Spring 5-1-2015

Selective Detection of Volatile Organic Compounds Using Metal Oxide Sensor Arrays

Anton Dmitrievich Netchaev
University of Southern Mississippi

The University of Southern Mississippi

SELECTIVE DETECTION OF VOLATILE ORGANIC COMPOUNDS USING METAL OXIDE SENSOR ARRAYS

by

Anton Dmitrievich Netchaev

An Abstract of a Dissertation
Submitted to the Graduate School
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

May 2015

ABSTRACT

SELECTIVE DETECTION OF VOLATILE ORGANIC COMPOUNDS USING METAL OXIDE SENSOR ARRAYS

by Anton Dmitrievich Netchaev

May 2015

Selective detection of organic contaminant using widely available and inexpensive metal oxide sensors has the potential to be used in various robotic platforms for navigation, harmful chemical leak detection, mobile environmental monitoring, etc. Selective gas detection in real world environments using easily available sensors has not been reported and can be used in many industries. A sensor system using only four commercially available sensors with accompanying signal conditioning and clustering algorithm capable of discriminatory detection of chemical marker is possible. Tests have shown that temperature, humidity and concentration fluctuations can be accounted for to produce systems for real world environments. An algorithm that accounts for sensor fouling and degradation is produced to achieve a repeatability rate of ninety three percent in a simulated real world environment.

The University of Southern Mississippi

SELECTIVE DETECTION OF VOLATILE ORGANIC COMPOUNDS USING METAL OXIDE SENSOR ARRAYS

by

Anton Netchaev

A Dissertation
Submitted to the Graduate School
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved:

Dr. Randy Buchanan
Committee Chair

Dr. Paige Buchanan

Dr. Amer Dawood

Dr. Zhaoxian Zhou

Dr. Andrew Sung

Dr. Karen Coats
Dean of the Graduate School

May 2015

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Randy Buchanan, and all the other members on my committee for their unrelenting support and guidance as my mentors. I especially would like to thank Dr. Buchanan for providing me and my peers with exciting, ever-challenging projects and mentoring throughout our tenure at The University of Southern Mississippi. His support and guidance have driven my academic success and helped me overcome personal tragedies encountered in the last couple of years.

I also would like to thank Dr. Paige Buchanan from Department of Chemistry and Biochemistry for support and opportunities to work with various interdisciplinary projects. This experience has made me more valuable as an addition to a team in the future and has thought me to communicate my ideas with people of different educational backgrounds.

Last but not least, I want to extend my acknowledgments to all the funding agencies that have supported me during my tenure. I'm absolutely certain that without the complete dedication and support of these people, funding agencies and all of the faculty and staff at the Computing Department of The University of Southern Mississippi, I would not have received as good of an education and completed this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF ILLUSTRATIONS

Figure

LIST OF ABBREVIATIONS

LEL	-	Lower explosive limit
SMO	-	Semiconductive metal oxide
MEMS	-	Micro-electromechanical systems
VOC	-	Volatile organic compounds
CNT's	-	Carbon nanotubes
DOAS	-	Differential optical absorption spectroscopy
TDLAS	-	Tunable diode laser absorption spectroscopy
LIDAL	-	Raman Light Detection and Ranging
DIAL	-	Differential absorption LIDAR
ICAS	-	Intra-cavity absorption spectrometry
ppth	-	Parts-per-thousand
GC	-	Gas chromatography
FPD	-	Flame photometric detection
PFPD	-	Pulse flame photometric detection
SCD	-	Sulphur chemiluminescence detection
AED	-	Atomic emission detection
TOF	-	Time-of-flight
MOX	-	Metal oxide
MOS	-	Metal oxide semiconductor
AAO	-	Anodic aluminum oxide
MS/s	-	Mega samples per second
AtoD	-	Analog to digital
EMA	-	Exponential moving average
ppth	-	Parts-per-thousand
ppm	-	Parts-per-million
ppb	-	Parts-per-billion
PR	-	Pattern recognition
P	-	Positive tuples
N	-	Negative tuples
TP	-	True Positives
TN	-	True Negatives
TF	-	False Positive
FN	-	False Negative
ROC	-	Receiver Operating Characteristic
PCA	-	Principal component analysis

Chapter I

Introduction

Various sensors have been used from the dawn of robotics for navigation and environment sensing. Robots use sensory systems to navigate the environment and to perform programed tasks without human intervention. Research in robotic navigation is primarily focused on vision and hearing methods and tend to ignore more basic sensory modalities such as smell, touch and taste. Regardless of their simplicity, olfactory sensory processes are remarkably adaptive and not sufficiently understood[4].

Navigation through the use of smell has several fundamental issues. The release of an organic or inorganic compound, or marker, in a natural environment produces a dispersion via microscopic and macroscopic processes[4]. Even without environmental factors on a micro scale, Brownian motion of molecules leads to a dispersion of the marker. The physics of the diffusion processes on the micro scale are well understood while on the macro scale, even with current advances in technology, are far less understood[4]. Diffusion of markers will produce unwanted changes in response of sensors and have to be accounted for in a robust navigation system that uses smell. Other factors like background contaminants also produce an environment in which it is hard to selectively identify individual organic compounds.

An array of organic sensors can be used to extract information from organic markers and be used for navigation in scenarios where traditional techniques are not viable. Markers can be used to relay navigational information to the robots in environments such as ships and metal buildings. Selection of organic sensors and signal processing algorithms that can selectively identify markers in a real world environment are required for navigation.

Chapter II

BACKGROUND

II.1 History of Chemical Gas Sensing

Chemical gas sensing became an important issue right after the discovery of harmful effects of gases in the beginning of the Industrial Revolution[2]. Coal mines where set up to acquire the energy resource driving the Industrial Revolution. Mines were primitive and required an immense amount of manual labor to dig up enough coal. Workers of all ages would wear helmets with candles for illumination of their work environment. Methane naturally present in these mines could not only displace the oxygen but also is a highly flammable gas. After understanding the danger of these gases, the miners found ways to detect them.

Figure II.1: Original sketch by H. Davey published in 1816 [3]

The first method of detecting methane without endangering humans was using a canary bird. Canaries are more vulnerable to low oxygen, methane gas, or CO gas. An incapacitated canary means it's time to evacuate the mine[2]. The next generation in gas detection technology was the Flame Safety Lamp (Davey's Lamp). This device was invented by Sir Humphry Davey in 1815. Davey's Lamp was a device with an oil flame that was adjusted to a specific height in fresh air so if methane is present the flame height would increase while low oxygen levels would reduce the flame height (Figure II.1) [3]. This technology still presented a danger of accidental exposure of the flame to the methane environment.

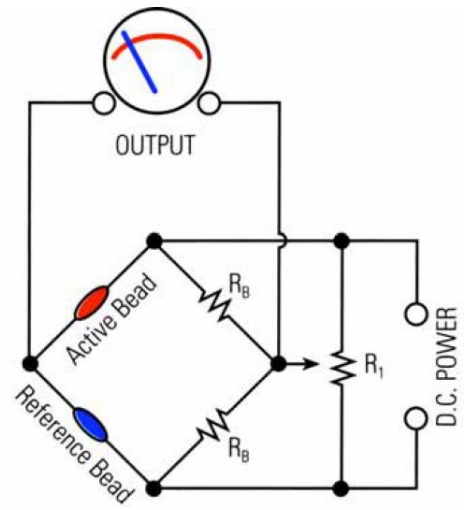


Figure II.2: Catalytic sensor circuit [24]

The next big step in gas sensing technology did not come until early 1930's when Dr. Oliver Johnson was working for Standard Oil Co.(now Chevron) and has developed the first intrinsically safe catalytic combustion lower explosive limit (LEL) sensor [24] . The sensor operated on the principle that a heated wire of platinum oxidizes flammable gas at lower levels than it would normally in air. The oxidation of gases around the wire causes an increase in temperature, thus increasing the resistance of the filament. A second filament which is not in volatile gas stream is used as a reference. Resistance change produced in this gas sensor is measured using a Wheatstone bridge connected to a deflection meter (Figure II.2) [24]. The first model (Model A) was built for demonstration purposes only but a second model (Model B) would become the first production model of LEL meter (Figure II.3). The Model B used an original idea of two filaments, one in a closed glass tube for reference and one with a tube and a flame arrester to allow the passage of sample air while preventing a flashback into the tested space.

The late 1920*s* also saw the creation of first interferometer gas detector. Initial design on the interferometer was conducted by Dr. Uzumi Doi in 1927 at the Institute of Physical & Chemical Research in Japan. The first prototype, however, was developed by Dr. Ziro Tsuji working at the same institute [24]. Interferometer employed the variability of light diffraction in different mediums. Light diffracting from presence of methane or gasoline vapors produced visible lines that shifted to indicate the concentration of gas contaminate. Sensors like these dramatically improved safety of operations in mines, oil wells, and

Figure II.3: Original sketch by H. Davey published in 1816 [24]

other industries dealing with hazardous gasses. These approaches where some of the first analytical methods to measure gas concentrations in air and paved the way for many other types of sensors.

II.2 Current Gas Sensing Methods

In the last fifty years industry and research have established various branches of gas sensing technologies. These technologies can be broken down into two groups: methods based on change of electrical properties, and methods that use a variation of other properties [22] (Figure II.4).

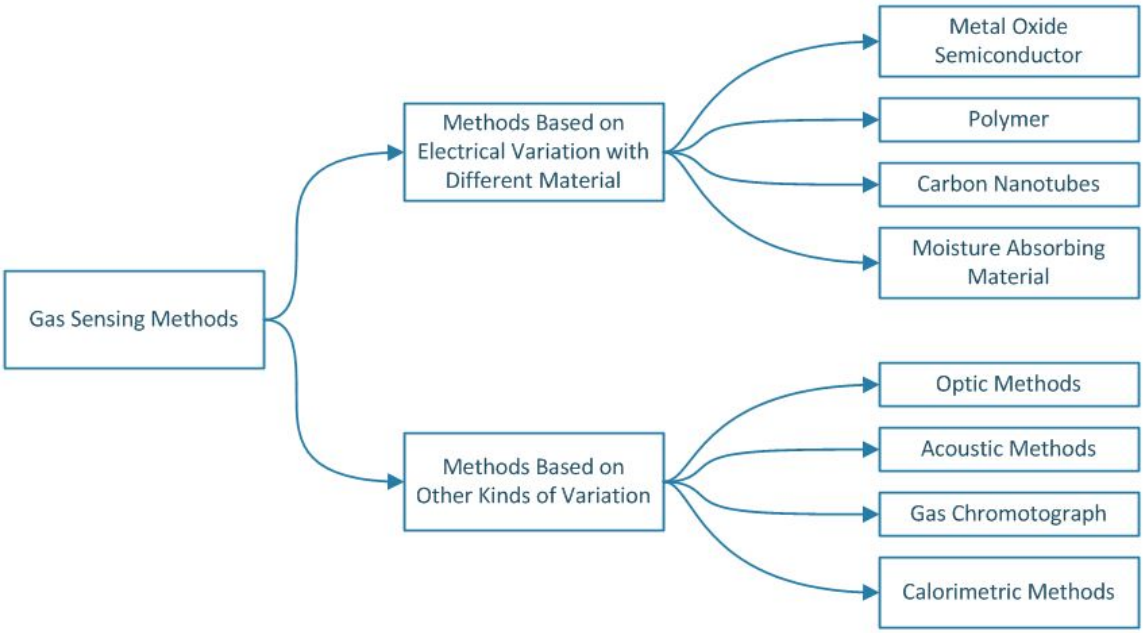


Figure II.4: Classification of gas sensing methods [22]

The overall advantages, disadvantages, and applications for the gas sensing technologies are summarized in Table II.1 . This table can be used for transducer selection for specific application and as a reference for further description of each of these technologies.

Table II.1: Summary of gas sensing methods [22]

Materials	Advantages	Disadvantages	Target Gasses and Application Fields
Metal Oxide Semiconductor	(a) Low cost; (b) Short response time; (c) Wide range of target gases; (d) Long lifetime.	(a) Relatively low sensitivity and selectivity; (b) Sensitive to environmental factors; (c) High energy consumption.	Industrial applications and civil use.
Polymer	(a) High sensitivity; (b) Short response time; (c) Low cost of fabrication; (d) Simple and portable structure; (e) Low energy consumption.	(a) Long-time instability; (b) Irreversibility; (c) Poor selectivity;	(a) Indoor air monitoring; (b) Storage place of synthetic products as paints, wax or fuels; (c) Workplaces like chemical industries.
Carbon Nanotubes	(a) Ultra-sensitive; (b) Great adsorptive capacity; (c) Large surface-area-to-volume ratio; (d) Quick response time; (e) Low weight.	(a) Difficulties in fabrication and repeatability; (b) High cost.	Detection of partial discharge (PD)
Moisture Absorbing Material	(a) Low cost; (b) Low weight; (c) High selectivity to water vapor.	(a) Vulnerable to friction; (b) Potential irreversibility in high humidity.	Humidity monitoring
Optical Methods	(a) High sensitivity, selectivity and stability; (b) Long lifetime; (c) Insensitive to environment change.	(a) Difficulty in miniaturization; (b) High cost.	(a) Remote air quality monitoring; (b) Gas leak detection systems with high accuracy and safety; (c) High-end market applications.
Calorimetric Methods	(a) Stable at ambient temperature; (b) Low cost; (c) Adequate sensitivity for industrial detection (ppth range).	(a) Risk of catalyst poisoning and explosion; (b) Intrinsic deficiencies in selectivity.	(a) Most combustible gases under industrial environment (b) Petrochemical plants; (c) Mine tunnels; (d) Kitchens.
Gas Chromatograph	(a) Excellent separation performance; (b) High sensitivity and selectivity.	(a) High cost; (b) Difficulty in miniaturization for portable applications.	Typical in laboratory analysis.
Acoustic Methods	(a) Long lifetime; (b) Avoiding secondary pollution.	(a) Low sensitivity; (b) Sensitive to environmental change.	Components of Wireless Sensor Networks.

II.2.1 Metal Oxide Semiconductor

This type of sensor is the most commonly used technology to detect gas due to its high sensitivity, low cost, and durability. Semiconductive metal oxide (SMO) is also one of the

most studied groups of chemiresistive gas sensors. In the past several decades SMO gas sensors where adopted as a prime technology in many industrial gas sensing systems. Two types of solid gas sensors are available on the market today. These sensors are categorized based on their electrochemical behavior, catalytic combustion, or resistance modulation of SMO [18]. Non-transition type sensor (e.g., Al2O3) contain elements with only one oxidation state while transition type sensors (e.g., Fe2O3) contain more than one oxidation state. Thus transition-metal oxides can form various oxidation states on the surface that are utilized as sensing mechanism compared to non-transition type sensor. Transition-metal oxides with d0 (e.g., TiO2, V2O2, WO3) and d10 (e.g., SnO2, ZnO) electron configuration can be used in gas sensing application [22]. Despite the simplicity of semiconductor metal oxide measurements for use as gas sensors, the exact detection mechanism is very complex and not yet fully understood. This complexity stems from numerous parameters that affect the interaction between the gas and sensing surface. These include absorption ability, electro-physical and chemical properties, catalytic activity, thermodynamic stability, as well as absorption/desorption properties of the surface [18]. A typical micro-electro-mechanical systems (MEMS) SMO type sensor consists of a heater able to maintain high temperatures ($\sim 350\,^{\circ}$C) during operation and a metal oxide coating (Figure II.5). When a SMO type sensor is heated in absence of oxygen, free electrons easily flow through the grain boundaries of metal oxide film, while in the presence of oxygen, a potential barrier is formed via the absorption of O2[13]. Interaction between atmospheric oxygen and a SMO surface forms a charged oxygen species which traps electrons from the sensitive layer of the sensor. This creates a region of depleted electrons which result in the potential barrier at the grain boundaries. This impedes the flow of electrons increasing the resistance of the material. Exposure to an atmosphere that contains a reducing gas enables the surface to free up trapped oxygen species and reduce the potential barrier thus reducing the electrical resistance of the sensor [18]. This and other interactions enable the sensor to act as a variable resistor whose value is a function of gas concentration.

II.2.2 Polymer

Polymer sensors typically used to detect species of volatile organic compounds (VOC's) that are not detectable using metal oxide semiconductors. Even though some studies do consider polymer-based sensors for detecting inorganic gases like CO2, they are most frequently used for detection of wide range of VOC's [22]. Similar to metal oxide semiconductors the exposure of the sensitive layer to gas contaminant will result in change of its dielectric properties upon gas absorption. The mechanism by which the VOC's interact with polymers

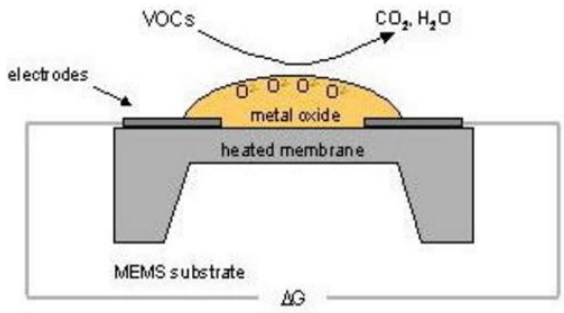


Figure II.5: Diagram of SMO sensor

include dipole/induced dipole interaction (also named London dispersion), dipole/dipole interaction and hydrogen bonds (Lewis acidity/basicity-concept) [12]. In addition to the differences between the interactions of active layers the polymer based sensors are able to operate in room temperature. Polymer based sensors can be further broken down into two categories of conduction and non-conducting polymers. Conducting polymers including gas sensing materials such as polypyrrole (PPy), polyaniline (PAni) and, polythiophene (PTh) exhibit changes in their electrical properties during exposure to diverse organic and inorganic gases. Non-conducting polymers do not change their electrical properties but instead can be used due to their physisorption mechanisms. For example, polymer layers that change their weight through absorption would change their resonant frequency which can be measured [22]. Polymer based sensors are still a widely researched topic and will continue to improve their characteristics as time goes by.

II.2.3 Carbon Nanotubes

Carbon nanotubes (CNT's) are attractive as a gas sensor due to the same reasons as polymer based sensors (Figure II.6). Unlike metal oxide semiconductors the CNT's do not exhibit poor sensitivity in room temperatures and provide unique properties making them promising for high-sensitivity gas sensing applications. CNT's have been found to be highly sensitive to extremely small quantities of gases such as alcohol, ammonia (NH3), carbon dioxide (CO2), and nitrogen oxide (NOx) at room temperature [30]. Like other gas sensing materials carbon nanotubes exhibit various properties when exposed to different target gasses. Some of the responses are irreversible due to chemisorption of chemicals like NH3. Due to irreversibility of sensor response after exposure to various chemicals, this technology currently uses other material to protect it from such effects and enhance the sensitivity and selectivity of the sensor [22]. Similar to the polymer sensors future research will enhance the characteristics of such system.

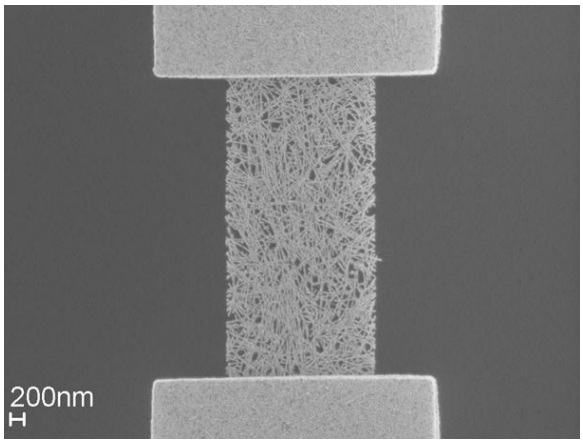


Figure II.6: Original sketch by H. Davey published in 1816 [3]

II.2.4 Optical Methods

Optical gas sensing methods are usually based on spectroscopy which enables the sensors to have higher sensitivity selectivity and stability than other methods. However, these types of sensors are restricted by their size and relatively high cost. Only a few commercially available sensors are based on optical techniques. Spectroscopic analysis is based on the principle of absorption spectrometry where the concentration-depended absorption (molar absorptivity ε) of the photons at specific gas wavelengths (i.e., Beer-Lambert law) [22] is measured. Precise wavelengths for specific gasses can be found in the HITRAN database [1]. Many other types of absorption spectroscopy exist including Differential Optical Absorption Spectroscopy (DOAS), Tunable Diode Laser Absorption Spectroscopy (TDLAS), Raman Light Detection and Ranging (LIDAL), Differential Absorption LIDAR (DIAL), Intra-Cavity Absorption Spectrometry (ICAS), etc. [22]. Sensors based on infrared source are the most widely used sensors and employ the basic absorption spectrometry. IR-based sensors have an infrared source that eliminates a chamber filled with gas, a detector opposite to the source is used to detect the light wavelets that have not been absorbed by the gas. An IR-based sensor diagram is shown in Figure II.7. The reliability and efficiency of such sensors has already proved itself useful in monitoring of remote areas for gas leaks where price is not an object.

II.2.5 Calorimetric Methods

Calorimetric methods of measuring gases consist of pellistors and account for a major share of all sensors on the market. Pellistors are solid-state devices used to detect combustible gasses and gasses with significantly different thermal conductivity from air [5]. As previously

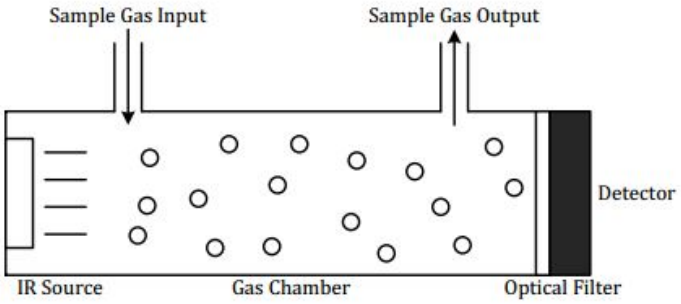


Figure II.7: IR-based sensor [22]

discussed the detecting elements of pellistors consists of a catalyst whose resistance changes with the presence of target gas. The limit of detection of such sensors lies in low parts-per-thousand (ppth) range which is usually only suitable for industrial applications. These sensors can be divided into catalytic and thermal conductivity types in each case the temperature variation created by gas is measured [22]. Figure 8 shows a diagram of a catalytic sensor which contains a platinum coil acting as a heater and calorimeter. The coil is heated to temperatures up too 500°C to burn combustible gases on the surface of the sensor, the heat generated by that reaction changes the resistance of the coil which can be measured. Currently, improvement of the performance of these types of sensors can be achieved by reducing power consumption, increasing life, and improving flameproof designs for prevention of explosions during measurement of certain types of gases.

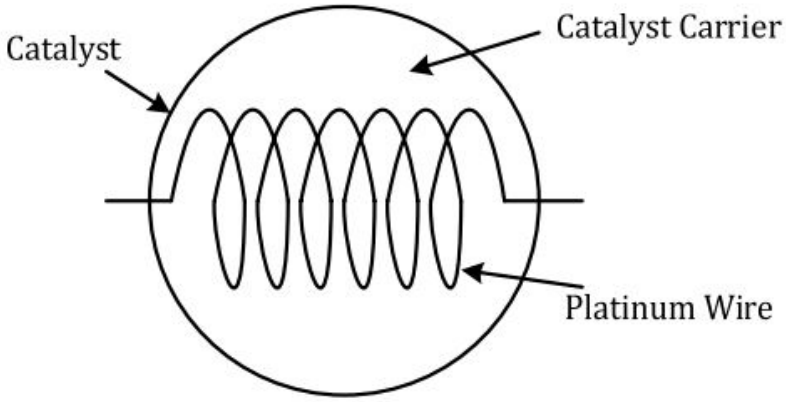


Figure II.8: Catalytic sensor configuration [22]

II.2.6 Acoustic Methods

Measurement of gas using acoustical methods can be broken down into three categories of sensors: (1) speed of sound; (2) attenuation; and (3) acoustic impedance. Measuring

the Time-of-Flight (TOF) of ultrasonic sound waves through different media is the best studied category of acoustical gas sensing techniques. Attenuation methods measure the loss of energy through the media. Acoustic impedance is usually used to determine the gas density, thus determining the gas the sound has traveled through. These methods are currently not used in the industry due to their lack of sensitivity, selectability, and high power consumption which makes this type of sensor also unusable for the research at hand [22].

II.3 Metal Oxide Based Sensor Array for Detection of Gasses

Semiconductor metal oxide type sensors are one of the most widely used groups of chemiresisitve gas sensors [19]. SMO sensors depending on the doping show a wide variety of response characteristics to volatile organics. Research as far back as 1998 has reported on the use of gas sensors in an array for use in VOC recognition and detection[15]. Reported devices use a large array of individually fabricated sensors in a lab environment to analyze and recognize the gas they are exposed to [22] [19][21]. Despite subtle differences in test stand setups in past studies the overall principle remains the same. Figure II.9 shows a general schematic of SMO sensor array used in previous studies. The test fixtures all use calibrated gasses mixed using mass flow controllers to maintain consistency between results and to eliminate error associated with real world operation. These studies yield favorable results and show that through proper selection of SMO sensors and appropriate software interpretation it is possible to design a system that will selectively detect and identify some VOC's.

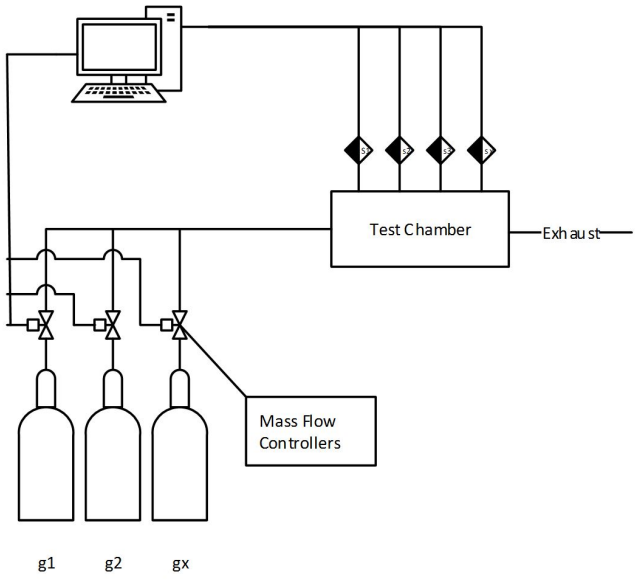


Figure II.9: General diagram for SMO array gas detector

Chapter III

Experimental Apparatus

III.1 Overview

The main objective of this research was to design a system that would identify various species of VOC's. A custom sensor array with power control and data acquisition circuitry had to be designed to properly measure the volatile organics in the air. The complete system is shown in Figure III.1. This chapter will give an overview of sensor array, signal conditioning hardware, data acquisition (DAQ) hardware/software, custom designed software and other parts that make up the system.

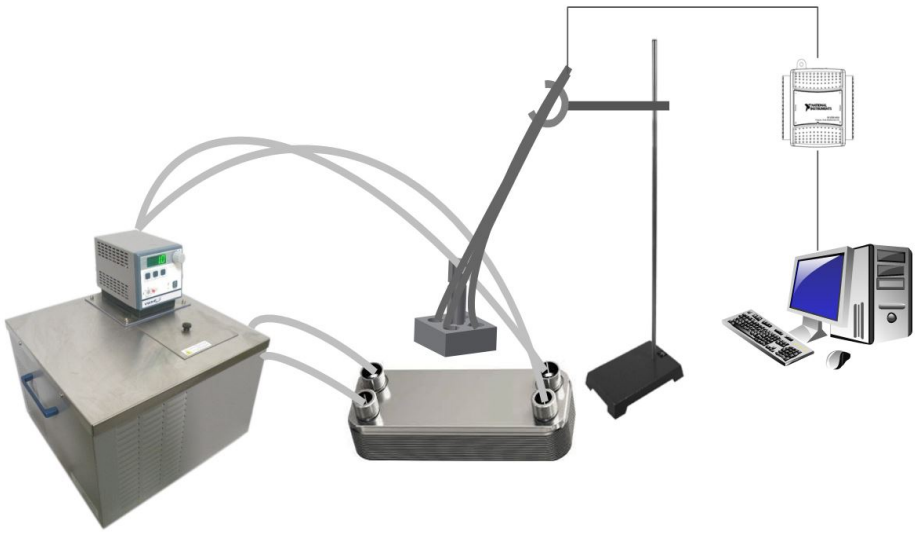


Figure III.1: General diagram for SMO array gas detector

III.2 Sensor Array

The sensors where selected with availability and cost in mind. The final system has to be easily put in production, thus custom made sensors would not be cost effective. Sensors selected where designed to respond to volatile organic compounds presence in the air. Each sensor was selected from independent companies holding a patent on their design, thus avoiding a chance of getting sensors with similar doping properties.

III.2.1 Synkera

Synkera MikroKera 4 VOC (Figure III.2) sensor part number (P/N) 725 is a metal oxide (MOX) type semiconductor sensor developed by Synkera Technologies, INC. Table III.1 shows the operational characteristics of Synkera 725 sensor. The sensor has a very low heater voltage and sensing voltage. The manufacturer states that this sensor has a strong response to a wide range of VOC's (Figure III.3) [28]. The gas sensitive metal oxide semiconductor (MOS) layer is made up of anodic aluminum oxide(AAO).

Table III.1: Synkera 725 Characteristics [28]

Property	Value
Environmental temperature range	$-20\,^{\circ}$C to $50\,^{\circ}$C
Environmental humidity range	0 to 95% RH
Heater power consumption	100mW
Heater Voltage	1.25VDC
Sensing Voltage	2.0VDC

Figure III.2: Synkera 725 sensor response [28]

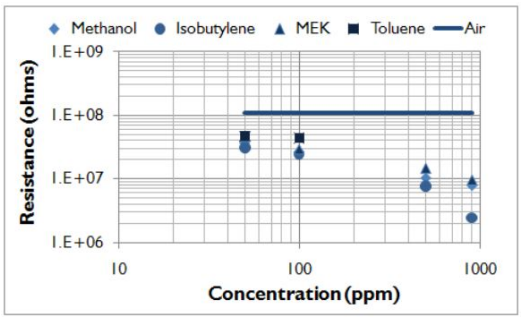


Figure III.3: Synkera 725 MOX sensor[28]

III.2.2 MiCS

E2v MiCS-5135 (Figure III.4) is a MOX type semiconductor sensor developed by e2v. Table III.2 shows the operational characteristics of a MiCS-5135 sensor. The sensor has a medium heater voltage and high sensing voltage. The manufacturer states that this sensor is documented to respond to carbon monoxide (CO), hydrocarbons (HC), ethanol, and volatile organic compounds (Figure III.5) [7] .

Table III.2: MiCS-5135 characteristics [7]

Property	Value
Environmental temperature range	$-40\,^{\circ}$C to $120\,^{\circ}$C
Environmental humidity range	5 to 95% RH
Heater power consumption	102mW
Heater Voltage	3.2VDC
Sensing Voltage	5.0VDC

Figure III.4: MiCS-5135 sensor response [7]

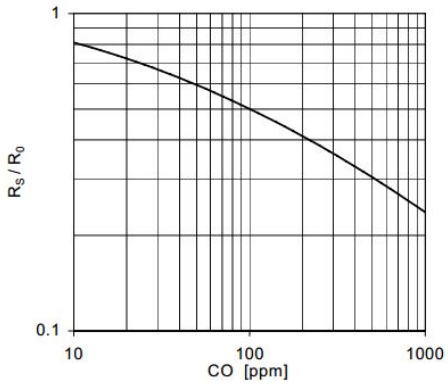


Figure III.5: MiCS-5135 sensor [7]

III.2.3 Figaro

TGS 2620 (Figure III.6) is a MOX type semiconductor sensor developed by Figaro. Table III.3 shows the operational characteristics of a TGS 2620 sensor. The sensor has a high heater voltage and high sensing voltage. The manufacturer states that this sensor is documented to have a high response to a variety of organic solvents and volatile vapors as well as combustible gasses such as carbon monoxide (Figure III.7) [10].

III.2.4 Grove

Grove-HCHO (Figure III.8) is a MOX type semiconductor sensor based on WPS2110 transducer. Table III.4 shows the operational characteristics of WPS2110 sensor. The sensor has a high heater voltage and high sensing voltage. The manufacturer states that this

Table III.3: TGS 2620 characteristics [10]

Property	Value
Environmental temperature range	$-20\,^{\circ}$C to $50\,^{\circ}$C
Environmental humidity range	5 to 95% RH
Heater power consumption	210mW
Heater Voltage	5.0VDC
Sensing Voltage	5.0VDC

Figure III.6: MiCS-5135 sensor response [7]

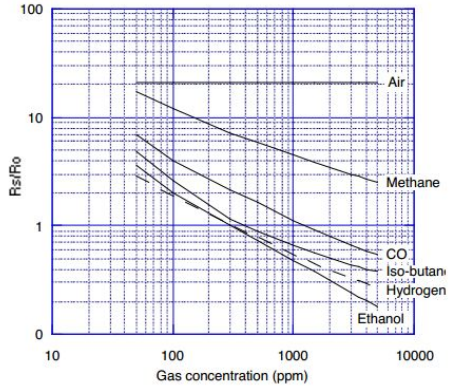


Figure III.7: MiCS-5135 sensor [7]

sensor is capable to detecting formaldehyde, benzene, toluene and other volatile compounds (Figure III.9) [8].

Table III.4: Grove-HCHO characteristics [8]

Property	Value
Environmental temperature range	$-20\,^{\circ}$C to $50\,^{\circ}$C
Environmental humidity range	5 to 95% RH
Heater power consumption	210mW
Heater Voltage	5.0VDC
Sensing Voltage	5.0VDC

III.2.5 Sensor Mount

A mount bracket for the sensor array was designed and printed in house. The bracket was designed using AutoCAD 3D design software to ensure proper fit of all sensors. The sensor array has to maintain equal distance from the surface to the sensors to reduce sensor response variation. Final design was printed on MakerBot Replicator 2X 3D printer using

Figure III.8: Grove sensor response [8]

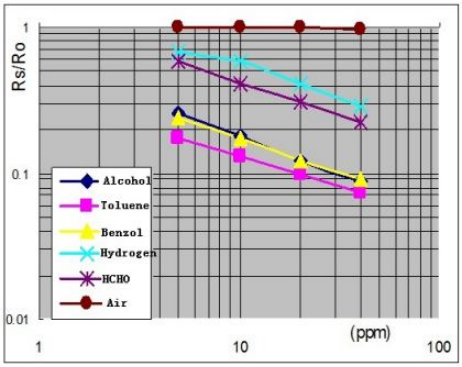


Figure III.9: Grove sensor [8]

MakerWare open source software (Figure III.10) . The final design included grooves to ensure proper sensor placement and a location for tying of cables.

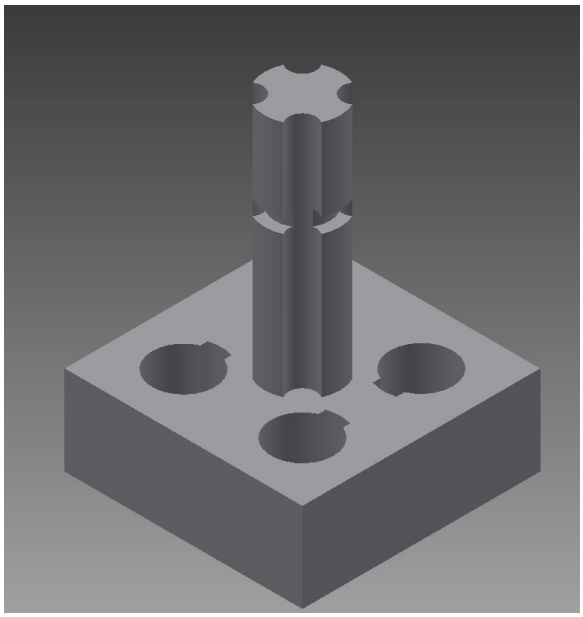


Figure III.10: CAD image of sensor array bracket

III.3 Power Distribution and Data Acquisition Hardware/Software

III.3.1 Sensor Power Supply

SMO sensors require various power supply circuits to function properly (Table III.1, Table III.2, Table III.3, and Table III.4). A custom power supply board was designed using Altium circuit design software. LM317 linear voltage regulators by National Semiconductor

are used to deliver power to the array (Figure III.11). To set a proper voltage on the regulator, R_H and R_L resistors have to be picked to produce V_{out} voltage that is desired (Equation III.1).

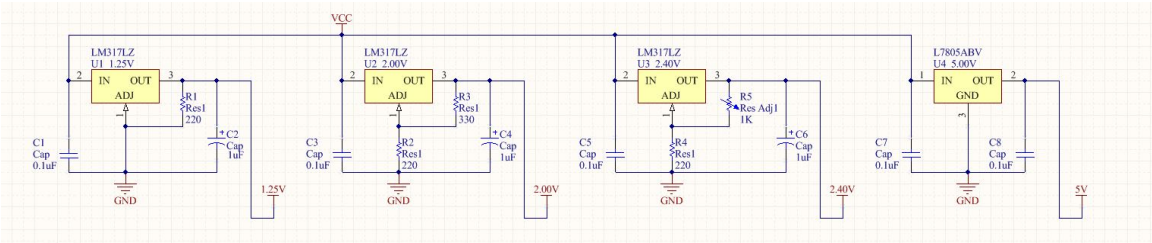


Figure III.11: Power supply circuit

$$V_{out} = V_{ref}(1 + \frac{R_L}{R_H}) + I_Q R_L \tag{III.1}$$

The power supply circuit can be powered by a single input voltage between 8V-35V. Such large input voltage range accommodates for commercially available power sources and ensures proper operation of the sensor array.

III.3.2 Data Acquisition

Data acquisition (DAQ) hardware from National Instruments was used to sample the output of the sensor array. NI-PXI 6251 (Figure 20) is a high-speed multifunctional data acquisition board optimized for superior accuracy at fast sampling rates. This DAQ has 16 analog inputs capable of sampling at 1.25 mega samples per second (MS/s) with 16 bit resolution all with a NIST-traceable calibration certificate[17]. The board can communicate with LabVIEW and MATLAB software packages with support of hardware time interrupts. This hardware was chosen due to its availability and not for its performance, a much cheaper DAQ can be used to acquire data without loss of system accuracy.

III.3.3 Software

Proper interpretation of data acquired by the sensors is crucial to the success of the proposed goal. Each sensor in the array has a unique response to various volatile organics. This unique response or signature can be interpreted to identify a chemical marker. MATLAB programing environment is used as it provides the necessary tools for fast prototyping and contains most of the needed libraries for hardware to software interaction. MathWorks MATLAB version R2011b (7.13.0.564) was used for programing. MATLAB is a markup programing language that allows matrix manipulations, plotting of live data streams, and

Figure III.12: NI PXI-6251 board[17]

creation of user interfaces as well as interfaces between languages like C, C++ and FORTRAN [16]. MATLAB contains libraries and drivers that enable it to communicate with many devices including NI-PXI 6251 which was used to sample the output of the sensors.

Chapter IV

Algorithm

IV.1 Introduction to Pattern Classification

Our brain is remarkably good at recognizing spoken words, handwritten characters, identifying people just by voice, picking out ripe fruits using just smell or sight, these being astoundingly complex processes underlie the act of pattern recognition. Pattern recognition is defined as the act of taking raw input data and making a decision based upon the category of the pattern presented. This recognition of a set of stimuli that is arranged in a certain pattern has been decisive in the survival of species and led to the evolution of highly sophisticated neural and cognitive systems [6].

As humans it is natural for us to design and build machines that can recognize faces, understand speech, interpret hand written data that mimic our own abilities. It is absolutely clear that accurate pattern analysis by a machine is extremely useful [6]. Using machines that can interact with the environment in a similar way as humans, it is possible to solve a myriad of problems that we are faced with today. For many problems in pattern recognition it is important to know how the problem is solved in nature to achieve the desired accuracy of the results.

Pattern classification can be viewed as a clustering method with a set of predefined classes. In a context of machine learning, classification is supervised learning and clustering is unsupervised learning [29]. Supervised learning refers to a task in which a an act of assigning a new data point to one of several known classes based upon information gained from a learning set. Pattern classification system can be broken down into steps shown in Figure IV.1.

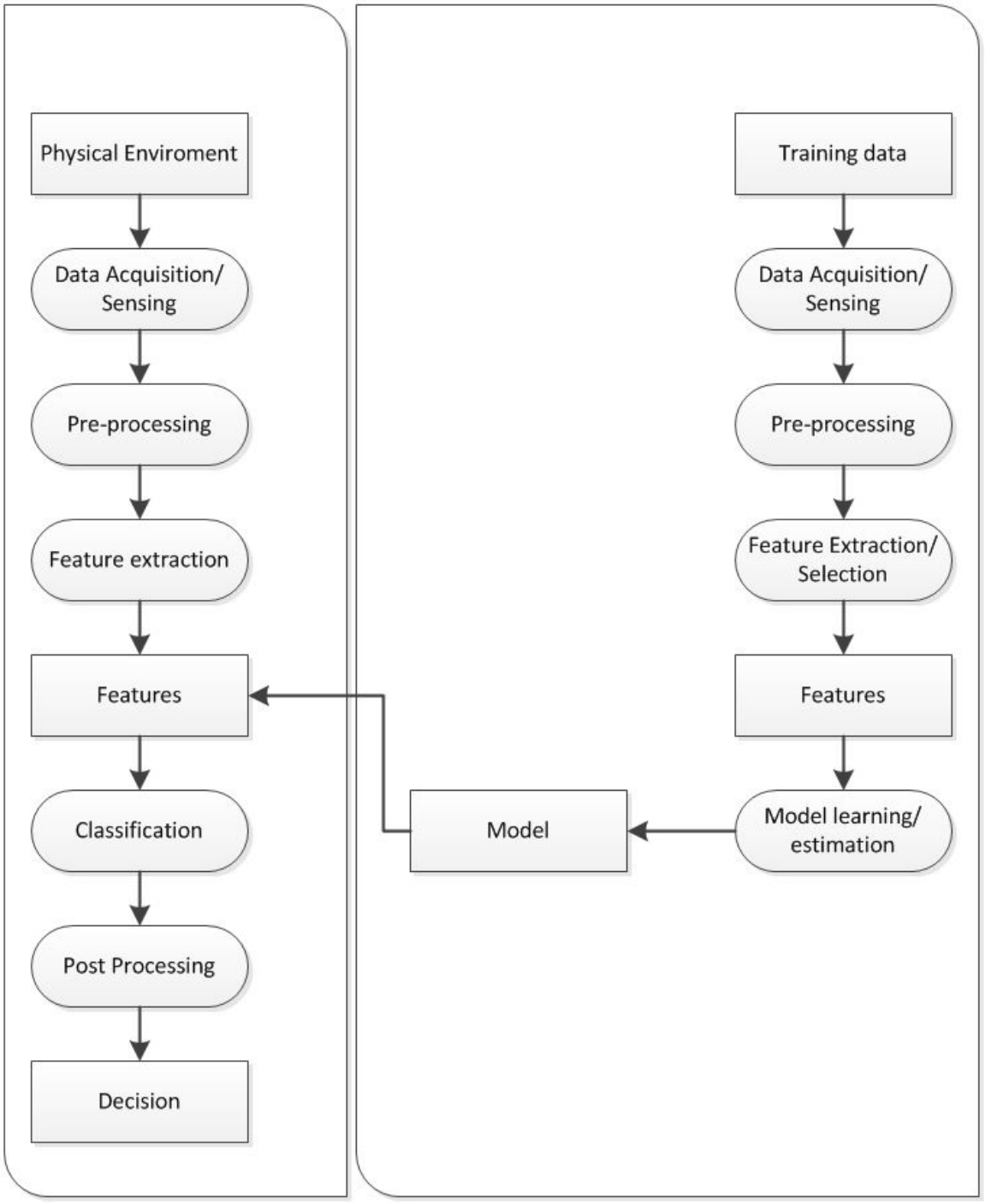


Figure IV.1: Object/process diagram of a pattern classification system

Pattern classification can be divided into two approaches. Decision-Theoretic approach represents the pattern as a vector in a vector space and uses a decision algorithm (mainly statistical) to assign a class to each pattern. Structural approach(syntactic approach) represents each pattern by its structure and uses parsing or graph matching to perform classification

[27]. Each of the described approaches can be divided into problem-dependent and problem-independent parts. Problem-dependent seeks how to convert the pattern into a points (vector) in space - feature extraction or how to represent such pattern by a desired structure. Problem independent part uses a decision algorithm for feature extraction or parsing algorithm, graph matching algorithm to classify extracted patterns [27].

IV.2 Implementation of Pattern Classification

Volatile organic markers were each tested using four sensors described previously. For the purpose of this research five chemicals where selected as markers. Sensors measured concentration of volatile contents in air. Each sensor based upon composition of the sensitive layer will produce a unique response to chemical markers (A,B,C,D,E). Pattern classification algorithm was developed to automatically transform each of the observations into a set of ω_i (for this research $i = 5$) classes. Supervised classification is required because newly measured markers are assigned a class label using a model developed from objects with a known class label (training or learning set). Steps taken to develop the algorithm are described in the subsequent sections and follow the process diagram IV.1

IV.2.1 Data Acquisition

The first step of classification algorithms for both the learning and execution stages is data acquisition as shown in Figure IV.1. To ensure the accuracy of incoming data, variable voltage regulators capable of maintaining the set output parameters while being subjected to noisy environments are implemented as described in section III.3.1. Hardware timer interrupt is used to sample data each second. To enable hardware interrupts using National Instruments hardware in MATLAB environment it is important to start a continuous session with hardware. This session can be terminated after a sufficient amount of data for classification has been collected. Interrupts are essential for data acquisition as they let other processes use the central processing unit (CPU) while waiting for the next sample. Sample code for configuring and enabling an interrupt for National Instruments hardware in MATLAB environment using session based interface is provided below.

```
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
%Function Name: Analog_input_setup
%input:     none
%output:
%           lh - interrupt handle
%           s  - daq session handle
%
%Function sets up data acquisition on an external PXI with external
%  hardware timer set to 1000msec with interrupts
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

function [lh,s] = Analog_input_setup
    global DATA
    s=daq.createSession('ni');
    s.IsContinuous = true;
    s.addAnalogInputChannel('PXI1SLot3',[0,1,2,3,5,6],'Voltage');
    s.NotifyWhenDataAvailableExceeds=1000;
    s.Rate = 1000;
    lh = s.addlistener('DataAvailable', @recordData);
    startBackground(s);
end

% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
%Function Name: recordData
%input:     src - source of the event
%           event - name of the event
%output:
%           none
%
%This function is called every time the set up interrupt is triggered
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

function recordData(src,event)
    global DATA
    data_a=mean(event.Data);
    time_p=round(event.TimeStamps(1000));
    DATA=vertcat(DATA,[time_p   data_a]);
end
```

IV.2.2 Pre-processing

The raw sensor responses contain high frequency noise due to electromagnetic interferences and analog to digital (AtoD) conversion. To remove such noise a moving average filter is implemented. Moving average filter (also called moving mean or rolling mean) is an infinite impulse response filter that averages number of points from the input signal to produce each point in the output signal [25]. An equation from this can be written:

$$y[i] = \frac{1}{M} \sum_{M-1}^{j=0} x[i+j] \tag{IV.1}$$

Due to the fact that the data from the sensors is continuous and not finite it is important

to create a method that extracts meaningful data from a continuous data stream. Data stream is defined as a sequence of real-time, continuous, and ordered items. Stream data can be converted to finite data using a window technique [11]. Window technique places boundaries on infinite data stream to analyze the data within the window. Most commonly used window techniques use a predefined window size to analyze streams. This technique is easy and efficient if it is possible to ensure the required data is residing in this window. Set window size is not efficient for the proposed design due to variability of evaporation rates of volatile organic compound markers. Figures IV.2 and IV.3 clearly show the difference in time required for data acquisition of these compounds.

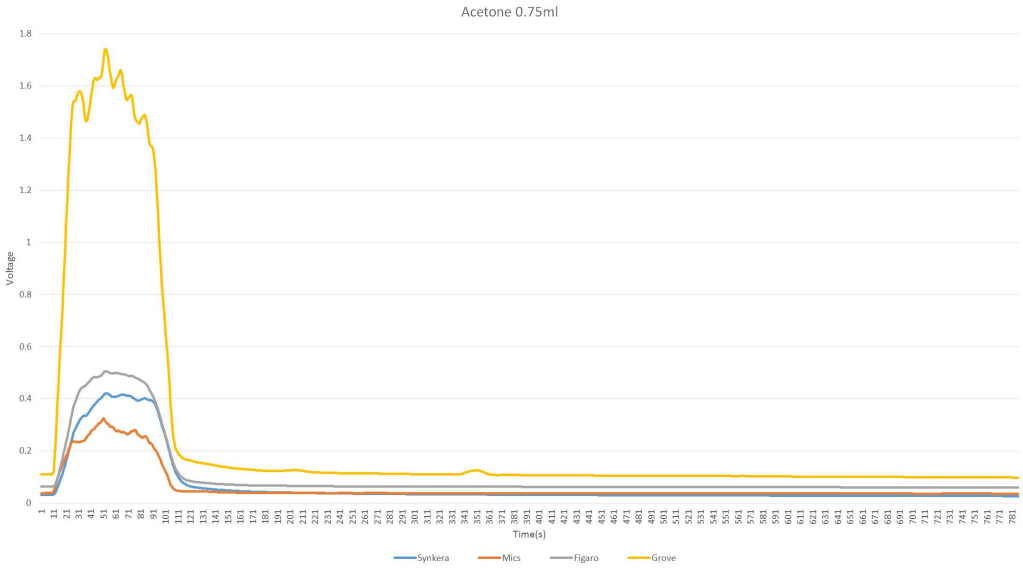


Figure IV.2: Ssensor output to 0.75ml acetone

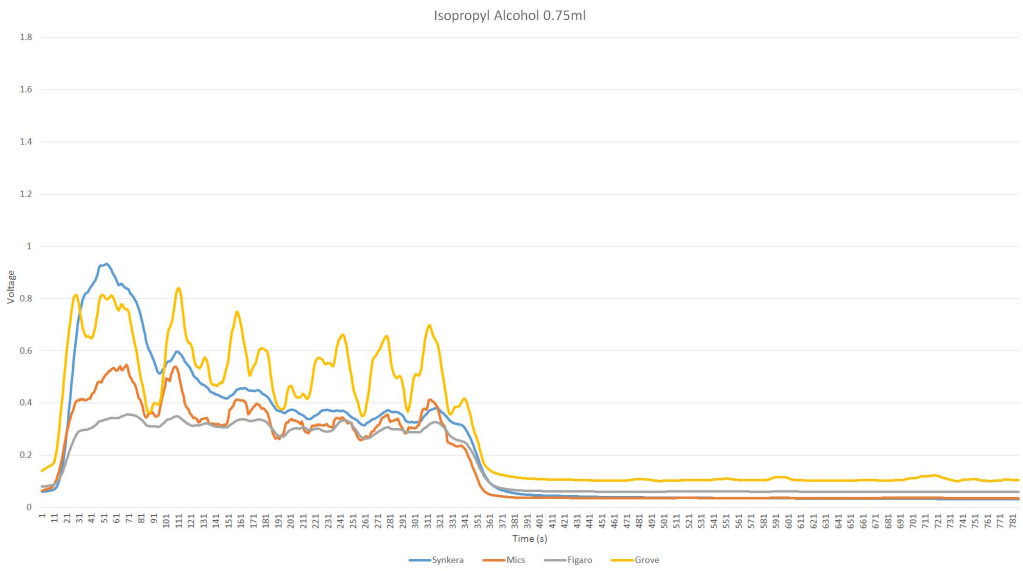


Figure IV.3: Sensor output to 0.75ml isotropyl alcohol

Thus a variable window technique had to be implemented. Selection of the beginning and ending of the window is highly important and will affect the final tuple size. The initial time and the length of the window are adjusted adaptively based on the information content in the measurement window. Start of the window will be selected as amplitude value for each sensor. This can be done because the sensor array is allowed to reach steady state between detections and thus the starting point of any measurement will be at a predefined voltage level. End of tuple is determined based on the sign of the determinant inside a smaller window (small window is variable for testing was set to 10 seconds).

Principal Component Analysis (PCA)

Principal component analysis (PCA) is a method of dimensionality reduction for visualization of complex data that was developed to capture as much variation in data as possible [14]. Conceptually the goal of PCA analysis is to reduce the number of variables of interest into a smaller set of components. To calculate the principal components of a data set with n independent observations taken on X_1, X_2, X_3, X_4 (Max output of each sensor), where the covariance between X_i and X_j is

$$cov(X_i, X_j) = \sum_{i,j} = \sigma^2 R_{i,j} \tag{IV.2}$$

for $i, j = 1, 2, 3, 4$ [20]. Let $\mu_1 < \mu_2 < \mu_3 < \mu_4 < 0$ be the eigenvalues of R and let z_1, z_2, z_3, z_4 be the corresponding eigenvectors normalized such that $z_j^T z_j = 1$. Define W_1 to be the

first principal component which is a linear combination of X with the largest possible variance IV.3[20]:

$$W_1 = a_1^T X = \sum_{i=1}^{4} a_{1_i} X_i \tag{IV.3}$$

where $a_1^T a_1 = 1$

$$var(W_1) = a_1^T \sum a_1 X = \sum_{i=1}^{4} a_{1_i} X_i \tag{IV.4}$$

Constrained maximization leads to the Lagrangian IV.5[20]:

$$L = a_1^T R a_1 + \lambda (a_1^T a_1 - 1) \tag{IV.5}$$

To solve assume a_1 is a unit vector satisfying $Ra_1 = -\lambda a_1$. To maximize $a_1^T R a_1 = \mu a_1^T a_1$ where μ is the eigenvalue which corresponds to the eigenvector a_1 [20]. Taking $a_1 = z_1$:

$$W_1 = z_1^T X = \sum_{i=1}^{4} z_{1_i} X_i \tag{IV.6}$$

Thus W_1 has the largest variance among all linear combination of the X's. Let W_2 be a second linear combination (second principal component) of the X's which has the largest possible variance [20]:

$$W_2 = a_2^T X = \sum_{i=1}^{4} a_{2_i} X_i \tag{IV.7}$$

but $corr(W_1, W_2) = 0$ in other words $a_2^T z_1 = 0$.Constrained maximization leads to the Lagrangian:

$$L = a_2^T R a_2 + \lambda_1 (a_2^T a_2 - 1) + \lambda_2 (a_2^T z_1) \tag{IV.8}$$

To solve [20]:

$$(R + \lambda_1 I) a_2 + \lambda_2 z_1 \tag{IV.9}$$

with $a_2^T a_2 = 1$, $Rz_1 = \mu_1 z_1$, and $a_2^T z_1 = 0$. Multiplying through by z_1^T to find that $\lambda_2 = 0$. Thus IV.10 a_2 is an eigenvector which is not z_1 due to orthogonality condition.

$$Ra_2 = -\lambda_1 a_2 \tag{IV.10}$$

Therefore, a_2 is an eigenvector (which cannot be z_1 due to orthogonality condition). Setting $a_2 = z_2$ will make $Var(W_2) = \sigma^2 \mu_2$ which is as large as possible under given constraints [20].

$$W_2 = z_2^T X = \sum_{i=1}^{4} z_{2_1} X_i \tag{IV.11}$$

Where W_2 has the largest variance among all linear combination of the X's which are orthogonal to W_1 [20].

By reducing the dimensionality of data generated by the sensors it is possible to reduce the data in such way as to retain information with largest variances. Reducing the number of dimensions in classification problem lends itself to simpler data interpretation. PCA analysis is used to construct a scree plot that shows the first two components explain over 90% of the total variance (Figure IV.4). There are clear breaks in variance between all three principal components with the first component explaining $\approx 60\%$ of the variance. A reasonable way to reduce the dimensionality of the problem in this case would be to consider only the first two principal components.

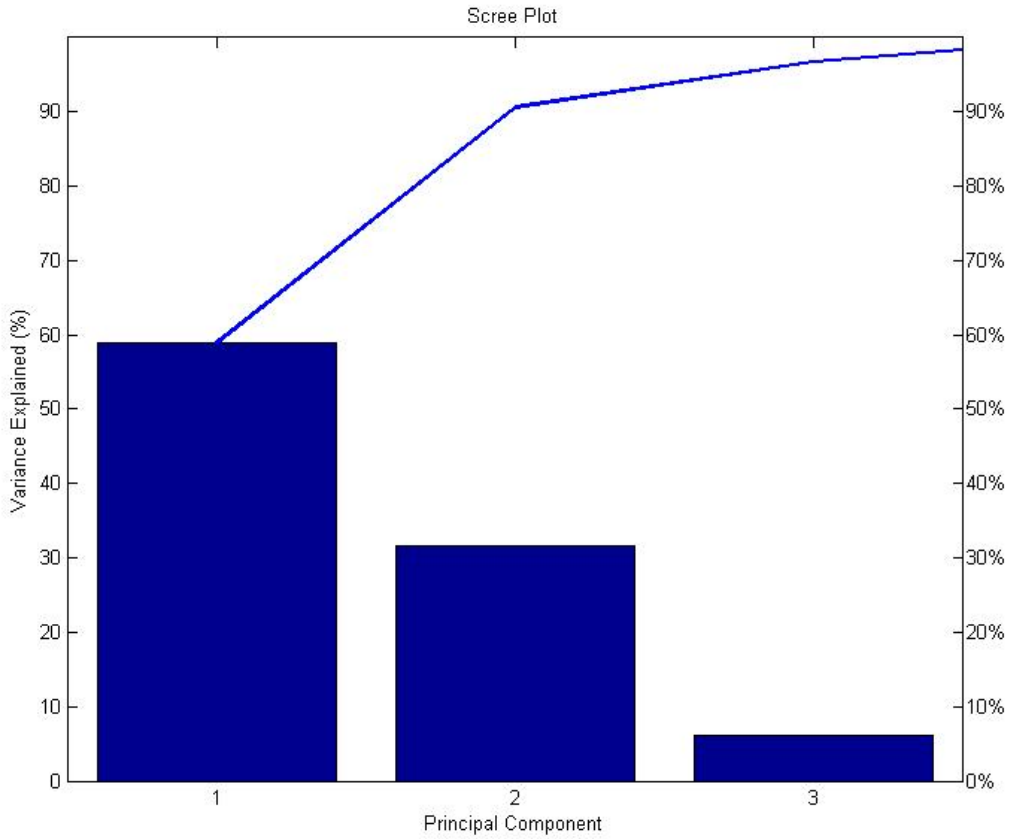


Figure IV.4: Scree Plot

Centered and scaled principal component scores for each observation are plotted in Figure IV.5. Each color represents a chemical used in preliminary testing. The figure clearly shows separation of chemicals but shows that the reduction in dimensions also introduced some uncertainty for future task of pattern classification. Chemical B is clearly in the zone of

chemical A as other chemicals are also intruding on each others zones except for Chemical E. This method of reducing the dimensionality and mapping the data on two dimensional space produces too many problems to the naked eye and should be considered if the number of sensors is above six.

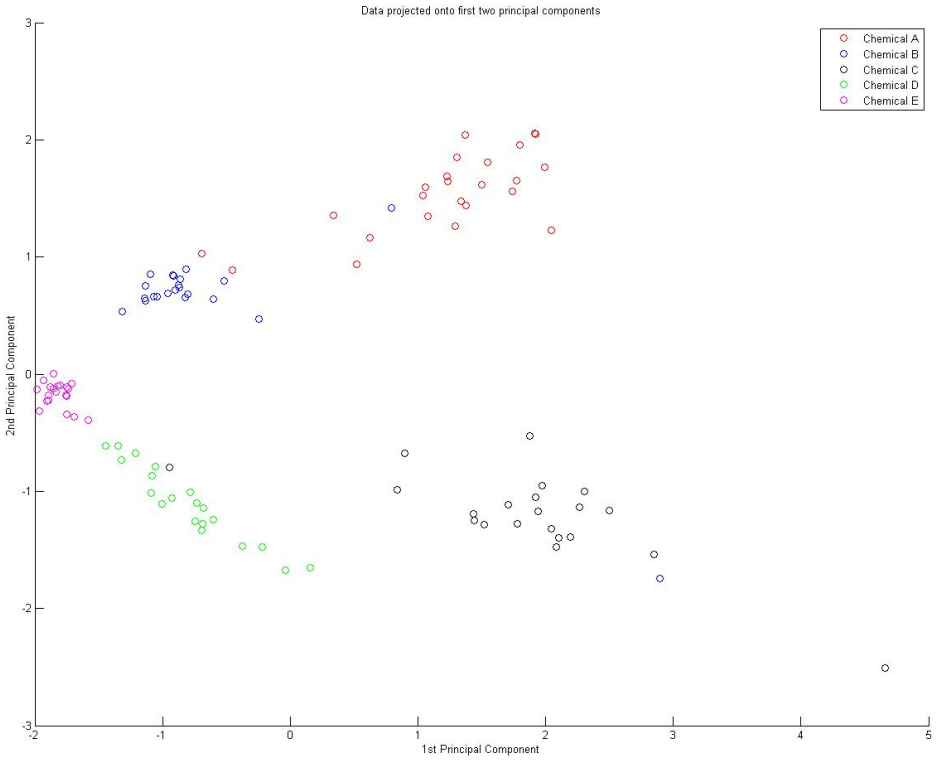


Figure IV.5: Principal Component Scores

Mapping Based on Sensor Ratios

For the problem at hand a simpler way of mapping data onto two dimensional space is considered. A peak point is computed using already filtered tuple. By computing the ratio between the peak points of sensor output (IV.12, IV.13)

$$Grove_{max}/Synkera_{max} = x_1 \tag{IV.12}$$

$$Figaro_{max}/Mics_{max} = x_2 \tag{IV.13}$$

it is possible to map this problem to a point or feature vector x in a two dimensional feature space where:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{IV.14}$$

This method of reducing dimensionality not only takes care of the data complexity but produces well separated areas with sample data. Figure IV.6 shows sample data mapped onto two dimensional space using incorrect ratios which results in data that is fairly hard to classify. On the other hand figure IV.7 uses the ratio that was found to be the best in separating instances to facilitate implementation of pattern classification algorithms

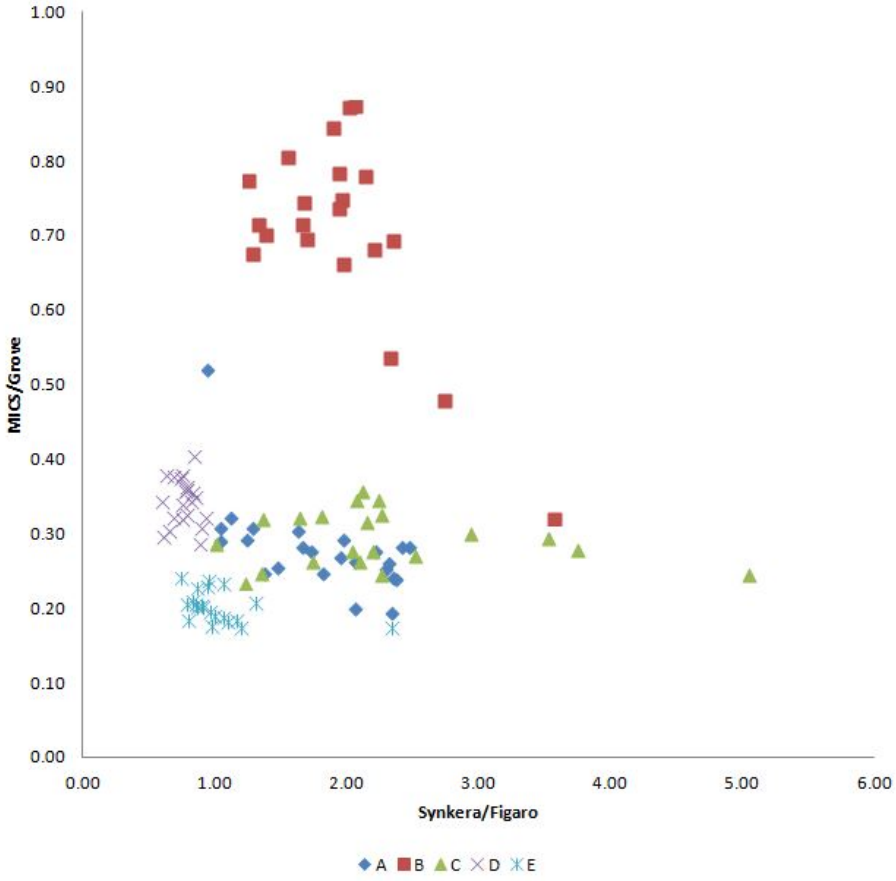


Figure IV.6: Ratio Selection Incorrect Results

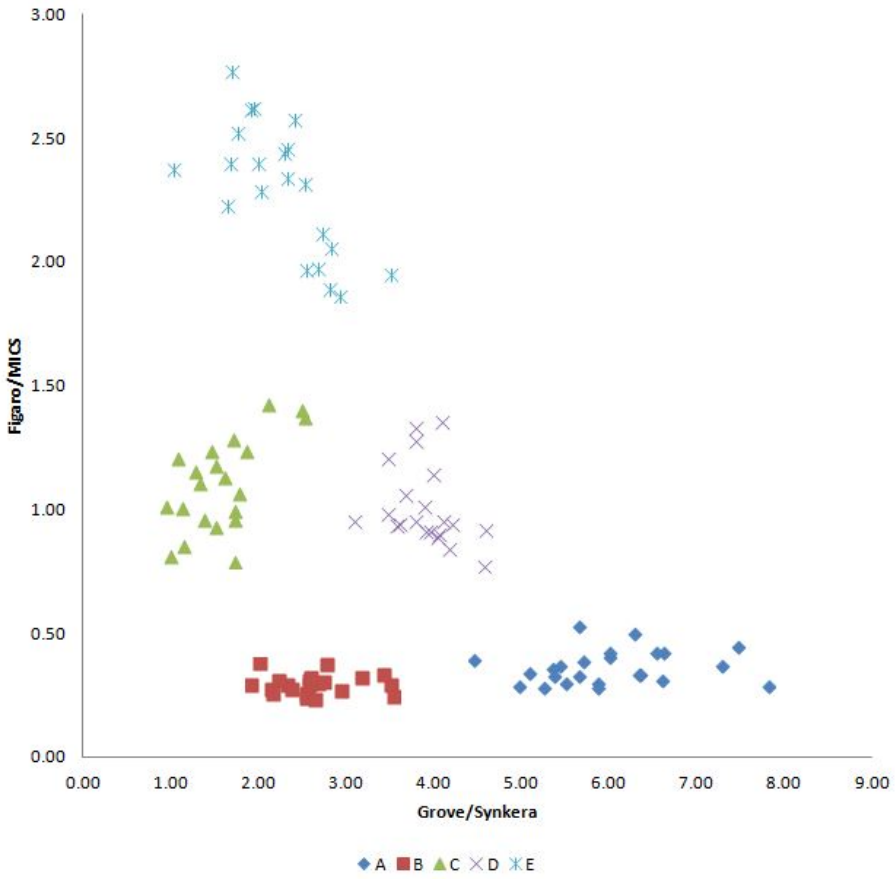


Figure IV.7: Ratio Selection Correct Results

Now the problem lies in partitioning the feature space into five regions that represent the chemical markers selected.

IV.2.3 Classifier

Mapping the feature vector x of sample data it is possible to obtain plot shown in figure IV.8. Ratio plot shows clear differences between the features of various VOC markers. It is unlikely that a complex decision boundary will provide a good generalization so a simple representation based on geometrical ellipse is chosen. The method will fall under partitional segmentation which implies a division of set of data objects into non overlapping subsets such that each data object is in exactly one subset[29]. This solution will also be considered as partial classification in which each object will not have to be assigned to a cluster thus avoiding misclassification of natural contaminants in the environment.

Inscribing an ellipse around a small set of points can be achieved through the use of the geometrical definition of the ellipse. Mathematically an ellipse with center at (h,k) and semiaxes a and b is formulated as such IV.15 :

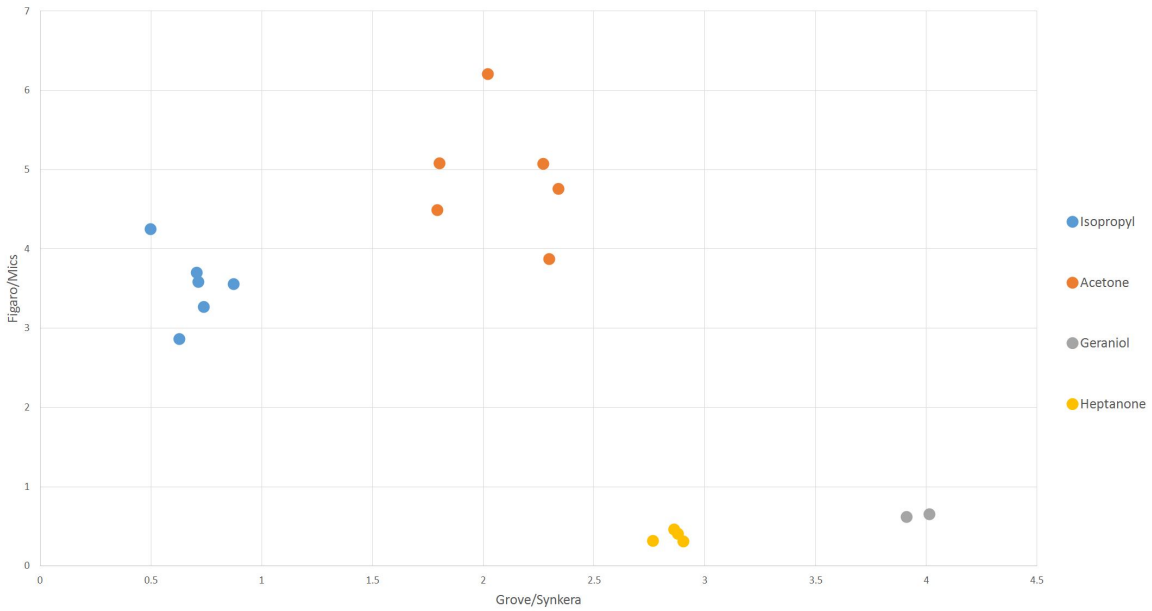


Figure IV.8: Feature vector scatter plot

$$\left(\frac{x-h}{a}\right)^2+\left(\frac{y-k}{b}\right)^2=1 \tag{IV.15}$$

i.e., the set of points $(x(t),y(t))$ where

$$x(t)=h+a\cos(t) \tag{IV.16}$$

$$y(t)=k+b\sin(t) \tag{IV.17}$$

and $0\leq t<2\pi$ [23]. This does not account for the fifth parameter of the ellipse which is tilt. To rotate the ellipse counter clockwise by θ radians an equation using matrix-vector can be used (Equation IV.18).

$$\begin{bmatrix}x(t)\\y(t))\end{bmatrix}=\begin{bmatrix}h\\k\end{bmatrix}+\begin{bmatrix}\cos(\theta)-\sin\theta)\\\sin(\theta)+\cos(\theta)\end{bmatrix}\begin{bmatrix}a\cos(t)\\b\cos(t)\end{bmatrix} \tag{IV.18}$$

Using this definition of an ellipse it is possible to construct an ellipse around n set of observations $(x_0,x_1,x_2,x_3,...,x_{n-1})$ where each of the observations is a feature vector on a two dimensional vector space. Computing a mean center for each individual marker in the scatter plot will produce a center of the ellipse (h,k) (Equation IV.19 IV.20) (Figure IV.9) :

$$\bar{h}=\frac{\sum_{i=1}^{n}x_{1_i}}{n} \tag{IV.19}$$

$$\bar{k} = \frac{\sum_{i=1}^{n} x_{2_i}}{n} \tag{IV.20}$$

Knowing the center of the ellipse and calculating the max distance of points in x and y direction plus the angle ϕ relative to the x axis an ellipse can be inscribed over a given set of points with a buffer to account for variability of results in simulated real world environments. The data for the ellipse is saved for each of the markers to act as a classifier in the form of parameters:

r_a, r_b - semi-major axis radius
ϕ - angle of semi-major axis from x-axis
(h,k) - center of the ellipse

Visualization of the ellipse constructed to inscribe sample set for an isopropyl marker is shown in Figure IV.10 where blue and red line represent computed ellipse and ten percent buffered computed ellipse, respectively. The constructed classifier performs partional exclusive classification of marker data mapped in 2-dimensional using previously discussed features (IV.12 IV.13).

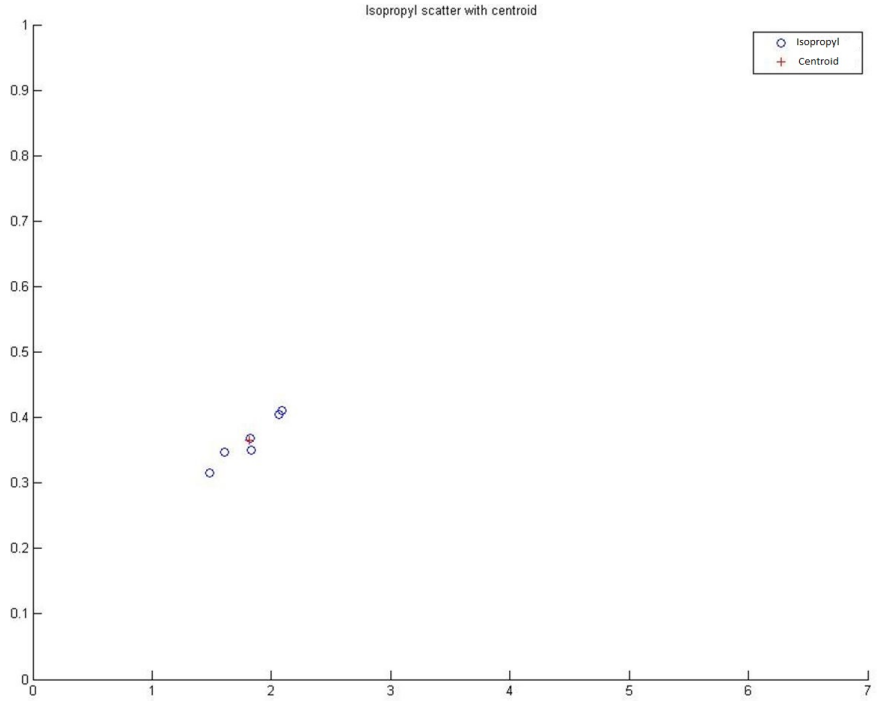


Figure IV.9: Isopropyl with computed center

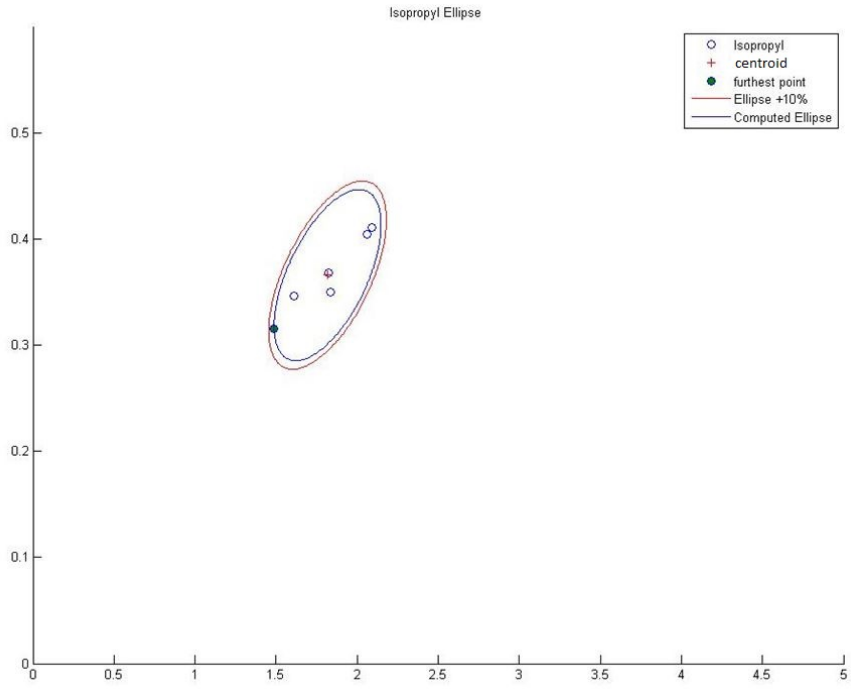


Figure IV.10: Isopropyl marker test data inscribed by an ellipse

IV.3 Receiver Operating Characteristic (ROC)

Receiver operating characteristics (ROC) graph is a visualization technique used for organizing and selecting classifiers based upon their performance. ROC curve has been used for signal detection theory to depict the trade of between hit rates and false alarm rates of classifiers [9]. ROC curves have been adopted for use in visualizing the behavior of diagnostic systems and are widely used in medical community for diagnostic testing [9]. The early adopter of ROC graphs in machine learning was Kent A. Spackman who demonstrated the inherit value of using receiver operating characteristics curves in evaluating algorithms [26]. Recent increase in use of ROC graphs is due in part to the realization that simple classification accuracy is often a poor metric for measuring the performance of a classifier.

ROC graph is a two-dimensional plot in which false positive rate is plotted versus true positive rate. ROC curve presents relative tradeoffs between benefits and costs associated with various constraints in the classifier [9]. This method is suited for analysis of discrete classifiers which ouputs only a class label with each discrete classifier producing a false positive and true positive pair corresponding to a single point on ROC space. Where in the space the point of (0,1) represents a perfect classification of input data [9].

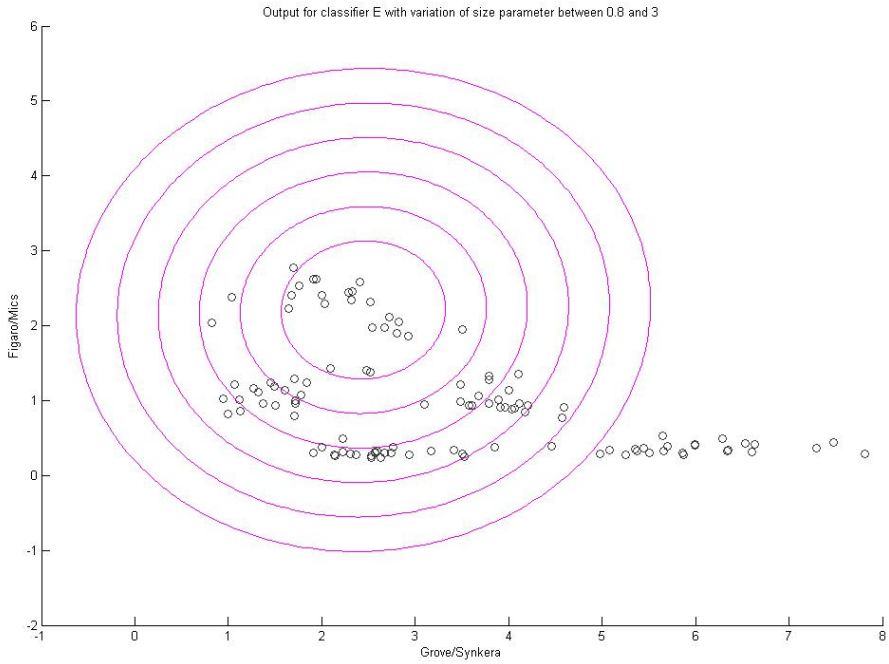


Figure IV.11: Output for classifier E with variation of size parameter between 0.8 and 3

Receiver operating characteristic curves are used to determine the proper ellipse ratio and the number of learning points to be used to achieve best classification results. To achieve this a program was written to analyze the test data varying the ellipse size and number of learning points parameters. By changing the buffer size of the ellipse it is possible to account for subtle variation of output measurements which are present when dealing with real world data acquisition. Figure IV.11 shows the output of the algorithm if the size parameter is varied between 0.8 to 2.8 in increments of 0.4. Variation of the parameter will size the ellipse and if not chosen properly will mis-classify or omit points from the result. As the idea of the classifier to detect the chemicals correctly it is of high importance to select the buffer parameter appropriately.

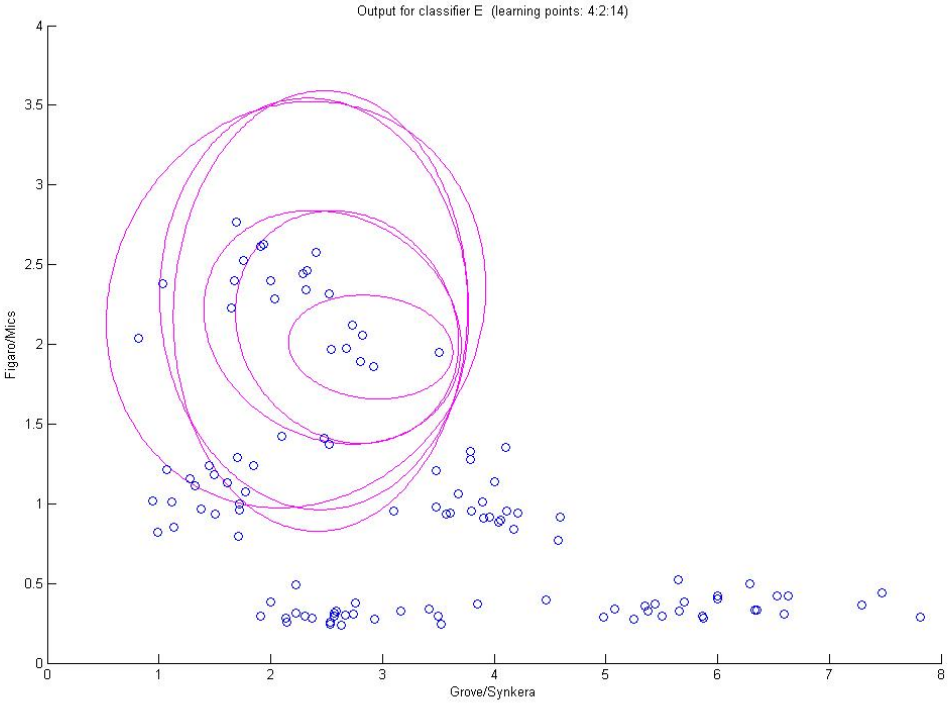


Figure IV.12: Output for classifier E varying learning points (4:2:14)

Number of learning points is also a very important metric in classification algorithm design. It is beneficial for usability of the system to have the lowest number of learning points as possible while still maintaining maximum achievable accuracy. Figure IV.12 shows the output of the algorithm if the learning points number is varied between four and fourteen in increments of two. Variation of both of the parameters will produce multiple curves each representing various number of learning points with points on each curve representing the size parameter.

To produce receiver operating characteristic curves for classifiers designed for this system data from a test dataset described in test method section of this paper each classifier was subject to variation of learning points from four to fourteen in steps of two and 0.8 to three in steps of 0.1 for the size the parameter. Each result is analyzed to determine the best parameters for classifier performance. These calculations can improve the efficiency and accuracy of developed algorithm for future system use. ROC curves for classifier A are shown in figure IV.13, the plot does not include points out to one of false positive rate to focus on the area of interest. Analysis of classifier specification shows four points are not enough to correctly classify all the tuples even with a size parameter of three. To achieve minimum number of learning points in conjunction with maximum achievable true positive

rate at least six learning points have to be used.

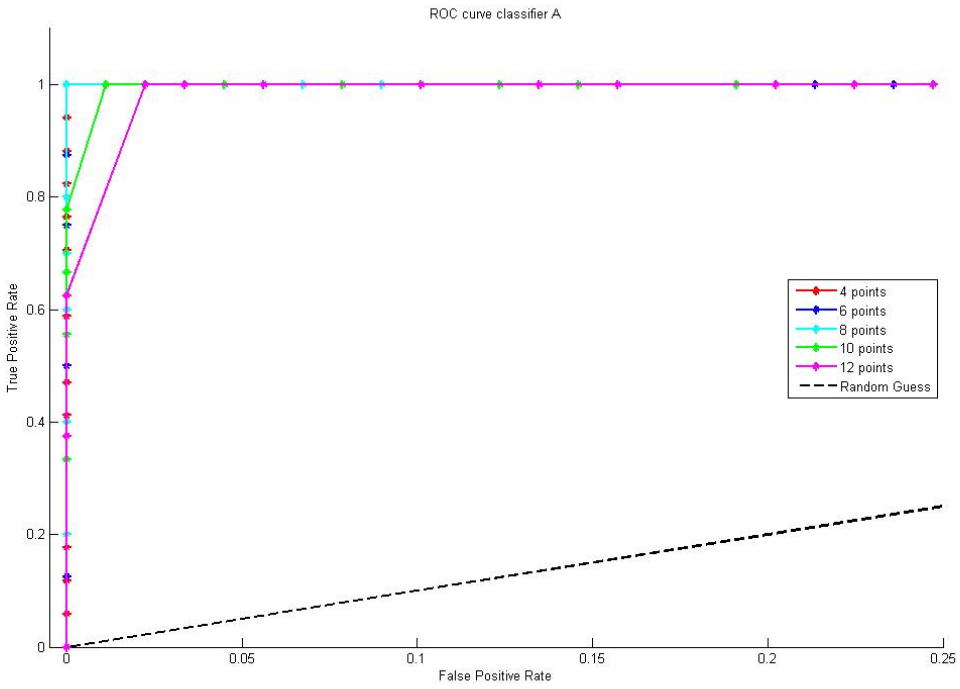


Figure IV.13: ROC curves for classifier A

Figure IV.14 is displaying the best case scenario for the classifier which is achieved by using 6 points for classifier training and 1.3 as ellipse size modifier. This classifier parameters will produce discrimination factor of one which is the area underneath the curve and describes the ability of the algorithm to correctly identify the tuples in the test data set.

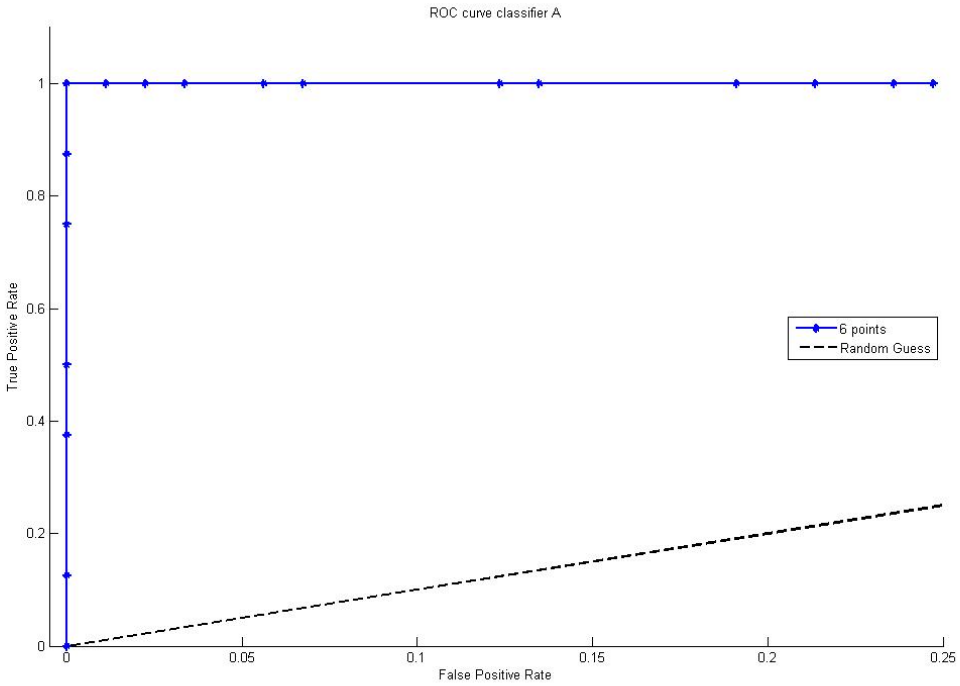


Figure IV.14: Best ROC curve for classifier A

Figure IV.15 shows the resulting curves for classifier B. The number of classifier training tuples has to be 6 or above to produce full recognition without false positive results. Best settings were chosen to be six learning points with 1.3 size factor. Testing data for the classifier has produced a discrimination factor of one by identifying all tuples correctly without false positive hits. Figure IV.16 shows the ROC curve for the best setting achieved with test data.

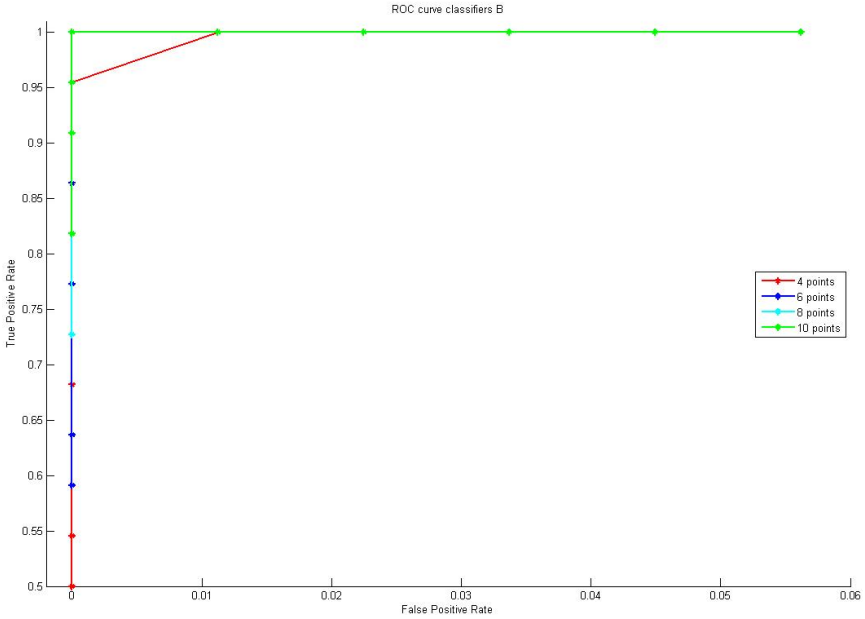


Figure IV.15: ROC curves for classifier B

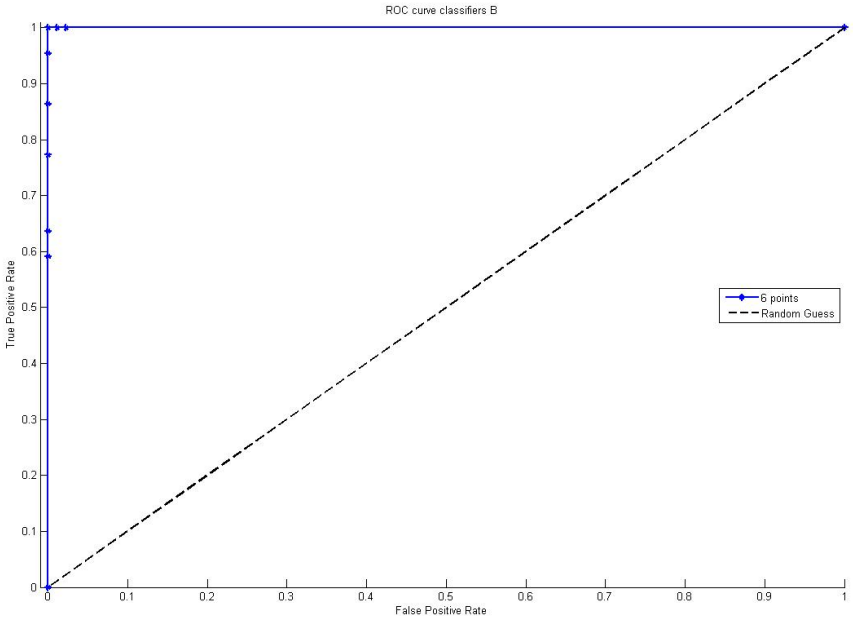


Figure IV.16: Best ROC curve for classifier B

Figure IV.17 shows the resulting curves for classifier C. The number of classifier training tuples for best algorithm performance is determined to be ten. Best settings where chosen to

be ten learning points with 1.0 size factor. These settings produces discrimination factor of 0.9457 calculated using trapezoidal numerical integration. Figure IV.18 shows the ROC curve for the best setting achieved with test data.

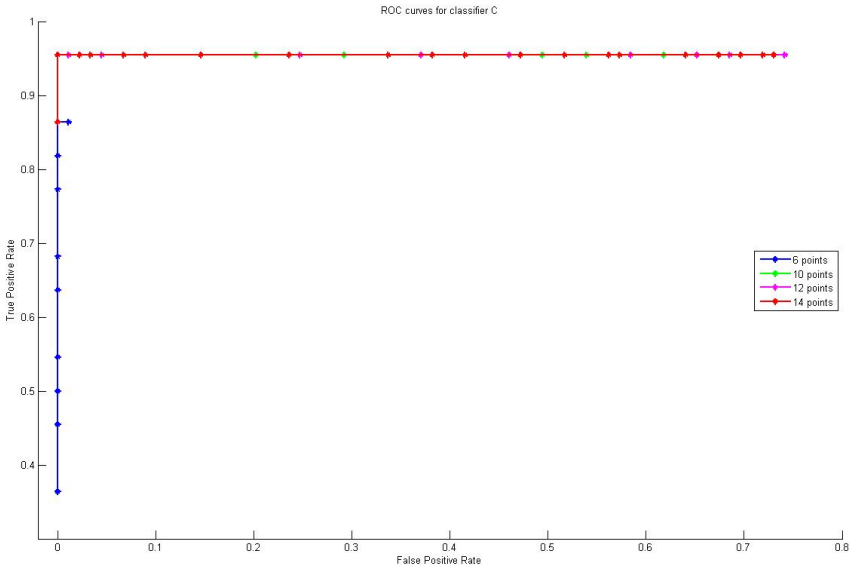


Figure IV.17: ROC curves for classifier C

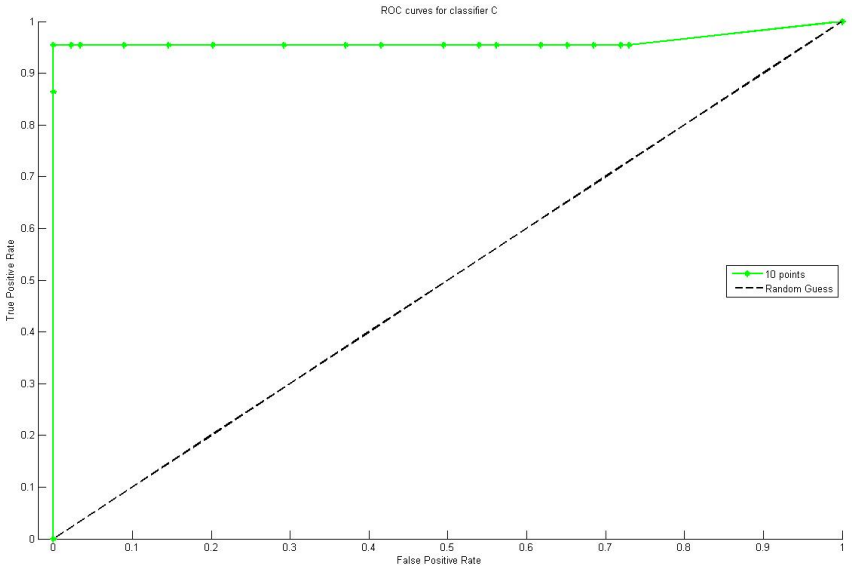


Figure IV.18: Best ROC curve for classifier C

Figure IV.19 shows the resulting curves for classifier D. The number of classifier training

tuples for best algorithm performance is determined to be ten. Best settings were chosen to be ten learning points with 1.0 size factor. These settings produces discrimination factor of 0.9724 calculated using trapezoidal numerical integration. Figure IV.18 shows the ROC curve for the best setting achieved with test data.

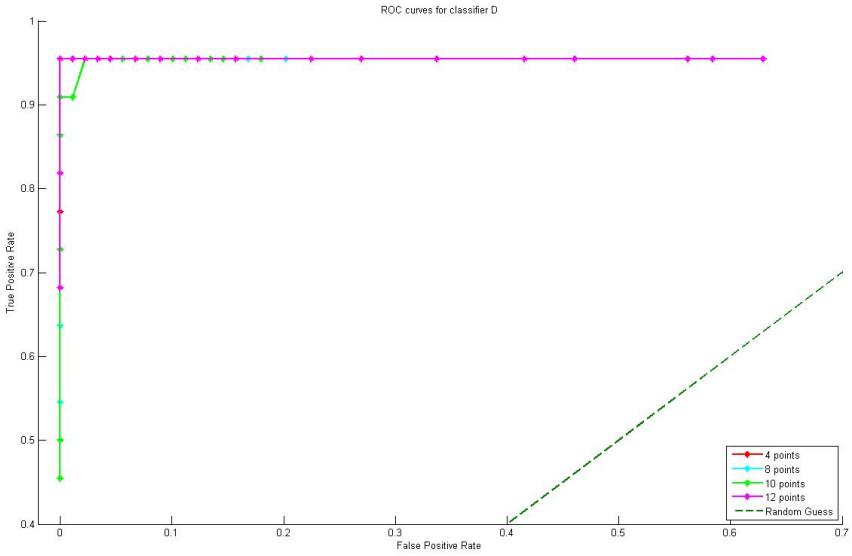


Figure IV.19: ROC curves for classifier D

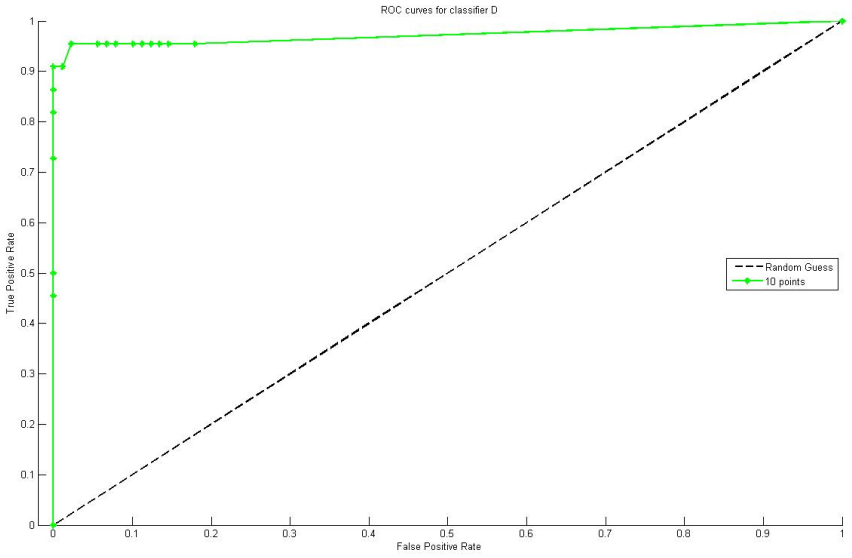


Figure IV.20: Best ROC curve for classifier D

Figure IV.21 shows the resulting curves for classifier E. The number of classifier training tuples for best algorithm performance is determined to be ten. Best settings were chosen to be 1.5 size factor. These settings produce discrimination factor of 0.9967 calculated using trapezoidal numerical integration. Figure IV.22 shows the ROC curve for the best setting achieved with test data.

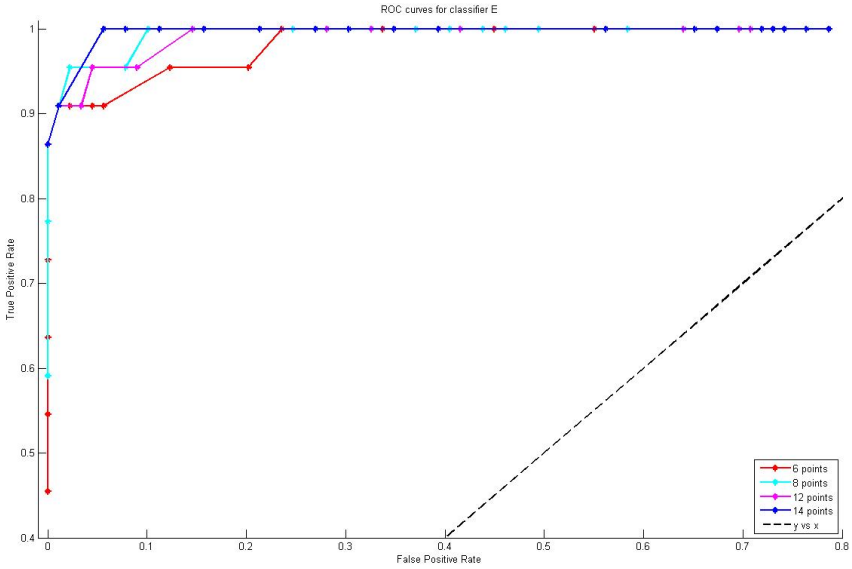


Figure IV.21: ROC curves for classifier E

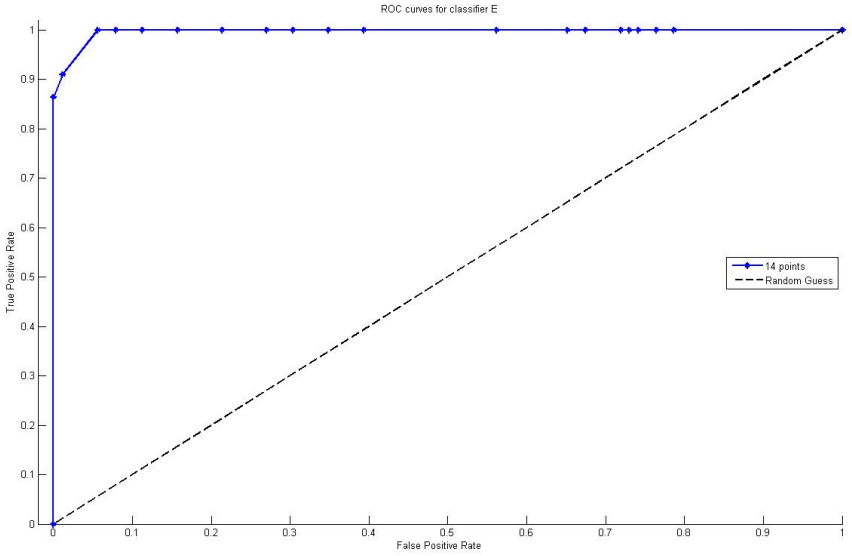


Figure IV.22: Best ROC curve for classifier E

IV.4 Preliminary Results of Classification Algorithm

Pattern classification techniques are designed to have the capability for selective detection of volatile organic markers in simulated real world environment. The classifier is trained for each marker with user specified training points. Each volatile organic marker is assigned a key under which information from learning data, individual responses, classifier parameters and chemical name are stored (Figure IV.24). These parameters were chosen to be saved to allow for post processing of the data. A chemical is considered as detected if the features mapped are located within the ellipse. In Figure IV.23 an ellipse was constructed around sample data (blue circles) with computed center shown with a black ring and two sample points (green) one inside the ellipse which we considered a detected marker and one outside which will be unassigned. This method is scalable up to two or more volatile organic markers as shown in Figure IV.25 in which isopropyl and methanol tests were used as sample data.

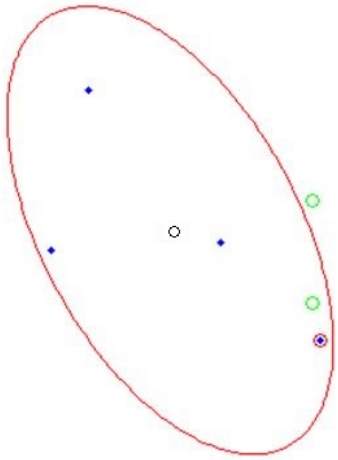


Figure IV.23: Sample data test

Based upon the classification of the algorithm presence of a prior learned marker can be detected. Decisions based on the outcome of the classifier can be made to control function of a larger system.

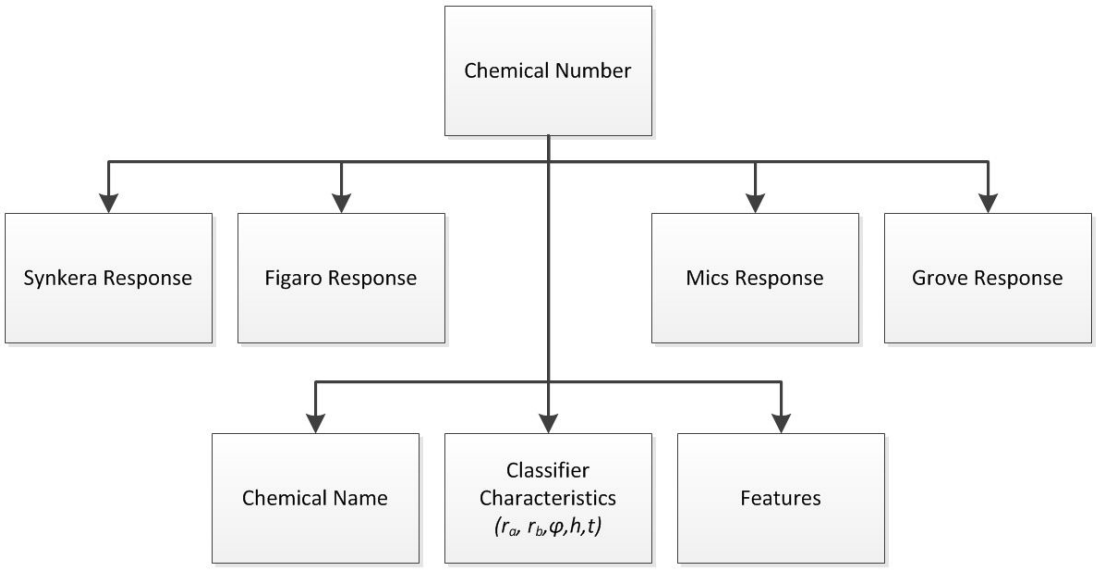


Figure IV.24: Structure diagram

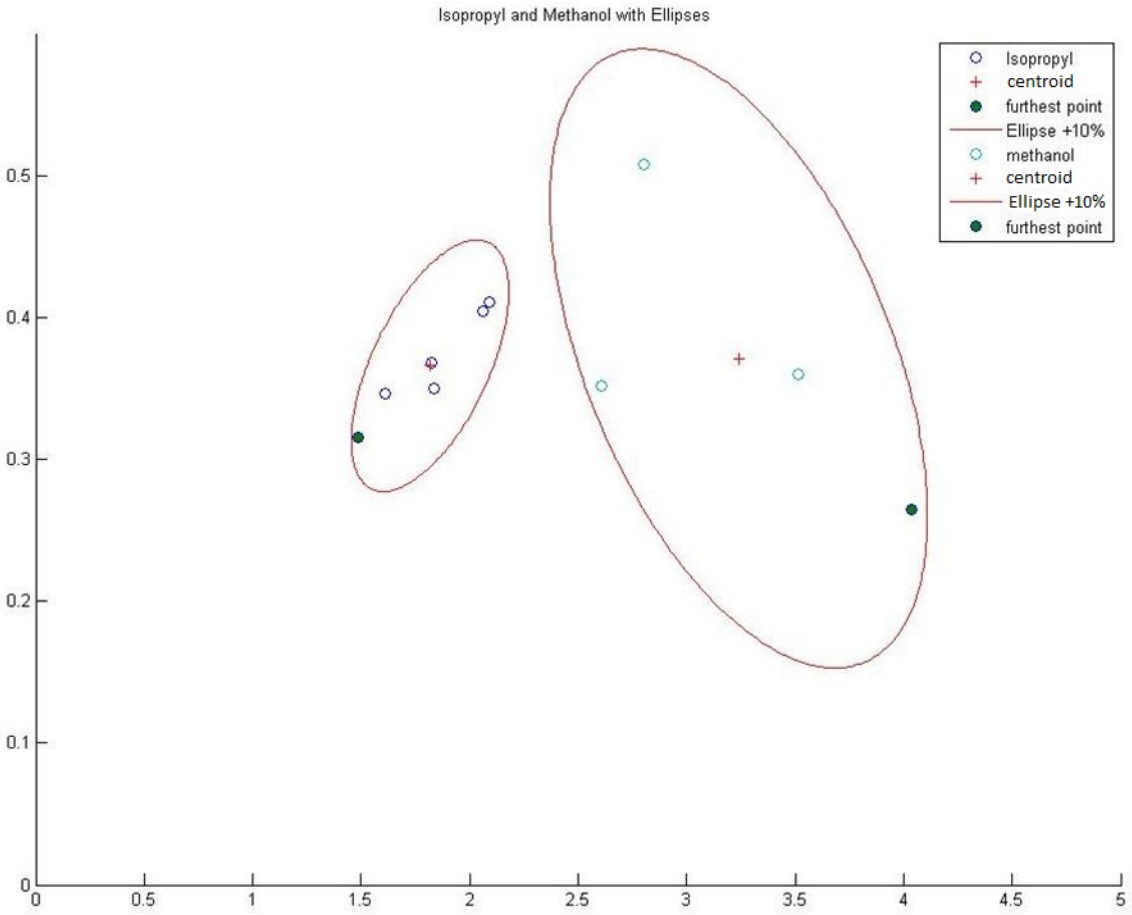


Figure IV.25: Structure diagram

Chapter V

Results

V.1 Test Method

To test the accuracy of the proposed system an experiment was designed as follows. Chemicals were selected to represent a wide range of VOC with various vapor pressures and compositions. Chemicals used for testing are shown in Table V.1 .

Table V.1: Chemical Characteristics

Name	Chemical Forlmula	Melting Point	Boiling Point	Vapor Pressure @20°C
Acetone	CH_3COCH_3	-94°C	56°C	184 mmHg
Methanol	C_3H_8O	-89.5°C	82°C	33 mmHg
Acetonitrile	C_2H_3N	-48°C	81°C	73.99 mmHg
2-Heptanone	$CH_3(CH_2)_4COCH_3$	-35°C	149°C	2.14 mmHg
Geraniol	$(CH_3)_2C=CHCH_2CH_2C(CH_3)=CHCH_2OH$	-15°C	229°C	0.2 mmHg
Isopentyl Acetate	$CH_3CO_2CH_2CH_2CH(CH_3)_2$	-78°C	142°C	3.8 mmHg
Ethanol	CH_3CH_2OH	-114°C	78°C	44.6 mmHg

All chemicals were chosen due to their safety and availability. Chemicals were injected onto a 2”x4”Q-panel (Figure V.1) using an ordinary 5ml syringe to produce variable exposure volumes. The panel was placed on a temperature controlled heat exchanger to maintain temperature during testing. VWR 1140s Refrigerated Bath system was used to control the temperature of the test stand (Figure V.2). VWR 1140s is capable of maintaining temperatures between -20°C to 150°C by pumping liquid through a thermal exchange plate at 9 to 15 liters/min.

Figure V.1: MiCS-5135 sensor response [7]

Figure V.2: MiCS-5135 sensor [7]

A LONC LC6015B fan was used to create an airstream across the test stand to reproduce real life environment. LONC fan was operated at 4V out of maximum 12V to ensure a gentle air flow across the Q-panel. The system was warmed up for minimum of 30 minutes before each test. The warm-up period insured for stable operation of the sensors. Each chemical had its own Q-panel and syringe to avoid cross contamination during testing. After conducting each test the system was allowed to return to its stable zero state condition before running the next test. Time to achieve zero state varies depending on the compound being tested. These guidelines where used for all tests conducted on the system.

V.2 Test Results

The chemical detection system was subjected to various tests to ensure proper operation. MOX sensors have various responses to change of temperature, concentration, and humidity. In spite of this, variations the sensors displayed high variability and selectability when exposed to selected target gases. Data for all tests was collected and saved in database files for future review.

Initial data acquired from sensor array shows the remarkable differences in sensor response to various contaminants (Figure V.3). Highly volatile compounds like acetone, isopropyl alcohol, and ethanol show a higher response than the compounds with low volatility like 2-heptane and gerniol. One exception to this is the acetonitrile which shows lower sensor array response than the others. Normalizing the responses shows the distinct signatures of each gas (Figure V.4). System algorithm uses this information to predict the gas it is exposed to.

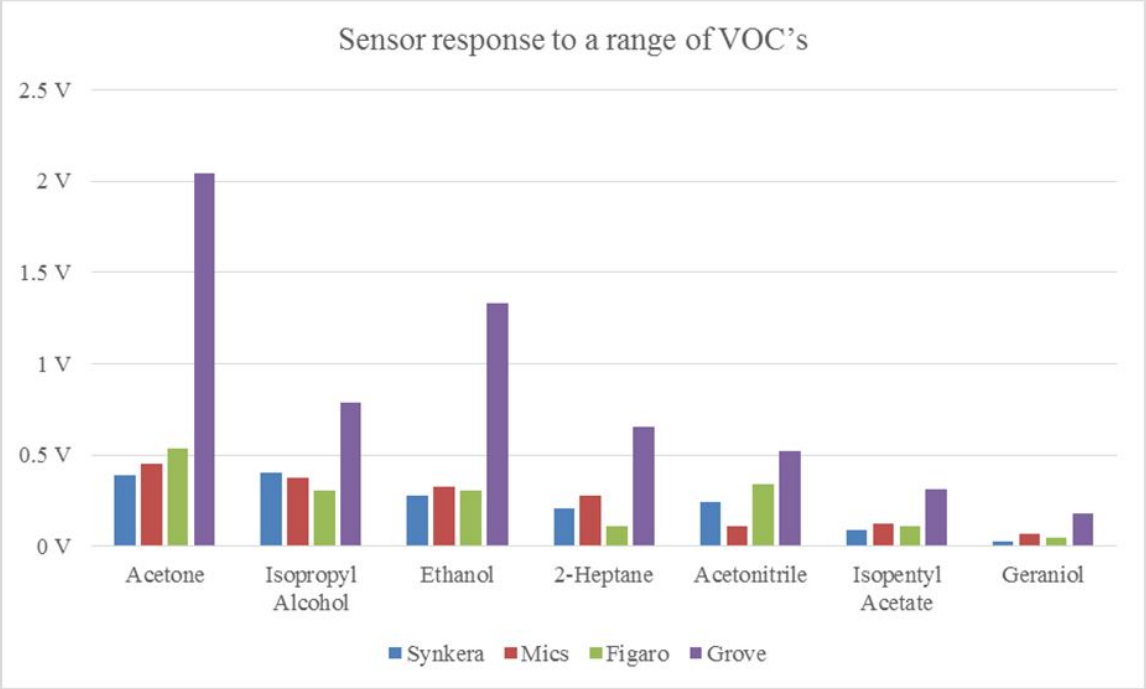


Figure V.3: Sensor response to a range of VOC's

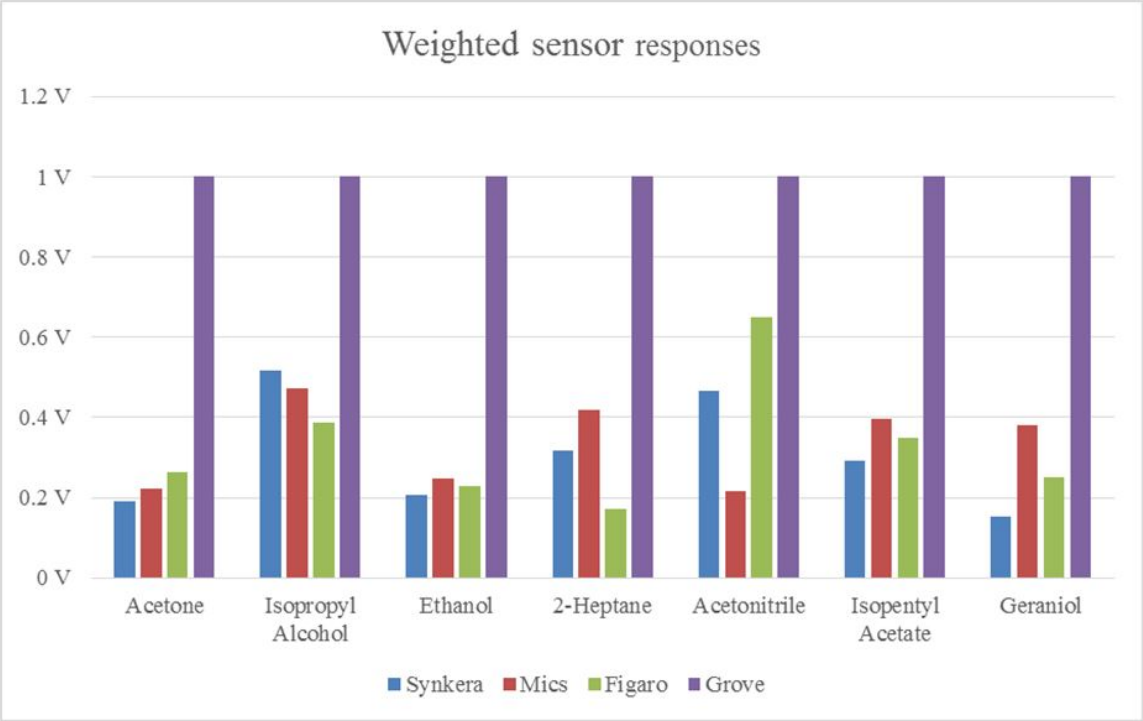


Figure V.4: Weighted sensor responses

For real world operation the system has to be able to detect and identify chemicals

in various weather conditions. While the scope of this project was a feasibility study, it is important to see if the sensor array can detect chemicals if it is employed in different temperatures. The relationship between vapor pressure and temperature is not linear, thus presenting a real challenge to the detection abilities of the system. To simulate temperature variations a test using a heat transfer plate was conducted. The temperature of the plate was changed in $10°C$ increments from $20°C$ to $80°C$. As expected the sensor response increases with increase of temperature (Figure V.5). The increase is not linear but a linear fit can be calculated to approximate the rate of change (Figure V.6). The variation of vapor pressure has a significant effect on the output of the sensors and needs to be taken into account if the device is used on surfaces with high temperature fluctuations.

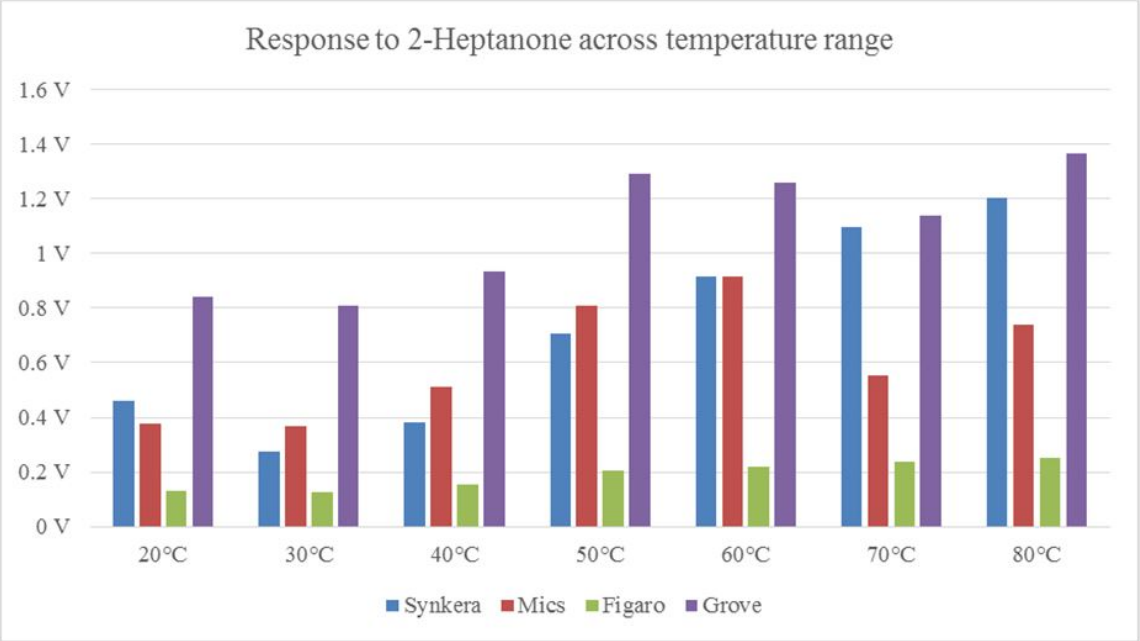


Figure V.5: 2-Hepatanone response across a temperature range

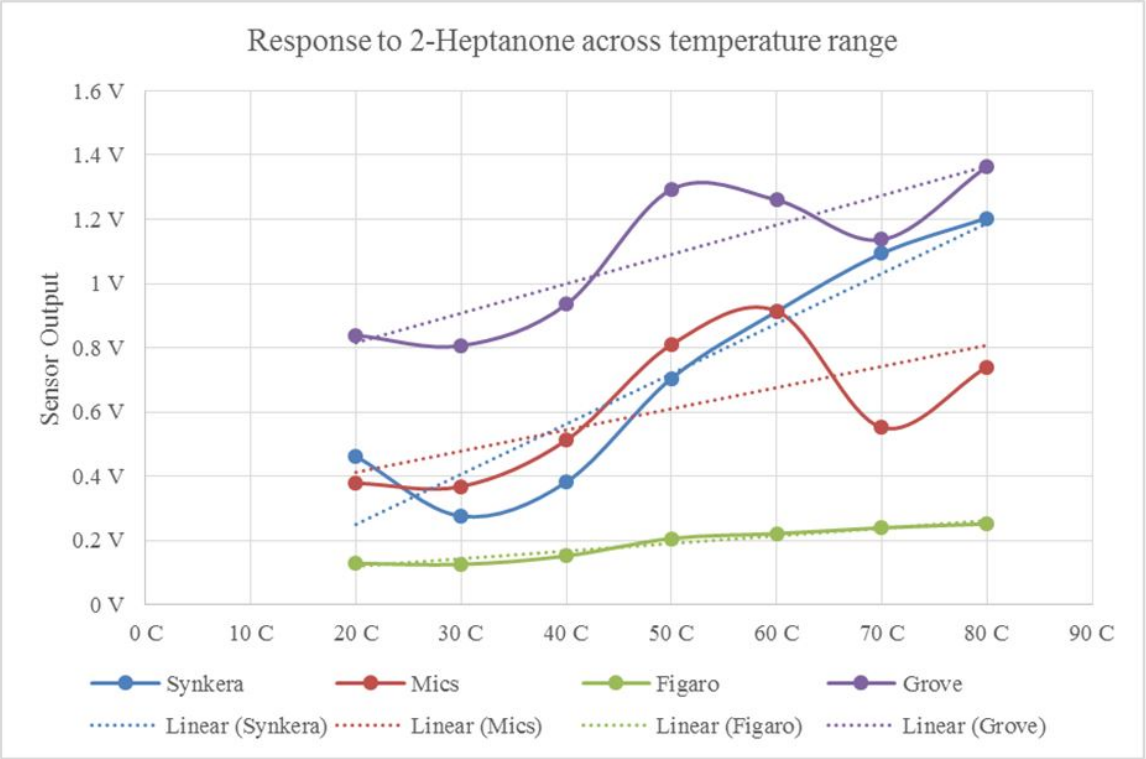


Figure V.6: 2-Heptanone temperature response with linear approximation

To correctly gauge how this affects the algorithm's repeatability the data has to be processed and compared to other compounds. Figure V.7 shows the response of several chemicals vs the response of 2-heptanone measured at different temperatures. The temperature affects the readings by shifting the detection points in one or both axes. Even though temperature produces highly variable data in comparison to stable temperature data it would still be possible to detect various VOC's if they are properly chosen. For example if 2-heptanone and acetonitrile are used there would not be an overlap. Temperature study shows that even in presence of highly variable temperature it is possible to distinguish between various VOC's using the proposed method.

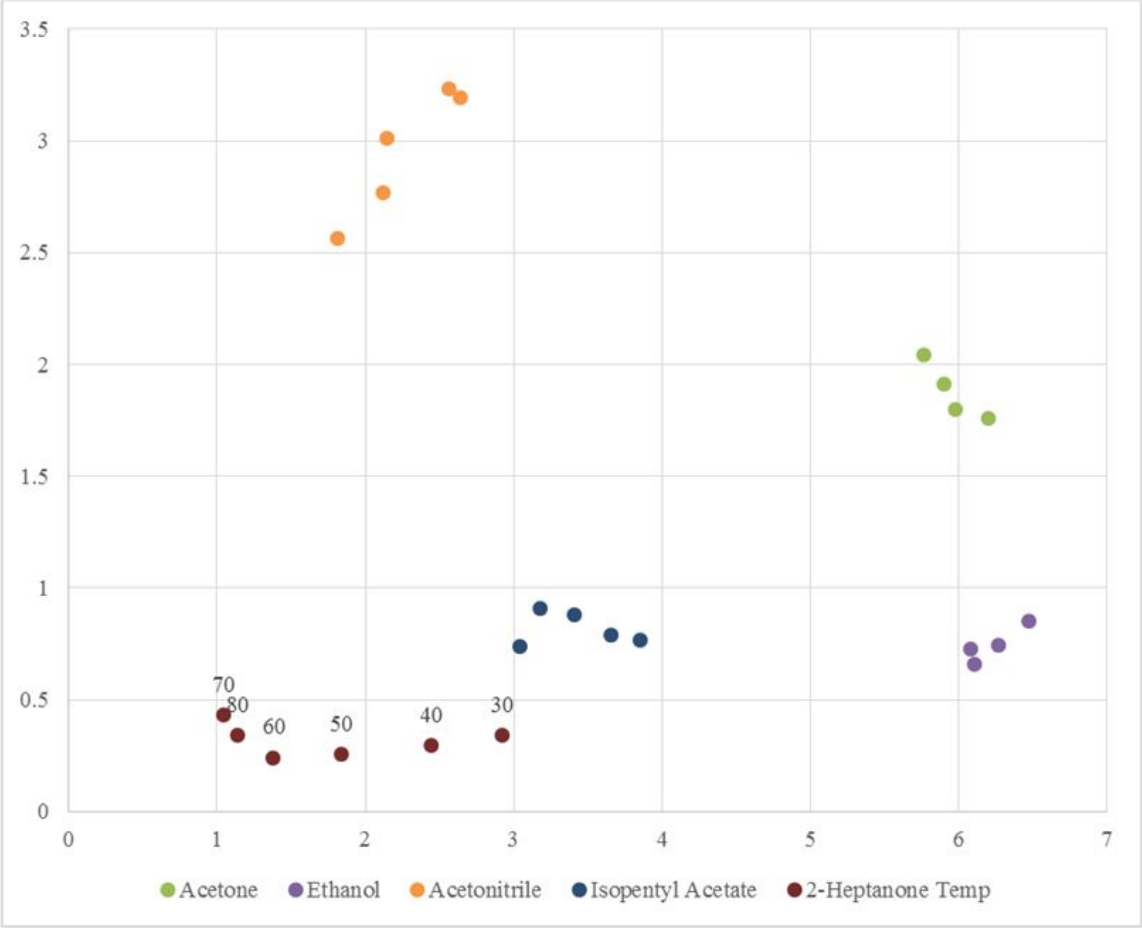


Figure V.7: 2-Heptanone response to temperature plotted with other markers

Another variable that affects the repeatability of recognition is the concentration of contaminant in gas phase. To simulate the effects of vaporization of VOC's, volumes from 0.2ml to 1ml are used to simulate a range of contaminant concentrations. It is known that the concentration of target gases affects the response of MOX type sensors. A concentration study was conducted to understand the effects of concentration on the response of the system. Change of concentration will affect the outputs of the sensors in various ways as shown in Figure V.8. As in temperature study the system will still be able to detect the chemical markers if proper VOC markers are chosen and used.

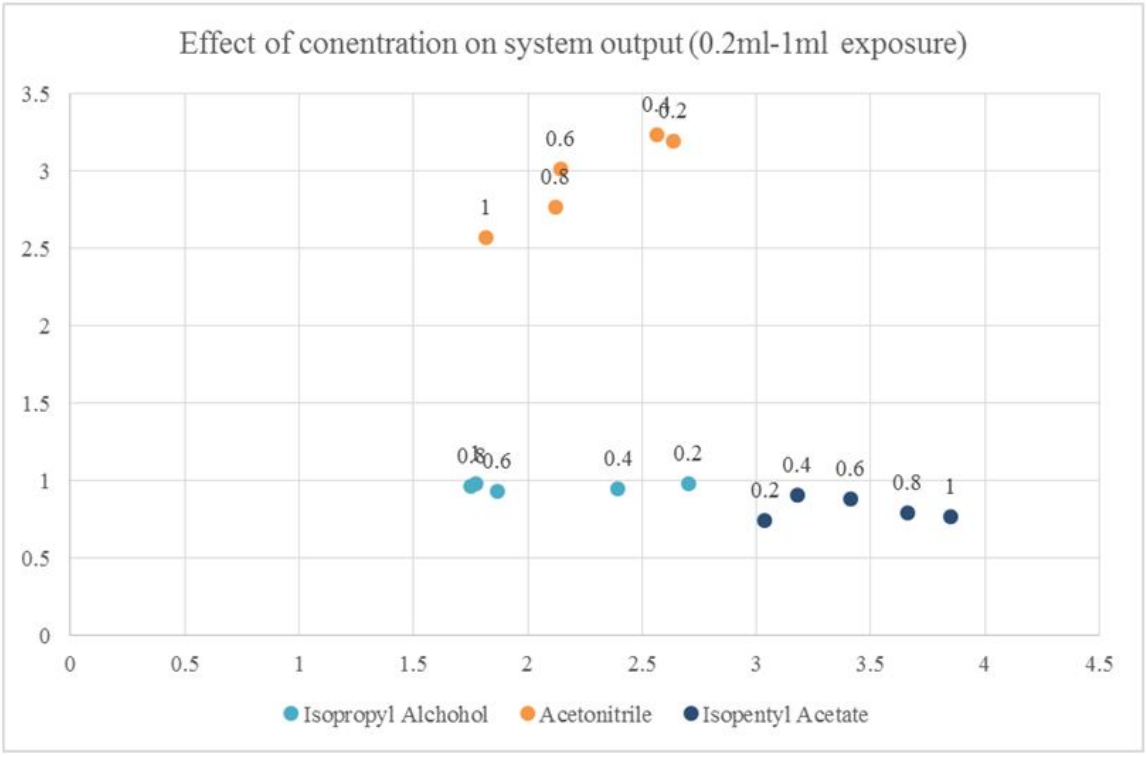


Figure V.8: Concentration effect on system response

The scope of the research required an investigation into feasibility of chemical markers for use in navigation.A system designed to perform this task requires learning points to set up baselines for the markers chosen. The markers where chosen based upon prior experimental data and included chemicals that display similar response characteristics as well as various evaporation times. For the purposes of this document the chemicals are going to be referred to as A, B, C, D, and E. Five learning test sets for each marker are performed on the test system with each chemical having its own syringe and panel for the duration of the tests. The volume of the injected markers are varied between 0.2ml and 0.8ml. All the tests are conducted on a heat conducting plate that is maintained at a temperature of $40^{\circ}C$. Between each test the panel is wiped clean with a paper towel to avoid leftover residue on the surface of the q-panel. Figure V.9 represents the data acquired during a learning process with ellipses enclosing each of the target markers.

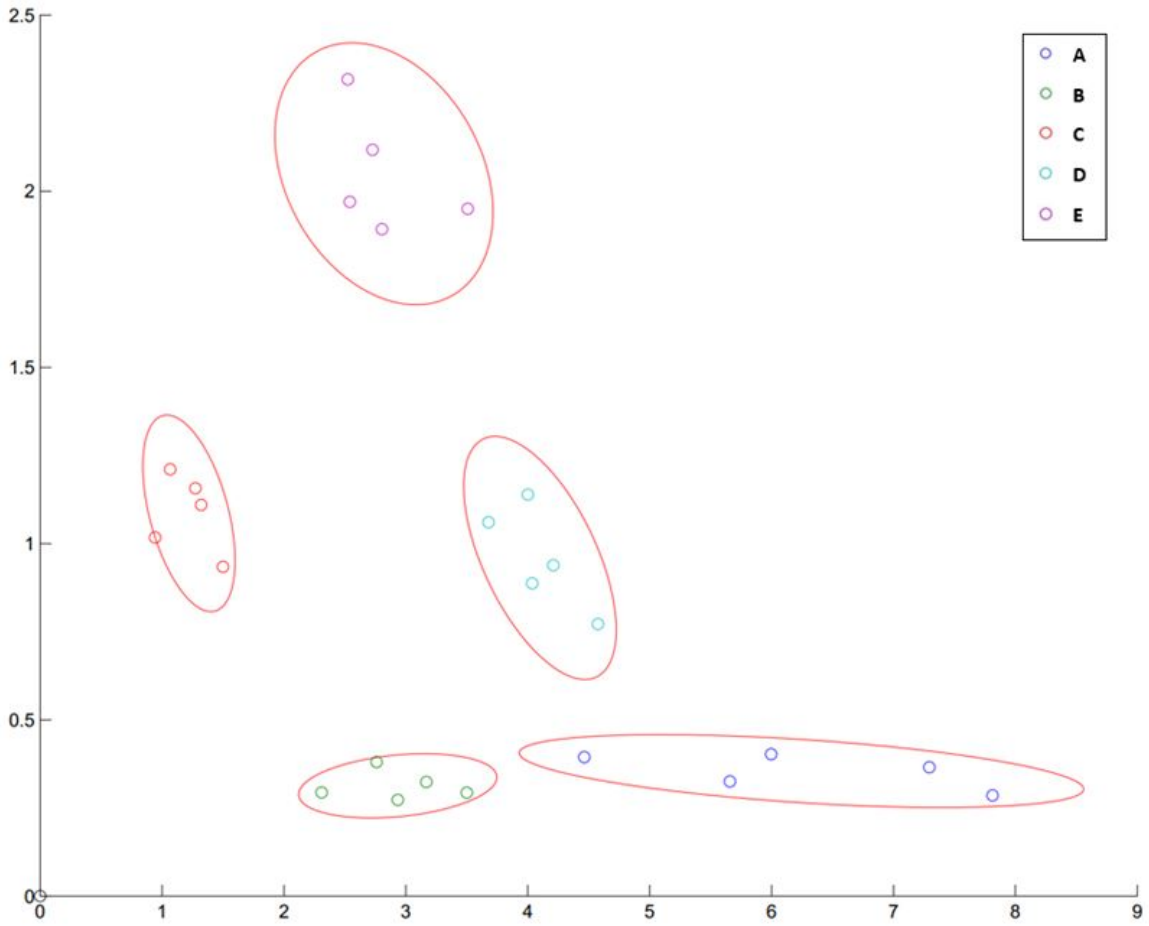


Figure V.9: Learning dataset

For maximum repeatability of the model the datasets adjust throughout the test by incorporating additional points into the dataset used to calculate the enclosing ellipse. Additional points incorporated into the calculation of the model affect the calculated centroid and the major and minor axis of the ellipse thus transforming it throughout the test (Figure V.10). During this test the memory was not limited thus all points are used to calculate the ellipse region without a moving window.

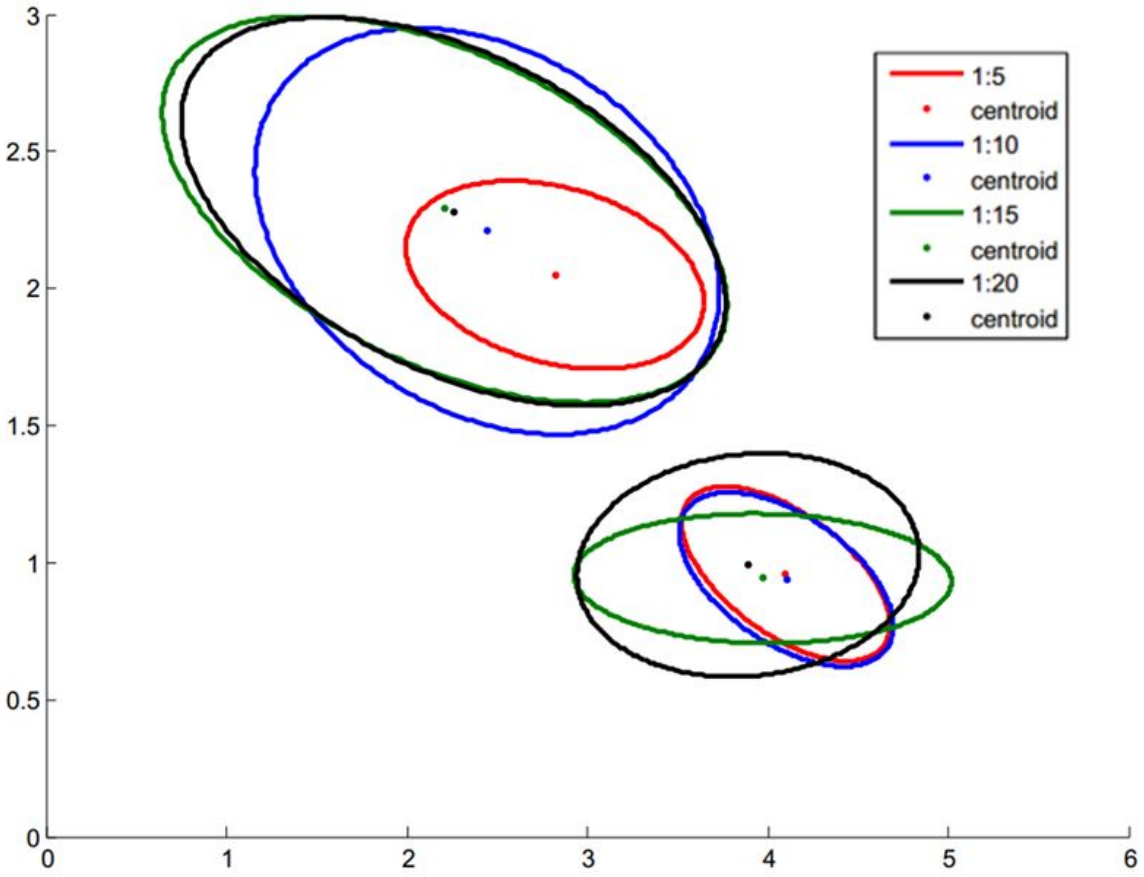


Figure V.10: Ellipse transformations

To test the repeatability of the system, a test is constructed to subject the system to a random chemical. Selecting a random chemical ensures that the system is capable of switching between various chemicals during operation without the introduction of error. Markers are tested in sets of two and a total of fifty individual tests are performed (Table V.2).

Table V.2: Order of tests performed

Test	Chemical	Test	Chemical
1	4	26	2
2	4	27	1
3	2	28	1
4	2	29	4
5	1	30	4
6	1	31	4
7	3	32	4
8	3	33	5
9	1	34	5
10	1	35	2
11	5	36	2
12	5	37	3
13	3	38	3
14	3	39	4
15	2	40	4
16	2	41	1
17	5	42	1
18	5	43	5
19	1	44	5
20	1	45	4
21	3	46	4
22	3	47	3
23	5	48	3
24	5	49	2
25	2	50	2

The system is allowed to return to a steady state between tests. If a chemical is misidentified by the system or does not fit into any zone it is disregarded. Results of the tests are recorded into an excel file with one representing correct identification and zero assigned to an unresolved marker.

Results of the repeatability tests produced a set of ellipses enclosing correctly identified markers (Figure V.11). The final repeatability of detection for this data set was determined to be 93% with errors occurring during recognition of markers A, B, and E (Table V.3).

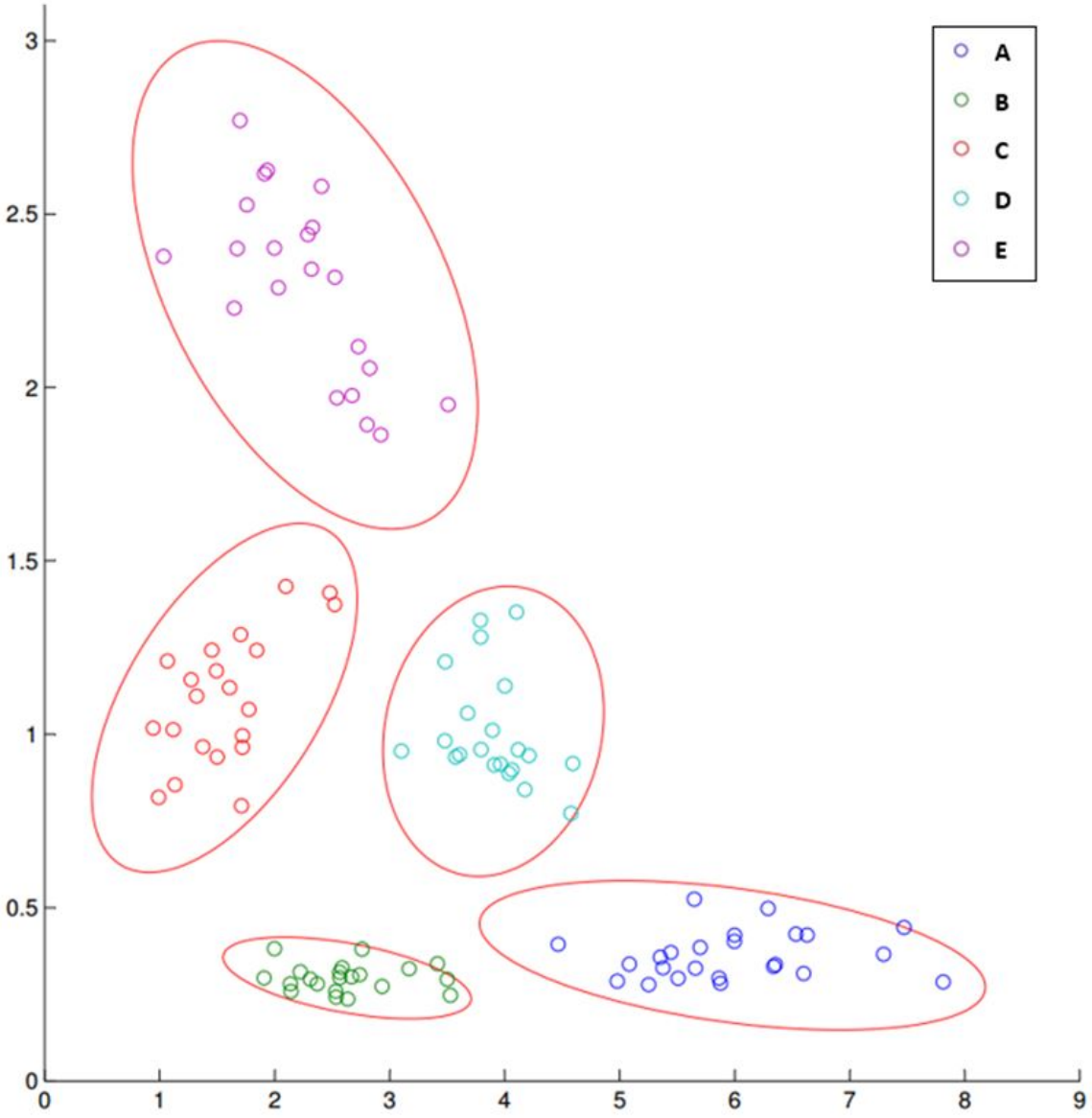


Figure V.11: Final output of the system including all detected markers

Table V.3: Test Results

Marker	A	B	C	D	E
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	0
4	1	1	1	1	1
5	1	1	1	1	1
6	1	1	1	1	1
7	0	0	1	1	1
8	1	1	1	1	1
9	1	1	1	1	1
10	1	1	1	1	1
Sum	9	9	10	10	9
%Correct	90%	90%	100%	100%	90%
Total % Correct	94%				

V.3 Results Analysis

Using training data to derive the accuracy of the learned model can result in misleading overoptimistic estimates so for measurement of classifier's accuracy, a test set of class-labeled tuples that was not used in derivation of the model was used. Metrics for assessing how "accurate" the chosen classifier is at predicting the class label of tuples include evaluation measures in Table V.4 [14].

Table V.4: Evaluation Measures [14]

Measure	Formula
Accuracy, recognition rate	$\frac{TP+TN}{P+N}$
Error rate, misclassification rate	$\frac{FP+FN}{P+N}$
Sensitivity, true positive rate, recall	$\frac{TP}{P}$
Specificity, true negative rate	$\frac{TN}{N}$
Precision	$\frac{TP}{TP+FP}$

With various terms used in the table defined as follows:

- **P**: Number of positive tuples.
- **N**: Number of negative tuples.
- **True Positives (TP)**: Positive tuples that were correctly labeled by the classifier.

- **True Negatives (TN)**: Negative tuples that were identified correctly by the classifier.
- **False Positives (FP)**: Negative tuples that were labeled incorrectly as positive by the classifier.
- **False Negatives (FN)**: Positive tuples that where labeled incorrectly as negative by the classifier.

This terms can be summarized in a confusion matrix which is also known as a contingency table or an error matrix. This allows for visualization of the performance of a pattern recognition algorithm. Using the data acquired from Table V.3 a confusion matrix for each of the classifiers can be constructed in the form of Table V.5 [14]. In table V.6 the standard confusion matrix was filed in with the results acquired during repeatability testing.

Table V.5: Standard Confusion Matrix

		Predicted class		
		Yes	No	Total
Actual class	Yes	TP	FN	P
	No	FP	TN	N
	Total	P	N	P+N

Table V.6: Confusion Matrix for Classifiers Designed to Detect Chemicals A,B,C,D,E

	A	A'	Total
A	9	1	10
A'	0	40	40
Total	9	41	50

	B	B'	Total
B	9	1	10
B'	0	40	40
Total	9	41	50

	C	C'	Total
C	10	0	10
C'	0	40	40
Total	10	40	50

	D	D'	Total
D	10	0	10
D'	0	40	40
Total	10	40	50

	E	E'	Total
E	9	1	10
E'	0	40	40
Total	9	41	50

The accuracy of a classifier on a given test set is the percentage of test tuples that are correctly identified ty the classifier:

$$accuracy = \frac{TP + TN}{P + N} \tag{V.1}$$

Using the equation V.1 and the data in table V.6 the accuracy of the developed classifiers on the data set are:

$$\begin{aligned} accuracy_A &= \frac{9+40}{10+40} * 100\% \quad = 98\% \\ accuracy_B &= \frac{9+40}{10+40} * 100\% \quad = 98\% \\ accuracy_C &= \frac{10+40}{10+40} * 100\% \quad = 100\% \\ accuracy_D &= \frac{9+40}{10+40} * 100\% \quad = 100\% \\ accuracy_E &= \frac{9+40}{10+40} * 100\% \quad = 98\% \end{aligned} \tag{V.2}$$

Test data produced a great accuracy on the test dataset using the testing parameters overviewed in test method section. Precision and recall measures are also widely used in classification. Precision is thought of as a measure of exactness or what percentage of tuples are labeled as positive are actually such, whereas recall is a measure of completeness or what percentage of positive tuples are labeled as such. Recall is the same as sensitivity (or true positive rate). Calculation required for both measures is shown in table V.6. Recall and sensitivity are calculated for the test data across all of the classifiers in table V.7. The perfect score of 1 for classifiers C and D show that the system has identified all of the tuples correctly.

Table V.7: Evaluation of Precision and Recall

Classifier	Precision	Recall
A	0.18	0.9
B	0.18	0.9
C	0.2	1
D	0.2	1
E	0.18	0.9

V.3.1 Bootstrap Method

B.Efron (1979) has introduced the Bootstrap method[14]. Bootstrap method samples a given training set uniformly with replacement such that each time a tuple is selected it is equally likely to be chosen again and re-added to the training set. The most commonly used method is called the .632 bootstrap. The data set is sampled d times with replacement which yields a bootstrap sample or training set of d in this sample. The probability of original data tuples to occur more than once in this tuple is very high [14]. Data tuples that were not included in the training set form the test set. On average 63.2% of the original data will end up in the bootstrap sample hence the name .632 bootstrap. Repeating the sampling

procedure k times with each iteration increasing the accuracy of the model estimate obtained from this method produces an overall estimated accuracy for the model, M, as [14]:

$$Acc(M) = \frac{1}{k}\sum_{i=1}^{k}(0.632 * Acc(M_i)_{test\_set} + 0.368 * Acc(M_i)_{train\_set}) \tag{V.3}$$

where $Acc(M_i)_{test\_set}$ is the accuracy of the model obtained with boostrap sample i when it is applied to test set i. $Acc(M_i)_{train\_set}$ is the accuracy of the model obtained with bootstrap sample i when it is applied to the original set of data tuples [14]. Bootstrapping is considered to be overly optimistic and it works best with small data sets.

Computing correlation coefficient using bootstrap re-sampling of the test data set yields results shown in Table V.3. Correlation coefficient are statistics that quantify the relation between input and output in unit-free terms. Histogram in Figure V.12 shows the variation of the correlation coefficient across all the bootstrap samples. Bootstrap standard error for the estimated correlation coefficient is calculated using:

$$s = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{1}{2}} \tag{V.4}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i \tag{V.5}$$

and n is the number of elements in the sample. Standard error for the estimated correlation coefficients is 0.0734.

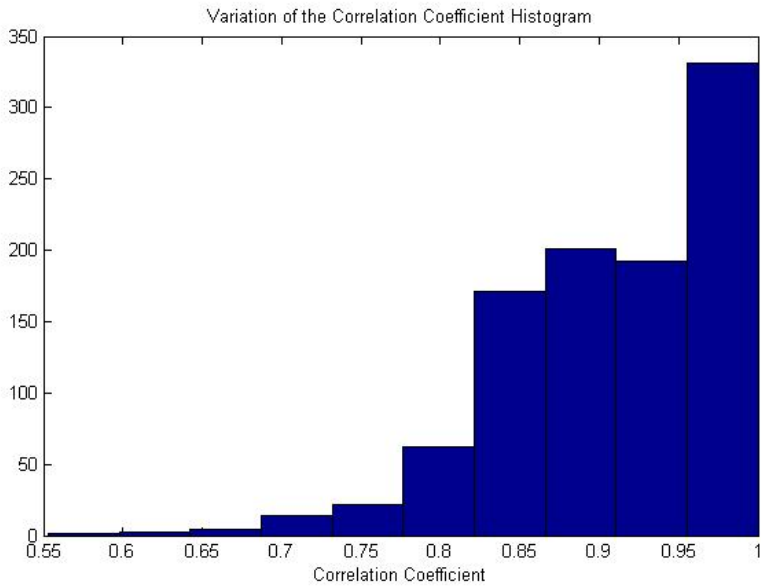


Figure V.12: Variation of the Correlation Coefficient

Chapter VI

Summary and Conclusion

VI.1 Conclusion

This document shows that it is possible to selectively detect volatile organic chemicals using of the shelf metal oxide sensor arrays available in the industry. Various tested markers prove to have unique response characteristics that can be used for identification (Figure V.3). These unique responses can be used to selectively identify VOC markers using computational methods. Preprocessing of sensor signal output using principal component analysis resulted in less than desirable two dimensional representation of the data (Figure IV.5). Use of sensor signal ratio for two dimensional mapping of data produced distinct clusters for various volatile organic species (Figure IV.7). It is highly interesting that a certain ratio produces such distinct clustering for each of the markers used (Figure IV.7, Figure IV.7). The underlying cause of such behavior is outside the scope of this research, but is of utmost importance for further understanding of such sensor systems.

An algorithm using mathematical definition of an ellipse (Equation IV.15) was used as a classifier of chemical markers. Classifiers were tuned using receiver operating characteristic analysis by varying the number of learning points and buffer applied to the ellipse (Figure IV.14, Figure IV.16, Figure IV.18, Figure IV.20 and Figure IV.22). Tuning the classifiers produced parameters with highest accuracy and lowest number of learning points for the sample dataset. Effects of environment temperature and concentration variability has shown effect on the system in spite of signal conditioning and use of sensor ratios for data mapping (Figure V.5, Figure V.6 and Figure V.8). Surface temperature has shown to have the largest effect on data samples (Figure V.7). Temperature increase causes an concentration rise in the air due to evaporation of VOC's. This concentration change is dependent on the vapor pressure of chemical used (Table V.1). These effects are highest in extreme ranges of temperature and concentration can be accounted for in future research or by specially selected markers that do not present overlapped responses through range of condition presented.

System testing using a test set of five chemicals with fifty overall exposures has produced a combined repeatability of 94%. Classifier's accuracies have been calculated to be at or above 98% (Equation V.2) for the data set used in the scope of this paper. Sensitivity

measures are calculated to be between 0.9 and 1 for all test markers (Table V.7). Bootstrap method with a thousand iterations has produced correlation coefficient average of around 90% with standard error for the estimated correlation coefficients of 0.0734 (Figure V.12). The use of commercially available mass produced sensors also contributes to the repeatability of the system due to low amount of variation between sensors produced in this environment. Finished product using the described approach can create a cheap effective way to selectively measure and detect various VOC compounds in the commercial and industry applications.

VI.2 Summary

The purpose of this dissertation was to develop a method for discriminant detection of volatile organic compounds in air. A system with such capability can help advance robotic navigation in environments where traditional methods such as electromagnetic waves are unavailable. The system has to be robust and economically feasible. Previous research into this subject yielded systems designed around large number of custom built sensors that don not provide the repeatability needed for this task. The scope of this dissertation is full system design of a device that can selectively detect various volatile organic compounds present in the environment.

A complete system including transducer level sensors, signal conditioning, data acquisition, and pattern classification algorithm was designed to observe and understand the unique signal signatures that are present with VOC compounds. System level control and communication software necessary for automated testing of the contaminants was developed and implemented first using LabVIEW then MATLAB environment. The system used PXI to PCI bus for communications between software and data acquisition hardware. The raw measurements were stream processed to determine the VOC and saved onto a hard drive for post processing. Software developed in MATLAB used real time hardware interrupts to sample the data for processing while a pattern classification algorithm made a decision based upon the model on which VOC compound it is exposed to. By creating a database of real data measurements the classifier can be tuned to provide the best results.

A suite of preprocessing procedures is implemented to facilitate extraction of information from data stream. The first step of preprocessing the data is to filter the incoming stream data to remove any unwanted noise. The next step requires an automated selection of window size to determine which window of data the classifier is running on and to remove any unwanted data from the dataset. A preset number of highest points are selected for each sensor to be used as the inputs for the classifier. The data is linearized based on the humidity and temperature of the environment. A ratio of the measurements is taken to eliminate the

effects of concentration and to map the data onto two dimensional space.

The custom designed algorithm is then trained using a learning data set. The parameters of the classifiers are adjusted based on the results of receiver operating characteristics curve to maximize the sensitivity and reduce fall-out. The trained classifiers are then able to selectively identify volatile organics in the environment.

VI.3 Contributions

Several contributions are made to the field of applications of computational tools and methods applied to volatile organic sensing by this research. It is the first of a kind system for selective detection of volatile organics in non-controlled environments. The approach to remove the effects of VOC concentration using specific sensor ratios is also first of a kind. In addition the classifier was custom designed with a data centric approach to minimize complexity and increase accuracy for easy implementation across micro controllers. Linearization of sensor data to account for humidity and temperature variation further improves the repeatability of the system. The computational methods used in this research resulted in the most effective approach for selective volatile organic detection.

The research work yielded a system for further understanding into methods for selective VOC detection. The work also has sparked high interest in the department of naval command for its application in autonomous robotic navigation on ships. All software written in this work is ready for use in further refinement of the concept presented.

VI.4 Limitation of the Research

However, the existing experimental setup is designed and built to perform best at a given setup, it has certain limitation as well, which are listed below.

i. Due to inherent sensor characteristics detection of multiple volatile organics simultaneously was not investigated.

ii. Sensors tend to foul and wear over time requiring recalibration or continuous adjustment as they age.

iii. To get adequate air movement across a sensor a fan was used to simulate air flow.

iv. Very high recognition accuracy can only be obtained with proper chemical and sensor selection.

v. System requires a certain amount of time in an area with high enough concentration to make a decision.

VI.5 Future Research

To optimize the system for use in real world environment testing has to be done on contaminant effects. For the system to have markers that have a predetermined evaporation time various chemicals have to be tested to produce a workable set of markers. More sensors with different doping can be added to the system to not only increase the selectivity but to gain detection accuracy of the classification algorithm. Some improvement to the algorithm can include the addition of a continuous output classifier to reduce the time needed for a measurement and user friendly interface for ease of data acquisition. Integrated system design will reduce the probability of hardware failure due to accidental contact of wires or sensors.

Appendix A

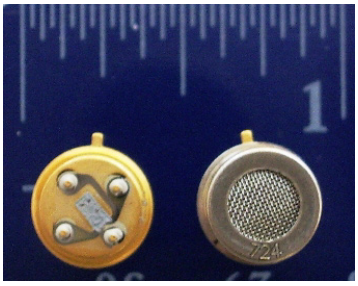
List of Sensors and its Specifications

SYNKERA TECHNOLOGIES, INC.

MikroKera 4L VOC Sensor (P/N 725)

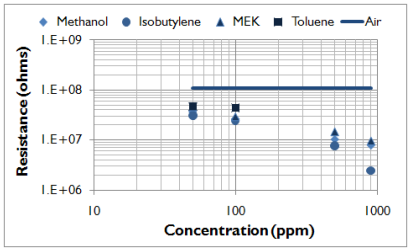
SENSOR FEATURES:

- Strong response to a wide range of VOCs
- Fast response time (T_{90} <15 seconds at 100 ppm ethanol)
- Environmental temperature range of –20 to 50°C
- Thermistor heater allows active control of sensor temperature based on environmental temperature
- Environmental humidity range of 0 to 95% RH, non-condensing

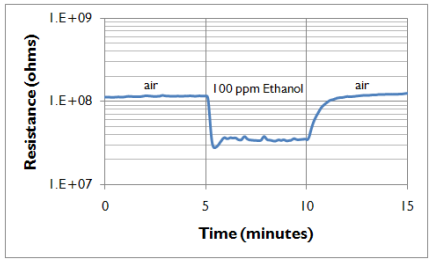


SENSOR RESPONSE CHARACTERISTICS:

The information below represents typical behavior for sensors operated in clean, dry gas.



Sensor resistance versus concentration for various VOCs.



Sensor response to 100 ppm ethanol in humid air. Ethanol applied at 5 min and removed at 10 min.

Cross sensitivity – ppm Isobutylene equivalents.

Vapor	ppm Isobutylene	Vapor	ppm Isobutylene
Methane – 1000 ppm	1	Nitrogen Dioxide – 5 ppm	negative response
Chlorine – 1 ppm	0	Sulfur Dioxide – 5 ppm	negative response
Hydrogen Sulfide – 15	940	Carbon Monoxide - 100 ppm	0

ELECTRICAL CHARACTERISTICS:

The properties below are typical for MikroKera 4L VOC Sensors. Circuits are available that are preset to the appropriate values.

Property	Symbol	Value	Remarks
Heater Power Consumption	P_H	~ 100 mW	Continuous at V_H = 1.25
Heater Voltage	V_H	1.25 VDC	T_{sensor} ~160°C
Heater Resistance	R_H	10Ω ± 0.5 Ω	At room temperature
Sensing Voltage	V_C	2.0 VDC	Recommended

SYNKERA TECHNOLOGIES, INC., 2605 Trade Centre Ave., Ste. C, Longmont, CO 80503
720-494-8401 e-mail: info@synkera.com www.synkera.com 720-494-8402 (fax)

- For information on warranty, please refer to Synkera Technologies, Inc. Standard Terms and Conditions.
- Information on this data sheet represents typical values. Actual values from sensor to sensor can vary slightly.

June 2010

e2v

MiCS-5135
VOC Sensor

This datasheet describes the use of the MiCS-5135 in VOC measurement applications. The package and the mode of operation illustrated in this document target the detection of reducing gases such as carbon monoxide (CO), hydrocarbons (HC), ethanol, and volatile organic compounds (VOC).

FEATURES

- Low heater current
- Wide detection range
- High sensitivity
- Fast thermal response
- Electrostatic discharge protected
- Miniature dimensions
- High resistance to shocks and vibrations

IMPORTANT PRECAUTIONS

Read the following instructions carefully before using the MiCS-5135 sensor described in this document to avoid erroneous readings and to prevent the device from permanent damage.

- The sensor must not be wave soldered without protection, or exposed to high concentrations of organic solvents, ammonia, or silicone vapours, to avoid poisoning the sensitive layer.
- Heating powers above the maximum rating of 120 mW can destroy the sensor due to overheating.
- This sensor is to be placed in a filtered package that protects it against any water or dust projection.
- For any additional questions, email enquiries@e2v.com or telephone +44 (0)1245 493493.

OPERATING MODE

The recommended mode of operation is a constant power mode. A heater power of P_H = 102 mW is applied. This causes the temperature of the sensing resistor (R_S) to reach about 450 °C.

Detection of the pollution gases is achieved by measuring the sensing resistor R_S during operation.

SENSOR RESPONSE

The sensor response to CO in air is represented in Fig. 1. The sensor resistance R_S is normalised to the resistance under air (R_0).

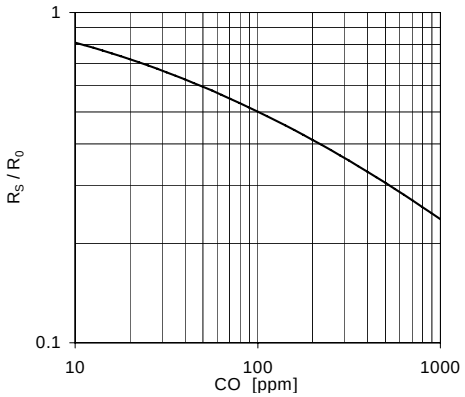


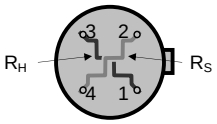
Fig. 1: R_S/R_0 as a function of gas concentration at 50% RH and 25 °C.

e2v technologies (uk) limited, Waterhouse Lane, Chelmsford, Essex CM1 2QU United Kingdom Telephone: +44 (0)1245 493493 Facsimile: +44 (0)1245 492492
e-mail: enquiries@e2v.com Internet: www.e2v.com Holding Company: e2v technologies plc
e2v technologies inc. 4 Westchester Plaza, PO Box 1482, Elmsford, NY10523-1482 USA Telephone: (914) 592-6050 Facsimile: (914) 592-5148 e-mail: enquiries@e2vtechnologies.us

MEASUREMENT CIRCUIT

Fig. 2 shows the pin connections of the MiCS-5135 gas sensor. A simple circuit to measure the pollution level is proposed in Fig. 3. The heating voltage V_H is applied to pins 3 and 1. A load resistor R_L is connected in series with R_S to convert the resistance R_S to a voltage V_S between pins 2 and 4. R_S can then be calculated by the following expression:

$$R_S = R_L / (V_{CC} - V_S) \times V_S$$



Pin	Connection
1	Heater ground
2	Sensor pin
3	Heater power
4	Sensor pin

Fig. 2: Equivalent circuit of MiCS-5135 (top view)

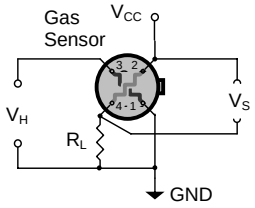


Fig. 3: Measurement circuit for pollution gas detection

ELECTRICAL SPECIFICATIONS

Maximum Ratings

Rating	Symbol	Value/ Range	Unit
Maximum sensor supply voltage	V_{CC}	5	V
Maximum heater power dissipation	P_H	120	mW
Maximum sensor power dissipation	P_S	1	mW
Relative humidity range	R_H	5 – 95	%RH
Ambient operating temperature	T_{amb}	-40 – 120	°C
Storage temperature range	T_{sto}	-40 – 120	°C
Storage humidity range	RH_{sto}	5 – 95	%RH

Operating Conditions

Parameter	Symbol	Typ	Min	Max	Unit
Heating power (see note 1)	P_H	102	85	120	mW
Heating voltage	V_H	3.2	-	-	V
Heating current	I_H	32	-	-	mA
Heating resistance (see note 2)	R_H	97	85	110	Ω

Sensitivity Characteristics

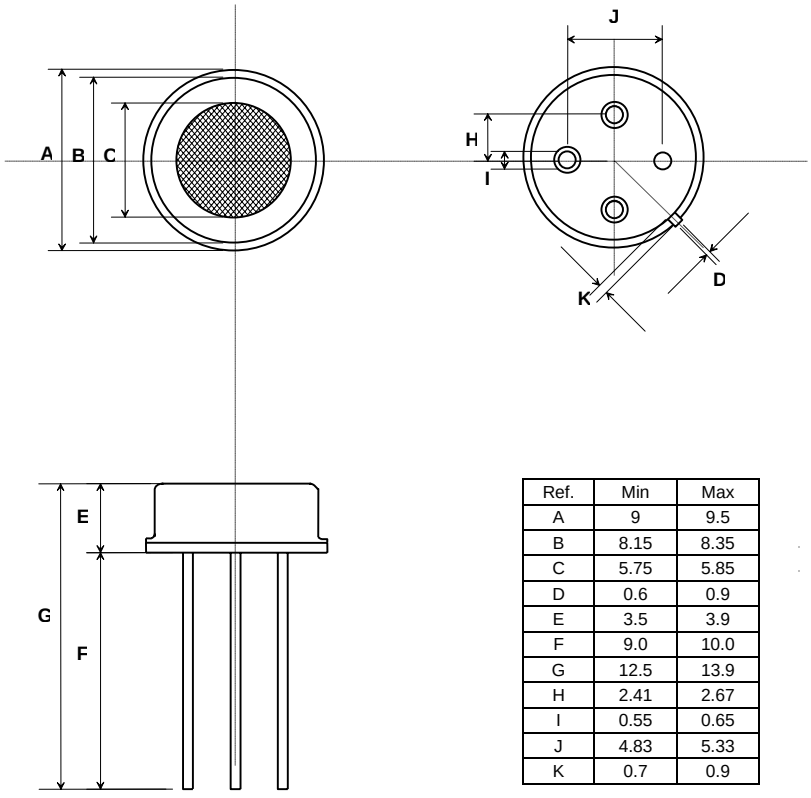
Characteristic	Symbol	Typ	Min	Max	Unit
CO detection range	FS		10	1000	ppm
Sensing resistance in air (see note 3)	R_0	100	20	400	kΩ
Sensitivity factor (see note 4)	S_R	2.2	1.5	3.0	-

Notes:

1. A power of 85 mW might provide sufficient sensitivity to certain gases. Heating powers above 120 mW can cause permanent damage to the sensor when ambient temperatures exceed 120 °C.
2. Heating resistor values from sensors out of production range between 85 and 110 Ω. Due to material properties of the heating resistor, its value increases during operating life. This behaviour has to be taken into account in the application design.
3. Sensing resistance in air (R_0) is measured under ambient air at 23 ± 5 °C and 50 ± 10%RH. These values are representative of most sensors, but some sensors could present R_0 from 1 kΩ to 1 MΩ.
4. Sensitivity factor (S_R) is defined as R_S at 60 ppm of CO. Test conditions are 23 ± 5 °C and 50 ± 10% RH. The S_R values are indicative values only.

PACKAGE AND FILTER OUTLINE

(All dimensions nominal and in millimetres)



Ref.	Min	Max
A	9	9.5
B	8.15	8.35
C	5.75	5.85
D	0.6	0.9
E	3.5	3.9
F	9.0	10.0
G	12.5	13.9
H	2.41	2.67
I	0.55	0.65
J	4.83	5.33
K	0.7	0.9

Outline Note:

A perfect pin alignment is not guaranteed.

e2v semiconductor gas sensors are well suited for leak detection and applications requiring limited accuracy. Their use for absolute gas concentration detection is more complicated because they typically require temperature compensation, calibration, and sometimes as well, humidity compensation. Their base resistance in clean air and their sensitivity can vary overtime depending on the environment they are in. This effect must be taken into account for any application development (311-4.0).

FIGARO

PRODUCT INFORMATION

TGS 2620 - for the detection of Solvent Vapors

Features:

* Low power consumption
* High sensitivity to alcohol and organic solvent vapors
* Long life and low cost
* Uses simple electrical circuit

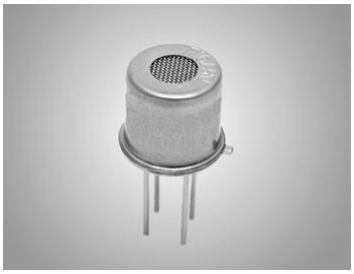
Applications:

* Alcohol testers
* Organic vapor detectors/alarms
* Solvent detectors for factories, dry cleaners, and semiconductor industries

The sensing element is comprised of a metal oxide semiconductor layer formed on an alumina substrate of a sensing chip together with an integrated heater. In the presence of a detectable gas, the sensor's conductivity increases depending on the gas concentration in the air. A simple electrical circuit can convert the change in conductivity to an output signal which corresponds to the gas concentration.

The **TGS 2620** has high sensitivity to the vapors of organic solvents as well as other volatile vapors. It also has sensitivity to a variety of combustible gases such as carbon monoxide, making it a good general purpose sensor.

Due to miniaturization of the sensing chip, TGS 2620 requires a heater current of only 42mA and the device is housed in a standard TO-5 package.

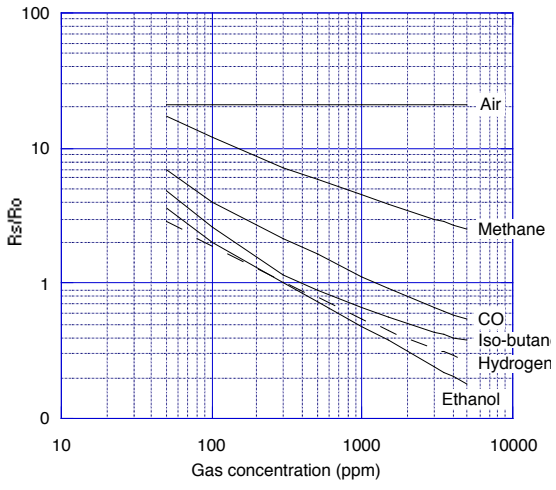


The figure below represents typical sensitivity characteristics, all data having been gathered at standard test conditions (see reverse side of this sheet). The Y-axis is indicated as sensor resistance ratio (Rs/Ro) which is defined as follows:

Rs = Sensor resistance in displayed gases at various concentrations

Ro = Sensor resistance in 300ppm of ethanol

Sensitivity Characteristics:

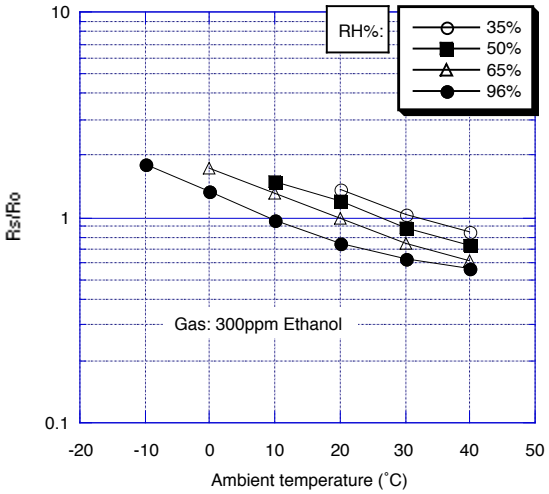


The figure below represents typical temperature and humidity dependency characteristics. Again, the Y-axis is indicated as sensor resistance ratio (Rs/Ro), defined as follows:

Rs = Sensor resistance in 300ppm of ethanol at various temperatures/humidities

Ro = Sensor resistance in 300ppm of ethanol at 20°C and 65% R.H.

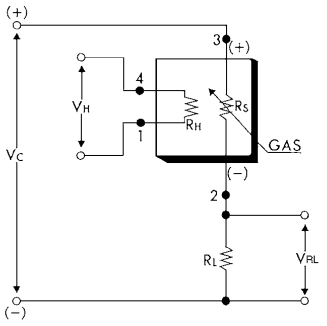
Temperature/Humidity Dependency:



Basic Measuring Circuit:

The sensor requires two voltage inputs: heater voltage (V_H) and circuit voltage (V_C). The heater voltage (V_H) is applied to the integrated heater in order to maintain the sensing element at a specific temperature which is optimal for sensing. Circuit voltage (V_C) is applied to allow measurement of voltage (V_{RL}) across a load resistor (R_L) which is connected in series with the sensor.

A common power supply circuit can be used for both V_C and V_H to fulfill the sensor's electrical requirements. The value of the load resistor (R_L) should be chosen to optimize the alarm threshold value, keeping power consumption (P_S) of the semiconductor below a limit of 15mW. Power consumption (P_S) will be highest when the value of Rs is equal to R_L on exposure to gas.



Specifications:

Model number			TGS 2620-C00	
Sensing element type			D1	
Standard package			TO-5 metal can	
Target gases			Alcohol, Solvent vapors	
Typical detection range			50 ~ 5,000 ppm	
Standard circuit conditions	Heater Voltage	V_H	5.0±0.2V DC/AC	
	Circuit voltage	V_C	5.0±0.2V DC/AC	Ps ≤ 15mW
	Load resistance	R_L	Variable	0.45kΩ min.
Electrical characteristics under standard test conditions	Heater resistance	R_H	83Ω at room temp. (typical)	
	Heater current	I_H	42 ± 4mA	
	Heater power consumption	P_H	approx. 210mW	
	Sensor resistance	Rs	1 ~ 5 kΩ in 300ppm ethanol	
	Sensitivity (change ratio of Rs)		0.3 ~ 0.5	Rs (300ppm) / Rs (50ppm)
Standard test conditions	Test gas conditions		Ethanol vapor in air at 20±2°C, 65±5%RH	
	Circuit conditions		V_C = 5.0±0.01V DC V_H = 5.0±0.05V DC	
	Conditioning period before test		7 days	

The value of power dissipation (P_S) can be calculated by utilizing the following formula:

$$P_S = \frac{(V_C - V_{RL})^2}{R_S}$$

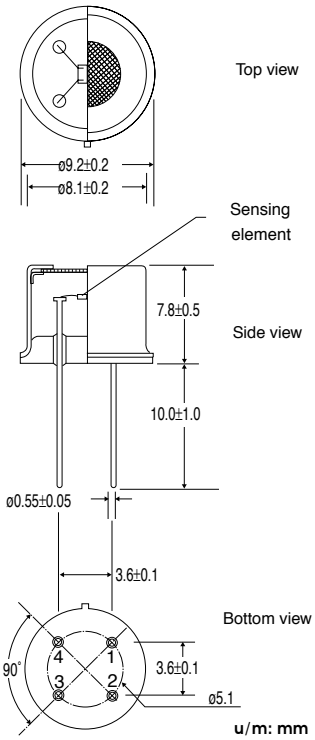
Sensor resistance (Rs) is calculated with a measured value of V_{RL} by using the following formula:

$$R_S = \frac{V_C - V_{RL}}{V_{RL}} \times R_L$$

For information on warranty, please refer to Standard Terms and Conditions of Sale of Figaro USA Inc.

REV: 01/05

Structure and Dimensions:



Pin connection:
1: Heater
2: Sensor electrode (-)
3: Sensor electrode (+)
4: Heater

FIGARO USA, INC.
121 S. Wilke Rd. Suite 300
Arlington Heights, Illinois 60005
Phone: (847)-832-1701
Fax: (847)-832-1705
email: figarousa@figarosensor.com

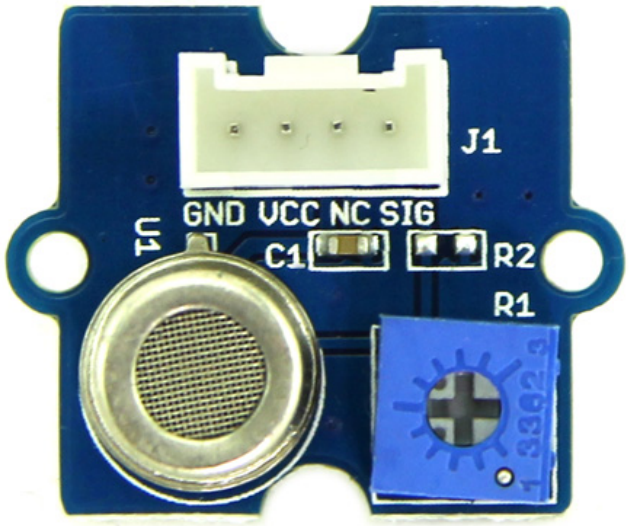
Grove - HCHO Sensor

From Wiki

Contents

Introduction

The Grove - HCHO Sensor is a semiconductor VOC gas sensor. Its design is based on WSP2110 whose conductivity changes with the concentration of VOC gas in air. Through the circuit, the conductivity can be converted to output signal that corresponding to the gas concentration. This sensor has a very high sensitivity and stability, it can detect the gas whose concentration is up to 1ppm. It's suitable for detecting formaldehyde, benzene, toluene and other volatile components. This product can be used to detect harmful gas in the home environment. Therefore, it's a good assistant for you to improve indoor environment quality of life.
Model:SEN01500P (http://www.seeedstudio.com/depot/grove-hcho-sensor-p-1593.html)

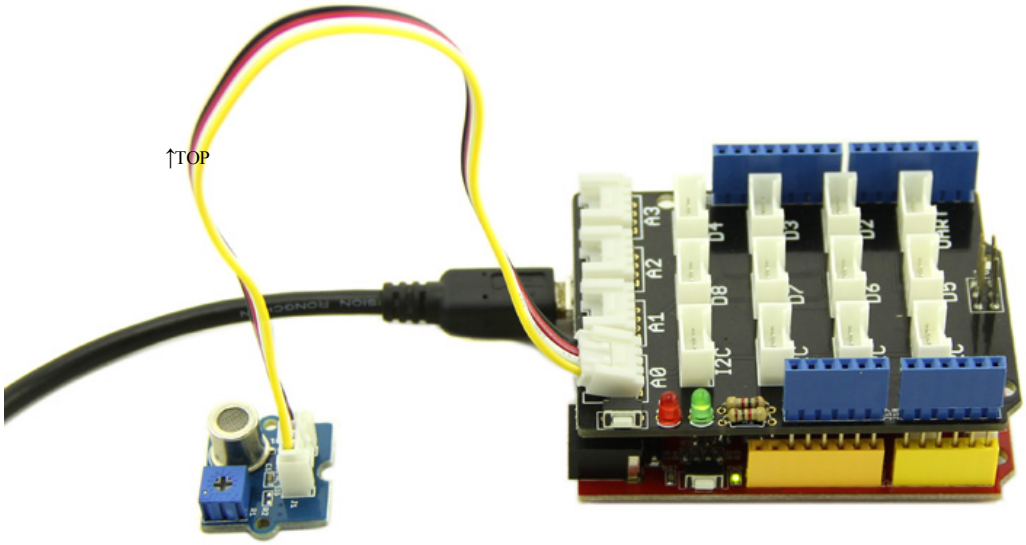


Specification

Operating Voltage: 5.0V ± 0.3V
Target Gases: HCHO, Benzene,Toluene,Alcohol
Concentration Range: 1~50 ppm
Sensor Resistance Value(Rs): 10KΩ-100KΩ(in 10ppm HCHO)
Sensitivity: Rs(in air)/Rs(10ppm HCHO)≥5

Demonstration

The Grove - HCHO Sensor can be used to detect VOCs, such as HCHO,toluene, benzene, alcohol. Here we take alcohol for an example to demonstrate how to use this sensor.



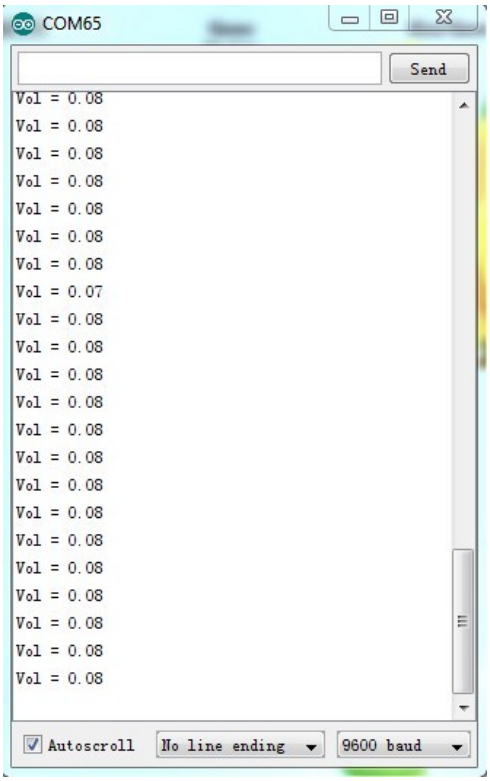

```
// demo of Grove - HCHO Sensor

void setup()
{
    Serial.begin(9600);
}

void loop()
{
    int sensorValue=analogRead(A0);
    float Vol=sensorValue*4.95/1023;
    Serial.print("Vol = ");
```

```
    Serial.println(Vol);
    delay(500);
}
```

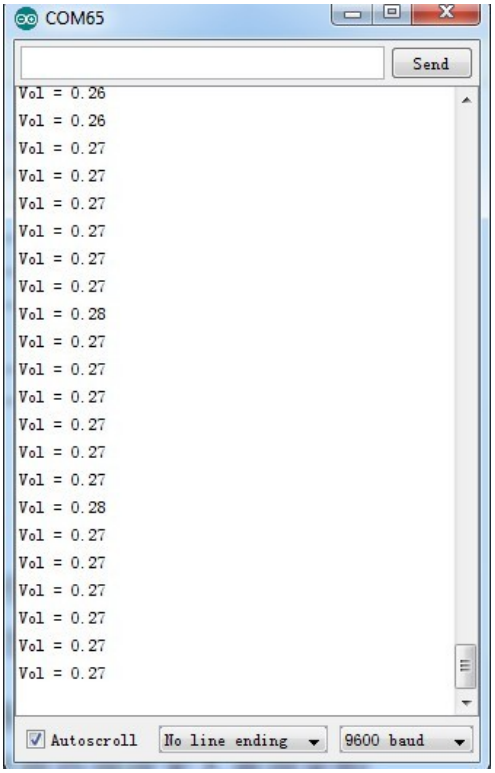
After uploading the code, open the serial monitor to get the voltage(Vol) under normal condition.



Now list out the formula describing the relationship of Vol and R0:

R0=(Vc/Vol-1)×R1 (Vc=4.95V) ①

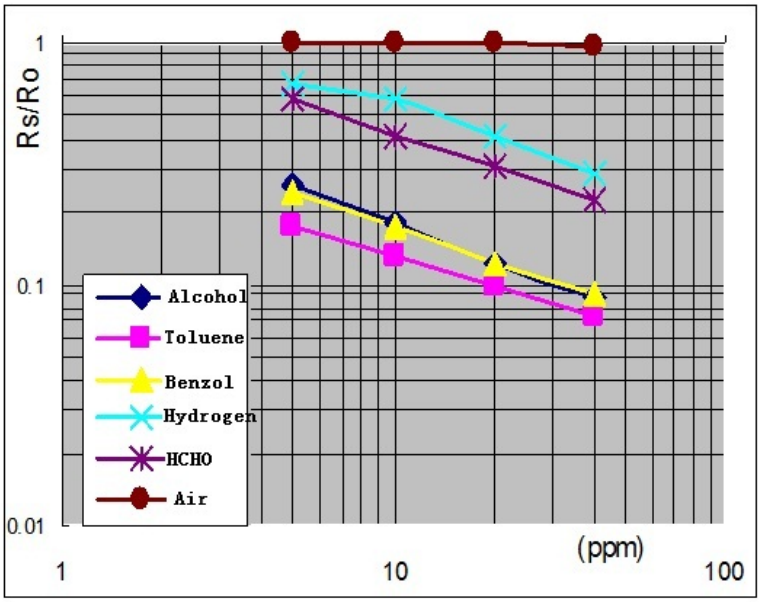
Then put a bottle of alcohol near the sensor, and read again the sensor value:



And we get the Rs:

$$Rs = (Vc/Vol-1)\times R1 \ (Vc=4.95V) \ ②$$

Now calculate Rs/R0. Here we get 0.285. Then refer to the sensitivity characteristic diagram below and find the alcohol concentration is about 5 ppm.



Resources

Grove - HCHO Sensor Eagle File (http://www.seeedstudio.com/wiki/images/f/ff/Grove_-_HCHO_Sensor_Eagle_File.zip)

Support

If you have questions or other better design ideas, you can go to our forum (http://www.seeedstudio.com/forum) or wish (http://wish.seeedstudio.com) to discuss.

Retrieved from "http://www.seeedstudio.com/wiki/index.php?title=Grove_-_HCHO_Sensor&oldid=38408"

BIBLIOGRAPHY

[1] Air Force Cambridge Research Laboratories (AFCRL). The hitran database. Web. 8 October 2014. <http://www.cfa.harvard.edu/hitran/>.

[2] Jeremy Davalle. Gas detection through the ages. Web. 20 August 2014. <http://ehstoday.com/industrial_hygiene/instrumentation/gas-detection-ages-0501>.

[3] H. Davy. On the fire-damp of coal mines, and on methods of lighting the mines so as to prevent its explosion. *Philosophical Transactions of the Royal Society of London*, 106:1–22, 1816. Web. 25 August 2014. <http://rstl.royalsocietypublishing.org/content/106/1.1.short>.

[4] Reimundo Deveza, David Thiel, Andrew Russell, and Alan Mackay-Sim. Odor sensing for robot guidance. *The International Journal of Robotics Research*, 13(3):232–239, 1994.

[5] A.M. Doncaster and T.D. Brown. Gas sensors, April 15 2003. Web. 5 September 2014. <http://www.google.com/patents/US6548024>.

[6] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern classification*. Pattern Classification and Scene Analysis: Pattern Classification. Wiley, 2001.

[7] e2v. *MiCS-5135*. Web. 5 September 2014. <http://datasheet.octopart.com/MICS-5135-SGX-Sensortech-datasheet-8617098.pdf>.

[8] Elefantasy. *Grove - HCHO Sensor*. Seed Wiki. Web. 20 August 2014. <http://www.seeedstudio.com/wiki/index.php?title=Grove_−_HCHO_Sensor&oldid=38408>.

[9] Tom Fawcett. An introduction to {ROC} analysis. *Pattern Recognition Letters*, 27(8):861 – 874, 2006. {ROC} Analysis in Pattern Recognition.

[10] Figaro. *TGS 2620 - for detection of Solvent Vapors*. Web. 5 September 2014. <http://www.figarosensor.com/products/2620pdf.pdf>.

[11] Minyi Guo. *Parallel and distributed processing and applications international symposium, ISPA 2003, Aizu-Wakamatsu, Japan, July 2-4, 2003 : proceedings*. Springer, Berlin New York, 2003.

[12] C. Hagleitner, D. Lange, A Hierlemann, O. Brand, and H. Baltes. Cmos single-chip gas detection system comprising capacitive, calorimetric and mass-sensitive microsensors. *Solid-State Circuits, IEEE Journal of*, 37(12):1867–1878, Dec 2002.

[13] Hahn. *Metal oxide nanostructures and their applications*. American Scientific Publishers, Stevenson Ranch, Calif, 2010.

[14] Jiawei Han. *Data mining : concepts and techniques*. Elsevier/Morgan Kaufmann, Amsterdam Boston, 2012.

[15] MC Horrillo, J Getino, L Ares, JI Robia, I Sayago, and FJ Gutierrez. Measurements of vocs with a semiconductor electronic nose. *Journal of the Electrochemical Society*, 145(7):2486–2489, 1998.

[16] B.R. Hunt, R.L. Lipsman, and J. Rosenberg. *A Guide to MATLAB: For Beginners and Experienced Users*. Cambridge University Press, 2006. Web. 20 October 2014. <http://books.google.com/books?id=ijCKngEACAAJ>.

[17] National Instruments. Ni pxi-6251. Web. 21 August 2014. <http://sine.ni.com/nips/cds/view/p/lang/en/nid/14125>.

[18] Sofian M. Kanan, Oussama M. El-Kadri, Imad A. Abu-Yousef, and Marsha C. Kanan. Semiconducting metal oxide based sensors for selective gas pollutant detection. *Sensors*, 9(10):8158–8196, 2009.

[19] Sofian M. Kanan, Oussama M. El-Kadri, Imad A. Abu-Yousef, and Marsha C. Kanan. Semiconducting metal oxide based sensors for selective gas pollutant detection. *Sensors*, 9(10):8158–8196, 2009.

[20] Shonda Kuiper and Jeff Sklar. *Practicing Statistics: Guided Investigations for the Second Course*. Pearson, 2012.

[21] Dae-Sik Lee, Jong-Kyong Jung, Jun-Woo Lim, Jeung-Soo Huh, and Duk-Dong Lee. Recognition of volatile organic compounds using sno2 sensor array and pattern recognition analysis. *Sensors and Actuators B: Chemical*, 77(1 - 2):228 – 236, 2001. Proceeding of the Eighth International Meeting on Chemical Sensors IMCS-8 - Part 2.

[22] Xiao Liu, Sitian Cheng, Hong Liu, Sha Hu, Daqiang Zhang, and Huansheng Ning. A survey on gas sensing technology. *Sensors*, 12(7):9635–9665, 2012.

[23] Charles F. Van Loan. Using the ellipse to fit and enclose data points. Web. 20 August 2014. <http://www.cs.cornell.edu/cv/OtherPdf/Ellipse.pdf>.

[24] Inc RKI Instruments. Gas detection history. Web. 10 September 2014. <http://www.rkiinstruments.com/history/index.htm>.

[25] Steven Smith. *The scientist and engineer's guide to digital signal processing*. California Technical Pub, San Diego, Calif, 1997.

[26] Kent A. Spackman. Signal detection theory: Valuable tools for evaluating inductive learning. In *Proceedings of the Sixth International Workshop on Machine Learning*, pages 160–163, San Francisco, CA, USA, 1989. Morgan Kaufmann Publishers Inc.

[27] Sargur Srihari. Introduction to pattern classification. Department of Computer Science, The University of Hong Kong, 2009. Web. 20 August 2014. <http://i.cs.hku.hk/ kpchan/CSIS8502/Notes/1.introduction.pdf>.

[28] INC Synkera Technologies. *MikroKera 4L VOC Sensor (P/N 725)*. Web. 12 September 2014. <http://www.synkerainc.com/images/pdf/Synkera_DataSheet_VOC.pdf>.

[29] P.N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Pearson Education, Limited, 2013.

[30] T.T. Thai, Li Yang, G.R. DeJean, and M.M. Tentzeris. Nanotechnology enables wireless gas sensing. *Microwave Magazine, IEEE*, 12(4):84–95, June 2011.