Feature Extraction by Wavelet Decomposition of Surface Acoustic Wave Sensor Array Transients

Prashant Singh and R.D.S. Yadava*

Faculty of Science, Banaras Hindu University, Varanasi-221 005
*E-mail: ardius@gmail.com, ardius@bhu.ac.in

ABSTRACT

The paper presents a new approach to surface acoustic wave (SAW) chemical sensor array design and data processing for recognition of volatile organic compounds (VOCs) based on transient responses. The array is constructed of variable thickness single polymer-coated SAW oscillator sensors. The thickness of polymer coatings are selected such that during the sensing period, different sensors are loaded with varied levels of diffusive inflow of vapour species due to different stages of termination of equilibration process. Using a single polymer for coating the individual sensors with different thickness introduces vapour-specific kinetics variability in transient responses. The transient shapes are analysed by wavelet decomposition based on Daubechies mother wavelets. The set of discrete wavelet transform (DWT) approximation coefficients across the array transients is taken to represent the vapour sample in two alternate ways. In one, the sets generated by all the transients are combined into a single set to give a single representation to the vapour. In the other, the set of approximation coefficients at each data point generated by all transients is taken to represent the vapour. The latter results in as many alternate representations as there are approximation coefficients. The alternate representations of a vapour sample are treated as different instances or realisations for further processing. The wavelet analysis is then followed by the principal component analysis (PCA) to create new feature space. A comparative analysis of the feature spaces created by both the methods leads to the conclusion that both methods yield complimentary information: the one reveals intrinsic data variables, and the other enhances class separability. The present approach is validated by generating synthetic transient response data based on a prototype polyisobutylene (PIB) coated 3-element SAW sensor array exposed to 7 VOC vapours: chloroform, chlorobenzene o-dichlorobenzene, n-heptane, toluene, n-hexane and n-octane.

Keywords: SAW sensor transients, wavelet decomposition, feature extraction, VOC identification, machine olfaction, surface acoustic wave, discrete wavelet transformer

1. INTRODUCTION

Machine olfaction is a biomimetic approach to detect and identify odourant chemicals in vapour phases[1]. It comprises three parts: sniffer, sensor, and processor. The sniffer samples the target gaseous ambient for collecting the odour causing molecules (odourants) and directing them to interact with the sensor. The sensor generates electrical signal via change in some property of its material (e.g. conductivity or mass density) upon exposure to odourants. The processor usually consists of two parts: pre-processor and pattern recognition. The pre-processing is to prepare the data (that is, the record of sensor output) by applying appropriate procedures of denoising, shifting, scaling, normalisation, and transformation such that the influence of all non-information bearing parameters are removed[2,3]. The pattern recognition combines methods of feature extraction and classification[2]. The features are mathematical descriptors of odourants. The feature extraction is to determine a suitable set of features that represents odour identity in a unique manner. The classifier maps feature sets to class identities. The set of features representing an odour are usually referred to as odourprint or vapourprint or chemical fingerprint. The machine olfaction is therefore chemical fingerprinting implemented by electronic sensors and pattern recognition[1-5].

Polymer-coated surface acoustic wave (SAW) oscillators make an important class of chemical sensors which have been extensively experimented to develop odour detection and identification instruments (popularly referred to as electronic nose)[6]. The SAW electronic nose instruments employ an array of sensors where each sensor is coated with a different polymer[7]. The selection of polymer coatings is optimised according to the application target. An effort is made to employ broad selectivity sensors such that individual sensors respond preferentially to different chemical constituents in the target odour. This produces an array response pattern in which the vapour identity information is hidden. The feature extraction methods of pattern recognition attempt to decipher this. This is done by creating a set of mathematical descriptors (or features) using data transformation and statistical estimation techniques such that the representation of each vapour is unique. Different

vapour classes occupy separate regions in feature space. A suitable classifier can map these regions into class identities, provided it is trained *a priori* with known samples[6-9].

The sensor array-based vapour recognition system depends most commonly on the steady-state responses obtained after thermal equilibration between vapour molecules and polymer matrix.[2,9,10] In several studies[11-17] however it was pointed out that the sensor transients during sensing or purge cycles (starting from the time the vapours are coming in contact with sensor to the time taken to reach equilibrium or from the equilibrated state to the return to the base-level respectively) carry additional information about molecular diffusion in polymer coatings. The transient regions are therefore richer in information for vapour discrimination. In some of the earlier works, it was suggested that by analysing sensor transients, the vapour discrimination can be achieved by a single sensor.[12-14] In these works, wavelet analysis-based set of wavelet expansion coefficients are taken to be the features. In later studies,[15-17] multiresolution wavelet representation of transients has been employed either to reduce data or to select a feature subset for processing by a classifier. Recently, several interesting variations of transient data generation and processing methodologies have been reported that suggest more rigorous exploitation of the transient features-based electronic nose instrumentation[17-23].

In this paper, a sensor array design consisted of single polymer coatings of different thickness on different sensors for transient response generation and a novel method of feature extraction by PCA (principal component analysis) processing of discrete wavelet transforms has been proposed. The polymer thickness are selected in such a way that the vapour in-diffusion is terminated at different stages before completion of the equilibration process. The usefulness of the method is demonstrated by analysing synthetic data for six volatile organic compounds (VOCs) generated by polymeric sensor array with three different thickness.

2. TRANSIENT RESPONSE OF POLYMER-COATED SURFACE ACOUSTIC WAVE CHEMICAL SENSOR

2.1 SAW Oscillator Sensor

Figure 1 shows schematic of a polymer-coated SAW delay line oscillator sensor. The SAW device is used as phase sensitive element in the feedback network of an amplifier circuit. The circuit achieves stable resonance oscillations when the Barkhausen's criterion for the loop gain $|A\beta| \geq 1$ and the loop phase shift $\Delta\varphi = 2n\pi$ are satisfied. The amplifier gain A is usually adjusted to be a little more than that needed to offset the insertion loss due to SAW device β. The phase change around the loop is controlled by the SAW device. The oscillations are set up at that frequency in the passband of SAW device at which loop phase becomes integer multiple of 2π. The SAW propagation path between the input and output interdigital transducers (IDTs) is coated with chemical selective polymer film. Under exposure to vapour sample the polymer film sorbs chemical constituents in vapour phase in selective manner. This is accompanied by change in SAW velocity, hence the propagation delay between the two IDTs. A change in delay generates shifts in loop phase by $2\pi f \Delta\tau$. Therefore, to satisfy the loop phase condition, the oscillations are established at a different frequency. The shift in frequency of oscillations under vapour loading defines the sensor response (or signal)[24-28].

2.2 Vapour Sorption Kinetics and Sensor Response

The vapour sorption refers to thermodynamic partitioning of chemical molecules at the vapour-polymer interface[6-8,29,30]. The moment molecules in vapour phase come in contact with the polymer surface a fraction of them gets adsorbed. This is immediately followed by diffusion into the film. The rear interface between the polymer film and SAW substrate is impermeable to diffusing gaseous species. Therefore, when the gas molecules reach this barrier, they start accumulating, consequently the diffusive influx starts decreasing. Under continuous exposure to vapour, this process continues till the entire film achieves a uniform concentration level and no net diffusive flow occurs. In this condition, the film is said to have reached equilibrium or steady state. Figure 2 shows the evolution of diffusion profile schematically. The equilibrium partitioning is defined by the ratio of concentrations of analyte species in the polymer phase C_p to that in the vapour phase C_v, and is referred to as the partition coefficient $K = C_p/C_v$. The time taken to reach steady state depends on the diffusion coefficient of the gaseous analytes in the polymer material and the film thickness. Over the time scales from start to steady-state, the analyte concentration varies with depth below the film surface. The sensor loading at any instant is determined by the integrated concentration profile across the film thickness. Therefore, the time-dependence of sensor loading is determined by the time-evolution of diffusion profile which depends on three parameters: partition coefficient K, analyte diffusion coefficient D, and film thickness h[29,30].

The sensor signal consists of contributions from vapour loading-induced changes in the polymer mass density,
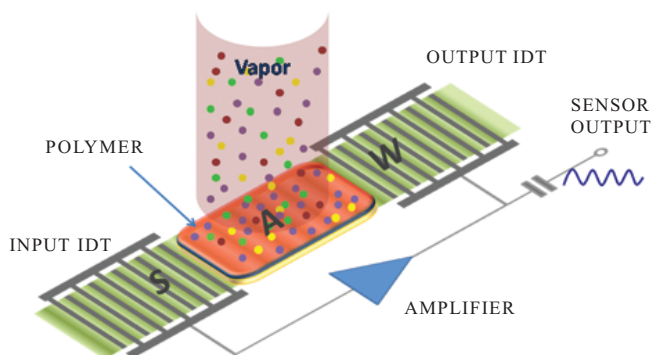

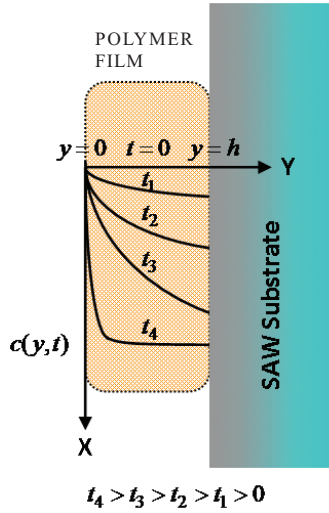**Figure 1. Schematic of a polymer-coated SAW delay line oscillator sensor.**

Figure 2. Schematic of the vapour diffusion profile into the polymer film at different times after the initial exposure at time $t = 0$.

swelling and viscoelastic coefficients. These effects are described by the Martin-Frye-Senturia model of SAW velocity perturbation.[31,32] The model relates the changes in SAW velocity Δv and attenuation $\Delta\alpha$ due to polymer coating through the following relation

$$\frac{\Delta\alpha}{k_0} - j\frac{\Delta v}{v_0} = \sum_{i=1}^{3} c_i \frac{\beta_i M_i}{\omega} \tanh(j\beta_i h) \qquad (1)$$

where k_0, v_0 and c_i denote the unperturbed SAW propagation vector, SAW velocity and normalised surface particle velocities on the piezoelectric surface, respectively, and

$$\beta_i = \omega\left(\frac{\rho - E_i/v_0^2}{M_i}\right)^{1/2}, \ M \qquad (2)$$

where $M_i = M_i' + jM_i''$ is complex shear modulus $G = G' + jG''$ for $i = 1$ or 3 and complex bulk modulus $K = K' + jK''$ for $i = 2$, $\omega = 2\pi f$ is the SAW angular frequency, h is the film thickness, ρ is the film mass density, $E_i = E_i' + jE_i''$ is the film complex elastic modulus. The latter is related to G and K as $E_1 = G$, $E_2 \approx 0$, and $E_3 \approx 4G$. In deriving Eqn (1), the SAW propagation is taken to be in the z-direction, the film thickness along the y-direction, and in-plane normal along the x-direction. The v_0, c_1, c_2 and c_3 are constants specific to SAW substrate and propagation direction. For SAW device on ST-X quartz substrate, $v_0 = 3.158 \times 10^5$ cms⁻¹, and $c_1 = 0.013 \times 10^{-7}$, $c_2 = 1.142 \times 10^{-7}$ and $c_3 = 0.615 \times 10^{-7}$ in units of cm²sg⁻¹.

The change in SAW velocity due to polymer coating is obtained from Eqn (1) by equating the imaginary parts on both the sides. The associated shift in oscillator frequency is related to the shift in SAW velocity as[24,25]

$$\Delta f / f_0 = \Delta v / v_0 \qquad (3)$$

The Eqn (1) tells that the sensor signal due to vapour sorption in polymer film will be generated through changes in polymer mass density ρ, thickness h (swelling), and viscoelastic coefficients M_i. Therefore, to calculate vapour induced sensor output, these changes must be modelled. Here, it is assumed that the vapour-induced film swelling and viscoelastic effects are much smaller than the mass loading effect. Therefore, the sensor signal can be calculated through a model of the change in polymer mass density. That is, one needs to model the time-dependent vapour sorption to calculate the transient response of the sensor. At any instant t after the start of vapour exposure, the variation in polymer mass density will be given as

$$\rho(t) = \rho_p + \frac{1}{h}\int_0^h c(y,t)dy \qquad (4)$$

where ρ_p denotes the polymer mass density before vapour sorption, and $c(y,t)$ as shown in Fig. 2 denotes the vapour-concentration profile at time t. The vapour induced changes in the frequency of the polymer-coated SAW sensor can be obtained from the instantaneous frequency at time t by subtracting from it the change due to polymer coating before the onset of vapour sorption. Let $\Delta f(t)$ denote the frequency change with $\rho = \rho(t)$ as given by Eqn (4) and Δf_p the frequency change with $\rho = \rho_p$, then the sensor transient signal will be given as

$$s(t) = \Delta f(t) - \Delta f_p \qquad (5)$$

2.3 Vapour Sorption-Desorption Model

The theory of sorption and desorption in thin films has been developed on the basis of solution of one-dimensional Fickian diffusion equation under the conditions of transient[33]. Since the lateral dimensions of the film are much larger than the film thickness, the diffusion of species only along the y-axis is considered. The film is assumed to be initially unloaded, therefore the concentration of vapour species in the film is zero at $t < 0$. At time $t = 0$ suppose the vapour concentration is increased by step C_v, the thermodynamic partitioning will set up a constant concentration $C_p = KC_v$ on polymer side of the vapour-polymer interface for all $t > 0$. The surface concentration will be maintained while the diffusion of vapour species progresses in depth. If M_t denotes the total amount of diffusing species which has entered by time t, and M_∞ denotes the corresponding quantity after infinite time (practically, the time taken to reach steady state) the theory yields[33]

$$\frac{M_t}{M_\infty} = 1 - \sum_{n=0}^{\infty} \frac{8}{(2n+1)^2\pi^2} \exp\left[\frac{-D(2n+1)^2\pi^2 t}{h^2}\right] \qquad (6)$$

where D denotes the diffusion coefficient and h is as before, the film thickness. It can be noted that the ratio M_t/M_∞ is the same as the ratio of thickness-averaged concentration of sorbed vapour species introduced in Eqn (4) at respective times. Therefore Eqn (6) can be used to calculate the second term on the right hand side of Eqn (4). Let $C(t)$ denotes this term, then

$$\rho(t) = \rho_p + C(t) \tag{7}$$

After steady state is reached, the concentration across the film thickness becomes uniform and equal to the surface concentration C_p. Therefore, from $C(t)/C_p = M_t/M_\infty$, one can obtain

$$C(t) = KC_v \frac{M_t}{M_\infty} \tag{8}$$

Thus, using Eqns (6) through (7), the transient sensor signal defined by Eqn (5) can be calculated.

To calculate the desorption transients, the same equations are applicable with the initial conditions redefined according to the film state at which the desorption starts. If the film is exposed to vapour for duration t_d before the onset of desorption, a new time variable $t' = t - t_d$ can be redefined to describe desorption transient. If the film has reached steady state before the onset of desorption then $C(t' = 0) = KC_v$. The concentration in the vapour phase is set to be zero for all $t' > 0$. The desorption transient is obtained with

$$C(t') = KC(t'=0) \sum_{n=0}^{\infty} \frac{8}{(2n+1)^2 \pi^2} \exp\left[\frac{-D(2n+1)^2 \pi^2 t'}{h^2}\right] \tag{9}$$

3. MACHINE OLFACTION BASED ON SENSOR ARRAY TRANSIENTS

In electronic nose operation an array of sensors, in which individual sensors possess range of selectivity and sensitivity towards target group of chemicals, provide variability in information generation. The information content in both the transient and the steady-state data are used for feature extraction and vapour classification. The major developments, however, have taken place on the basis of steady-state responses. The approach based on transient responses started early,[12-14] but real sophistication in data processing picked up only recently[15-18]. The initial developments were mostly based on modelling sensor transients and estimating model parameters (e.g., rise time or multi-exponential time constants) for being taken as features for pattern recognition[12,15].

Later, however, more advanced methods that include wavelet analysis, principal component analysis in wavelet space, support vector machine, and neural network classification were employed[16-23]. It is widely recognised that the transient parts of sensor responses are richer in information content. The data processing methods for efficient utilisation of this region are however still in the stage of evolving[21,23]. The transient response analysis also offers the possibility of using only a single sensor for vapour identification because the kinetic parameters such as diffusion coefficients, sorption affinities, and relaxation processes, which determine the shape of transients, may provide adequate variability from one molecular species to the other for discrimination. Though, in some previous studies this has been exploited[12-15], however, most of the works have utilised sensor arrays with heterogeneous polymer functionalities. The different polymers add to the selectivity variation in the transient array data, provide richer data source to mine for robust vapour identification.

3.1 Single Polymer Multiple Thickness SAW Sensor Array–A Novel Approach

In the present work, a new approach based on using only a single polymer-based sensor array, whose individual sensing elements are coated with different film thicknesses, is proposed. The variation in film thickness across the sensor array will introduce different time scales of equilibration. The latter is of the order of h^2/D, that is equilibration times vary in proportion to the square of film thickness[29,30]. Therefore, the vapour sorption and desorption (or sensing and purge) cycles can be adjusted to impart additional vapour discriminating variability in transient response data. Of several approaches to transient analysis, the methods based on wavelet transform are perhaps the most efficient[16-18]. In this approach, the wavelet expansion coefficients of transients are taken to represent vapour samples, and class discrimination is sought in the wavelet space. If a single sensor is used, the different vapour classes will be represented by different sets of expansion coefficients, and discrimination can be achieved. However, if multiple sensors are used, then a vapour sample will be represented by a combined set of wavelet expansion coefficients of all transients of the sensor array. In effect, a vapour sample is represented by a wider set of features, though the dimensionality is enormously increased. The latter can be managed by applying dimensionality reduction algorithms such as principal component analysis, and selecting only a few largest eigenvalue components[16,17].

It is worth clarifying the difference a traditional multiple-polymer multiple (MPM)-sensor array and the proposed single-polymer multiple (SPM)-thickness sensor array would make in information content and its expected impact on vapour classification. For brevity, we may refer to these as MPM- and SPM-sensor array, respectively. In MPM sensor array, the kinetic parameters of vapour species (K and D) vary from sensor to sensor, therefore, the transient features obtained from their wavelet analysis are alternate representations of the vapour sample in terms of different independent sets of wavelet coefficients. In SPM sensor array, the values of the kinetic parameters K and D remain the same for each sensor, however the thickness variation forces different diffusion profiles to emerge for any sensing and purge cycle duration. The sets of wavelet expansion coefficients of individual sensors of the array are not independent; rather, they are the alternate realisations of the same feature set. The PCA of combined features would thus be the reapportion of the vapour information over the same set of independent principal components, hopefully in reinforced manner. One can expect therefore a more robust set of features in comparison to the MPM sensor array as well as the single sensor transient cases. The proposed methodology will effectively reduce uncertainty

associated with variations in chemical interactions in polymer matrix and accentuate vapour-specific features prominently.

3.2 Design of SPM Sensor Array and Transient Response Generation

The validation of proposed sensor array architecture has been done on the basis of simulated transient responses of a multiple thickness SAW sensor array. The response calculation was done as described in Section 2. The SAW oscillators were assumed to be controlled by ST-X quartz delay line devices operating at nominal frequency $f_0 = 100$ MHz. The selected polymer was poly-isobutylene (PIB). A 3-element sensor array coated with 100 nm, 500 nm, and 1000 nm thicknesses of PIB was considered. The transient responses corresponding to seven VOCs exposure at 1 ppm concentration in vapour phase were calculated. The seven vapours were: chloroform, chlorobenzene, o-dichlorobenzene, n-heptane, toluene, n-hexane and n-octane (representatives of). The selection of vapours was made to represent typically the alkanes, aromatic, and halogenated family of compounds, and also for which the partition coefficient K and diffusion coefficient D are available from the published sources. The other parameters needed for the calculation are listed in Table 1.

Table 1. Material characteristics and the vapour-polymer partition coefficient and diffusion coefficient available from the mentioned references. In different papers these values are reported for the temperature range 298-325 K

Vapour	Partition coefficient		Diffusion coefficient	
	K		D ($\times 10^{-11}$ cm^{2}s^{-1})	
Chloroform	200	ref. 34	260	ref. 38
Chlorobenzene	4680	ref. 35	230	ref. 38
o-Dichlorobenzene	22500	ref. 35	5.49	ref. 39
n-Heptane	1200	ref. 36	48	ref. 40
Toluene	1000	ref. 34	35	ref. 40
n-Hexane	180	ref. 34	160	ref. 40
n-Octane	955	ref. 37	38	ref. 40

Figure 3 shows the calculated transient responses of the sensor array for toluene vapours. For 10 s sensing duration, all the sensors were in transient state, Fig. 3(a); whereas, for 60 s sensing duration, at least one sensor has still not reached steady state. In a given MPM type sensor array, and for a chosen sensing and purge cycle, it may be possible that all sensors do not reach steady state for all vapours. Therefore, the data analysis assuming steady-state response may lead sometimes to erroneous conclusion. Moreover in making MPM sensor array, often polymer coatings thicknesses are adjusted to be nearly identical by keeping the coating-induced frequency shifts to be nearly equal. This will inevitably introduce thickness variation. In SPM sensor array, this uncertainty is made to be a certainty by design, and the uncertainty associated
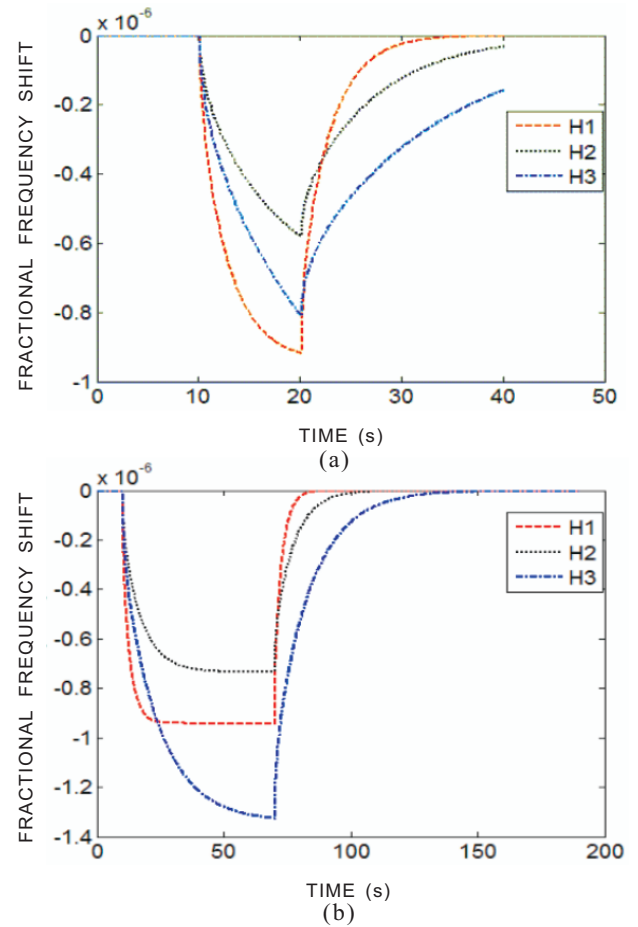

Figure 3. Calculated transient responses of PIB-coated 3-element prototype SAW sensor array under exposure to toluene vapours at 1ppm: (a) for sensing duration 10 s and purge duration 20 s, and (b) for sensing duration 60 s and purge duration 120 s. H1, H2, and H3 denote sensors corresponding to PIB film thicknesses 10 nm, 50 nm, and 100 nm, respectively.

with polymer matrix is removed by working with only one polymer.

The Figure 4 shows all the transients from 3 SPM sensors corresponding to 7 vapours listed in Table 1. The data is generated at 0.01 s interval for 30 s sensing (sorption) and 60 s purge (desorption) durations, and the x-axis indicates data points instead of actual time. This has been done keep data record in discrete time format suitable for discrete wavelet transform (DWT) discussed in the following section. The computed array response can therefore be denoted as $s_{pq}(t = n\Delta t)$ where $p = 1,2,...,P$ represents vapour sample number, $q = 1,2,..., M$ represents sensor number, and $n = 1,2,..., N$ represents discrete time points spaced at Δt with N being the data length. Alternatively, the discrete time transient data can simply be denoted as $s_{p,q,n}$. The real continuous sensor transients are actually sampled at data acquisition stage for DWT analysis. The real system transients are acquired by data acquisition systems at a specified rate (samples per unit time) for a specified duration. The sampling rate is chosen to capture transients with
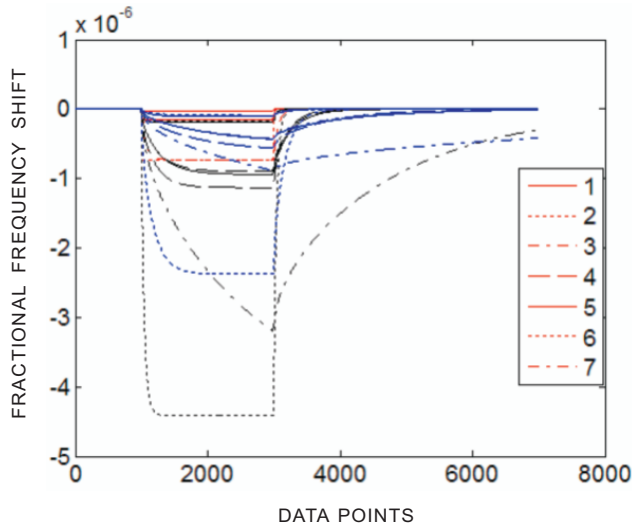
Figure 4. Calculated transient responses of the PIB coated 3-element SPM SAW sensor array with thicknesses 10 nm, 50 nm, and 100 nm when exposed to 1ppm concentration of seven vapours (1: chloroform, 2: chlorobenzene, 3: o-dichlorobenzene, 4: *n*-heptane, 5: toluene, 6: *n*-hexane, 7: *n*-octane).

adequate resolution so that characteristic features are not missed; and the sampling duration is optimised by studying the dynamics of processes involved. The output data from each sensor is a discrete set of sample values.

3.3 Feature Extraction by Wavelet Decomposition

The theory of wavelet transform provides a powerful mathematical tool to analyse time-varying finite energy signals. A signal $s(t)$ is modelled as linear combination of orthonormal basis functions which are created from a mother wavelet by scaling and shifting operations. The fundamentals of wavelet transforms can be found in several books[41-43]. However, an excellent summary is given by Distante[16], *et al.* in their paper on feature extraction for electronic nose. The paper[16] describes a discrete wavelet transform (DWT)-based method of feature extraction from the transient responses of an array of tin-oxide gas sensors. The shapes of individual transients are represented by separate sets of DWT coefficients which are taken as features. A vapour sample is then represented by the combined set of features across the sensor array. The principal component analysis can be used additionally to reduce the dimensionality of feature space in a usual manner. In a recent paper[17], this approach is further extended for feature subset selection based on genetic algorithm (GA). In another recent paper[18], this approach is supported by comparing the results with more generalised method of feature extraction based on wavelet packet analysis.

For analysing the synthetic data in the present paper using the wavelet toolbox in Matlab, DWT has been implemented. The function 'wavedec' using db1 wavelet was employed for the multilevel 1-D wave decomposition. The function implements Mallat algorithm[44]. The scales

and shifts at every level of decomposition in this algorithm are selected on the basis of power of 2. That is, at j^{th} level, the scale factor is 2^j. This is referred to as dyadic scale. Starting from $s_{pq,n}$ input data set for each (p, q) transient containing $n = 1, 2, .. N$ data points, the algorithm in the first level of decomposition produces two sets of coefficients $cA1$ and $cD1$ called the 'approximation' and 'detail' coefficients, respectively. These coefficients are obtained by convolving signal $s_{pq,n}$ with low-pass and high-pass digital filters of length N followed by dyadic decimation. The approximation coefficients $cA1$ follow from the low-pass filter and the detail coefficients $cD1$ from the high-pass filter. At the next level, the detailed coefficients $cD1$ are rejected and the approximation coefficients $cA1$ replace $s_{pq,n}$ with the limit on n redefined according to the number of $cA1$ coefficients. The same procedure is repeated to produce $cA2$ and $cD2$ coefficients of level 2 decomposition. The process is continued till the desired level of decomposition is reached. The final set of approximation coefficients cAj represents the transient features. Figure 5 shows wavelet decomposition of transient responses shown in Fig. 4.

Suppose each sensor transient produces a set of N' features after wavelet decomposition. Then, a vapour sample will be represented by the total MN' features after combining the features produced by each of the M sensors in the array. The original 3-dimensional $P \times M \times N$ data matrix will thus be represented by 2-dimensional $P \times MN'$ feature matrix wherein a row defines a feature vector $[cAj, j = 1, 2, ..., MN']$ corresponding to a vapour sample. As mentioned before, in past two works, PCA has been used for next level feature generation and dimensionality reduction[16], and GA has been used for feature subset selection[17] based on the feature matrix $[cAj]$. In the present
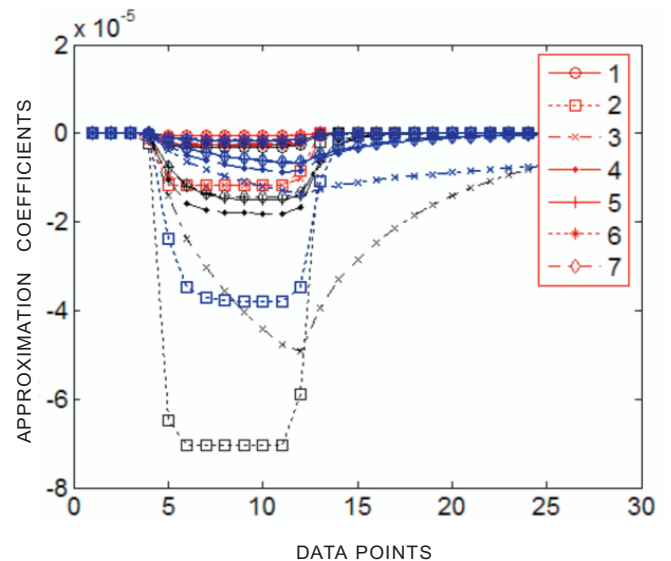

Figure 5. Approximation coefficients at 6th level of discrete wavelet decomposition of the SAW sensor array transients shown in Fig. 4. Each colour denotes a separate sensor.

work, an alternate approach for the next level feature generation by PCA has been explored, and the results have been compared with the earlier approach. In the new approach, the feature set across the sensor array at each data point after the wavelet decomposition is taken to represent an instance or a realisation of the vapour sample. Thus, a vapour sample can be represented by any of the N' alternate realisations each containing M components. All possible realisations can be represented by $N' \times M$ feature matrix wherein each row corresponds to a data point after wavelet decomposition. This is unlike the row vector in the earlier method. Stacking these $N' \times M$ feature matrices corresponding to all vapour samples vertically in a tower matrix, one would obtain $PN' \times M$ feature matrix for the whole data set. Doing PCA on this matrix produces $PN' \times M$ matrix of principal components wherein each vapour sample is represented by a block N' rows with M new extracted feature components.

Based on these new feature vectors, two approaches can be adapted for classification. In one, the best among N' alternate representations of a sample is chosen for input to the classifier. This, however, requires some criterion for defining the best. A possible way could be to link the goodness of the feature vector to the performance of a classifier, for example, the output of a neural network. In the latter, the feature set no. which yields maximum output for the true class in training phase of the network can be selected. In the second approach, the classifier output for each alternate feature vector is processed in some way for the class prediction. For example, maximum of a probabilistic neural network output for all alternate representation can be made the basis for class prediction.

4. VALIDATION

Figures 6 and 7 show the PC score plots of the data shown in Fig. 4 after 6th level of wavelet decomposition. Figure 6 displays the projection of feature space created by the new method, and Fig. 7 by the old method (that is, PCA by combining all features from all transients). In Fig. 6, all the alternate representations of the vapour sample are shown by one symbol in the same colour. It is interesting to note that each vapour class traces a separate path in the new feature space. These traces comprise two branches, the upper branch starting from rarer-to-denser points in clockwise direction is associated with the sorption transient, and the lower branch in the same manner is associated with the desorption transient. It can be noted that the class separation in the initial part of the sorption transients and in the later portion of the desorption transient is poor. The later part of the sorption and the initial part of desorption transients produce substantially enhanced separation. The present method of analysis may therefore provide a basis for choosing those portions of the sensor transients that produce best discrimination among classes.

Another interesting observation is made by examining the eigenvalues associated with both the methods of analyses. These are shown in Table 2. It can be seen that the first two components carry 99 per cent of the total variance of the data in both methods, however, apportioning of the variance values are more contrasting in the new method than in the old method. That is, the new method concentrates information over fewer independent variables, therefore, can lead to more compact class formation in the feature space than the old method. It was noted however that the old method results in only first 5 eigenvalues having finite values; the eigenvalues of all components beyond 5 are zero. This was found to be true for all levels of wavelet decomposition. This implies that the old method can reveal intrinsic dimensionality of the problem. The two methods therefore appear to reveal complimentary information in
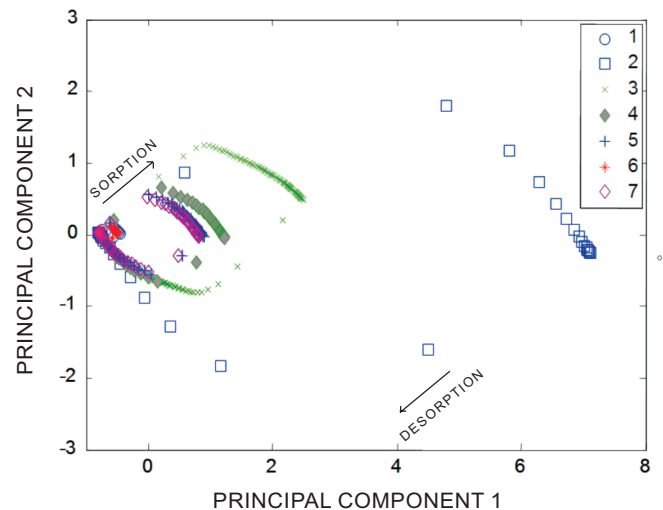

Figure 6. Score plots of the principal components obtained by the new method after 6th level wavelet decomposition of the data shown in Fig. 4. The results shown here are obtained by PCA of the results in Fig. 5.
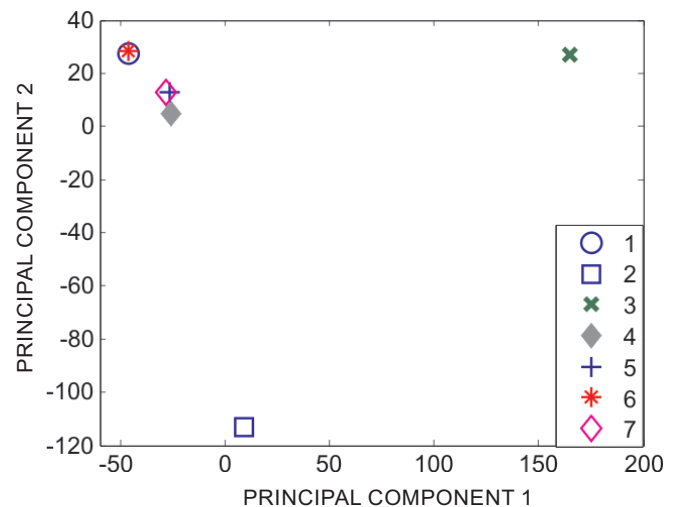

Figure 7. Score plots of the principal components obtained by the old method after 6th level of wavelet decomposition of the data shown in Figs 4 and 5.

Table 2. **Eigenvalue of principal components after 6ᵗʰ level of wavelet decomposition**

PC	New method		Old method	
	Eigen value	Cumulative variance (%)	Eigen value	Cumulative variance (%)
1	2.74	91.2	177.0	67.9
2	0.236	99.1	80.7	98.9
3	0.0281	100	2.51	99.9
4			0.230	99.99
5			0.00884	99.999
6			0.00126	100
7			0	100

some way. For example, the old method can be used to analyse number of hidden variables, and the new method can be used to optimise class separation by focussing on most potent parts of transients.

It may be noted that the new method produces several alternate feature sets for each vapour sample. The point is then, how to use them for classification. A strategy for this has been outlined in the last para of the preceding section. However, without going into any specific method, the consistency of these feature sets for each vapour, by using a fraction of them for training and the rest for testing a neural network classifier, has been examined. Alternate data points at 6ᵗʰ level of decomposition of the sorption transient were divided into training and test sets. A 3 x 2 x 7 neural network architecture using 'newff' function in Matlab with 'logsig' activation and 'trainlm' training algorithm was used for this purpose. It was seen that 100 per cent correct vapour identity could be predicted for all the vapour samples. That means all the alternate representations are equally good. However, it may be premature to conclude this because the present data set is simulated one, and it does not include noisy components that are always present in real data. Further effort would be made to fully explore the advantages of the present method.

5. CONCLUSIONS

Using single polymer multiple thickness coatings on individual sensors in SAW sensor array brings variability in transient patterns that are specific to vapour species. The present work explored this for extraction of vapourprints (or signatures) by doing principal component analysis of the wavelet approximation coefficients. Validation is done by considering a prototype 3-element polyisobutylene (PIB)-coated SAW sensor array under exposure to 7 volatile organic compounds. The model calculation to generate synthetic transient responses is based on well-established SAW perturbation theory due to vapour partitioning and diffusion in polymer films. A new approach to sensor array design based on single polymer is defined and validated. A new method for transient feature extraction is proposed that enhances class separability. In brief, the present analysis defines a new approach for design and development of efficient odour-sensing system based on SAW sensor array and pattern recognition.

REFERENCES

1. Pearce, T.C.; Schiffman, S.S.; Nagle, H.T. & Gardner, J.W. *In* Handbook of machine olfaction: Electronic nose technology. WILLY-VHC Verlag GmbH & Co, Weinheim, 2003.

2. Osuna, R.G. Pattern analysis for machine olfaction: a review. *IEEE Sensors J.*, 2002, **2**(3), 189-02.

3. Osuna, R.G. & Nagle, H.T. A method for evaluating data pre-processing techniques for odour classification with an array of gas sensors, *IEEE Trans. Syst. Man Cybern. B*, 1999, **29**(5), 626-32.

4. Scott, S.M. Data analysis for electronic nose systems, *Microchimica Acta*, 2007, **156**, 183-207.

5. Jurs, P.C.; Bakken, G.A. & McClelland, H.E. Computational methods for the analysis of chemical sensor array data from volatile analytes, *Chemical Review,* 2000, **100**, 2649-678.

6. Grate, J.W. Acoustic wave microsensor arrays for vapour sensing. *Chemical Review*, **100**(7), 2000, 2627-2648.

7. Yadava, R.D.S. & Chaudhary, R. Solvation, transduction and independent component analysis for pattern recognition in SAW electronic nose. *Sensors Actuators B: Chemical*, 2006, **113**, 1-21.

8. Zellers, E.T.; Batterman, S.A.; Han, M. & Patrash, S.J. Optimal coating selection for the analysis of organic vapour mixtures with polymer-coated surface acoustic wave sensor arrays. *Analytical Chemistry*, 1995, **67**(6), 1092-106.

9. Hines, E.L. Pattern analysis for electronic noses. *In* Handbook of Machine Olfaction, edited by T.C. Pearce, S.S. Schiffman, H.T. Nagle & J.W. Gardner. Wiley-VCH, Weinheim, 2003. pp. 133-60.

10. Ho, C.K.; Lindgren, E.R.; Rawlinson, K.S.; McGrath. L.K. & Wright, J.L. Development of a surface acoustic wave sensor for *in-situ* monitoring of volatile organic compounds. *Sensors*, 2003, **3**, 236-47.

11. Zhao, W.; Bhusan, A.; Santamaria, A.D.; Simon, M.G. & Davis, C.F. Machine learning: A crucial tool for sensor design. *Algorithms*, 2008, **1**, 130-52.

12. Vilanova, X.; Llobet, E.; Alcubilla, R.; Sueiras, J.E. & Correig, X. Analysis of the conductance transient in thick-film tin oxide gas sensors. *Sensors Actuators B*, 1996, **31**, 175- 80.

13. Llobet, E.; Brezmes, J.; Vilanova, X.; Fondevila, L. &

Correig, X. Quantitative vapour analysis using the transient response of non-selective thick-film tin oxide gas sensors. *In* International Conference on Solid-state Sensors and Actuators, Chicago, June 16-19, 1997. pp. 971-74.

14. Osuna, R.G.; Nagle, H.T. & Schiffman, S.S. Transient response analysis of an electronic nose using multi-exponential models. *Sensors Actuators B*, 1999, **61**, 170-82.

15. Islam, A.K.M.S.; Ismail, Z.; Ahmad, M.N.; Saad, B.; Othman, A.R.; Shakaff, A. Y. M.; Daud, A. & Ishak, Z. Transient parameters of a coated quartz crystal microbalance sensor for the detection of volatile organic compounds (VOCs). *Sensors Actuators B*, 2005, **109,** 238-43.

16. Distante, C.; Leo, M.; Siciliano, P. & Persaud, K.C. On the study of feature extraction methods for an electronic nose. *Sensors Actuators B,* 2002, **87**, 274-88.

17. Phaisangittisagul, E. & Nagle, H. T. Sensor selection for machine olfaction based on transient feature extraction. *IEEE Trans. Instrum. Meas.*, 2008, **57**(2), 369-378.

18. Yin, Y.; Yu, H. & Zhang, H. A feature extraction method based on wavelet packet analysis for discrimination of Chinese vinegars using a gas sensors array, *Sensors Actuators B,* 2008, **134**,1005-009.

19. Nucci, C.D.; Fort, A.; Rocchi, S; Ulivieri, N.; Vignoli, V. & Catelani, M. Feature extraction techniques for QCM sensors dynamic responses. *In* Instrumentation and Measurement Technology Conference (IMTC 2004), Como, Italy, 18-20 May 2004. pp. 605-09.

20. Leone, A.; Distante, C.; Ancona, N.; Persaud, K.C.; Stella, E. & Siciliano, P. A powerful method for feature extraction and compression of electronic nose responses. *Sensors Actuators B*, 2005, **105**(2), 378-92.

21. Vergara, A.; Llobet, E.; Martinelli, E.; Natale, C.D.; D'Amico, A. & Correig, X. Feature extraction of metal oxide gas sensors using dynamic moments. *Sensors Actuators B,* **122**, 2007, 219-26.

22. Hui, D.; Jun-hua,L. & Zhong-ru, S. Drift reduction of gas sensor by wavelet and principal component analysis. *Sensors Actuators B,* **96**, 2003, 354-63.

23. Gouws, G. J. & Gouws, D. J. Analyte identification using concentration modulation and wavelet analysis of QCM sensors. *Sensors Actuators B,* 2003, **91**, 326-32.

24. Wholtjen, H. & Dessy, R. Surface acoustic wave probe for chemical analysis *Analytical Chemistry,* 1979, **51**, 1458-464.

25. Yadava, R.D.S. Enhancing mass sensitivity of SAW delay line sensors by chirping transducers. *Sensors Actuators B*, 2006, **114**, 127-31.

26. Yadava, R.D.S. SAW sensor technology for detection of CW agents–An appraisal. *In* Proceedings 8 th National Seminar on Physics and Technology of Sensors (NSPTS-8), IGCAR, Kalpakkam, India, 27 Feb-1 Mar 2001. pp.

I-12.1-12.7.

27. Khaneja, M. & Yadava, R.D.S. Optimising transducer design for SAW chemical and biological sensors. *In* Proceedings 8th National Seminar on Physics and Technology of Sensors (NSPTS-8), IGCAR, Kalpakkam, India, 27 Feb -1 Mar 2001. pp. C-36.1-36.4.

28. Deobagkar, D.; Limaye, V.; Sinha, S. & Yadava, R.D.S. Acoustic wave immunosensing of *E. Coli* in water. *Sensors Actuators B*, 2005, **104**, 85-89.

29. Liron, Z.; Kaushansky, N.; Frishman, G.; Kaplan, D. & Greenblat, J. The polymer-coated SAW sensor as a gravimetric sensor. *Analytical Chemistry*, 1997, **69**(14), 2848-854.

30. Fischerauer, G. & Dickert, F. L. An analytic model of the dynamic response of mass-sensitive chemical sensors. *Sensors Actuators B,* 2007, **123**, 993-1001.

31. Martin, S.J.; Frye, G.C. & Senturia, S.D. Dynamics and response of polymer-coated surface acoustic wave devices: effect of viscoelastic properties and film resonance. *Analytical Chemistry*, **66**, 1994, 2201-219.

32. Yadava, R.D.S.; Kshetrimayum, R. & Khaneja, M. Multifrequency characterisation of viscoelastic polymers and vapour sensing based on SAW oscillators. *Ultrasonics*, 2009. doi:10.1016/j.ultras.2009.03.006.

33. Crank, J. The mathematics of diffusion. Clarendon Press, Oxford, Sec 4.3, Eqn (4.18), 1986.

34. Grate, J.W.; Kaganove, S.N. & Bhethanabotla, V.R. Examination of mass and modulus contributions to thickness shear mode and surface acoustic wave vapour sensor responses using partition coefficients. *Faraday Discussion*, 1997, **107**, 259-83.

35. Patrash, S.J. & Zellers, E.T. Characterisation of polymeric surface acoustic wave sensor coatings and semiempirical models of sensor responses to organic vapours, *Analytical Chemistry*, **65**, l993, 2055-066.

36. Fu, Y. & Finklea, H.O. Quartz crystal microbalance sensor for organic vapour detection based on molecularly imprinted polymers. *Analytical Chemistry*, 2003, **75**, 5387-393.

37. Houser, E.J.; Mlsna, T.E; Nguyen, V.K.; Chung, R.; Mowery, R.L. & McGill, R.A. Rational materials design of sorbent coatings for explosives: applications with chemical sensors. *Talanta*, 2001, **54,** 469-485.

38. Gobel, R.; Seitz, R.W.; Tomellini, S.A.; Krska, R. & Kellner, R. Infrared attenuated total reflection spectroscopic investigations of the diffusion behaviour of chlorinated hydrocarbons into polymer membranes. *Vibrational Spectroscopy*, 1995, **8**, 141-49.

39. Howley, R.; MacCraith, B.D.; O'Dwyer, K.; Kirwan, P. & McLoughlin, P. A study of the factors affecting the diffusion of chlorinated hydrocarbons into polyisobutylene and polyethylene-co-propylene for evanescent wave sensing.*Vibrational Spectroscopy*, 2003, **31**, 271-78.

40. Jiang, W.H.; Liu, H.; Hu, H.J. & Han, S.J. Infinite dilution diffusion coefficient of n-hexane, *n*-heptane

and *n*-octane in polyisobutylene by inverse gas chromatographic measurements, *Euro. Poly. J.*, 2001, **37**, 1705-712.

41. Rao, R.M. & Boparkidar, A.S. Wavelet transform: Introduction to theory and applications. Addision-Wiesley, USA, 1998.

42. Mallat, S.A. Wavelet tour of signal processing. Ed. 2, Academic, San Diego, CA, 1999.

43. Burrus, C.S.; Gopinath, R.A. & Guo, H. Introduction to wavelets and wavelet transforms: A primer. Prentice-Hall, New Jersey, 1998.

44. Mallat, S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Pattern Anal. Mac. Intell.*, 1989, **11**(7), 674-93.

Contributors

Mr Prashant Singh received his BSc and MSc (Physics) from Tilak Dhari PG College, Jaunpur (UP) in 2003 and 2005, respectively. He is working as a PhD scholar in the Department of Physics, Banaras Hindu University, Varanasi. His research area include: Dynamic response modelling and analysis of polymer functionalised surface acoustic wave (SAW) sensors, and transient signal-based electronic nose application. He is currently working on the development of wavelet analysis-based feature extraction methods, sensor design, and signal denoising methods.

Dr R.D.S. Yadava   (See page 376)