

Human Identification with their VOC Distribution Through CMS – SEN Model

Prathyusha Kanakam (✉ prathyusha.kanakam@gmail.com)

JNTUK: Jawaharlal Nehru Technological University Kakinada

ASN Chakravarthy

JNTU Kakinada: Jawaharlal Nehru Technological University Kakinada

Research Article

Keywords: Odor printing, Smell printing, VOC, Bio-analysis, Cogno Monitoring System, SEN Model.

Posted Date: June 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-655909/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Human Identification with their VOC distribution through CMS – SEN Model

Prathyusha Kanakam*, Research Scholar, Department of Computer Science and Engineering, JNTUK University, Kakinada, Andhra Pradesh – 533003, India.

ASN Chakravarthy, Department of Computer Science and Engineering, JNTUK University College of Engineering Vizianagaram, Vizianagaram, Andhra Pradesh – 535003, India.

*Corresponding author Mailid : prathyusha.kanakam@gmail.com

Abstract

Smell printing or odor printing is a novel morphological characteristic that an object can be defined by its odor. Human body odor is one such biological trait that yields less error rate of 15% among other biometrics. The human odor printing or smell printing possesses significance against the world towards screening of security checkpoint, searching for survivals under rubbles, investigating criminals, and many more. Cogno-monitoring system (CMS) is a specific prototype to furnish two essential processes -odor analysis and odor encoding through the Sensing-Encoding-Notifying (SEN) model to give the sensitivity and specificity score among the individuals. Human body odor can be interpreted as the alliance of various volatile organic compounds (VOCs) and they are recognized, classified in the encoding process. This article exhibits a detailed analysis of the traditional detection methods including bioanalysis concerning the human body odor experimented with 6 people. By applying principal component analysis along with random forest classifier, the VOCs distribution of the individuals is measured. This work calculates that 18.7% of VOCs are having a match with all the individuals which become the plinth for the identification of humans.

Keywords

Odor printing, Smell printing, VOC, Bio-analysis, Cogno Monitoring System, SEN Model.

1. Introduction

Odor printing or smell printing elongated its direction towards developing the olfactory detection system to recognize the aroma of an individual which will be differentiated from others and persists as one's unique biological property. Research is carrying out over the last five decades on these types of systems to make a machine learn about the identification of human odor. They strongly perceived as Electronic Noses or E-Noses, which are assisting for several applications for inspecting agriculture [1], measuring air quality index[2,3], environmental monitoring [4], food quality assessment[5,6], medical diagnosis[7] and many more.

E-Noses have extended its applicative scenario towards the detection and recognition of humans by their redolence. Authors employed a collection of metal oxide sensors in a networked E-Nose [8] to detect a human odor gathered from armpit, which is a primary source of odorants. To achieve patterns and discernment, Principal Component Analysis (PCA) is applied with a proposed scheme to correct the sensor drift. Likewise, equipment of E-nose [9] has been designed to measure the sensitivity of each metal oxide sensor by using a voltage divider resistor to classify and identify volatile organic compounds from human armpit body odor, which is controlled by in-house developed software for fixing humidity noise over a compact USB data acquisition card with PCA.

VOC analyzed through sweat that is concealed from the surface of the skin of the patient employed to detect, diagnose and monitor disease through an e-nose [10]. To distinguish the patient groups, PCA with discriminant function analysis is applied to achieve 87% accuracy. VOCs sources include breath samples used to diagnose illness and identifying pathogens from culture through an e-nose [11] used for analyzing biological oxidation of organic compounds which is called cell metabolism.

2. Bio-analysis of human body odor:

A human body odor Bio-analysis comprises the absolute study of odor that provides a characterization of each chemical compound of the mixture which will be evaporated at room temperature. According to the research [12], Bio-analysis of odor is categorized as sweat (produced from the skin odor), odor released from human excreta (e.g.,

urine) and the odor released from the oral cavity (e.g., breathing). It also depends on various factors based on diet of an individual, pressure, immune status and their genetics may allow the discharge of the complex group of non-volatile and volatile molecules by humans. Human body odor is secreted through apocrine, eccrine, sebaceous glands and produces volatile organic compounds. These volatile organic compounds deliver primitive sense in human body odor which brings the uniqueness of an individual. Basic human body odor possesses low or moderate vapor pressure with the chemically stable feature. The epidemiological studies review that body odor of individual influences, the physiological features like heart beat rate, blood pressure in certain scenarios, respiratory condition rate. Skin acts as conductance and causes significantly impact physiological outcomes like irritant symptoms, mood, and cognition (during task performance). The behavioral information of individuals can be used to inquire about their health and emotional status that prove many kinds of research regarding human discrimination by their odor. The odor is a perfect mixture of the odorants or chemical compounds emitted from different parts of the body like armpit, feet, breath, and stomach. Body odor chemistry has revealed these facts in Table 1 [13]. Each scent is basically in two forms- 'smells like', 'feels like'. Among these different secretions of chemical compounds from multiple body parts, the one which secreted from armpit determines a unique pattern helps to identify and detection the individual in various scenarios. Out of all these 12 compounds, 8 compounds are distinct and 4 are of predominant nature of their existing over different parts of the body. The volatile sulphur composites are detected by a human nose with a threshold of 0.00047 parts per million. whereas utmost, the human scent is dominated by volatile organic compounds. It is observed that the source of these VOCs is from the axillae of individuals. Due to the environmental conditions, internal body conditions an individual body odor may change from time to time.

Table 1. Compounds and its nature based on the Odor

| Type of odor | Location | Compounds | Smell-like |
|---------------|-----------------|-----------------------------------|---------------------------|
| Underarm odor | Armpit sweat | (E)-3-Methyl-2-Hexenoic acid | Goat |
| | | 3-Hydroxy-3-methylhexanoic acid | Onion |
| | | (S)-3-Methyl-3-SulfanylHexan-1-ol | Cumin |
| Halitosis | Hair Sweat | Methanethiol | Sulfur, garlic |
| | | Hydrogen sulphide | Sulfur rotten-eggs |
| | | Dimethyl sulphide | Cabbage , Sulfur, Sweet |
| Flatulence | Abdominal sweat | Hydrogen sulphide | Sulfur rotten-eggs |
| | | Methanethiol | Sulfur, garlic |
| | | Methanethiol | Cabbage , Sulfur, Sweet |
| Foot odor | Foot | Methanethiol | Sulfur, garlic |
| | | Propanoic acid | Pungent, rancid, sour |
| | | Isovaleric acid | Cheesy, fermented, rancid |

However, it is still uncertain to explain the diversification of odorants with their peculiar characteristics. From the previous work of the authors [14], all the odorants that are volatile in nature will be evaporated at room temperature and weighed from 30 to 300 gm/moles depending on the distribution of molecules in a specific compound. It has been observed[15] that, (E)-3-Methyl-2-Hexenoic is highly ranked compound found in most of the individuals and also the odorants are unsaturated acids, alcohols, aldehydes, ketones, steroids and carbonyls that form as either straight or branched chain of C₆-C₁₁.

3. Analyzing and detecting techniques:

Various techniques are analyzed qualitatively (where the presence of the compound is detected among mixture), quantitatively (the quantity of the compounds are determined) and structurally (how the molecules are distributed among the mixture). Determining the fingerprint of the human body odor can be detected through different qualitative analytical techniques with their predominance in identifying chemical signatures of human body odor depending on the unique information of VOCs in body odor. The procedure begins with the odor sampling for separating VOC compounds in any given mixture. The probability of detection of compounds in a mixture relies on the sampling and pre-concentration processes of particular technologies [15]. Even though a single method of analyzing technique is enough to detect the molecules, the purpose of using the combination of techniques is to predict the unpredictable nature of VOCs in a human body odor.

3.1 Solid Phase Micro-Extraction (SPME):

This technology that produces the sample to be unobstructedly dissolvable operates in a quick, efficient and adaptable manner. SPME uses a fiber coated polymer or sorbent, where polymer utilised to manage liquid sample and sorbent

for handling solid sample or a mix of both. Most of the researches were used di-vinyl benzene (DVB)/ carboxen (CAR)/ poly dimethyl siloxene (PDMS), where DVB, CAR were used as solid sorbents and PDMS used as liquid sorbent.

Then the compounds in each liquid sample or solid sample are absorbed to their respective coating by the extraction of fiber. Then the SPME fiber is inserted into chromatograph for further investigation. The applications of this technique have extended to numerous domains to detect flavors, scents, forensics, toxicology, environmental and organic matrices.

3.2 Gas Chromatography (GC):

It's an analytic technique to separate the molecules which evaporates at room temperature. These molecules are well noted as volatile organic compounds as a mixture like methanol, acetone, and heptane. This system associated with a column, mobile phase and stationary phase components. Before analysis, the sample that is extracted can be passed to the column through an inlet where the temperature of the column is maintained uniformly and should be 50°C more than that of room temperature to achieve efficient detection of the chemical compounds. Separation of the molecules takes place between the mobile and stationary phases whereas all the light-weight and more weight molecules will interact with the stationary phase and mobile phase respectively and transfer quickly from the column. Once the separation is completed, the detector will be attached at the end to detect the compounds in a sample.

3.3 High-Performance Liquid Chromatography (HPLC):

It is a column modified analytic system performed at high pressure of 40mPa (megapascals) used to pump mobile phase molecules into the column with absorbent material coated stationary phase. The column is filled with a small size absorbent material particle. It gives a large surface area with high resolution to these molecules for the ease of interaction.

3.4 Mass Spectrometry:

The analytic technique in which the molecules are converted to ions based on their mass to charge ratio. It deals with production, acceleration, separation and detection ions by using ionization chamber, vacuum pump and analyzer to calculate mass to charge ratio. The pump passes the sample to the ionization chamber with a pressure of 1 μ torr, where the sample of molecules get separated as ions with either filament functionality or Ammonia/methane ions and inert gases (Argon/Helium/Xenon) interactions. These ions are analyzed through a mass analyzer where mass to charge ratio is analyzed and detected through a detector.

The entire process of a qualitative analytic task depends primarily on extraction, analyzing, detecting of the compounds in the mixture. The Table 2 lists various techniques involved in different sub-tasks as mentioned. For each analytical method, compounds and their type are defined.

Table 2: Description of analyzing and detecting techniques

| | | |
|--------------------|--|--|
| Extraction methods | Solid Phase Extraction (SPE) | |
| | Supercritical Fluid Extraction (SFE) | |
| | Liquid-Liquid Extraction (LLE) | |
| | Solid Phase Micro Extraction (SPME) | |
| Analytical Methods | Gas Chromatography(GC) | Gases Volatile Compounds |
| | Gas Chromatography and Mass Spectroscopy (GC/MS) | Complex Structures, Un-known Compounds |
| | Liquid Chromatography and Mass Spectroscopy (LC/MS), High-Performance Liquid Chromatography (HPLC) | Non-volatile organic compounds |
| | Atomic Absorption Sampling (AAS)/ Inductively Coupled Plasma (ICP) | Toxic metals |
| Detectors | GC Detectors | |
| | Mass Spectroscopy | |
| | Optical Spectroscopy | |
| | Electrochemical detectors | |

4. Cogno-Monitoring System (CMS):

Currently, the pattern recognition system and sensor arrays cloning the human senses is the potential advancement in the technical and commercialization direction. Research is in an on-going process to develop technologies for detecting, recognizing the primary compounds in numerous odors and flavors. The established receptor systems human olfaction, canine olfaction comprises several steps of recognition process includes identifying, comparing, quantifying with data storage and retrieval process. Moreover, knowing that it is relevant to individual opinion, hedonic perception is a characteristic of the human nose. Before-mentioned instruments have undergone substantial development and have not met industrial needs at present.

CMS is an artificial sensing system used for chemical fingerprinting or odor printing where the specific chemicals are identified from a complex compound. These chemicals are specialized as monomolecular compounds that either odor dominated by strong principal component or gives a mixed aroma by the combination of all the odors. Authors in their previous work described [14][16] the traditional receptor systems that imitate olfaction process to project four distinct methods- pre-processing followed by feature extraction, classification and ultimately, decision-making process consolidated into SEN methodology.

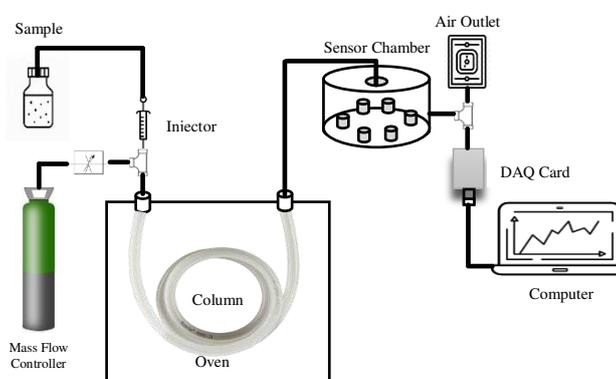


Fig 1. Lab view of Cogno Monitoring System

CMS is the base Electronic Nose (E-Nose) prototype that implemented based on two essential units- Sensing-Encoding, Pattern Recognition Schemas as shown in Fig. 1. It is the qualitative analytic system based on the quantity of the compound that includes sample extraction, analyzing and detection of chemical compounds in human body odor. These VOC from human body odor is sampled and injected through the electric valves where the analytical sample preparation is carried out. To perform any extraction technique, two phases which lead to sample (a mixture of VOC from human body odor) and extractant (analyte used for separation of compounds). The primary objective of sample extraction is the isolation of the analyte of interest, the cleaning of the sample, and the pre-concentration of the analyte of interest. The mass flow controller resembles the mobile phase in GC where it holds the inert gas collection to carry the samples and connected to the column where the uniform temperature is maintained from 150°C to 300°C using an oven to isolate the compounds from sample. Sensor chamber contains an array of sensors so that each isolated compound is distinguished by its respective sensor. The information related to the odor compounds are accumulated into Data Acquisition (DAQ) card that acts as an analyser. As the smell printing deals with organic compounds which are more volatile in nature may vary their distribution. In the pattern recognition unit, feature extraction, classification and decision making is carried out as the sub-tasks. The feature extraction deals with the process of qualitative analysis of compounds in a mixture that coincides with the service of the detector in GC. All the units are integrated to printed circuitry board with power supply unit.

4.1 Sensing-Encoding-Notifying (SEN) model:

It furnishes the standardized approach to bring out step by step tasks described in the CMS. The pivotal philosophy of SEN model lies in projecting the overlapping sensitivities of traditional olfaction systems that contain olfactory epithelium and olfactory receptors that relate to the response of chemical compounds after pre-processing, sensing and encoding process which is acquired from a sensor chamber and analyzed and stored the information in DAQ card. This SEN Methodology includes three basic processes includes sensing, encoding, and notification. SEN

algorithm indicates the executable script to be performed step by step to operate 3 primary functions- Sensing in which analyte selection, sampling, filtering, preconditioning of column is executed, whereas Encoding deals with ions distribution and their recognition, then concluded with notifying the results obtained. Features extracted are identified based on their peak area or peak width which directly points to the concentration of the compound. There are two parameters need to be measured after completion of the subtasks. During the sensing phase, sensitivity is attained as output and during the encoding phase, specificity is attained as output. The detailed procedure will be explained in the experimental section.

Algorithm for SEN Model:

Start

Step-I: Sensing of the chemical compounds found in human body odor.

Step-1: Selection of analyte material for the interactions with the sample in a given mixture of VOC.

Step-2: Sampling can be done through various injection methods into electric values.

Step-3: Filtering with the mass flow controller of the carrier to make the separation of compounds.

Step-4: Precondition the temperature of the column greater than room temperature in the sensor chamber.

Parameter: sensitivity

Definition 1: Sensitivity (S_n) is the absolute quantity which gives the smallest absolute amount of change of concentration of a chemical compound of an individual that can be detected through CMS.

$$S_n = \ll |\Delta(Q_{ai+1}, Q_{ai})|$$

Where ' Q ' is the odor threshold value, ranges between 1 to n i.e., $\{Q_{a1}, Q_{a2}, \dots, Q_{an}\}$ are the odor thresholds of a particular compound that were gathered by an individual 'a' examined different times

Step-II: Encoding of the chemical compounds found in human body odor.

Step-5: Releasing of ions through detectors collected through DAQ card.

Step-6: Feature extraction through DAQ card

$$W \propto C$$

Where w is the peak width of compound distribution observed in a detector and C is the concentration of the compound.

Step-7: comparison with memorized data (comparing the peak values concerning retention time of a compound)

Step-8: Feature Extraction with PCA and find covariance of the compounds among different individuals.

Parameter: specificity

Definition 2: Specificity (S_p) is the comparative measure of the magnitude of two output signal produced. It gives the ration of odor threshold of non-homogeneous compounds between any two individuals.

$$S_p = \frac{Q_{ca}}{Q_{cb}}$$

Where ' Q_c ' is the odor threshold value of a particular chemical compound concerning two individuals 'a' and 'b'

Step-III: Notifying the results obtained.

End

5. Experimental Discussion:

Most of the investigations employed the direct or indirect collection of a sample of human body odors [17] where direct collection involves the interactions of cotton/gauze pads with the skin (from a prime source of their body odor, armpit) [18] after continuous monitoring of volunteers throughout 5 to 7 consecutive days. During this treatment, they are instructed to control their ordinary life activities such as the use of cosmetics, sexual activities to obtain the qualitative form of sample. Indirect sample collection encompasses the method of collecting VOC by using flow rates of air sorbent and the design of the material used for the sample collection. As the compounds have the drift of evaporation at room temperature, either massive pumping or most limited pumping leads to inappropriate results. The sorbent role is to break the VOC sample into separate molecules but not to react with them to form new ones. The qualitative and quantitative measures employed in selecting sorbent or analyte applied in these injection procedures. Supercritical Fluid Extractor (SFE) was mostly used pre-treatment process found in several considerations. The consequent extraction and identification of compounds can be done through the Solid Phase Micro Extraction (SPME) method that uses either DVB/CAR or PDMS fibers [19]. The thermal sorbent Tenax/carbonograph tube is employed in another study [18] to extract VOCs which is collected through cotton pads by heating the sample up to 90°C. A Nalophan sampling bag is employed for the collection of samples of headspace SPME, contact SPME, liquid-liquid extraction and the other dynamic sorbent methods-Dynamic headspace sorbent tube sampling (DHS). Out of all these styles, SPME has a potent position in providing fast and efficient sample extraction. However, the lower amount of VOC is isolated efficiently by DHS [20, 21].

5.1 CMS Setup:

The GC setup has made to build CMS with the silicon tube column of 30 meter long with 0.25 millimeter diameter. Helium acts as the carrier gas for the sample mixture (mass flow controller) which is used in the mobile phase to transfer the sample within the column employing the stationary phase. The column is packed with silica gel to give biosensor prototype for separating and identifying the VOC from the mixture. The initial temperature of the room is at 25°C and 75 mm of Hg is considered as vapor pressure and temperature in the column should be 50°C more than that of room temperature. Sample collected is injected through a 5v diaphragm pump to isolate the compounds from the mixture. This experiment is taken among 6 individuals (4 males and 2 females of average age (20-40 yrs)). SEN model parameters (sensitivity and specificity) is calculated and graphically shown in fig. The observed and analyzed compounds are categorized as classes depend on their chemical formula. From the synthetic data taken, 80 compounds are observed totally among 6 individuals out of which 15 compounds are identified as frequent in all the volunteers and the groups of which the compounds are categorized in the Table 3. 5 aldehydes, 5 ketones, 3 non-carbon mixed compounds, 1 acid and 1 alcohol type. It is observed that 14 compounds are of group1 (c2-c10) unsaturated fatty acids and one is sulphur compound where the carbon is absent and are common in all the individuals. The sensitivity of male 4 is calculated and shown in Fig. (2a), whereas specificity ratio is drawn and shown in Fig. (2b), among the male1 and female 2 for all the Common Class VOCs compounds of the sample.

Table .3 Observed classes of dataset.

| Non-Common Classes | | Common Classes | | Observed Classes | |
|----------------------|----|----------------|----|----------------------|----|
| Acid | 3 | Other | 3 | Acid | 4 |
| Alcohol | 13 | Acid | 1 | Esters | 3 |
| Aldehyde | 15 | Alcohol | 1 | Alcohol | 14 |
| Aromatic-hydrocarbon | 22 | Aldehyde | 5 | Aldehyde | 20 |
| Esters | 3 | Ketone | 5 | Aromatic-hydrocarbon | 22 |
| Alkanes | 5 | | | Halogens | 1 |
| Halogens | 1 | | | Others | 3 |
| Ketones | 3 | | | Ketones | 8 |
| | | | | Alkanes | 5 |
| | 65 | | 15 | | 80 |

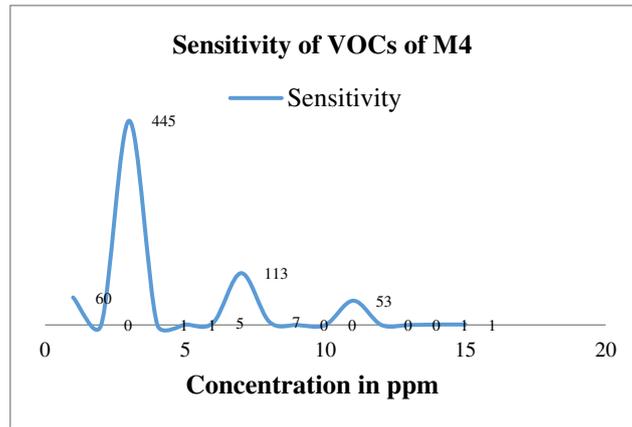


Fig. 2(a). Sensitivity of VOCs of M4

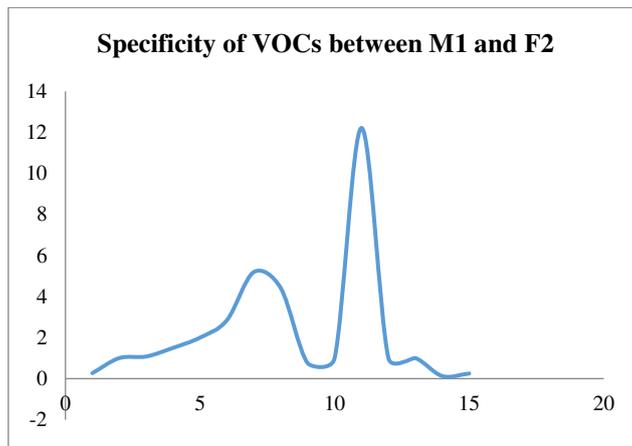


Fig. 2(b). Specificity of VOCs between M1 and F2

It is observed that 80 compounds listed in the Table 4 are analyzed from 6 individuals. All the compounds categorized as separate classes and out of which 15 compounds are common (highlighted) with all the individuals. The presence of the compound in their respective body odor tabulated.

5. Results:

According to the parameters specified by the authors, the synthetic dataset is taken and employed machine learning algorithms to prove the similarity in the body odor components as 25 % which is used for the application of identification of individuals under rubbles.

From the observations made, the peak height of the compound is considered as one of the parameters.

Table. 4 Distribution of VOC among individuals.

| Compound Details | M1 | M2 | M3 | M4 | F1 | F2 |
|--|----|----|----|----|----|----|
| 1-octanol | | √ | | √ | | √ |
| (E)-2-nonenal | √ | | √ | √ | | |
| (E)-2-octenal | √ | | √ | | √ | |
| beta.-Pinene | | √ | √ | | | √ |
| 1,2,4-Trimethylbenzene | √ | | √ | √ | | |
| 1,3-Diacetyloxypropan-2-yl acetate (triacetin) | | √ | | √ | | √ |
| 1,3-dimethyl-benzene | | √ | √ | | | √ |
| 1-Phenyl-ethanone | √ | | √ | | √ | |
| 2-(4-Tert-butylphenyl)propanal(p-tert-butyl dihydrocinnamaldehyde) | √ | | √ | √ | | |
| 2,2,4-Trimethyl-1,3-pentanediol diisobutyrate | | √ | | √ | | √ |
| 2,2-Dimethylbutane | | √ | √ | | | √ |
| 2,3 Butanedione | √ | √ | √ | √ | √ | √ |
| 2-Decanone | | √ | √ | | | √ |
| 2-Furancarboxaldehyde | √ | | √ | √ | | |
| 2-Furanmethanol | √ | | √ | | √ | |
| 2-methyl-2-butenal | | √ | | √ | √ | |
| 2-Pentanone | √ | √ | √ | √ | √ | √ |
| 2-Propenal | √ | √ | √ | √ | √ | √ |
| 3,4-Dimethyl-1-pentanol | | √ | | √ | √ | |
| 3,7-dimethyl-1,6-octadien-3-ol | √ | | √ | √ | | |
| 3,7-dimethyl-6-octen-1-ol | | √ | | √ | | √ |
| 3-Dodecene | √ | | √ | | √ | |

| | | | | | | |
|------------------------------------|---|---|---|---|---|---|
| 3-Hexene-2,5-diol | √ | | √ | √ | | |
| 4-Cyanocyclohexene | | √ | √ | | | √ |
| 4-methyl-3-Pentene-2-One | √ | √ | √ | √ | √ | √ |
| 5-Methyl-2-isopropyl cyclohexanol | √ | | √ | | √ | |
| 6,10-dimethyl-5,9-undecadien-2-one | | √ | √ | | √ | √ |
| Acetaldehyde (*) | √ | √ | √ | √ | √ | √ |
| Acetic acid | | √ | | √ | | √ |
| Acetone | √ | √ | √ | √ | √ | √ |
| Benzaldehyde | √ | √ | √ | √ | √ | √ |
| Benzene,1,2,3-trimethyl- | √ | | √ | √ | | |
| Benzene,1,3,5-trimethyl- | √ | | √ | | √ | |
| Benzene,1-ethyl-2-methyl- | | √ | √ | | | √ |
| Benzene,1-ethyl-3-methyl- | √ | | √ | √ | | |
| Benzyl alcohol | √ | | √ | | √ | |
| Benzylchloride | | √ | | √ | | √ |
| Butanal | √ | √ | √ | √ | √ | √ |
| Cyclotetradecane | | √ | √ | | | √ |
| Decanal | | √ | | √ | √ | |
| Decanoic acid | √ | | √ | √ | | |
| Decene | | | | | | |
| Dimethyl Sulfide | √ | √ | √ | √ | √ | √ |
| D-Limonene | √ | | √ | √ | | |
| Dodecanal | | √ | | √ | | √ |
| Dodecane | √ | √ | | √ | | √ |
| Dodecanoic Acid | √ | | √ | √ | | |
| E-2-decenal | √ | | √ | | √ | |
| Eicosane | | √ | | √ | | √ |
| Ethanol | √ | √ | √ | √ | √ | √ |
| Ethylene glycol | | √ | √ | | | √ |
| Heptadecane | | √ | | √ | √ | |
| Heptanal | √ | | √ | √ | | |
| Heptane | √ | | √ | | √ | |
| Hexadecane | | √ | | √ | | √ |
| Hexanal | √ | | √ | √ | | |
| Hydrogen Sulfide | √ | √ | √ | √ | √ | √ |
| Methanethiol | √ | √ | √ | √ | √ | √ |
| Methyl-2-Butanone | √ | √ | √ | √ | √ | √ |
| Naphthalene | | √ | | √ | √ | |
| Nonanal | √ | √ | √ | | √ | √ |
| Nonane | √ | | √ | √ | | |
| Nonanol | √ | | √ | | √ | |
| Octane | | √ | | √ | | √ |
| p-Cresol | √ | | √ | √ | | |
| Pentadecane | | √ | √ | | | √ |
| Phenol | √ | √ | | √ | √ | √ |
| Propanal | √ | √ | √ | √ | √ | √ |
| Propanoic acid | √ | √ | √ | √ | √ | √ |
| p-Xylene | | √ | | √ | √ | |
| Pyridine | √ | | √ | √ | | |
| Styrene | √ | | √ | | √ | |
| Tetradecanal | | √ | | √ | | √ |
| Tetradecane | √ | | √ | √ | | |
| Thiazolidine | | √ | √ | | | √ |
| Toluene | √ | | √ | | √ | |
| Tridecanal | √ | | √ | √ | | |
| Tridecane | | √ | | √ | | √ |
| Undecanal | √ | | √ | √ | | √ |
| Undecane | | √ | √ | | | √ |

Definition 3: Peak height (*Pk. Ht.*) is the ration of vapor pressure to the odor threshold. The peak height gives the altitude of the peak of the particular compound analyzed through the analyzer.

$$Pk. Ht. = \frac{V.P.}{O.Th.}$$

Where 'VP' is the vapor pressure of that compound at 25°C room temperature, the compound will be in the evaporated state and 'O.Th.' is the odor threshold or the odor concentration of the compound. Generally, under constant temperature and pressure factors the concentration of chemicals that resides in the air is measured in either milligrams/micrograms/nano-grams/pico-grams or parts per million (ppm)/parts per billion (ppb) depend on the molecular weight of the respective compound.

The pattern analysis is adopted to identify the odor components among different individuals to analyze the common components of the humans that served for recognizing the individuals under rubbles. The dimensionality reduction

feature of Principal Component Analysis (PCA) plays a critical role in reducing the non-correlated dimensions to correlated dimensions in a statistical way. PCA is a powerful technique that achieves less time and space complexities and stores the information in low memory through dimensionality reduction. Of all these functionalities PCA visualizes different classes of the compounds and their distribution transparently.

Algorithmic Logic:

Step-1: Importing the dataset ' $D_{m \times n}$ ' of 80 compounds with respective 5 parameters derived by authors in their previous work under different classes. Splitting of compounds based on the classes that the compounds belong to.

Step-2: Normalize non-correlated features into unit scale using 'Standard Scale' (where the mean=0 and variance =1), which is the optimal performance requirement. (' $Norm(D_{m \times n})$ ').

Step-3: Co-variance matrix (' $CoV(D_{m \times n})$ ') is formulated with the same dimensions along with standard scaling of non-correlated features for deriving eigen values '*Eg.Val.*' and eigen vectors '*Eg.Vec.*'.

Step-4: Project the data point onto selected eigen vectors. i.e., rearranging of co-variance matrix from maximum to minimum eigen values (' $CoV(D_{m \times n})_{max \rightarrow min}$ ').

Step-5: PCA result ('PC1 and PC2') is obtained by multiplying the two matrices (' $CoV(D_{m \times n})_{max \rightarrow min} \times Norm(D_{m \times n})$ ').

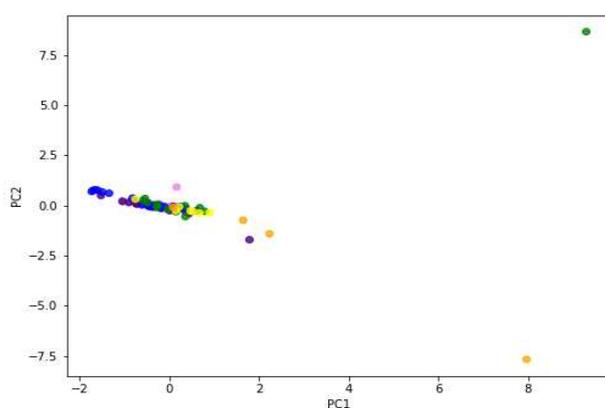


Fig. 3. Visualization of depending classes

By reducing the dimensions, no loss of data is remarked such that PC1 contains 57.3% variance and PC2 contributes 42.6% variance and together will contain 99.9% information is preserved. Fig 3 visualizes the PCA data depending on the classes through which the compound belongs to. The common and non-common compounds are separated and transformed to calculate the percentage of missing information. Both the sets retain 99.9% information as preserved and loss of data is negligible. Fig shows the PCA 2-D scattering for common ((a)) and non-common compounds ((b)).

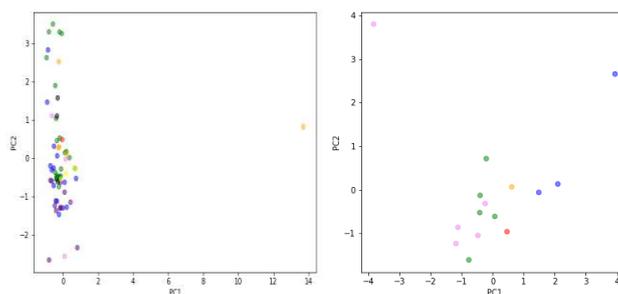


Figure 4(a): PCA 2D Scattering of non-common compounds

Figure 4(b): PCA 2D Scattering of common compounds

Unlike other classifiers, CMS used random forest classifier which is a meta estimator that usually fits several decision tree classifiers namely – hedonic tone in which the compounds are categorized based on the feel of pleasure of odor (pleasant/neutral/un-pleasant), compounds classified based on the organic group (group1(C2-C10)/group2(C11-C13)/group3(C14-C18)), the class parameter that sorts the compounds according to their chemical nature(Acid/Esters/Alcohol/Aldehyde/Aromatic-hydrocarbon/Halogens/Others/Ketones/alkanes). Depending on the rankings made by the classifier, it is observed that the score of the classifier with 72.5% accuracy. It is observed that 18.7% of the compounds resides in human body odor has matched. These compounds will help to identify humans under rubbles.

Conclusion:

Identification of humans under rubbles is one of the core tasks in the application of Smell printing or odor printing, which is a unique feature where an individual is identified by their body odor. CMS gives the resemblance of gas chromatograph setup for identifying an object which can be defined by its odor. As the human odor is the versatile combination of different VOCs, and each VOC gives different values for each of the odor parameters is defined by author. Depending on the synthetic data taken, PCA is applied to analyze the compounds statistically. As each of the odor compounds possesses different dimensionality, CMS achieves a negligible missing of data with 99.9% scattering of the compounds over 2D space. 72.5% of accurate predication score is given by random forest classifier. CMS derives that 18.7% matching style in the odorants of the individuals which is used to apply for the recognition of individuals under rubbles.

Author Contribution: All authors have equally contributed and all authors have read and agreed to the published version of manuscript.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Jansen, Roel, et al. "Automated signal processing applied to volatile-based inspection of greenhouse crops." *Sensors*, vol. 10, no. 8, pp. 7122-7133, 2010
- [2] Zampolli, S., et al. "An electronic nose based on solid state sensor arrays for low-cost indoor air quality monitoring applications," *Sensors and Actuators B: Chemical 101*, vol.1, no. 2, pp. 39-46, 2004.
- [3] Prathyusha, Kanakam, and A. S. N. Chakravarthy. "Chemical Sensing Through Cogno-Monitoring System for Air Quality Evaluation," *Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough*. Springer, pp. 27-37, 2020.
- [4] Cai, C.-J.; Geng, F.-H.; Tie, X.-X.; Yu, Q.; Peng, L.; Zhou, G.-Q. "Characteristics of Ambient Volatile Organic Compounds (Vocs) Measured in Shanghai," *China. Sensors*, vol. 10, pp. 7843–7862, 2010
- [5] Peris, M.; Escuder-Gilabert, L. "A 21st Century Technique for Food Control: Electronic Noses," *Anal. Chim. Acta*, vol. 638, pp. 1–15, 2009.
- [6] Zhang, H.; Chang, M.; Wang, J.; Ye, S., "Evaluation of Peach Quality Indices Using an Electronic Nose by MLR, QPST and BP Network," *Sens. Actuators B Chem.*, vol. 134, pp. 332–338, 2008.
- [7] Kateb, B.; Ryan, M.A.; Homer, M.L.; Lara, L.M.; Yin, Y.; Higa, K.; Chen, M.Y., "Sniffing out Cancer Using the JPL Electronic Nose: A Pilot Study of a Novel Approach to Detection and Differentiation of Brain Cancer," *NeuroImage*, vol. 47, pp. T5–T9, 2009.
- [8] Wongchoosuk, Chatchawal, et al. "Identification of people from armpit odor region using networked electronic nose," *2011 Defense Science Research Conference and Expo (DSR). IEEE*, 2011.
- [9] Wongchoosuk, Chatchawal, Mario Lutz, and Teerakiat Kerdcharoen. "Detection and classification of human body odor using an electronic nose," *Sensors* vol. 9, no. 9, pp. 7234-7249, 2009.
- [10] Voss, A., et al. "E6. 2-Development of an Electronic Nose System to Diagnose Heart Failure from Human Skin Odor", *Proceedings SENSOR 2013*, pp. 648-649, 2013.

- [11] Gardner, Julian W., Hyun Woo Shin, and Evor L. Hines. "An electronic nose system to diagnose illness", *Sensors and Actuators B: Chemical* 70, vol. 1, no. 3, pp. 19-24, 2000.
- [12] Pandey, Sudhir Kumar, and Ki-Hyun Kim. "Human body-odor components and their determination", *TrAC Trends in Analytical Chemistry*, vol. 30, no. 5, pp. 784-796, 2011.
- [13] "The Chemistry of Body Odours – Sweat, Halitosis, Flatulence & Cheesy Feet", *Compound Interest*, 2020. [Online]. Available: <https://www.compoundchem.com/2014/04/07/the-chemistry-of-body-odours-sweat-halitosis-flatulence-cheesy-feet/>. [Accessed: 06- Mar- 2020].
- [14] Kanakam, Prathyusha, S. Mahaboob Hussain, and A. S. N. Chakravarthy. "Cogno-Detective System for Body Odor Sensing", *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCCIC)*. IEEE, 2017.
- [15] Li, Sichu. "Recent developments in human odor detection technologies", *Journal of Forensic Science & Criminology*, vol. 1, no.1, pp. 1-12, 2014
- [16] Kanakam, Prathyusha, S. Mahaboob Hussain, and A. S. N. Chakravarthy. "Smart-Trace: A Future Security Aid Using Body Odor Sensing of an Individual with IoT", *Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIOTCT)*, 2018.
- [17] Cuzuel, Vincent, et al. "Origin, analytical characterization, and use of human odor in forensics", *Journal of forensic sciences*, vol. 62, no.2, pp. 330-350. 2017.
- [18] Sabri, Ahmed Qusay, and Rayner Alfred. "Volatile Organic Compounds (VOCs) Feature Selection for Human Odor Classification", *International Conference on Computational Science and Technology*. Springer, Singapore, 2017.
- [19] Curran, Allison M., et al. "The frequency of occurrence and discriminatory power of compounds found in human scent across a population determined by SPME-GC/MS", *Journal of Chromatography B* 846, vol. 1, no. 2. Pp. 86-97, 2007.
- [20] Dormont L, Bessere JM, McKey D, Cohuet A. "New methods for field collection of human skin volatiles and perspectives for their application in the chemical ecology of human-pathogen-vector interactions", *J Exp Biol*, pp. 2783–2788, vol. 216, no. 15, 2013.
- [21] Zhang ZM, Cai JJ, Ruan GH, Li GK. "The study of fingerprint characteristics of the emanations from human arm skin using the original sampling system by SPME-GC/MS", *J Chromatogr B Analyt Technol Biomed Life Sci*, vol. 822, no. 2, pp. 244-252, 2005.