm the identity of suspect wildlife parts which can not be identified during the inspection activities [84, 254, 260, 261]. However, these tests are time consuming, destructive and cannot be conducted on-site. A non-destructive headspace solid phase microextraction (HS-SPME) coupled with comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC×GC-TOFMS) method has been proposed to identify different species of wildlife products like rhinoceros horns and ivory samples [84, 262]. However, it also requires laboratory analysis as the GC×GC-TOFMS instrumentation is non-portable, expensive and requires expert knowledge for operation. Although wildlife detector dogs are sometimes employed on-site, there are natural interindividual variations in their response to odours [2, 263]. Moreover, there are individual deviations in the olfactory perception based on genetic differentiation [2]. Nevertheless, the GC×GC-TOFMS and detector dog methods have already proved the effectiveness of detecting and identifying wildlife contraband by analysing the volatile organic compounds (VOCs) or odours of the target items [84]. Therefore, electronic nose (e-nose) technology, which is also based on the detection of these volatile compounds, is introduced in this chapter as an alternative and cost-effective solution for wildlife products identification [43–45].

E-nose technology is capable of detecting chemical compounds by sensing and analysing odour molecules. As a type of machine olfaction, e-nose technology has the potential to be applied in different fields [1, 3–6]. Inspired by the knowledge of the mammals olfactory system (especially for the canine olfactory system) [2, 75, 79, 80], Figure 7.1 shows the similarities between the canine olfactory system and the e-nose system. A sensor array is used to mimic the olfactory receptor (OR) to transform the odorants information to the electrical signals. These signals can be processed by the "olfactory bulb" of an e-nose which is the micro-controller. Moreover, to recognise the odorants information, a computer is used to represent the brain, and the pattern recognition algorithms are the "content of the brain" in this computer.

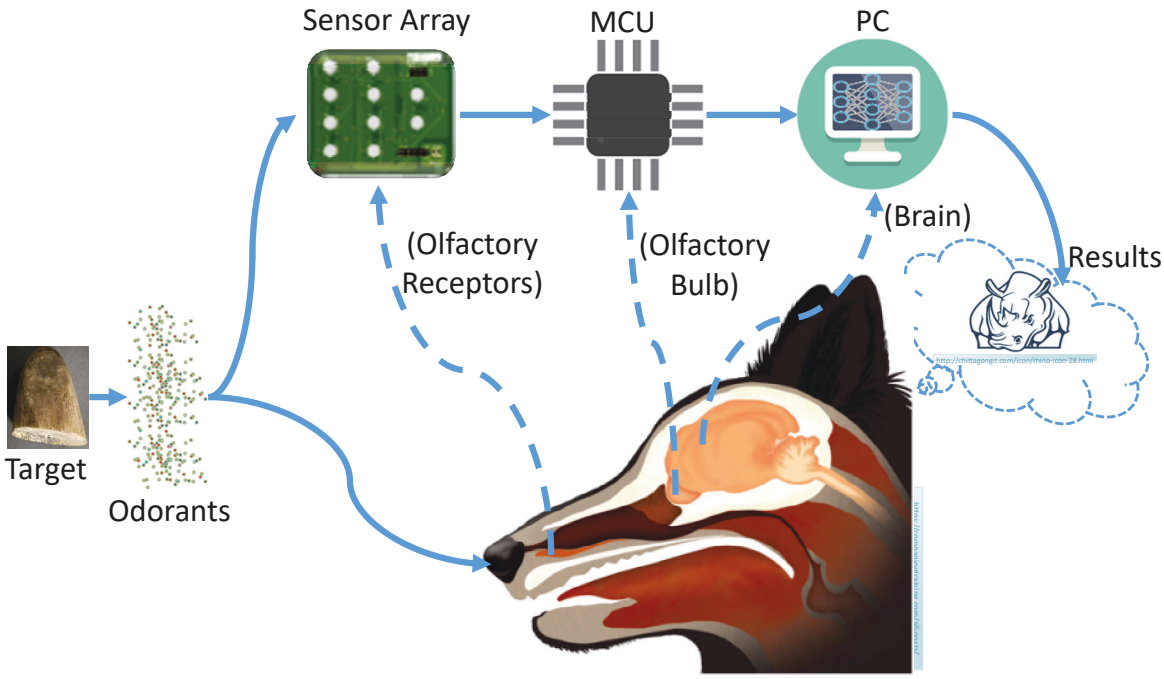


**Fig. 7.1** The similarities between the canine olfactory system and the e-nose system.

This study aims to develop an efficient and reliable e-nose system (named NOS.E) for the identification of illegal wildlife parts. As a proof of concept test, this study uses water buffalo horn and rhinoceros horn samples as the test targets to distinguish legal from illegal wildlife parts. Rhinoceros horns were selected as they are difficult to distinguish morphologically (often sold as powders, small fragments and sculptures). Moreover, they represent one of the most endangered mammals globally, and the international trading of rhinoceros horns are strictly prohibited as they are listed under

CITES (the Convention on International Trade in Endangered Species of Wild Fauna and Flora) appendix I (and II in the case of C. s. simum South African and Swaziland populations) [84, 264–266]. Water buffalo horn is a common "fake" marketed and improperly sold as rhinoceros horn [84, 267]. The NOS.E prototype also has potential to be used globally to differentiate a diverse range of trafficked species including large cats (e.g. tigers, leopards), elephants (i.e. ivory), pangolins, bears, sea turtles, sharks, and a range of exotic birds and reptiles, all of which have distinct odour signatures [253].

## 7.2 Equipment Design

### 7.2.1 Sensor chamber design

In a mammalian olfactory system, the air inhalations play an important role [80, 139, 140]; and the relationship among the nasal airflow, the number of odorant molecules, and the olfactory response is complex. It has been reported that three primary variables are used to determine the response of the olfactory system: the number of molecules ($N$), the duration of the air inhalations ($T$), and the volume of the air inhalations ($V$). These primary variables, in turn, define the three additional derived variables of concentration ($C = N/V$), delivery rate ($D = N/T$), and airflow velocity ($F = V/T$). Together, these six variables characterise the nature of the response in the olfactory system. They are not independent variables, such that an increase in $N$ will also increase $C$ and $D$, if $T$ and $V$ remain unchanged. So, if an increase in $N$ results in an increase in the olfactory response, it may also be attributable to the changes in $C$ and $D$ [79].

Since the functional organization of the e-nose system is similar to a mammal's olfactory system, the response of the e-nose system also follows the definition discussed above. Therefore, this chapter proposed an airflow design by combining the mixing chamber and sensor chamber and reducing the dimensions of a new sensor chamber (Dimensions:95mm*74mm*45mm). Additionally, the airflow velocity and the duration of the test can also be modified to fit the test target application. While a lower air flow speed and longer test time was used for a low concentration target odour, a higher

air flow speed and shorter test time are applied for a high concentration target odour respectively [2, 79, 80].

The airflow is pumped into the gas mixing and distribution chamber through the gas input port with the airflow velocity from $220mL/min$ to $2.2L/min$. The gas mixing and distribution chamber (shown in Figure 7.2) is designed as an air buffer to distribute one input airflow to four airflows. Moreover, in order to ensure uniform contact of the airflow with gas sensors, the layout of the sensor array is designed symmetrically and related to four airflows which come from the gas mixing and distribution chamber.

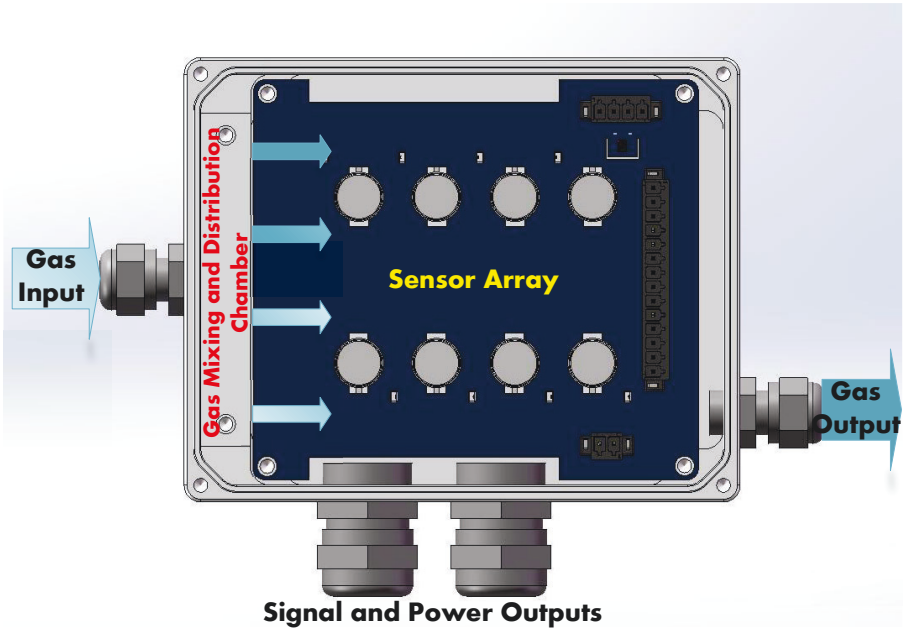


**Fig. 7.2** NOS.E Sensor Chamber.

### 7.2.2 Instrumentation design

The instrumentation design was based on the NOS.E I prototype developed by the UTS NOS.E team( see Chapter 3). This chapter is focused on a new e-nose prototype, NOS.E II (shown in Figure 7.3), by optimizing the airflow and automated control logic designs. Compared with the NOS.E I, the NOS.E II is more portable and efficient in terms of dimensions, weight, as well as test time.

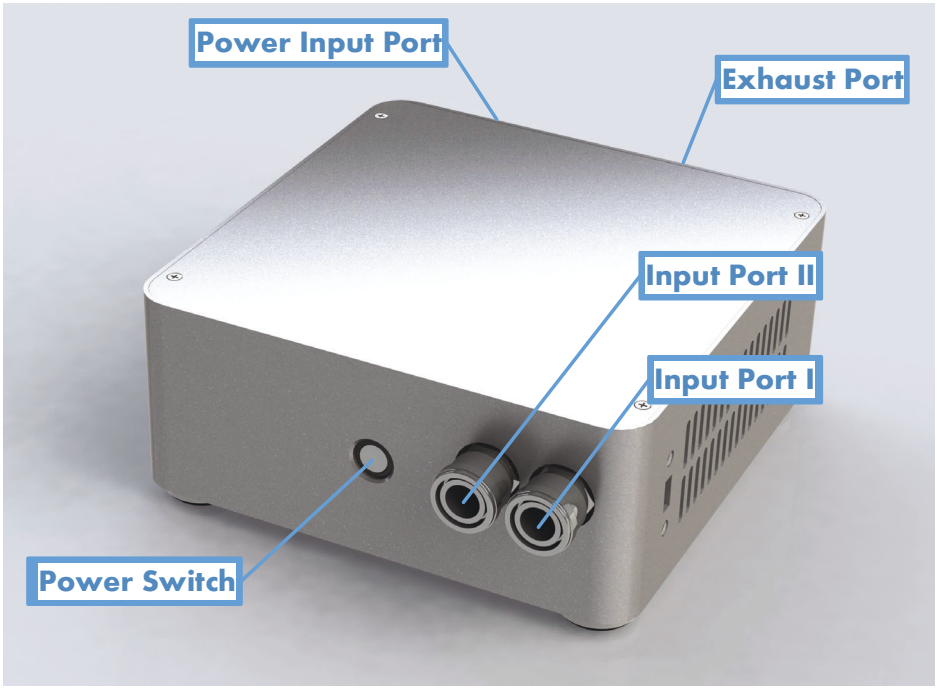


**Fig. 7.3** NOS.E Prototype II.

Figure 7.4 shows the block diagrams for the design of the NOS.E I and the NOS.E II. Considering the sensor array is designed based on eight MOX gas sensors (FIGARO ENGINEERING INC, Mino, Osaka JAPAN), the number of data acquisition channels was changed from 10 channels to 8 channels in the NOS.E II. The other designs of the host computer and slave computer are the same compared with the NOS.E I. Since the mixing chamber and the sensor chamber are combined as one chamber which is "Mixing & Sensor Chamber", Valve III is deleted from the design of NOS.E II and Valve IV is renamed as Valve III in the design of NOS.E II. Accordingly, the automated airflow control logic is redesigned as shown in Table 7.1. In this table, each row indicates one of eight working phases in the control logic of the NOS.E II. The control signals of Valve I, Valve II, Valve III, Pump I and Pump II are represented as $V_1, V_2, V_3, P_1$, and $P_2$. $T_1$, $T_2$ and $T_3$ are the control phase ID signals sent by the data acquisition and control module. In addition, the specific status of the different actuators is represented by using $H$ and $L$ in this tale. The actuators are enabled to work under the high logic level control signal ($H$) for the actuator driver circuits. Vice versa, for the low logic level control signal ($L$) for the actuator driver circuits the actuators are used to disable the corresponding actuators.

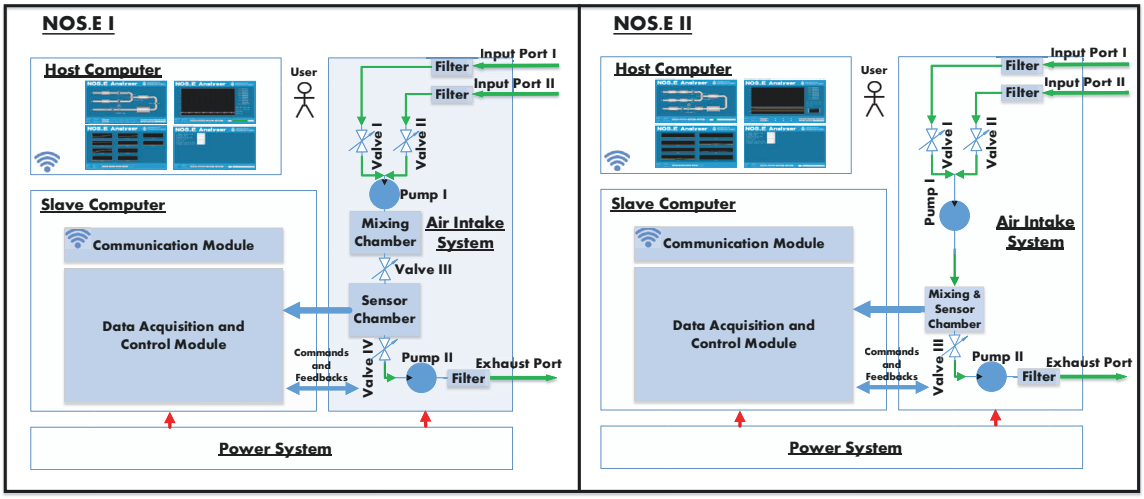


**Fig. 7.4** Block Diagram of NOS.E Prototype I and NOS.E Prototype II. Solid blue line represents the signal chain; Solid green line represents the airflow chain; Solid red line represents the power chain.

**Table 7.1** The Automated Control Logic for NOS.E II

| Phase | $V_1$ | $V_2$ | $V_3$ | $P_1$ | $P_2$ | $T_1$ | $T_2$ | $T_3$ |
|---|---|---|---|---|---|---|---|---|
| Warming up & Washing | H | L | H | H | H | L | L | L |
| Vacuuming | L | L | H | L | H | L | L | H |
| PAUSE | L | L | L | L | L | L | H | L |
| Baseline Setup | H | L | H | H | H | L | H | H |
| Vacuuming | L | L | H | L | H | H | L | L |
| Testing | L | H | H | H | H | H | L | H |
| Baseline Recovery | H | L | H | H | H | H | H | L |
| Washing | H | L | H | H | H | H | H | H |

## 7.3 Material and Methods

### 7.3.1 Rhinoceros horn and water buffalo horn samples

Seven rhinoceros horn and two water buffalo horn samples were obtained from the Australian Museum collections (shown in Figure 7.5). All the samples are registered in the Australian Museum with the unique collection number.

### 7.3.2 NOS.E II test condition

The testing system used in this chapter (shown in Figure 7.6) was set in a $MC6^{®}SF-TA$ 1200×900-900 secuflow fume (Waldner, Germany). The rhinoceros horn and water buffalo horn samples were put in individual sampling tins, and heated for 30 min at 90 ℃ using a sand bath and hotplate (IKA, Staufen, Germany). A thermometer was used


**Fig. 7.5** Rhinoceros Horn and Water Buffalo Samples.

to monitor the temperature during the sampling period. The NOS.E II collected the VOC signatures through the sampling needle and hose connected to the sampling tin. The reticulate oxygen provided by the secuflow fumehood was used as the reference gas through a cross-connection airflow adaptor. The NOS.E Analyser was used to configure the NOS.E II equipment, display the real-time responses of the sensor array and record the test data. The test duration for each round (including the chamber washing time) was 500 seconds.

### 7.3.3 Data Processing for NOS.E II

#### Kernel based modelling method

A new non-parametric kernel based modelling (KBM) algorithm was used as the pre-processing method to improve the reliability and stability of the test data by building the model for gas sensor response [218]. Since the sensor response ($y$) can be described by a single input single output (SISO) dynamic system, the finite impulse response (FIR) of this system can be expressed by the following equation:

$$\boldsymbol{Y}_N = \boldsymbol{\phi}_N \boldsymbol{\eta} + \boldsymbol{\varepsilon}_N, \quad \boldsymbol{\eta} = [g_1, g_2, \cdots, g_m]^T, \tag{7.1}$$

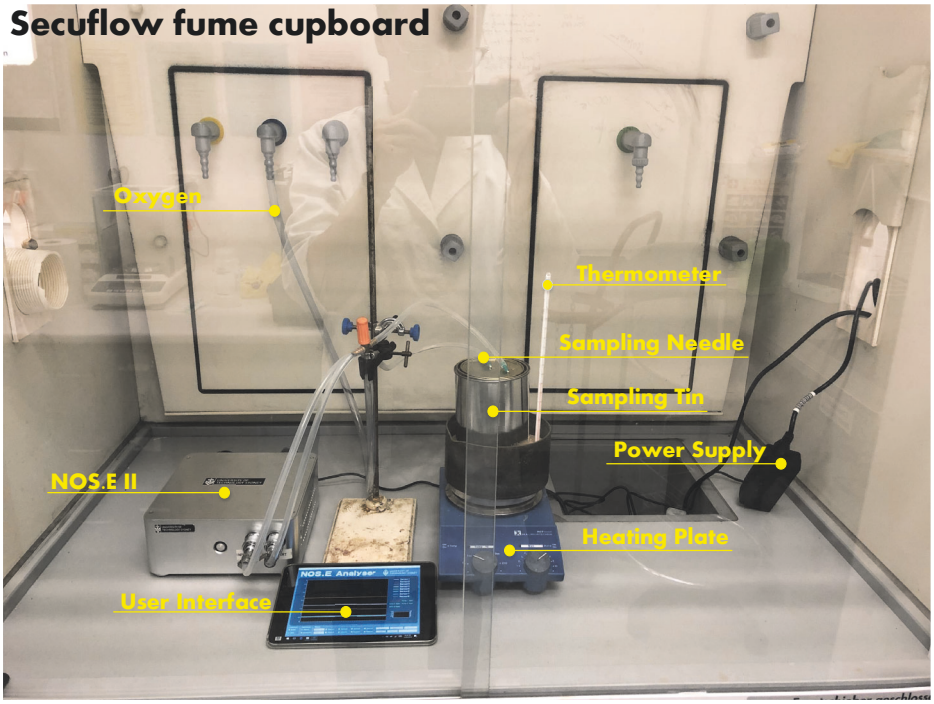


**Fig. 7.6** Diagram of the Rhinoceros Horn and Water Buffalo Horn Samples Testing System.

where $N = M - m$, and $M$ is the number of data points that we collected, $m$ is the order of FIR. $\boldsymbol{Y}_N$ represents a vector representation of the sensor's responses, and $\boldsymbol{\phi}_N$ is the vector representation of the input information of the system [218].

Therefore, this model can be identified by minimising the cost function:

$$\hat{\boldsymbol{\eta}} = arg\min_{\boldsymbol{\eta}\in\mathbb{R}^m} \boldsymbol{Z}(\boldsymbol{\eta}), \tag{7.2}$$

$$\boldsymbol{Z}(\boldsymbol{\eta}) = ||\boldsymbol{Y}_N - \boldsymbol{\phi}_N\boldsymbol{\eta}||^2 + \gamma\boldsymbol{\eta}^T\boldsymbol{K}^{-1}\boldsymbol{\eta}, \tag{7.3}$$

where $\gamma$ is a positive scalar, $\boldsymbol{K}$ represents the Stable Spline (SS) kernel matrix [218].

The cost function $\boldsymbol{Z}(\boldsymbol{\eta})$ can be written as:

$$\boldsymbol{Z}(\boldsymbol{\eta}) = \boldsymbol{Y}_N^T\boldsymbol{Y}_N - \boldsymbol{Y}_N^T\boldsymbol{\phi}_N\boldsymbol{\eta} - \boldsymbol{\eta}^T\boldsymbol{\phi}_N^T\boldsymbol{Y}_N + \boldsymbol{\eta}^T\boldsymbol{\phi}_N^T\boldsymbol{\phi}_N\boldsymbol{\eta} + \gamma\boldsymbol{\eta}^T\boldsymbol{K}^{-1}\boldsymbol{\eta}. \tag{7.4}$$

The equation below can be achieved by taking the partial derivative of cost function $\boldsymbol{Z}(\boldsymbol{\eta})$ with respect to $\boldsymbol{\eta}$,

$$\frac{\partial \boldsymbol{Z}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = 2\boldsymbol{\phi}_N^T \boldsymbol{\phi}_N \boldsymbol{\eta} - 2\boldsymbol{\phi}_N^T \boldsymbol{Y}_N + 2\gamma \boldsymbol{K}^{-1} \boldsymbol{\eta} = 0. \tag{7.5}$$

Then, the estimated model $\hat{\boldsymbol{\eta}}$ is obtained as:

$$\hat{\boldsymbol{\eta}} = (\boldsymbol{\phi}_N^T \boldsymbol{\phi}_N + \gamma \boldsymbol{K}^{-1})^{-1} \boldsymbol{\phi}_N^T \boldsymbol{Y}_N. \tag{7.6}$$

As a result, $\hat{\boldsymbol{\eta}}$ can be adapted as

$$\hat{\boldsymbol{\eta}} = \boldsymbol{K}\boldsymbol{\phi}_N^T (\boldsymbol{\phi}_N \boldsymbol{K} \boldsymbol{\phi}_N^T + \gamma \boldsymbol{I}_N)^{-1} \boldsymbol{Y}_N, \tag{7.7}$$

$$= \boldsymbol{K}(\boldsymbol{\phi}_N^T \boldsymbol{\phi}_N \boldsymbol{K} + \gamma \boldsymbol{I}_m)^{-1} \boldsymbol{\phi}_N^T \boldsymbol{Y}_N. \tag{7.8}$$

where $\boldsymbol{I}_N$ and $\boldsymbol{I}_m$ are identity matrices.

**Kernel principal components analysis method**

Nonlinear kernel principal components analysis (KPCA) was applied as a visual tool to verify the effectiveness of the KBM data pre-processing method, show groupings and separations within datasets as well as an important feature selection tool to improve the performance of the further classification tasks [268]. In this study, polynomial kernel and Gaussian kernel were applied to map the test data to the high-dimensional space.

The Polynomial kernel used is:

$$K(x_1, x_2) = (x_1^T x_2 + p)^d, p \geq 0, \tag{7.9}$$

where $x_1$ and $x_2$ are the key features extracted from the odour datasets; $p$ is a free parameter trading off the influence of higher-order versus lower-order terms in the polynomial; $d$ is the degree of polynomial.

The Gaussian ($RBF$) kernel used in this study is:

$$K(x_1, x_2) = exp\left(-\frac{(x_1 - x_2)^2}{2\sigma^2}\right). \tag{7.10}$$

where $x_1$ and $x_2$ are the key features extracted from the odour datasets; $\sigma$ is a free parameter.

**Support vector machine method**

Finally, as the popular classification solution, support vector machine (SVM) was implemented as the "content of the brain" to enable the NOS.E II to recognise the test samples.

The classification function used is:

$$F(x) = \sum_i y_i \alpha_i K(x, x_i) + b, \tag{7.11}$$

where the $\alpha_i$ are the solutions of the dual problem and are non-zero only for a subset of vectors $x_i$ called support vectors [175]; $K(x_1, x_2)$ is the RBF kernel (Eq. 7.12) implemented in this study, shown as:

$$K(x_1, x_2) = exp\left(-\frac{(x_1 - x_2)^2}{2\sigma^2}\right). \tag{7.12}$$

GAs was applied to select $\sigma$=6.01 as the optimal value for the parameter of the RBF kernel in the SVM classifier.

## 7.4 Results and Discussion

### 7.4.1 KBM pre-processing result

Of the eight sensors being installed in NOS.E II, four of them ($TGS\ 2611-E, TGS\ 2602, TGS\ 2600, and$ were selected for further data pre-processing as they gave clear responses to the test items. The remaining four sensors did not show a response to the particular test subject

used during the current study and were therefore omitted. Rhinoceros Horn Sample M47156-18 was used to show the differences between NOS.E II raw data waveforms (Figure 7.7a) and KBM pre-processed data waveforms (Figure 7.7b). Compared with the raw data collected by the NOS.E II instrument the KBM pre-processed data waveforms were smoother and provide more reliable traditional features (especially for the derivative-related features) used for further data analysis [218].

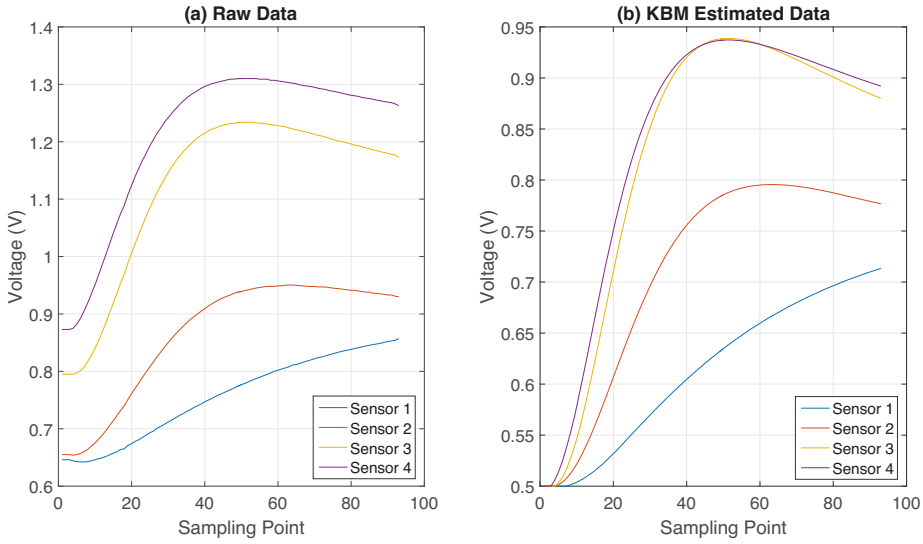


**Fig. 7.7** KBM Data pre-processing Results for Rhinoceros Horn Sample M47156-18.

### 7.4.2 PCA result

The Polynomial kernel PCA and Gaussian kernel PCA were used to analyse the wildlife part datasets, and the results are shown in Figure 7.8. Both the polynomial kernel and Gaussian kernel PCA can successfully separate the wildlife part samples into two groups. The water buffalo horn samples (highlighted with the black circle, WSH represents for the small water buffalo horn sample, WBH represents for the big water buffalo horn sample) were tightly gathered together compared with the rhinoceros horn samples (highlighted with the red circle) which were puffed gathered in these figures.

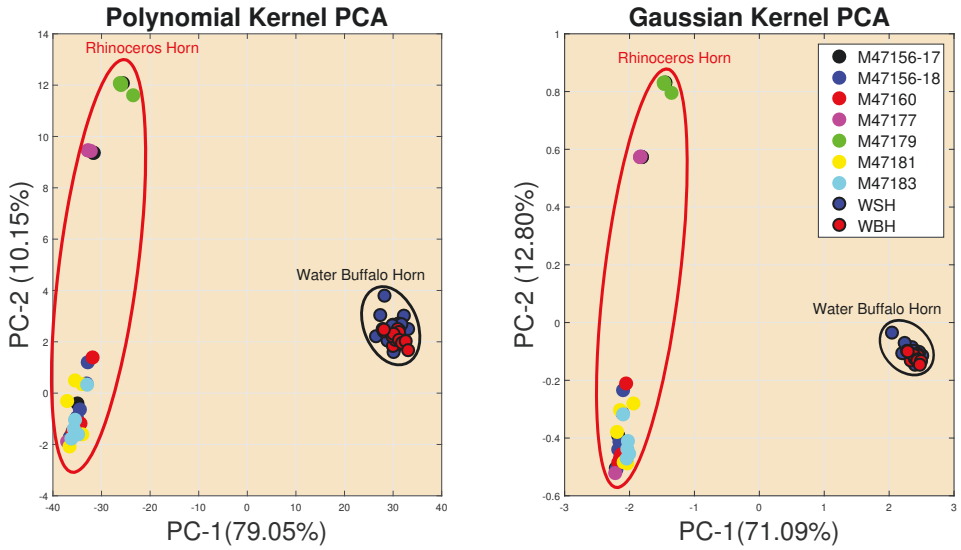


**Fig. 7.8** Rhinoceros Horn Samples and Water Buffalo Horn Samples KPCA Analysis Results.

### 7.4.3 SVM result

The support vector machine (SVM) with RBF kernel ($C = 901, \sigma = 6.01$) is utilised in this chapter as the odour classifier. A five-fold cross-validation method was implemented for the SVM classifier and run twenty times in this study, the average accuracy was used to assess the performance of the NOS.E II wildlife parts recognition ability. The classification results are listed in Table 7.2. A 93.33% average accuracy of classification was achieved based on the raw data collected by the NOS.E II. By applying the KBM method, we achieved 98.33% average classification accuracy. According to these classification results, the NOS.E II has the ability to recognise the real and fake Rhinoceros horns with a high successful recognition rate.

**Table 7.2** Classification Results

| Data Type | Parameters of Pre-processing Method | Average Accuracy of SVM Classification (%) |
|---|---|---|
| Raw Data | $N/A$ | 93.33 |
| KBM modelling Data | $Kernel = SS$, $\rho = 1000$, $\alpha = 0.98$ | 98.33 |

## 7.5 Conclusion and Future Works

This chapter developed an efficient and reliable e-nose prototype (NOS.E II) which was used to identify legal from illegal wildlife parts. Based on the mammalian olfactory

system, a unique sensor chamber was developed which combined the mixing chamber and the sensor chamber into one to improve the sensitivity of the NOS.E II. According to the novel air intake system, the newly automated control logic is designed and implemented to improve the test efficiency of the NOS.E II. Since e-noses have not previously tested rhinoceros horn and water buffalo horn samples, they are used to analyse the performance of the NOS.E II. According to the test results, the NOS.E II system can successfully differentiate the legally traded water buffalo horn samples from illegal rhinoceros horn samples. The rapid detection of wildlife products at points of entry would greatly assist law enforcement in seizing forensic evidence and prosecuting offenders [253]. The e-nose prototype proposed in this study could be used by frontline staff to rapidly detect illegal wildlife products onsite without the requirement for sophisticated laboratory analysis, and could be deployed at airports, seaports, mail centres, and other border crossings to identify imported and exported wildlife contraband. Future works will test more species including large cats (e.g. tigers, leopards), elephants, pangolins, bears, sea turtles, sharks, and a range of exotic birds and reptiles to verify the effectiveness of the NOS.E II [253]. The next generation NOS.E prototype (shown in Figure 7.9) will also be developed to create a portable, user-friendly, and cost-effective product for border enforcement departments throughout the world.


**Fig. 7.9** Conceptual design of Portable NOS.E For the Detection of Illegal Wildlife Parts.

# Chapter 8

# Conclusions and Future Work

## 8.1 Conclusions

Inspired by the mammalian olfactory system, this dissertation designed and developed an efficient and reliable machine olfactory system which is based on the e-nose technique.

The proposed e-nose instrument, NOS.E, was used for odour detection and identification purposes. Based on the specific control logic, an automated air intake design and the related fault detection and alarming design were equipped in the proposed NOS.E system. To validate the design of the automated airflow control and fault detection design, the NOS.E system was run 90 times over three days. The status of the airflow system was monitored by the NOS.E user interface. According to the test results, the NOS.E system can successfully execute the designed control logic. Moreover, the fault detection and monitoring system is able to provide the insurance to avoid wasting resources by timely terminating the system when the fault status of actuators is detected. The performance of the designed NOS.E system has been tested by three different alcohols and compared with the results of GC×GC-TOFMS by using PCA pattern recognition techniques. In summary, compared with the time consuming and expensive GC×GC-TOFMS odour analysis method, the proposed e-nose system is a suitable odour analysis platform in terms of its low cost, high efficiency and reliability features.

As the novel data processing and analysis methods for NOS.E systems, the non-parametric kernel based modelling (KBM) method is also investigated in this dissertation. The KBM method can overcome the noise introduced by the unexpected responses (caused by the gas interference or fluctuating environmental parameters) thus providing high quality derivative-related key features for the odour classifier. According to the test results, when extracting derivative-related features, the proposed non-parametric KBM method provides more reliable and stable pre-processing results and higher accuracy of classification compared with the other pre-processing methods.

As the novel odour pattern analysis algorithm, a hybrid of genetic algorithm (GA) and supervised fuzzy support vector machine (FSVM) is applied in the NOS.E system in this dissertation. The key features and the model parameters of FSVM were selected by the GA. To reduce the outlier effects to provide a robust classifier which has a steady classification accuracy, FSVM is introduced as the odour classifier. This hybrid algorithm can significantly improve the classification accuracy by comparing with some popular machine learning algorithms, such as support vector machine, the k-nearest neighbours and other combination algorithms.

Based on the NOS.E system we developed, this dissertation mainly utilised the NOS.E in two different applications: health monitoring based on the analysis of the human breath and the rapid wildlife parts identification according to the recognition of their odour signatures.

The study of exhaled breath for health diagnosis and monitoring is becoming an increasingly popular area of research. Unlike most traditional health monitoring and diagnosis through the use of bodily fluids (e.g. through blood, urine), breath collection is non-invasive and convenient. Furthermore, studies have shown there are many metabolomic compounds in human breath that could be used for health monitoring and disease diagnosis. Metabolomics are influenced by an individual's lifestyle, diet, and the environment. These factors make breath analysis an attractive option for personalised health care. To explore the potential biomedical applications, the NOS.E was used for the rapid assessment of human health conditions in this dissertation. By detecting the changes in the composition of an individual's respiratory gases, which have been shown to be linked to changes in metabolism, e-nose systems can be used

to characterise the physical health condition. We demonstrated our system's viability with a simple dataset consisting of breath collected under three different scenarios from one volunteer. The preliminary results show that the popular classifier SVM can discriminate NOS.E's responses under the three scenarios with high performance and the NOS.E could be applied in the human health monitoring area.

The scale of illegal wildlife trade has rapidly increased due to the huge profit margin. At the same time, this trend is also threatening biodiversity and accelerating the extinction for many species (such as rhinoceros, elephants, sea turtles etc.). In order to combat the illegal wildlife trade, law enforcement has cooperated with researchers by applying many strategies to inspect and identify the legal from illegal wildlife parts. Generally, the suspect wildlife parts are analysed in the laboratory by using isotopic fingerprinting and DNA testing. However, these tests are time-consuming, sophisticated and cannot be arranged on-site. Although wildlife detection dogs are used to detect illegal wildlife parts on-site, there are many factors such as emotional and individual deviations that can affect the test results. Based on the NOS.E system, which used an efficient and reliable method to identify illegal wildlife parts. The novel mechanical and airflow designs, as well as the KBM data preprocessing methods, were implemented in the NOS.E system to improve its sensitivity and portability. As a proof of concept test, rhinoceros horn and water buffalo horn samples were selected as the test targets to identify legal from illegal wildlife parts. According to the test results, the NOS.E system can successfully recognise the legal water buffalo horn samples from illegal rhinoceros horn samples. It is considering that rapid detection of wildlife products at points of entry would greatly assist law enforcement in seizing forensic evidence and prosecuting offenders [253]. The e-nose prototype proposed in this dissertation can be used by frontline staff to rapidly detect illegal wildlife products onsite without the requirement for sophisticated laboratory analysis and can be deployed at airports, seaports, mail centres, and other border crossings to identify imported and exported wildlife contraband.

## 8.2 Future Work

In order to fulfill the future commercialisation plan, the NOS.E design would be further optimised to improve its performance and make it more cost effective as well.

In order to reduce the power consumption of the NOS.E system, future works will continue to develop the NOS.E platform by optimising the design of the power system, airflow control system, as well as the data acquisition and processing system. Low power consumption management strategies will be considered according to different user scenarios. In addition, some components would be re-designed by balancing the cost and performance of the entire NOS.E design. Moreover, more NOS.E prototypes would be built to validate its repeatability and prepare the further electric performance tests such as Electro Magnetic Compatibility (EMC) test and Electro-Static discharge (ESD) test.

Additionally, the following areas will also be focused: 1. Collecting more different samples to validate the implementation of the NOS.E system across various odour identification and classification applications; 2. Improving the efficiency and functionality of the NOS.E system by more research outputs with the proposed non-parametric KBM data pre-processing method and NOS.E hardware design optimisation; 3. Considering discriminative models for feature extraction [226]); 4. Applying some cost-sensitive classification algorithms to improve the performance of NOS.E system [200].

As for the future work in the health monitoring area, we will recruit a small group of healthy volunteers and collect samples under a variety of scenarios and at different times, which will allow us to test NOS.E's capability. We envisage that the NOS.E system can be deployed as part of a cloud-based health monitoring system. The hardware e-nose equipment will be installed on site to monitor the exhaled breath of patients. The data collected can be uploaded to a remote server, which will monitor or diagnosis the health conditions. Users can check these data and control the e-nose equipment through a cell phone or personal computer. Furthermore, health professionals can also monitor the same data and provide their feedback to the users. This cloud-based comprehensive system would allow continuous and efficient monitoring of users' health. By introducing doctors to the remote server in this new e-nose health monitoring

system, its application could be expanded to medical areas such as general health screening and disease diagnosis. Ultimately, we aim to build and test a remote NOS.E system that can be used to monitor human health through the breath.

By implementing the NOS.E in the illegal wildlife parts identification area, future works will also test more species including large cats (e.g. tigers, leopards), elephants, rhinoceros, pangolins, bears, sea turtles, sharks, and a range of exotic birds and reptiles to verify the effectiveness of the NOS.E II [253]. The next generation NOS.E prototype will also be developed to make it a portable, user-friendly, and cost-effective product for border enforcement departments throughout the world.

The portable smart NOS.E system which is based on a low-cost hardware platform and contains the novel data processing algorithm and various communication technologies (e.g. Bluetooth, WIFI, Narrowband Internet of Things (NB-IoT))will be applied in more applications, especially for health monitoring, wildlife products identification, food quality assessment, illicit drug detection, etc.

Finally, based on the optimised NOS.E system, a marketing research task will be started, then the NOS.E business plan will be carried out toward the NOS.E commercialisation target.

# Appendix A

# NOS.E User Manual

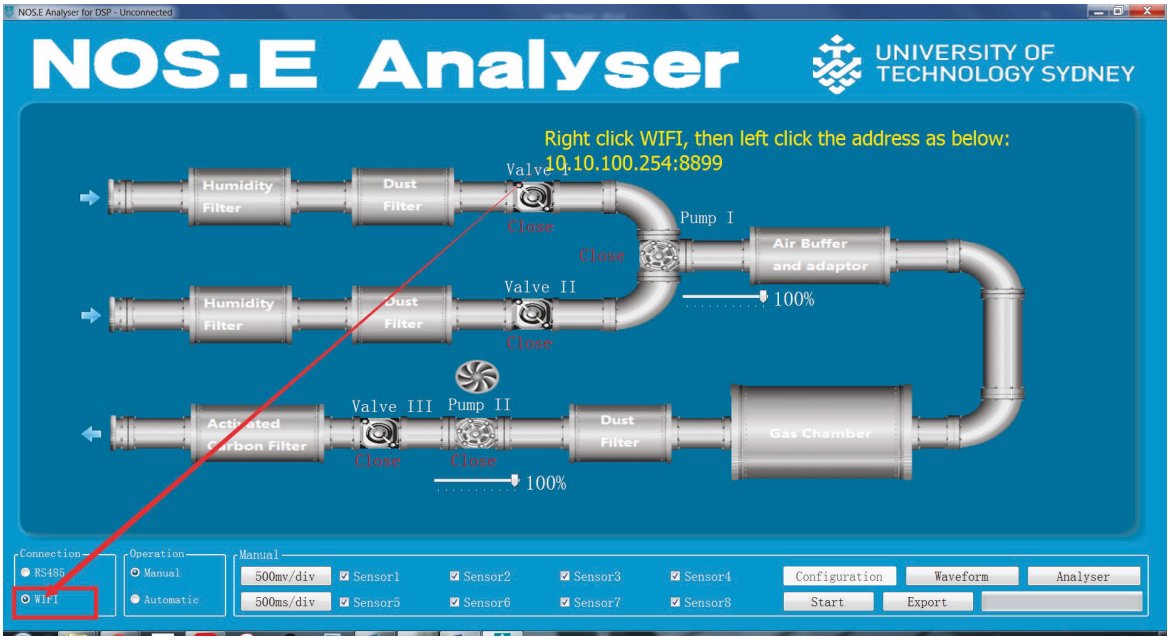


**Fig. A.1** The Communication Setup for the NOS.E.

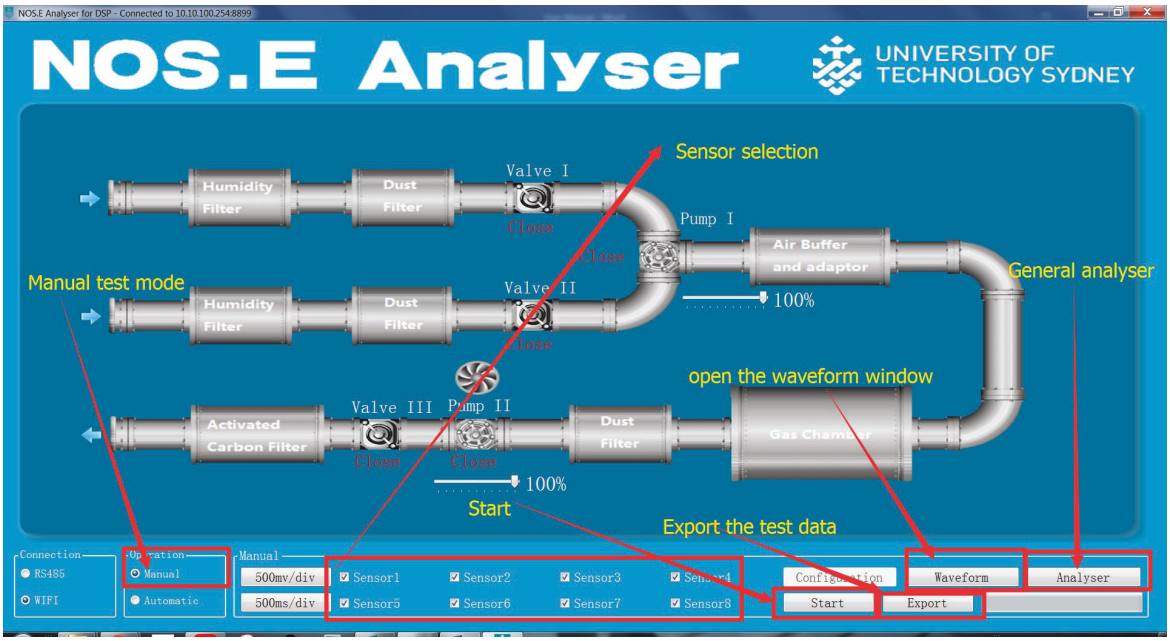


**Fig. A.2** The Manual Mode Setup for the NOS.E.

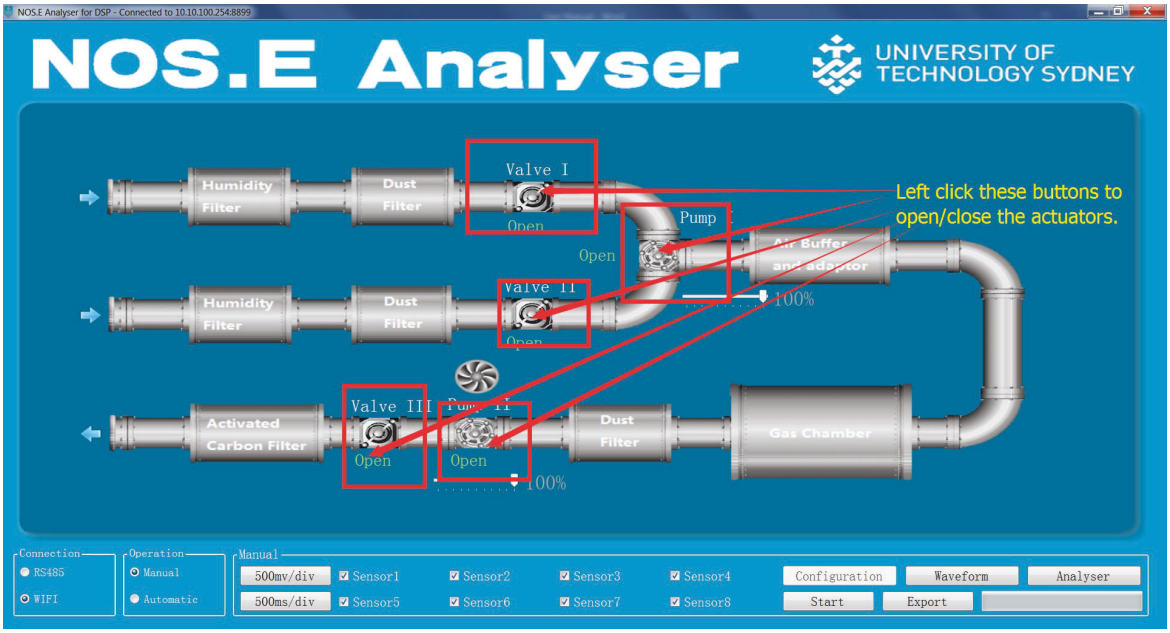


**Fig. A.3** The Actuators Setup for the NOS.E.

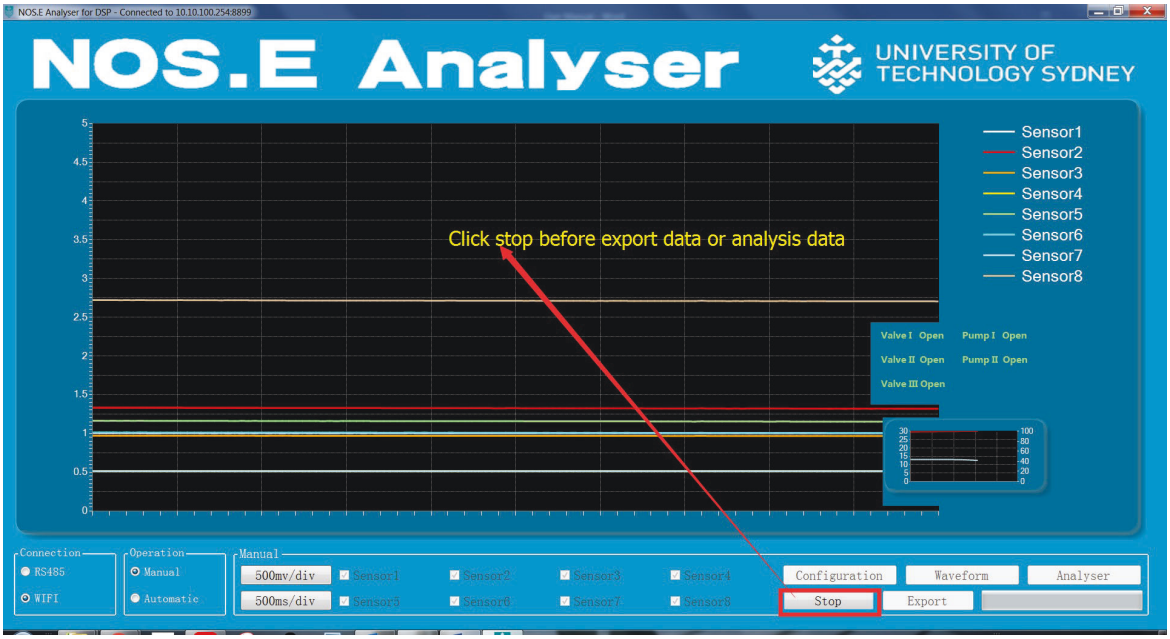


**Fig. A.4** The Waveform Display Window Setup for the NOS.E.

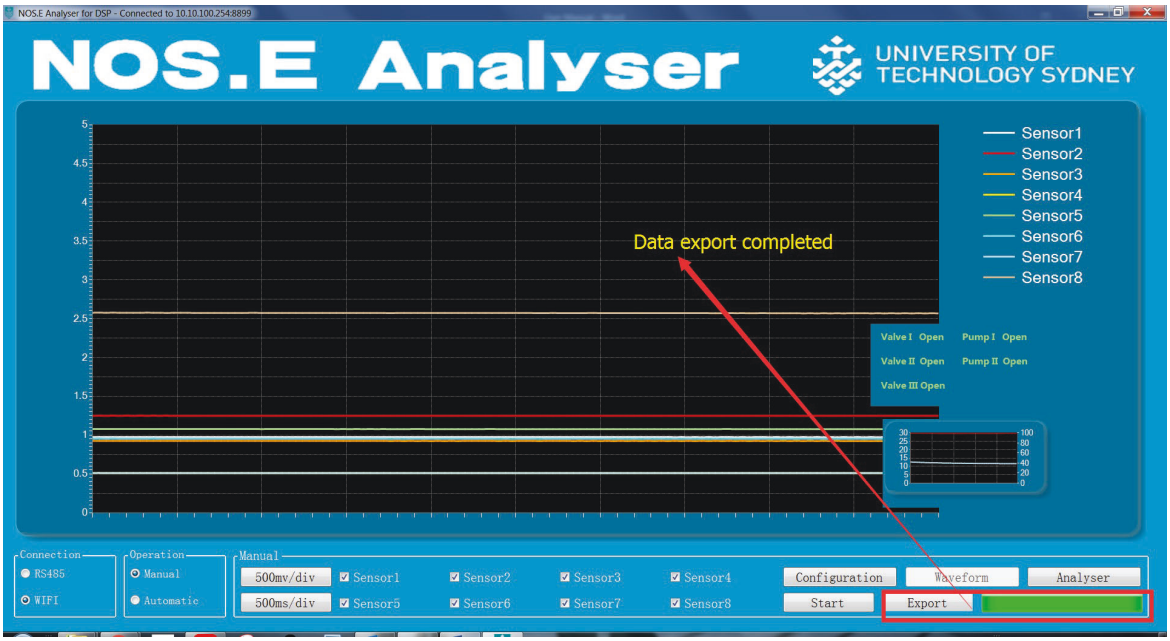


**Fig. A.5** The Export Data Setup for the NOS.E.

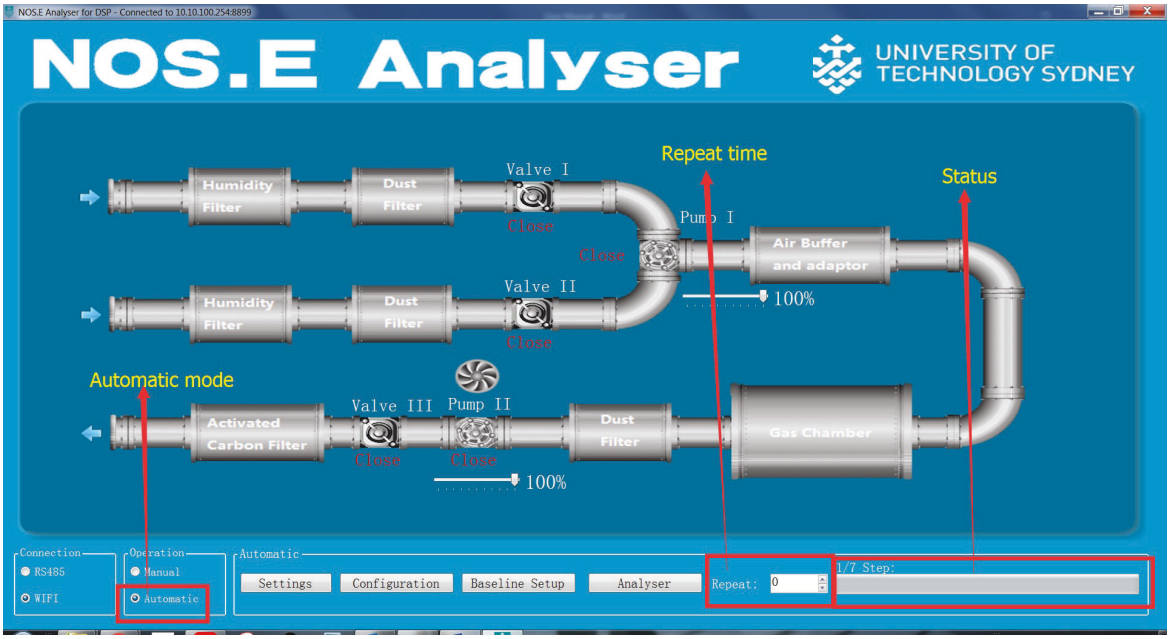


**Fig. A.6** The Automatic Mode Setup for the NOS.E.

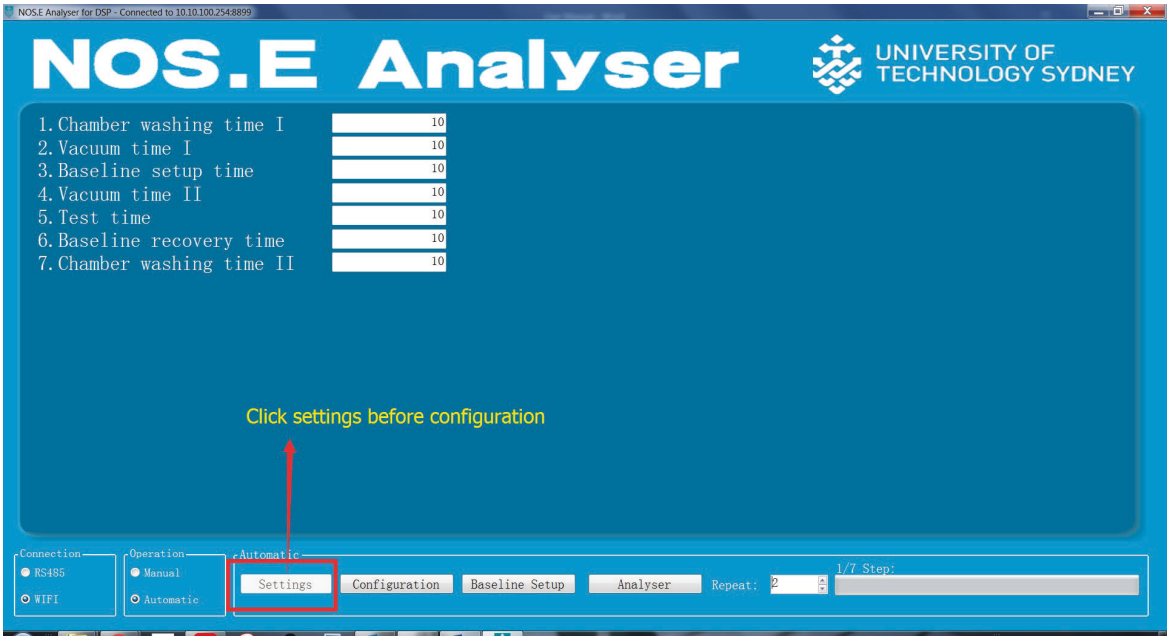


**Fig. A.7** The Automatic Mode Operation Time Parameters Setup for the NOS.E.

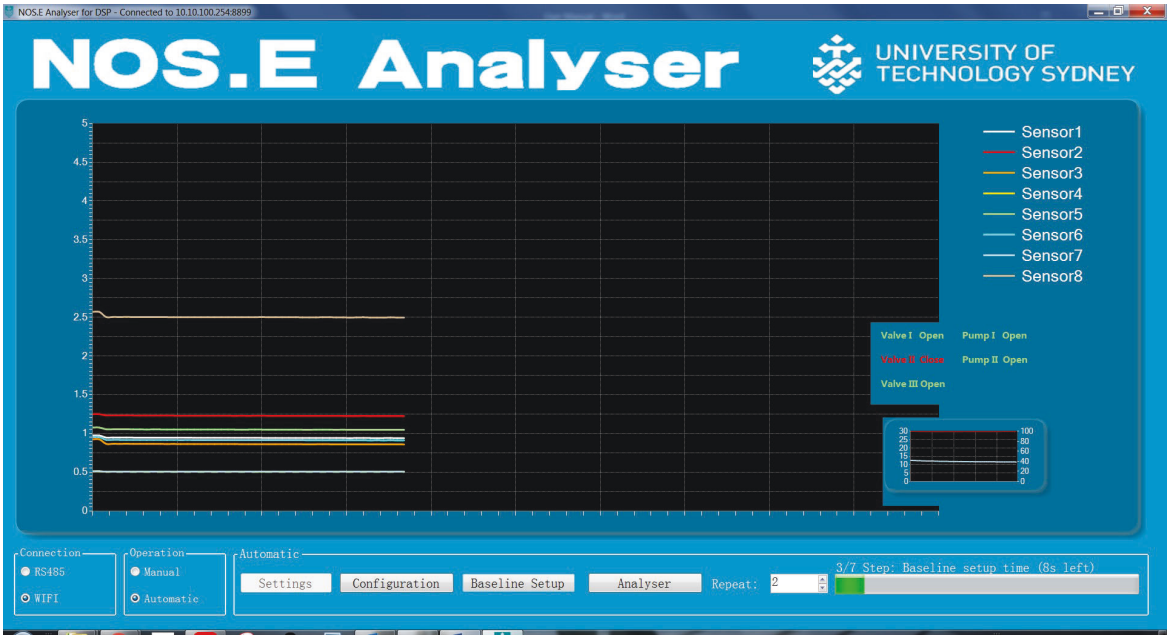


**Fig. A.8** The Baseline Setup for the NOS.E.

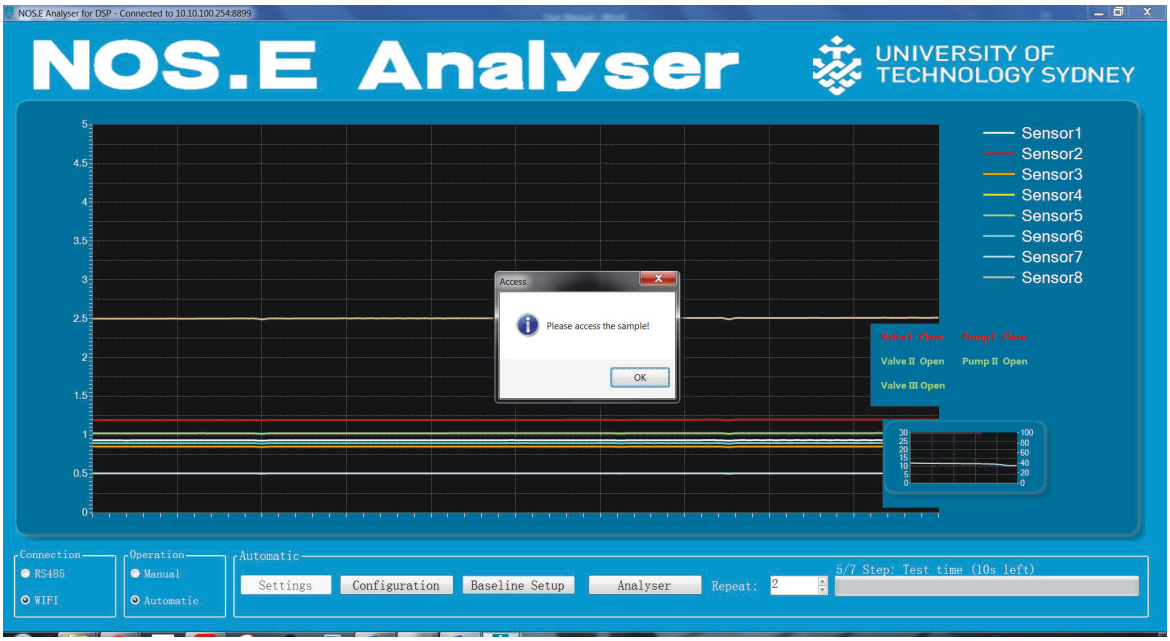


**Fig. A.9** The Testing Phase for the NOS.E.

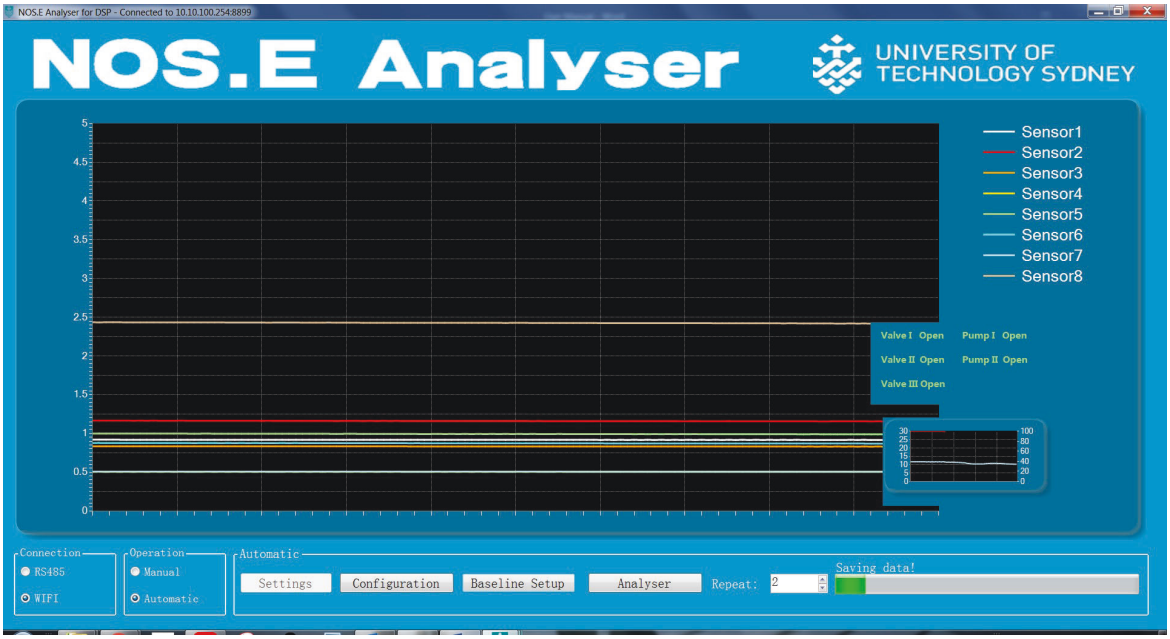


**Fig. A.10** The Saving Data Phase for the NOS.E.

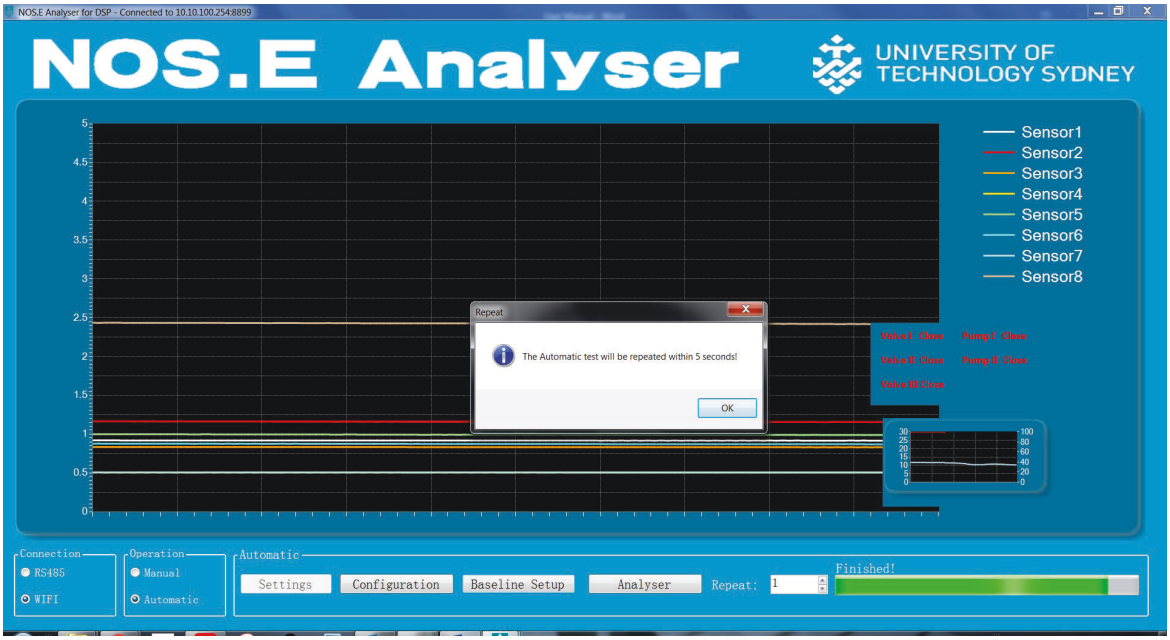


**Fig. A.11** The Automatic Test Mode for the NOS.E Test Round 2.

# Appendix B

# NOS.E Sensor Array Specifications

**Table B.1** The Sensitivity Characteristics of Sensor Array in the NOS.E

| Sensor Model | Sensitivity characteristics[1] |
|---|---|
| TGS 2600 | Hydrogen, Ethanol, Iso-butane, CO, Methane, etc. |
| TGS 2602 | Toluene, Hydrogen sulfide, Ethanol, Ammonia, Hydrogen, etc. |
| TGS 2603 | Trimethyl amine, Ethanol, Methyl mercaptan, $H_2S$, Hydrogen, etc. |
| TGS 2610C | Propane, Iso-butane, Methane, Hydrogen, Ethanol, etc. |
| TGS 2610D | Propane, Iso-butane, Methane, Hydrogen, etc. |
| TGS 2611C | Methane, Iso-butane, Hydrogen, Ethanol, etc. |
| TGS 2611E | Methane, Hydrogen, Iso-butane, etc. |
| TGS 2612 | Propane, Iso-butane, Methane, Ethanol, etc. |
| TGS 2620 | Ethanol, Hydrogen, Iso-butane, CO, Methane, etc. |
| TGS 2444 | Ammonia, $H_2S$, Ethanol, etc. |
| MQ 2 | LPG, Propane, Hydrogen, Ethanol, $CH_4$, CO, etc. |
| MQ 3 | Ethanol, Benzine, Hexane, LPG, CO, $CH_4$, etc. |
| MQ 8 | Hydrogen, Ethanol, LPG, $CH_4$, CO, etc. |
| MQ 135 | Acetone, Toluene, Ammonia, Ethanol, CO, etc. |

**Note 1**: *These sensors may also sensitive to other gases, which are not list in this table.*

# References


[1] J. W. Gardner and P. N. Bartlett, *Electronic noses: principles and applications*, vol. 233. Oxford university press New York, 1999.

[2] T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. W. Gardner, *Handbook of machine olfaction: electronic nose technology*. John Wiley & Sons, 2006.

[3] A. D. Wilson and M. Baietto, "Applications and advances in electronic-nose technologies," *Sensors*, vol. 9, pp. 5099–5148, Jun. 2009.

[4] J. W. Gardner and P. N. Bartlett, "A brief history of electronic noses," *Sensors and Actuators B: Chemical*, vol. 18, pp. 210–211, Mar. 1994.

[5] A. Loutfi, S. Coradeschi, G. K. Mani, P. Shankar, and J. B. B. Rayappan, "Electronic noses for food quality: A review," *Journal of Food Engineering*, vol. 144, pp. 103–111, 2015.

[6] M. Peris and L. Escuder-Gilabert, "A 21*st* century technique for food control: Electronic noses," *Analytica chimica acta*, vol. 638, pp. 1–15, Apr. 2009.

[7] I. Rodriguez-Lujan, J. Fonollosa, A. Vergara, M. Homer, and R. Huerta, "On the calibration of sensor arrays for pattern recognition using the minimal number of experiments," *Chemometrics and Intelligent Laboratory Systems*, vol. 130, pp. 123–134, Jan. 2014.

[8] E. Gobbi, M. Falasconi, G. Zambotti, V. Sberveglieri, A. Pulvirenti, and G. Sberveglieri, "Rapid diagnosis of enterobacteriaceae in vegetable soups by a metal oxide sensor based electronic nose," *Sensors and Actuators B: Chemical*, vol. 207, pp. 1104–1113, 2015.

[9] A. Caron, N. Redon, F. Thevenet, B. Hanoune, and P. Coddeville, "Performances and limitations of electronic gas sensors to investigate an indoor air quality event," *Building and Environment*, vol. 107, pp. 19–28, 2016.

[10] K. Brudzewski, S. Osowski, and A. Dwulit, "Recognition of coffee using differential electronic nose," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 6, pp. 1803–1810, 2012.

[11] L. Zhang, F.-C. Tian, X.-W. Peng, and X. Yin, "A rapid discreteness correction scheme for reproducibility enhancement among a batch of mos gas sensors," *Sensors and Actuators A: Physical*, vol. 205, pp. 170–176, 2014.

[12] L. Zhang and D. Zhang, "Domain adaptation extreme learning machines for drift compensation in e-nose systems," *IEEE Transactions on instrumentation and measurement*, vol. 64, no. 7, pp. 1790–1801, 2015.

[13] J. L. Herrero, J. Lozano, J. P. Santos, and J. I. Suárez, "On-line classification of pollutants in water using wireless portable electronic noses," *Chemosphere*, vol. 152, pp. 107–116, Jun. 2016.

[14] F. Röck, N. Barsan, and U. Weimar, "Electronic nose: current status and future trends," *Chemical reviews*, vol. 108, no. 2, pp. 705–725, 2008.

[15] X. R. Wang, J. T. Lizier, A. Z. Berna, F. G. Bravo, and S. C. Trowell, "Human breath-print identification by e-nose, using information-theoretic feature selection prior to classification," *Sensors and Actuators B: Chemical*, vol. 217, pp. 165–174, 2015.

[16] Z. Zhao, F. Tian, H. Liao, X. Yin, Y. Liu, and B. Yu, "A novel spectrum analysis technique for odor sensing in optical electronic nose," *Sensors and Actuators B: Chemical*, vol. 222, pp. 769–779, 2016.

[17] D. Wilson, "Recent progress in the design and clinical development of electronic-nose technologies," *Nanobiosensors in Disease Diagnosis*, vol. 5, pp. 15–27, 2016.

[18] J. Li, H. Feng, W. Liu, Y. Gao, and G. Hui, "Design of a portable electronic nose system and application in k value prediction for large yellow croaker (pseudosciaena crocea)," *Food Analytical Methods*, vol. 9, no. 10, pp. 2943–2951, 2016.

[19] J. Chilo, J. Pelegri-Sebastia, M. Cupane, and T. Sogorb, "E-nose application to food industry production," *IEEE Instrumentation & Measurement Magazine*, vol. 19, no. 1, pp. 27–33, 2016.

[20] S. Cui, J. Wu, J. Wang, and X. Wang, "Discrimination of american ginseng and asian ginseng using electronic nose and gas chromatography–mass spectrometry coupled with chemometrics," *Journal of ginseng research*, vol. 41, no. 1, pp. 85–95, 2017.

[21] D. Zhang, D. Guo, and K. Yan, "A novel medical e-nose signal analysis system," in *Breath Analysis for Medical Applications*, pp. 281–299, Springer, 2017.

[22] H. Sun, F. Tian, Z. Liang, T. Sun, B. Yu, S. X. Yang, Q. He, L. Zhang, and X. Liu, "Sensor array optimization of electronic nose for detection of bacteria in wound infection," *IEEE Transactions on Industrial Electronics*, Apr. 2017.

[23] E. Westenbrink, R. P. Arasaradnam, N. O'Connell, C. Bailey, C. Nwokolo, K. D. Bardhan, and J. Covington, "Development and application of a new electronic nose instrument for the detection of colorectal cancer," *Biosensors and Bioelectronics*, vol. 67, pp. 733–738, 2015.

[24] J.-H. Kim, J. Chun, J. W. Kim, W. J. Choi, and J. M. Baik, "Self-powered, room-temperature electronic nose based on triboelectrification and heterogeneous catalytic reaction," *Advanced Functional Materials*, vol. 25, no. 45, pp. 7049–7055, 2015.

[25] L. Gil-Sánchez, J. Soto, R. Martínez-Máñez, E. Garcia-Breijo, J. Ibáñez, and E. Llobet, "A novel humid electronic nose combined with an electronic tongue for assessing deterioration of wine," *Sensors and Actuators A: Physical*, vol. 171, no. 2, pp. 152–158, 2011.

[26] D. Li, T. Lei, S. Zhang, X. Shao, and C. Xie, "A novel headspace integrated e-nose and its application in discrimination of chinese medical herbs," *Sensors and Actuators B: Chemical*, vol. 221, pp. 556–563, 2015.

[27] A. Somov, E. F. Karpov, E. Karpova, A. Suchkov, S. Mironov, A. Karelin, A. Baranov, and D. Spirjakin, "Compact low power wireless gas sensor node with thermo compensation for ubiquitous deployment," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1660–1670, 2015.

[28] L. Zhang, Y. Liu, and P. Deng, "Odor recognition in multiple e-nose systems with cross-domain discriminative subspace learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1679–1692, 2017.

[29] X. Peng, L. Zhang, F. Tian, and D. Zhang, "A novel sensor feature extraction based on kernel entropy component analysis for discrimination of indoor air contaminants," *Sensors and Actuators A: Physical*, vol. 234, pp. 143–149, 2015.

[30] X. Sun, L. Liu, Z. Wang, J. Miao, Y. Wang, Z. Luo, and G. Li, "An optimized multi-classifiers ensemble learning for identification of ginsengs based on electronic nose," *Sensors and Actuators A: Physical*, vol. 266, pp. 135–144, 2017.

[31] S. Jiang, J. Wang, Y. Wang, and S. Cheng, "A novel framework for analyzing mos e-nose data based on voting theory: Application to evaluate the internal quality of chinese pecans," *Sensors and Actuators B: Chemical*, vol. 242, pp. 511–521, 2017.

[32] Y.-Q. Jing, Q.-H. Meng, P.-F. Qi, M.-L. Cao, M. Zeng, and S.-G. Ma, "A bioinspired neural network for data processing in an electronic nose," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2369–2380, 2016.

[33] X. Jiang, P. Jia, R. Luo, B. Deng, S. Duan, and J. Yan, "A novel electronic nose learning technique based on active learning: Eqbc-rbfnn," *Sensors and Actuators B: Chemical*, vol. 249, pp. 533–541, 2017.

[34] H. Jamalabadi, A. Mani-Varnosfaderani, and N. Alizadeh, "Detection of alkyl amine vapors using ppy-zno hybrid nanocomposite sensor array and artificial neural network," *Sensors and Actuators A: Physical*, vol. 280, pp. 228–237, 2018.

[35] S. Cui, J. Wu, J. Wang, and X. Wang, "Discrimination of american ginseng and asian ginseng using electronic nose and gas chromatography–mass spectrometry coupled with chemometrics," *Journal of ginseng research*, vol. 41, no. 1, pp. 85–95, 2017.

[36] J. de Swart, N. van Gaal, D. J. Berkhout, T. G. de Meij, and N. K. de Boer, "Smoking influences fecal volatile organic compounds composition," *Clinical Gastroenterology and Hepatology*, vol. 16, no. 7, pp. 1168–1169, 2018.

[37] S. Dragonieri, G. Pennazza, P. Carratu, and O. Resta, "Electronic nose technology in respiratory diseases," *Lung*, vol. 195, no. 2, pp. 157–165, 2017.

[38] W. Yang, J. Yu, F. Pei, A. M. Mariga, N. Ma, Y. Fang, and Q. Hu, "Effect of hot air drying on volatile compounds of flammulina velutipes detected by hs-spme–gc–ms and electronic nose," *Food chemistry*, vol. 196, pp. 860–866, 2016.

[39] S. Grassi, E. Casiraghi, S. Benedetti, and C. Alamprese, "Effect of low-protein diets in heavy pigs on dry-cured ham quality characteristics," *Meat science*, vol. 131, pp. 152–157, 2017.

[40] L. Xu, X. Yu, L. Liu, and R. Zhang, "A novel method for qualitative analysis of edible oil oxidation using an electronic nose," *Food chemistry*, vol. 202, pp. 229–235, 2016.

[41] F. Röck, N. Barsan, and U. Weimar, "Electronic nose: current status and future trends," *Chemical reviews*, vol. 108, no. 2, pp. 705–725, 2008.

[42] S. Kiani, S. Minaei, and M. Ghasemi-Varnamkhasti, "Application of electronic nose systems for assessing quality of medicinal and aromatic plant products: A review," *Journal of Applied Research on Medicinal and Aromatic Plants*, vol. 3, no. 1, pp. 1–9, 2016.

[43] A. Sanaeifar, H. ZakiDizaji, A. Jafari, and M. de la Guardia, "Early detection of contamination and defect in foodstuffs by electronic nose: A review," *TrAC Trends in Analytical Chemistry*, vol. 97, pp. 257–271, 2017.

[44] A. Wilson, "Application of electronic-nose technologies and voc-biomarkers for the noninvasive early diagnosis of gastrointestinal diseases," *Sensors*, vol. 18, no. 8, p. 2613, 2018.

[45] A. Wilson, "Applications of electronic-nose technologies for noninvasive early detection of plant, animal and human diseases," *Chemosensors*, vol. 6, no. 4, p. 45, 2018.

[46] L. Zhang and D. Zhang, "Domain adaptation extreme learning machines for drift compensation in e-nose systems," *IEEE Transactions on instrumentation and measurement*, vol. 64, no. 7, pp. 1790–1801, 2015.

[47] J. Fonollosa, L. Fernandez, A. Gutiérrez-Gálvez, R. Huerta, and S. Marco, "Calibration transfer and drift counteraction in chemical sensor arrays using direct standardization," *Sensors and Actuators B: Chemical*, vol. 236, pp. 1044–1053, 2016.

[48] J. Wang, S. Rathi, B. Singh, I. Lee, S. Maeng, H.-I. Joh, and G.-H. Kim, "Dielectrophoretic assembly of pt nanoparticle-reduced graphene oxide nanohybrid for highly-sensitive multiple gas sensor," *Sensors and Actuators B: Chemical*, vol. 220, pp. 755–761, 2015.

[49] H. Jamalabadi, A. Mani-Varnosfaderani, and N. Alizadeh, "Detection of alkyl amine vapors using ppy-zno hybrid nanocomposite sensor array and artificial neural network," *Sensors and Actuators A: Physical*, vol. 280, pp. 228–237, 2018.

[50] Y. Yin, Y. Bai, F. Ge, H. Yu, and Y. Liu, "Long-term robust identification potential of a wavelet packet decomposition based recursive drift correction of e-nose data for chinese spirits," *Measurement*, 2019.

[51] G. Korotcenkov and B. Cho, "Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey)," *Sensors and Actuators B: Chemical*, vol. 156, no. 2, pp. 527–538, 2011.

[52] S. K. Jha and R. Yadava, "Denoising by singular value decomposition and its application to electronic nose data processing," *IEEE Sensors Journal*, vol. 11, no. 1, pp. 35–44, 2011.

[53] Z. Liang, F. Tian, C. Zhang, H. Sun, A. Song, and T. Liu, "Improving the robustness of prediction model by transfer learning for interference suppression of electronic nose," *IEEE Sensors Journal*, 2017.

[54] E. Hines, E. Llobet, and J. Gardner, "Electronic noses: a review of signal processing techniques," *IEE Proceedings-Circuits, Devices and Systems*, vol. 146, no. 6, pp. 297–310, 1999.

[55] Z. Sun and C. Chang, "Structural damage assessment based on wavelet packet transform," *Journal of structural engineering*, vol. 128, no. 10, pp. 1354–1361, 2002.

[56] N. El Barbri, J. Mirhisse, R. Ionescu, N. El Bari, X. Correig, B. Bouchikhi, and E. Llobet, "An electronic nose system based on a micro-machined gas sensor array to assess the freshness of sardines," *Sensors and Actuators B: Chemical*, vol. 141, no. 2, pp. 538–543, 2009.

[57] J. Lozano, J. Santos, and M. Horrillo, "Classification of white wine aromas with an electronic nose," *Talanta*, vol. 67, no. 3, pp. 610–616, 2005.

[58] S. Capone, M. Epifani, F. Quaranta, P. Siciliano, A. Taurino, and L. Vasanelli, "Monitoring of rancidity of milk by means of an electronic nose and a dynamic pca analysis," *Sensors and Actuators B: Chemical*, vol. 78, no. 1-3, pp. 174–179, 2001.

[59] N. El Barbri, E. Llobet, N. El Bari, X. Correig, and B. Bouchikhi, "Electronic nose based on metal oxide semiconductor sensors as an alternative technique for the spoilage classification of red meat," *Sensors*, vol. 8, no. 1, pp. 142–156, 2008.

[60] A. D'Amico, G. Pennazza, M. Santonico, E. Martinelli, C. Roscioni, G. Galluccio, R. Paolesse, and C. Di Natale, "An investigation on electronic nose diagnosis of lung cancer," *Lung cancer*, vol. 68, no. 2, pp. 170–176, 2010.

[61] G. Pillonetto and G. De Nicolao, "A new kernel-based approach for linear system identification," *Automatica*, vol. 46, no. 1, pp. 81–93, 2010.

[62] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung, "Kernel methods in system identification, machine learning and function estimation: A survey," *Automatica*, vol. 50, no. 3, pp. 657–682, 2014.

[63] L. Ye, A. Argha, H. Yu, B. G. Celler, H. T. Nguyen, and S. Su, "Dynamic characteristics of oxygen consumption," *Biomedical engineering online*, vol. 17, no. 1, p. 44, 2018.

[64] T. Eklöv, P. Mårtensson, and I. Lundström, "Selection of variables for interpreting multivariate gas sensor data," *Analytica Chimica Acta*, vol. 381, no. 2, pp. 221–232, 1999.

[65] M. Pardo and G. Sberveglieri, "Comparing the performance of different features in sensor arrays," *Sensors and Actuators B: Chemical*, vol. 123, no. 1, pp. 437–443, 2007.

[66] S. Singh, E. L. Hines, and J. W. Gardner, "Fuzzy neural computing of coffee and tainted-water data from an electronic nose," *Sensors and Actuators B: Chemical*, vol. 30, no. 3, pp. 185–190, 1996.

[67] E. Llobet, E. L. Hines, J. W. Gardner, P. N. Bartlett, and T. T. Mottram, "Fuzzy artmap based electronic nose data analysis," *Sensors and Actuators B: Chemical*, vol. 61, no. 1, pp. 183–190, 1999.

[68] E. Esme and B. Karlik, "Fuzzy c-means based support vector machines classifier for perfume recognition," *Applied Soft Computing*, vol. 46, pp. 452–458, 2016.

[69] V. S. Kodogiannis and A. Alshejari, "A fuzzy-wavelet neural network model for the detection of meat spoilage using an electronic nose," in *Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on*, pp. 710–717, IEEE, 2016.

[70] C.-F. Lin and S.-D. Wang, "Fuzzy support vector machines," *IEEE Transactions on neural networks*, vol. 13, no. 2, pp. 464–471, 2002.

[71] T. Inoue and S. Abe, "Fuzzy support vector machines for pattern classification," in *Neural Networks, 2001. Proceedings. IJCNN'01. International Joint Conference on*, vol. 2, pp. 1449–1454, IEEE, 2001.

[72] R. Batuwita and V. Palade, "Fsvm-cil: fuzzy support vector machines for class imbalance learning," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 3, pp. 558–571, 2010.

[73] R. Dressler, R. S. Wood, and V. Alvarez, "R&d metrics: Evaluating r&d performance using the cost savings metric," *Research-Technology Management*, vol. 42, no. 2, pp. 13–14, 1999.

[74] C. Vaalma, D. Buchholz, M. Weil, and S. Passerini, "A cost and resource analysis of sodium-ion batteries," *Nature Reviews Materials*, vol. 3, no. 4, p. 18013, 2018.

[75] S. Firestein, "How the olfactory system makes sense of scents," *Nature*, vol. 413, no. 6852, p. 211, 2001.

[76] J. P. McGann, "Poor human olfaction is a 19th-century myth," *Science*, vol. 356, no. 6338, p. eaam7263, 2017.

[77] G. M. Shepherd, "Smell images and the flavour system in the human brain," *Nature*, vol. 444, no. 7117, p. 316, 2006.

[78] J. D. Mainland, Y. R. Li, T. Zhou, W. L. L. Liu, and H. Matsunami, "Human olfactory receptor responses to odorants," *Scientific data*, vol. 2, p. 150002, 2015.

[79] P. Som and T. Naidich, "The olfactory system: Part ii: How olfaction is processed in the olfactory epithelium and olfactory bulb," *Neurographics*, vol. 8, no. 2, pp. 136–153, 2018.

[80] J. Mainland and N. Sobel, "The sniff is part of the olfactory percept," *Chemical senses*, vol. 31, no. 2, pp. 181–196, 2005.

[81] L. Sherwood, *Human physiology: from cells to systems*. Cengage learning, 2015.

[82] K. Persaud and G. Dodd, "Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose," *Nature*, vol. 299, no. 5881, p. 352, 1982.

[83] S. S. Schiffman and T. C. Pearce, "Introduction to olfaction: perception, anatomy, physiology, and molecular biology," *Handbook of Machine Olfaction: Electronic Nose Technology*, pp. 1–32, 2003.

[84] M. Ueland, K. Ewart, A. N. Troobnikoff, G. Frankham, R. N. Johnson, and S. L. Forbes, "A rapid chemical odour profiling method for the identification of rhinoceros horns," *Forensic science international*, vol. 266, pp. e99–e102, 2016.

[85] M. A. Iqbal, K. D. Nizio, M. Ueland, and S. L. Forbes, "Forensic decomposition odour profiling: A review of experimental designs and analytical techniques," *TrAC Trends in Analytical Chemistry*, vol. 91, pp. 112–124, 2017.

[86] S. Cui, J. Wang, L. Yang, J. Wu, and X. Wang, "Qualitative and quantitative analysis on aroma characteristics of ginseng at different ages using e-nose and gc–ms combined with chemometrics," *Journal of pharmaceutical and biomedical analysis*, vol. 102, pp. 64–77, 2015.

[87] B. Ramoo, M. Funke, C. Frazee, and U. Garg, "Comprehensive urine drug screen by gas chromatography/mass spectrometry (gc/ms)," in *Clinical Applications of Mass Spectrometry in Drug Analysis*, pp. 125–131, Springer, 2016.

[88] D. James, S. M. Scott, Z. Ali, and W. T. O'hare, "Chemical sensors for electronic nose systems," *Microchimica Acta*, vol. 149, no. 1-2, pp. 1–17, 2005.

[89] K. Arshak, E. Moore, G. Lyons, J. Harris, and S. Clifford, "A review of gas sensors employed in electronic nose applications," *Sensor review*, vol. 24, no. 2, pp. 181–198, 2004.

[90] A. Dey, "Semiconductor metal oxide gas sensors: A review," *Materials Science and Engineering: B*, vol. 229, pp. 206–217, 2018.

[91] R. Bogue, "Detecting gases with light: a review of optical gas sensor technologies," *Sensor Review*, vol. 35, no. 2, pp. 133–140, 2015.

[92] T.-V. Dinh, I.-Y. Choi, Y.-S. Son, and J.-C. Kim, "A review on non-dispersive infrared gas sensors: Improvement of sensor detection limit and interference correction," *Sensors and Actuators B: Chemical*, vol. 231, pp. 529–538, 2016.

[93] A. Mirzaei and G. Neri, "Microwave-assisted synthesis of metal oxide nanostructures for gas sensing application: A review," *Sensors and Actuators B: Chemical*, vol. 237, pp. 749–775, 2016.

[94] O. Hotel, J.-P. Poli, C. Mer-Calfati, E. Scorsone, and S. Saada, "A review of algorithms for saw sensors e-nose based volatile compound identification," *Sensors and Actuators B: Chemical*, vol. 255, pp. 2472–2482, 2018.

[95] T. Wasilewski, J. Gkebicki, and W. Kamysz, "Bioelectronic nose: Current status and perspectives," *Biosensors and Bioelectronics*, vol. 87, pp. 480–494, 2017.

[96] J. H. Sohn, M. Atzeni, L. Zeller, and G. Pioggia, "Characterisation of humidity dependence of a metal oxide semiconductor sensor array using partial least squares," *Sensors and Actuators B: Chemical*, vol. 131, no. 1, pp. 230–235, 2008.

[97] N. El Barbri, A. Amari, M. Vinaixa, B. Bouchikhi, X. Correig, and E. Llobet, "Building of a metal oxide gas sensor-based electronic nose to assess the freshness of sardines under cold storage," *Sensors and Actuators B: Chemical*, vol. 128, no. 1, pp. 235–244, 2007.

[98] M. García, M. Aleixandre, J. Gutiérrez, and M. Horrillo, "Electronic nose for wine discrimination," *Sensors and Actuators B: Chemical*, vol. 113, no. 2, pp. 911–916, 2006.

[99] L. A. Horsfall, D. C. Pugh, C. S. Blackman, and I. P. Parkin, "An array of wo 3 and cto heterojunction semiconducting metal oxide gas sensors used as a tool for explosive detection," *Journal of Materials Chemistry A*, vol. 5, no. 5, pp. 2172–2179, 2017.

[100] M. Tohidi, M. Ghasemi-Varnamkhasti, V. Ghafarinia, M. Bonyadian, and S. S. Mohtasebi, "Development of a metal oxide semiconductor-based artificial nose as a fast, reliable and non-expensive analytical technique for aroma profiling of milk adulteration," *International dairy journal*, vol. 77, pp. 38–46, 2018.

[101] F. Tian, Z. Liang, L. Zhang, Y. Liu, and Z. Zhao, "A novel pattern mismatch based interference elimination technique in e-nose," *Sensors and Actuators B: Chemical*, vol. 234, pp. 703–712, 2016.

[102] X. Yin, L. Zhang, F. Tian, and D. Zhang, "Temperature modulated gas sensing e-nose system for low-cost and fast detection," *IEEE Sensors Journal*, vol. 16, no. 2, pp. 464–474, 2016.

[103] W. Wojnowski, T. Majchrzak, T. Dymerski, J. Gkebicki, and J. Namieśnik, "Portable electronic nose based on electrochemical sensors for food quality assessment," *Sensors*, vol. 17, no. 12, p. 2715, 2017.

[104] J. Wang, "Electrochemical sensing of explosives," in *Counterterrorist detection techniques of explosives*, pp. 91–107, Elsevier, 2007.

[105] W. Collier, D. Baird, Z. Park-Ng, N. More, and A. Hart, "Discrimination among milks and cultured dairy products using screen-printed electrochemical arrays and an electronic nose," *Sensors and Actuators B: Chemical*, vol. 92, no. 1-2, pp. 232–239, 2003.

[106] J. Gebicki, "Application of electrochemical sensors and sensor matrixes for measurement of odorous chemical compounds," *TrAC Trends in Analytical Chemistry*, vol. 77, pp. 1–13, 2016.

[107] H. Li, X. Mu, Y. Yang, and A. J. Mason, "Low power multimode electrochemical gas sensor array system for wearable health and safety monitoring," *IEEE Sensors Journal*, vol. 14, no. 10, pp. 3391–3399, 2014.

[108] P. K. Sekhar and F. Wignes, "Trace detection of research department explosive (rdx) using electrochemical gas sensor," *Sensors and Actuators B: Chemical*, vol. 227, pp. 185–190, 2016.

[109] J. Hatfield, P. Neaves, P. Hicks, K. Persaud, and P. Travers, "Towards an integrated electronic nose using conducting polymer sensors," *Sensors and Actuators B: Chemical*, vol. 18, no. 1-3, pp. 221–228, 1994.

[110] R. Stella, J. N. Barisci, G. Serra, G. G. Wallace, and D. De Rossi, "Characterisation of olive oil by an electronic nose based on conducting polymer sensors," *Sensors and Actuators B: Chemical*, vol. 63, no. 1-2, pp. 1–9, 2000.

[111] J. B. Chang, V. Liu, V. Subramanian, K. Sivula, C. Luscombe, A. Murphy, J. Liu, and J. M. Fréchet, "Printable polythiophene gas sensor array for low-cost electronic noses," *Journal of Applied Physics*, vol. 100, no. 1, p. 014506, 2006.

[112] J. N. Barisci, G. G. Wallace, M. K. Andrews, A. C. Partridge, and P. D. Harris, "Conducting polymer sensors for monitoring aromatic hydrocarbons using an electronic nose," *Sensors and Actuators B: Chemical*, vol. 84, no. 2-3, pp. 252–257, 2002.

[113] J. W. Gardner and P. Bartlett, "Application of conducting polymer technology in microsystems," *Sensors and Actuators-A*, vol. 51, pp. 57–66, 1996.

[114] C. H. A. Esteves, B. A. Iglesias, T. Ogawa, K. Araki, L. Hoehne, and J. Gruber, "Identification of tobacco types and cigarette brands using an electronic nose based on conductive polymer/porphyrin composite sensors," *ACS omega*, vol. 3, no. 6, pp. 6476–6482, 2018.

[115] Z. Zhao, F. Tian, H. Liao, X. Yin, Y. Liu, and B. Yu, "A novel spectrum analysis technique for odor sensing in optical electronic nose," *Sensors and Actuators B: Chemical*, vol. 222, pp. 769–779, 2016.

[116] C. Di Natale, D. Salimbeni, R. Paolesse, A. Macagnano, and A. D'Amico, "Porphyrins-based opto-electronic nose for volatile compounds detection," *Sensors and Actuators B: Chemical*, vol. 65, no. 1-3, pp. 220–226, 2000.

[117] J. R. Askim, M. Mahmoudi, and K. S. Suslick, "Optical sensor arrays for chemical sensing: the optoelectronic nose," *Chemical Society Reviews*, vol. 42, no. 22, pp. 8649–8682, 2013.

[118] K. J. Albert, D. R. Walt, D. S. Gill, and T. C. Pearce, "Optical multibead arrays for simple and complex odor discrimination," *Analytical Chemistry*, vol. 73, no. 11, pp. 2501–2508, 2001.

[119] A. T. Semeano, D. F. Maffei, S. Palma, R. W. Li, B. D. Franco, A. C. Roque, and J. Gruber, "Tilapia fish microbial spoilage monitored by a single optical gas sensor," *Food control*, vol. 89, pp. 72–76, 2018.

[120] D. Compagnone, M. Faieta, D. Pizzoni, C. Di Natale, R. Paolesse, T. Van Caelenberg, B. Beheydt, and P. Pittia, "Quartz crystal microbalance gas sensor arrays for the quality control of chocolate," *Sensors and Actuators B: Chemical*, vol. 207, pp. 1114–1120, 2015.

[121] X. Jin, Y. Huang, A. Mason, and X. Zeng, "Multichannel monolithic quartz crystal microbalance gas sensor array," *Analytical chemistry*, vol. 81, no. 2, pp. 595–603, 2008.

[122] W. Tao, P. Lin, Y. Ai, H. Wang, S. Ke, and X. Zeng, "Multichannel quartz crystal microbalance array: Fabrication, evaluation, application in biomarker detection," *Analytical biochemistry*, vol. 494, pp. 85–92, 2016.

[123] K. Zhang, R. Hu, G. Fan, and G. Li, "Graphene oxide/chitosan nanocomposite coated quartz crystal microbalance sensor for detection of amine vapors," *Sensors and Actuators B: Chemical*, vol. 243, pp. 721–730, 2017.

[124] S. Muñoz-Aguirre, A. Yoshino, T. Nakamoto, and T. Moriizumi, "Odor approximation of fruit flavors using a qcm odor sensing system," *Sensors and Actuators B: Chemical*, vol. 123, no. 2, pp. 1101–1106, 2007.

[125] F. Temel and M. Tabakci, "Calix [4] arene coated qcm sensors for detection of voc emissions: Methylene chloride sensing studies," *Talanta*, vol. 153, pp. 221–227, 2016.

[126] A. Mujahid and F. Dickert, "Surface acoustic wave (saw) for chemical sensing applications of recognition layers," *Sensors*, vol. 17, no. 12, p. 2716, 2017.

[127] F. Gao, F. Boussaid, W. Xuan, C.-Y. Tsui, and A. Bermak, "Dual transduction surface acoustic wave gas sensor for voc discrimination," *IEEE Electron Device Letters*, vol. 39, no. 12, pp. 1920–1923, 2018.

[128] J. Devkota, P. Ohodnicki, and D. Greve, "Saw sensors for chemical vapors and gases," *Sensors*, vol. 17, no. 4, p. 801, 2017.

[129] J. J. Caron, R. B. Haskell, C. J. Freeman, and J. F. Vetelino, "Surface acoustic wave mercury vapor sensors," Nov. 30 1999. US Patent 5,992,215.

[130] A. Marina, Y. Che Man, and I. Amin, "Use of the saw sensor electronic nose for detecting the adulteration of virgin coconut oil with rbd palm kernel olein," *Journal of the American Oil Chemists' Society*, vol. 87, no. 3, pp. 263–270, 2010.

[131] J. Santos, M. Fernandez, J. Fontecha, J. Lozano, M. Aleixandre, M. Garcia, J. Gutierrez, and M. Horrillo, "Saw sensor array for wine discrimination," *Sensors and Actuators B: Chemical*, vol. 107, no. 1, pp. 291–295, 2005.

[132] S. Trowell, A. Berna, H. Dacres, M. Tehseen, and A. Anderson, "Is cybernose® an instrument for measuring odor space?," *Flavour*, vol. 3, no. 1, p. O17, 2014.

[133] Q. Liu, H. Cai, Y. Xu, Y. Li, R. Li, and P. Wang, "Olfactory cell-based biosensor: a first step towards a neurochip of bioelectronic nose," *Biosensors and Bioelectronics*, vol. 22, no. 2, pp. 318–322, 2006.

[134] Q. Liu, W. Ye, H. Yu, N. Hu, L. Du, P. Wang, and M. Yang, "Olfactory mucosa tissue-based biosensor: a bioelectronic nose with receptor cells in intact olfactory epithelium," *Sensors and Actuators B: Chemical*, vol. 146, no. 2, pp. 527–533, 2010.

[135] A. Z. Berna, A. R. Anderson, and S. C. Trowell, "Bio-benchmarking of electronic nose sensors," *PloS one*, vol. 4, no. 7, p. e6406, 2009.

[136] T.-Z. Wu, "A piezoelectric biosensor as an olfactory receptor for odour detection: electronic nose," *Biosensors and Bioelectronics*, vol. 14, no. 1, pp. 9–18, 1999.

[137] N. Yamazoe and K. Shimanoe, "Basic approach to the transducer function of oxide semiconductor gas sensors," *Sensors and Actuators B: Chemical*, vol. 160, no. 1, pp. 1352–1362, 2011.

[138] NASA, "Electronic nose:nasa researchers are developing an exquisitely sensitive artificial nose for space exploration." Online, 2004. https://science.nasa.gov/science-news/science-at-nasa/2004/06oct_enose.

[139] R. Teghtsoonian and M. Teghtsoonian, "Testing a perceptual constancy model for odor strength: the effects of sniff pressure and resistance to sniffing," *Perception*, vol. 13, no. 6, pp. 743–752, 1984.

[140] E. Bocca, A. Antonelli, and O. Mosciaro, "Mechanical co-factors in olfactory stimulation," *Acta Oto-Laryngologica*, vol. 59, no. 2-6, pp. 243–247, 1965.

[141] B. Malnic, J. Hirono, T. Sato, and L. B. Buck, "Combinatorial receptor codes for odors," *Cell*, vol. 96, no. 5, pp. 713–723, 1999.

[142] M. Pardo, G. Niederjaufner, G. Benussi, E. Comini, G. Faglia, G. Sberveglieri, M. Holmberg, and I. Lundstrom, "Data preprocessing enhances the classification of different brands of espresso coffee with an electronic nose," *Sensors and Actuators B: Chemical*, vol. 69, no. 3, pp. 397–403, 2000.

[143] D. Wilson, S. Garrod, S. Hoyt, S. McKennoch, and K. Booksh, "Array optimization and preprocessing techniques for chemical sensing microsystems," *Sensors Update*, vol. 10, no. 1, pp. 77–106, 2002.

[144] S. K. Jha and R. Yadava, "Preprocessing of saw sensor array data and pattern recognition," *IEEE Sensors Journal*, vol. 9, no. 10, pp. 1202–1208, 2009.

[145] A. Reznik, A. Galinskaya, O. Dekhtyarenko, and D. Nowicki, "Preprocessing of matrix qcm sensors data for the classification by means of neural network," *Sensors and Actuators B: Chemical*, vol. 106, no. 1, pp. 158–163, 2005.

[146] W. Li, H. Leung, C. Kwan, and B. R. Linnell, "E-nose vapor identification based on dempster–shafer fusion of multiple classifiers," *IEEE Transactions on instrumentation and measurement*, vol. 57, no. 10, pp. 2273–2282, 2008.

[147] S. Di Carlo and M. Falasconi, "Drift correction methods for gas chemical sensors in artificial olfaction systems: techniques and challenges," in *Advances in Chemical Sensors*, InTech, 2012.

[148] W. Hidayat, A. Y. M. Shakaff, M. N. Ahmad, and A. H. Adom, "Classification of agarwood oil using an electronic nose," *Sensors*, vol. 10, no. 5, pp. 4675–4685, 2010.

[149] A. Leone, C. Distante, N. Ancona, K. Persaud, E. Stella, and P. Siciliano, "A powerful method for feature extraction and compression of electronic nose responses," *Sensors and Actuators B: Chemical*, vol. 105, no. 2, pp. 378–392, 2005.

[150] C. Zanchettin and T. B. Ludermir, "Wavelet filter for noise reduction and signal compression in an artificial nose," *Applied Soft Computing*, vol. 7, no. 1, pp. 246–256, 2007.

[151] J. Gardner, M. Craven, C. Dow, and E. Hines, "The prediction of bacteria type and culture growth phase by an electronic nose with a multi-layer perceptron network," *Measurement Science and Technology*, vol. 9, no. 1, p. 120, 1998.

[152] G. Pillonetto, A. Chiuso, and G. De Nicolao, "Prediction error identification of linear systems: a nonparametric gaussian regression approach," *Automatica*, vol. 47, no. 2, pp. 291–305, 2011.

[153] A. N. Tikhonov and V. I. Arsenin, *Solutions of ill-posed problems*, vol. 14. Vh Winston, 1977.

[154] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *Journal of the ACM (JACM)*, vol. 9, no. 1, pp. 84–97, 1962.

[155] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.

[156] D. Whitley, "A genetic algorithm tutorial," *Statistics and computing*, vol. 4, no. 2, pp. 65–85, 1994.

[157] D. S. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: A review," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 3, pp. 343–353, 1997.

[158] Y. Deng, Y. Liu, and D. Zhou, "An improved genetic algorithm with initial population strategy for symmetric tsp," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[159] X. Du, K. K. K. Htet, and K. K. Tan, "Development of a genetic-algorithm-based nonlinear model predictive control scheme on velocity and steering of autonomous vehicles," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 6970–6977, 2016.

[160] C. Li, P. H. Heinemann, and P. Reed, "Genetic algorithms (gas) and evolutionary strategy to optimize electronic nose sensor selection," *Transactions of the ASABE*, vol. 51, no. 1, pp. 321–330, 2007.

[161] J. Gardner, P. Boilot, and E. Hines, "Enhancing electronic nose performance by sensor selection using a new integer-based genetic algorithm approach," *Sensors and Actuators B: Chemical*, vol. 106, no. 1, pp. 114–121, 2005.

[162] B. Shi, L. Zhao, R. Zhi, and X. Xi, "Optimization of electronic nose sensor array by genetic algorithms in xihu-longjing tea quality analysis," *Mathematical and Computer modelling*, vol. 58, no. 3, pp. 752–758, 2013.

[163] X.-Y. Tian, Q. Cai, and Y.-M. Zhang, "Rapid classification of hairtail fish and pork freshness using an electronic nose based on the pca method," *Sensors*, vol. 12, no. 1, pp. 260–277, 2012.

[164] H. Yu and J. Wang, "Discrimination of longjing green-tea grade by electronic nose," *Sensors and Actuators B: Chemical*, vol. 122, no. 1, pp. 134–140, 2007.

[165] A. H. Gómez, J. Wang, G. Hu, and A. G. Pereira, "Electronic nose technique potential monitoring mandarin maturity," *Sensors and Actuators B: Chemical*, vol. 113, no. 1, pp. 347–353, 2006.

[166] R. Dutta, E. L. Hines, J. W. Gardner, and P. Boilot, "Bacteria classification using cyranose 320 electronic nose," *Biomedical engineering online*, vol. 1, no. 1, p. 4, 2002.

[167] J. Brezmes, E. Llobet, X. Vilanova, G. Saiz, and X. Correig, "Fruit ripeness monitoring using an electronic nose," *Sensors and Actuators B: Chemical*, vol. 69, no. 3, pp. 223–229, 2000.

[168] S. Dragonieri, R. Schot, B. J. Mertens, S. Le Cessie, S. A. Gauw, A. Spanevello, O. Resta, N. P. Willard, T. J. Vink, K. F. Rabe, *et al.*, "An electronic nose in the discrimination of patients with asthma and controls," *Journal of Allergy and Clinical Immunology*, vol. 120, no. 4, pp. 856–862, 2007.

[169] D. Melucci, A. Bendini, F. Tesini, S. Barbieri, A. Zappi, S. Vichi, L. Conte, and T. G. Toschi, "Rapid direct analysis to discriminate geographic origin of extra virgin olive oils by flash gas chromatography electronic nose and chemometrics," *Food chemistry*, vol. 204, pp. 263–273, 2016.

[170] A. Romero-Flores, L. L. McConnell, C. J. Hapeman, M. Ramirez, and A. Torrents, "Evaluation of an electronic nose for odorant and process monitoring of alkaline-stabilized biosolids production," *Chemosphere*, vol. 186, pp. 151–159, 2017.

[171] L.-Y. Chen, C.-C. Wu, T.-I. Chou, S.-W. Chiu, and K.-T. Tang, "Development of a dual mos electronic nose/camera system for improving fruit ripeness classification," *Sensors*, vol. 18, no. 10, p. 3256, 2018.

[172] T. Saidi, O. Zaim, M. Moufid, N. El Bari, R. Ionescu, and B. Bouchikhi, "Exhaled breath analysis using electronic nose and gas chromatography–mass spectrometry for non-invasive diagnosis of chronic kidney disease, diabetes mellitus and healthy subjects," *Sensors and Actuators B: Chemical*, vol. 257, pp. 178–188, 2018.

[173] K. Brudzewski, S. Osowski, and T. Markiewicz, "Classification of milk by means of an electronic nose and svm neural network," *Sensors and Actuators B: Chemical*, vol. 98, no. 2-3, pp. 291–298, 2004.

[174] H. Liu, Q. Li, B. Yan, L. Zhang, and Y. Gu, "Bionic electronic nose based on mos sensors array and machine learning algorithms used for wine properties detection," *Sensors*, vol. 19, no. 1, p. 45, 2019.

[175] M. Pardo and G. Sberveglieri, "Classification of electronic nose data with support vector machines," *Sensors and Actuators B: Chemical*, vol. 107, no. 2, pp. 730–737, 2005.

[176] Z. Haddi, A. Amari, H. Alami, N. El Bari, E. Llobet, and B. Bouchikhi, "A portable electronic nose system for the identification of cannabis-based drugs," *Sensors and Actuators B: Chemical*, vol. 155, no. 2, pp. 456–463, 2011.

[177] Q. Chen, J. Zhao, Z. Chen, H. Lin, and D.-A. Zhao, "Discrimination of green tea quality using the electronic nose technique and the human panel test, comparison of linear and nonlinear classification tools," *Sensors and Actuators B: Chemical*, vol. 159, no. 1, pp. 294–300, 2011.

[178] S. Qiu, L. Gao, and J. Wang, "Classification and regression of elm, lvq and svm for e-nose data of strawberry juice," *Journal of Food Engineering*, vol. 144, pp. 77–85, 2015.

[179] M. Trincavelli, S. Coradeschi, A. Loutfi, B. Soderquist, and P. Thunberg, "Direct identification of bacteria in blood culture samples using an electronic nose," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2884–2890, 2010.

[180] S. Buratti, S. Benedetti, M. Scampicchio, and E. Pangerod, "Characterization and classification of italian barbera wines by using an electronic nose and an amperometric electronic tongue," *Analytica Chimica Acta*, vol. 525, no. 1, pp. 133–139, 2004.

[181] M. Cosio, D. Ballabio, S. Benedetti, and C. Gigliotti, "Evaluation of different storage conditions of extra virgin olive oils with an innovative recognition tool built by means of electronic nose and electronic tongue," *Food Chemistry*, vol. 101, no. 2, pp. 485–491, 2007.

[182] J. Goschnick, I. Koronczi, M. Frietsch, and I. Kiselev, "Water pollution recognition with the electronic nose kamina," *Sensors and Actuators B: Chemical*, vol. 106, no. 1, pp. 182–186, 2005.

[183] X. Hong, J. Wang, and Z. Hai, "Discrimination and prediction of multiple beef freshness indexes based on electronic nose," *Sensors and Actuators B: Chemical*, vol. 161, no. 1, pp. 381–389, 2012.

[184] S. Panigrahi, S. Balasubramanian, H. Gu, C. Logue, and M. Marchello, "Design and development of a metal oxide based electronic nose for spoilage classification of beef," *Sensors and Actuators B: Chemical*, vol. 119, no. 1, pp. 2–14, 2006.

[185] Y. G. Martın, M. C. C. Oliveros, J. L. P. Pavón, C. G. Pinto, and B. M. Cordero, "Electronic nose based on metal oxide semiconductor sensors and pattern recognition techniques: characterisation of vegetable oils," *Analytica Chimica Acta*, vol. 449, no. 1-2, pp. 69–80, 2001.

[186] Q. Zhang, S. Zhang, C. Xie, D. Zeng, C. Fan, D. Li, and Z. Bai, "Characterization of chinese vinegars by electronic nose," *Sensors and Actuators B: Chemical*, vol. 119, no. 2, pp. 538–546, 2006.

[187] H. Yu, J. Wang, C. Yao, H. Zhang, and Y. Yu, "Quality grade identification of green tea using e-nose by ca and ann," *LWT-Food Science and Technology*, vol. 41, no. 7, pp. 1268–1273, 2008.

[188] M. Holmberg, F. Winquist, I. Lundström, J. Gardner, and E. Hines, "Identification of paper quality using a hybrid electronic nose," *Sensors and Actuators B: Chemical*, vol. 27, no. 1-3, pp. 246–249, 1995.

[189] L. Torri, N. Sinelli, and S. Limbo, "Shelf life evaluation of fresh-cut pineapple by using an electronic nose," *Postharvest biology and technology*, vol. 56, no. 3, pp. 239–245, 2010.

[190] S. Benedetti, N. Sinelli, S. Buratti, and M. Riva, "Shelf life of crescenza cheese as measured by electronic nose," *Journal of dairy science*, vol. 88, no. 9, pp. 3044–3051, 2005.

[191] Q. Li, X. Yu, L. Xu, and J.-M. Gao, "Novel method for the producing area identification of zhongning goji berries by electronic nose," *Food chemistry*, vol. 221, pp. 1113–1119, 2017.

[192] R. Dutta, K. Kashwan, M. Bhuyan, E. L. Hines, and J. Gardner, "Electronic nose based tea quality standardization," *Neural Networks*, vol. 16, no. 5-6, pp. 847–853, 2003.

[193] H. Singh, V. B. Raj, J. Kumar, U. Mittal, M. Mishra, A. Nimal, M. Sharma, and V. Gupta, "Metal oxide saw e-nose employing pca and ann for the identification of binary mixture of dmmp and methanol," *Sensors and Actuators B: Chemical*, vol. 200, pp. 147–156, 2014.

[194] J. Tan and W. L. Kerr, "Determining degree of roasting in cocoa beans by artificial neural network (ann)-based electronic nose system and gas chromatography/mass spectrometry (gc/ms)," *Journal of the Science of Food and Agriculture*, vol. 98, no. 10, pp. 3851–3859, 2018.

[195] H. Zhang, M. Chang, J. Wang, and S. Ye, "Evaluation of peach quality indices using an electronic nose by mlr, qpst and bp network," *Sensors and Actuators B: Chemical*, vol. 134, no. 1, pp. 332–338, 2008.

[196] B. G. Kermani, S. S. Schiffman, and H. T. Nagle, "Using neural networks and genetic algorithms to enhance performance in an electronic nose," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 4, pp. 429–439, 1999.

[197] M. S. Cosio, D. Ballabio, S. Benedetti, and C. Gigliotti, "Geographical origin and authentication of extra virgin olive oils by an electronic nose in combination with artificial neural networks," *Analytica Chimica Acta*, vol. 567, no. 2, pp. 202–210, 2006.

[198] L. Zhang and F. Tian, "Performance study of multilayer perceptrons in a low-cost electronic nose," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 7, pp. 1670–1679, 2014.

[199] H. Luo, P. Jia, S. Qiao, and S. Duan, "Enhancing electronic nose performance based on a novel qpso-rbm technique," *Sensors and Actuators B: Chemical*, vol. 259, pp. 241–249, 2018.

[200] L. Zhang and D. Zhang, "Evolutionary cost-sensitive extreme learning machine," *IEEE transactions on neural networks and learning systems*, vol. 28, no. 12, pp. 3045–3060, 2017.

[201] C. Peng, J. Yan, S. Duan, L. Wang, P. Jia, and S. Zhang, "Enhancing electronic nose performance based on a novel qpso-kelm model," *Sensors*, vol. 16, no. 4, p. 520, 2016.

[202] L. Zhang, D. Zhang, X. Yin, and Y. Liu, "A novel semi-supervised learning approach in artificial olfaction for e-nose application," *IEEE Sensors Journal*, vol. 16, no. 12, pp. 4919–4931, 2016.

[203] A. He, G. Wei, J. Yu, Z. Tang, Z. Lin, and P. Wang, "A novel dictionary learning method for gas identification with a gas sensor array," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 12, pp. 9709–9715, 2017.

[204] M. Li, H. Ruan, Y. Qi, T. Guo, P. Wang, and G. Pan, "Odor recognition with a spiking neural network for bioelectronic nose," *Sensors*, vol. 19, no. 5, p. 993, 2019.

[205] J. Lever, M. Krzywinski, and N. Altman, "Points of significance: Principal component analysis," 2017.

[206] M. Ghasemi-Varnamkhasti, S. S. Mohtasebi, M. Siadat, and S. Balasubramanian, "Meat quality assessment by electronic nose (machine olfaction technology)," *Sensors*, vol. 9, no. 8, pp. 6058–6083, 2009.

[207] F. Artoni, A. Delorme, and S. Makeig, "Applying dimension reduction to eeg data by principal component analysis reduces the quality of its subsequent independent component decomposition," *NeuroImage*, vol. 175, pp. 176–187, 2018.

[208] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[209] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[210] G.-m. Xian, "An identification method of malignant and benign liver tumors from ultrasonography based on glcm texture features and fuzzy svm," *Expert Systems with Applications*, vol. 37, no. 10, pp. 6737–6741, 2010.

[211] S. Haykin, "Neural networks, a comprehensive foundation second edition by prentice-hall," 1999.

[212] J. Gębicki and B. Szulczyński, "Discrimination of selected fungi species based on their odour profile using prototypes of electronic nose instruments," *Measurement*, vol. 116, pp. 307–313, 2018.

[213] M. Gancarz, J. Wawrzyniak, M. Gawrysiak-Witulska, D. Wiącek, A. Nawrocka, M. Tadla, and R. Rusinek, "Application of electronic nose with mos sensors to prediction of rapeseed quality," *Measurement*, vol. 103, pp. 227–234, 2017.

[214] S. Kiani, S. Minaei, and M. Ghasemi-Varnamkhasti, "Real-time aroma monitoring of mint (mentha spicata l.) leaves during the drying process using electronic nose system," *Measurement*, vol. 124, pp. 447–452, 2018.

[215] M. Ezhilan, N. Nesakumar, K. J. Babu, C. Srinandan, and J. B. B. Rayappan, "Freshness assessment of broccoli using electronic nose," *Measurement*, 2019.

[216] M. S. Amaral and P. J. Marriott, "The blossoming of technology for the analysis of complex aroma bouquets—a review on flavour and odorant multidimensional and comprehensive gas chromatography applications," *Molecules*, vol. 24, no. 11, p. 2080, 2019.

[217] S. E. Prebihalo, K. L. Berrier, C. E. Freye, H. D. Bahaghighat, N. R. Moore, D. K. Pinkerton, and R. E. Synovec, "Multidimensional gas chromatography: advances in instrumentation, chemometrics, and applications," *Analytical chemistry*, vol. 90, no. 1, pp. 505–532, 2017.

[218] W. Zhang, T. Liu, L. Ye, M. Ueland, S. L. Forbes, and S. W. Su, "A novel data pre-processing method for odour detection and identification system," *Sensors and Actuators A: Physical*, vol. 287, pp. 113–120, 2019.

[219] C. J. Gajanayake, S. D. G. Jayasinghe, S. Nadarajan, and A. K. Gupta, "Open switch fault detection and identification in a two-level voltage source power converter," May 8 2018. US Patent 9,964,600.

[220] B. S. J. Costa, P. P. Angelov, and L. A. Guedes, "Fully unsupervised fault detection and identification based on recursive density estimation and self-evolving cloud-based classifier," *Neurocomputing*, vol. 150, pp. 289–303, 2015.

[221] A. Abbaspour, P. Aboutalebi, K. K. Yen, and A. Sargolzaei, "Neural adaptive observer-based sensor and actuator fault detection in nonlinear systems: Application in uav," *ISA transactions*, vol. 67, pp. 317–329, 2017.

[222] M. M. Bordbar, J. Tashkhourian, and B. Hemmateenejad, "Qualitative and quantitative analysis of toxic materials in adulterated fruit pickle samples by a colorimetric sensor array," *Sensors and Actuators B: Chemical*, vol. 257, pp. 783–791, 2018.

[223] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.

[224] H. Lebret and S. Boyd, "Antenna array pattern synthesis via convex optimization," *IEEE transactions on signal processing*, vol. 45, no. 3, pp. 526–532, 1997.

[225] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale l1 regularized least squares," *IEEE journal of selected topics in signal processing*, vol. 1, no. 4, pp. 606–617, 2007.

[226] L. Zhang, X. Wang, G.-B. Huang, T. Liu, and X. Tan, "Taste recognition in e-tongue using local discriminant preservation projection," *IEEE Transactions on Cybernetics*, 2018.

[227] N. Bhattacharyya, R. Bandyopadhyay, M. Bhuyan, B. Tudu, D. Ghosh, and A. Jana, "Electronic nose for black tea classification and correlation of measurements with "tea taster" marks," *IEEE transactions on instrumentation and measurement*, vol. 57, no. 7, pp. 1313–1321, 2008.

[228] W. Cynkar, R. Dambergs, P. Smith, and D. Cozzolino, "Classification of tempranillo wines according to geographic origin: Combination of mass spectrometry based electronic nose and chemometrics," *Analytica Chimica Acta*, vol. 660, no. 1, pp. 227–231, 2010.

[229] P. S. Gromski, E. Correa, A. A. Vaughan, D. C. Wedge, M. L. Turner, and R. Goodacre, "A comparison of different chemometrics approaches for the robust classification of electronic nose data," *Analytical and bioanalytical chemistry*, vol. 406, no. 29, pp. 7581–7590, 2014.

[230] P.-Y. Hao, J.-H. Chiang, *et al.*, "A fuzzy model of support vector regression machine," *International Journal of Fuzzy Systems*, vol. 9, no. 1, 2007.

[231] Y. Xu, J. B. Weaver, D. M. Healy, and J. Lu, "Wavelet transform domain filters: a spatially selective noise filtration technique," *IEEE transactions on image processing*, vol. 3, no. 6, pp. 747–758, 1994.

[232] H. R. Estakhroueiyeh and E. Rashedi, "Detecting moldy bread using an e-nose and the knn classifier," in *Computer and Knowledge Engineering (ICCKE), 2015 5th International Conference on*, pp. 251–255, IEEE, 2015.

[233] S. Bedoui, H. Samet, M. Samet, and A. Kachouri, "Gases identification with support vector machines technique (svms)," in *Advanced Technologies for Signal and Image Processing (ATSIP), 2014 1st International Conference on*, pp. 271–276, IEEE, 2014.

[234] C.-L. Huang and C.-J. Wang, "A ga-based feature selection and parameters optimizationfor support vector machines," *Expert Systems with applications*, vol. 31, no. 2, pp. 231–240, 2006.

[235] M. Phillips, N. Altorki, J. H. Austin, R. B. Cameron, R. N. Cataneo, J. Greenberg, R. Kloss, R. A. Maxfield, M. I. Munawar, H. I. Pass, *et al.*, "Prediction of lung cancer using volatile biomarkers in breath," *Cancer biomarkers*, vol. 3, no. 2, pp. 95–109, 2007.

[236] E. R. Thaler and C. W. Hanson, "Medical applications of electronic nose technology," *Expert review of medical devices*, vol. 2, no. 5, pp. 559–566, 2005.

[237] A. P. Turner and N. Magan, "Electronic noses and disease diagnostics," *Nature Reviews Microbiology*, vol. 2, no. 2, pp. 161–166, 2004.

[238] J. W. Gardner, H. W. Shin, and E. L. Hines, "An electronic nose system to diagnose illness," *Sensors and Actuators B: Chemical*, vol. 70, no. 1, pp. 19–24, 2000.

[239] G. Gavazzi and K.-H. Krause, "Ageing and infection," *The Lancet infectious diseases*, vol. 2, no. 11, pp. 659–666, 2002.

[240] C. Di Natale, A. Macagnano, E. Martinelli, R. Paolesse, G. D'Arcangelo, C. Roscioni, A. Finazzi-Agrò, and A. D'Amico, "Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors," *Biosensors and Bioelectronics*, vol. 18, no. 10, pp. 1209–1218, 2003.

[241] M. S. Kirkman, V. J. Briscoe, N. Clark, H. Florez, L. B. Haas, J. B. Halter, E. S. Huang, M. T. Korytkowski, M. N. Munshi, P. S. Odegard, *et al.*, "Diabetes in older adults," *Diabetes care*, vol. 35, no. 12, pp. 2650–2664, 2012.

[242] A. K. Pavlou, N. Magan, C. McNulty, J. M. Jones, D. Sharp, J. Brown, and A. P. Turner, "Use of an electronic nose system for diagnoses of urinary tract infections," *Biosensors and Bioelectronics*, vol. 17, no. 10, pp. 893–899, 2002.

[243] A. Z. Berna, J. S. McCarthy, R. X. Wang, K. J. Saliba, F. G. Bravo, J. Cassells, B. Padovan, and S. C. Trowell, "Analysis of breath specimens for biomarkers of plasmodium falciparum infection," *The Journal of infectious diseases*, vol. 212, no. 7, pp. 1120–1128, 2015.

[244] L. C. Martins, M. R. D. de Oliveira Latorre, P. H. N. Saldiva, and A. L. F. Braga, "Air pollution and emergency room visits due to chronic lower respiratory diseases in the elderly: an ecological time-series study in sao paulo, brazil," *Journal of occupational and environmental medicine*, vol. 44, no. 7, pp. 622–627, 2002.

[245] A. Seaton, D. Godden, W. MacNee, and K. Donaldson, "Particulate air pollution and acute health effects," *The lancet*, vol. 345, no. 8943, pp. 176–178, 1995.

[246] S. Reinberg, "About 1 in 7 older adults has lung disease: Cdc." https://www.webmd.com/lung/copd/news/20150106/about-1-in-7-older-adults-has-some-form-of-lung-disease-cdc#1. Accessed Jan. 6, 2015.

[247] C. Deng, J. Zhang, X. Yu, W. Zhang, and X. Zhang, "Determination of acetone in human breath by gas chromatography–mass spectrometry and solid-phase microextraction with on-fiber derivatization," *Journal of Chromatography B*, vol. 810, no. 2, pp. 269–275, 2004.

[248] B. Buszewski, A. Ulanowska, T. Ligor, N. Denderz, and A. Amann, "Analysis of exhaled breath from smokers, passive smokers and non-smokers by solid-phase microextraction gas chromatography/mass spectrometry," *Biomedical Chromatography*, vol. 23, no. 5, pp. 551–556, 2009.

[249] M. J. Wilde, R. L. Cordell, D. Salman, B. Zhao, W. Ibrahim, L. Bryant, D. Ruszkiewicz, A. Singapuri, R. C. Free, E. A. Gaillard, *et al.*, "Breath analysis by two-dimensional gas chromatography with dual flame ionisation and mass spectrometric detection–method optimisation and integration within a large-scale clinical study," *Journal of Chromatography A*, vol. 1594, pp. 160–172, 2019.

[250] X. R. Wang, J. Cassells, and A. Z. Berna, "Stability control for breath analysis using gc-ms," *Journal of Chromatography B*, vol. 1097, pp. 27–34, 2018.

[251] M. Pardo and G. Sberveglieri, "Classification of electronic nose data with support vector machines," *Sensors and Actuators B: Chemical*, vol. 107, no. 2, pp. 730–737, 2005.

[252] Z. Wang and X. Xue, "Multi-class support vector machine," in *Support Vector Machines Applications*, pp. 23–48, Springer, 2014.

[253] W. C. T. CHALLENGE, "Thank you to the challenge prize winners for their outstanding work!." https://wildlifecrimetech.org/blog?article_id=. November. 6, 2017.

[254] R. Ogden, N. Dawnay, and R. McEwing, "Wildlife dna forensics—bridging the gap between conservation genetics and law enforcement," *Endangered Species Research*, vol. 9, no. 3, pp. 179–195, 2009.

[255] C. Nellemann, R. Henriksen, P. Raxter, N. Ash, E. Mrema, *et al.*, *The environmental crime crisis: threats to sustainable development from illegal exploitation and trade in wildlife and forest resources.* United Nations Environment Programme (UNEP), 2014.

[256] R. Martin, "Grey areas: temporal and geographical dynamics of international trade of grey and timneh parrots (psittacus erithacus and p. timneh) under cites," *Emu-Austral Ornithology*, vol. 118, no. 1, pp. 113–125, 2018.

[257] T. C. Haas and S. M. Ferreira, "Conservation risks: When will rhinos be extinct?," *IEEE transactions on cybernetics*, vol. 46, no. 8, pp. 1721–1734, 2016.

[258] J. Rademeyer, *Killing for profit: Exposing the illegal rhino horn trade*. Penguin Random House South Africa, 2012.

[259] A. Collins, C. Cox, and N. Pamment, "Culture, conservation and crime: regulating ivory markets for antiques and crafts," *Ecological Economics*, vol. 135, pp. 186–194, 2017.

[260] A. Hall-Martin, N. Van Der Merwe, J. Lee-Thorp, R. Armstrong, C. Mehl, S. Struben, and R. Tykot, "Determination of species and geographic origin of rhinoceros horn by isotopic analysis and its possible application to trade control," in *Proceedings of an international conference: rhinoceros biology and conservation (OA Ryder, ed.). Zoological Society of San Diego, San Diego, California*, pp. 123–135, 1993.

[261] S. K. Gupta, S. K. Verma, and L. Singh, "Molecular insight into a wildlife crime: the case of a peafowl slaughter," *Forensic science international*, vol. 154, no. 2-3, pp. 214–217, 2005.

[262] K. Perrault, B. Stuart, and S. Forbes, "A longitudinal study of decomposition odour in soil using sorbent tubes and solid phase microextraction," *Chromatography*, vol. 1, no. 3, pp. 120–140, 2014.

[263] Y. Soudry, C. Lemogne, D. Malinvaud, S.-M. Consoli, and P. Bonfils, "Olfactory system and emotion: common substrates," *European annals of otorhinolaryngology, head and neck diseases*, vol. 128, no. 1, pp. 18–23, 2011.

[264] H.-M. Hsieh, L.-H. Huang, L.-C. Tsai, Y.-C. Kuo, H.-H. Meng, A. Linacre, and J. C.-I. Lee, "Species identification of rhinoceros horns using the cytochrome b gene," *Forensic Science International*, vol. 136, no. 1-3, pp. 1–11, 2003.

[265] R. Amin, M. Bramer, and R. Emslie, "Intelligent data analysis for conservation: experiments with rhino horn fingerprint identification," in *Applications and Innovations in Intelligent Systems X*, pp. 207–222, Springer, 2003.

[266] T. Milliken, J. Shaw, and R. H. Emslie, "The south africa–vietnam rhino horn trade nexus: A deadly combination of institutional lapses, corrupt wildlife industry professionals and asian crime syndicates," 2012.

[267] S. T. Rhino, "Synthetic rhino horn: Will it save the rhino?." https://www.savetherhino.org/thorny-issues/synthetic-bio-fabricated-rhino-horn-will-it-save-the-rhino/. May. 11, 2015.

[268] P. Jia, F. Tian, Q. He, S. Fan, J. Liu, and S. X. Yang, "Feature extraction of wound infection data for electronic nose based on a novel weighted kpca," *Sensors and Actuators B: Chemical*, vol. 201, pp. 555–566, 2014.