



A simple electronic tongue

Ritesh Kumar, Amol P. Bhondekar*, Rishemjit Kaur, Saru Vig, Anupma Sharma, Pawan Kapur

CSIR-Central Scientific Instruments Organisation (CSIR-CSIO), Sector-30C, Chandigarh 160 030, India

ARTICLE INFO

Article history:

Received 6 March 2012

Received in revised form 4 May 2012

Accepted 12 June 2012

Available online 20 June 2012

Keywords:

Audio CODEC

Electronic tongue

Electrochemical impedance spectroscopy (EIS)

Fourier transform impedance spectroscopy

Impedance tongue

Support Vector Machine (SVM)

ABSTRACT

In this work a novel electronic tongue (ET) using Fourier transform impedance spectroscopy has been demonstrated. Odd random phase multisine waveform has been used as an excitation signal. Texas Instruments' PCM2900B USB audio CODEC chip has been used as a signal generation and an acquisition module for the ET. The acquired impedance features have been further subjected to Principal Component Analysis (PCA) for dimensionality reduction and Support Vector Machines (SVM) for pattern classification. A good classification accuracy has been achieved for single specie samples (taste samples), multi-species samples (water) and complex samples (tea). Also, the performance of the proposed system has been measured in terms of qualitative performance parameters namely, false positive rate, false negative rate, sensitivity rate and specificity rate.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Various electronic tongue (ET) concepts based upon potentiometry [1–4], voltammetry [5–7], amperometry [8], electrochemical impedance spectroscopy (EIS) [9–16] and hybrid approaches (potentiometry, voltammetry and conductivity) [17] have been reported earlier. The impedance measurements are advantageous in comparison to potentiometry and especially voltammetry, owing to the potential experimental simplicity and the reduction of the response times [13], and has been applied to qualitative discrimination of mineral water, tea, coffee and red wine [11–16]. However, implementation of these techniques requires sophisticated and expensive laboratory instruments such as electrochemical workstations, impedance analyzers and data acquisition systems. Valle and Riul have done exhaustive review on ETs and have discussed sensing methodologies, data treatment, applications and future challenges [18,19].

A novel ET concept based upon EIS, christened as an impedance-Tongue (iTongue) has been reported by Bhondekar et al. [14–16] recently for discrimination of tea samples. This work demonstrated a novel concept of creating an array of pseudo-sensors by exciting single working electrode at different frequencies thereby opening possibilities for development of simple and cost-effective ET systems. However, usually in an EIS system the sinusoidal excitation waveforms are sequentially applied in increasing or decreasing

order of frequencies. This increases the experimental time, particularly at low frequencies. In order to decrease the experimental time, a multisine technique is used, wherein, the frequencies of interest are simultaneously applied, and impedances at these frequencies are measured using data processing techniques (e.g. Fourier analysis). However, this increases the system complexity in terms of hardware as well as software. This encouraged us to look for a simple and low cost solution. Use of standard personal computer audio ports seemed to be a promising solution due to their ubiquity, full duplex operation (simultaneous recording and playback) and programming flexibility, but the non-uniformity in hardware and electrical characteristics of the audio ports across different makes and models of computer manufacturers, limits its generalization. Apart from it, noise issues related to the on-board sound ports and the cables (sound port to electrochemical setup), dissuaded us from using the on-board sound ports and therefore, a USB based audio CODEC chip proved to be a better option.

In this work, we demonstrate a simple and low cost ET using Texas Instruments' USB audio CODEC chip (PCM2900B). An odd random phase multisine waveform in the frequency range of 20 Hz to 20 kHz has been used as an excitation signal. Further, apparent impedance using Fourier transforms of excitation and measured signals are calculated, and a Support Vector Machine (SVM) in conjunction with Principal Component Analysis (PCA) uses these impedance features for pattern classification. A good classification accuracy has been achieved for single species samples (taste samples i.e. sweet (sucrose), sour (HCl), bitter (MgCl₂) and salty (NaCl) at different concentrations), multi-species samples (water) and complex samples (Kangra orthodox black tea (*Camellia sinensis*

* Corresponding author. Tel.: +91 172 2657811 489; fax: +91 172 2657820.
E-mail address: amolbhondekar@csio.res.in (A.P. Bhondekar).

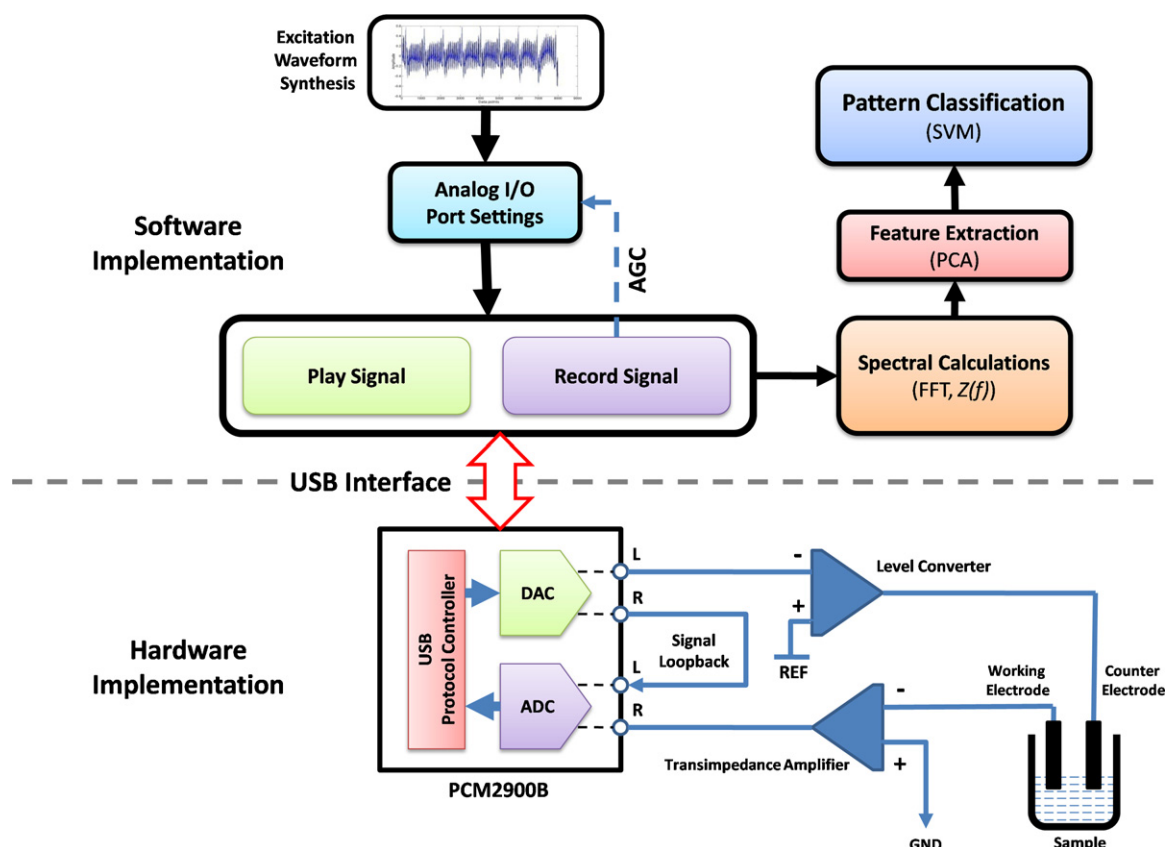


Fig. 1. Block diagram of hardware and software sections of the ET.

(L.) O. Kuntze)). The proposed low cost ET is capable of classifying the samples with concentration and manufacturing variability. Further, performance of the ET has been measured in terms of qualitative performance parameters namely, false positive rate (FPR), false negative rate (FNR), sensitivity rate (SER) and specificity rate (SPR). However, the performance of ET can be improved further by optimally selecting the frequencies and classifier parameters. The instrument provides a rapid analysis with comparatively easy implementation.

2. Implementation

2.1. Hardware

The overall system consists of a personal computer with USB port, electronics and a 2-electrode electrochemical cell. Standard voltammetric platinum and gold disk electrodes (solvent-resistant CTFE plastic body, 7.5 cm length \times 6 mm outer diameter, 1.6 mm disk diameter, supplied by BASi, USA) have been used as counter and working electrodes, respectively. The computer communicates via USB with the electronics. A GUI on the computer controls the electronics and acquires data. The electronic circuit is powered by the USB supply. Fig. 1 shows the block diagram of the complete system. Wherein, the hardware implementation section consists of PCM2900B (Texas Instruments' single-chip, USB, full-duplex stereo audio CODEC) [20], signal level-down converter and a transimpedance amplifier. The level converter steps down the excitation signal (connected via the left output channel) to appropriate voltage level (typically 5–10 mV) in order to avoid electrode polarization. The output of level converter then excites the platinum counter electrode of a 2-electrode electrochemical cell. The transimpedance amplifier converts the current through the gold

working electrode to appropriate voltage level. The output of transimpedance amplifier is acquired through the right input channel of PCM2900B. The excitation signal is duplicated on the left and right output channels and the right output channel is looped back via the left input channel for synchronization of the excitation and sense signals; and also for Automatic Gain Control (AGC) of the excitation signal. Fig. 2 shows the prototype of the system. This prototype has been fabricated using a 3D rapid prototyping system (Stratasys Inc., USA; Model: uPrint Plus) available in our laboratory.



Fig. 2. ET prototype set up along with stand.

2.2. Software

In the software implementation section, excitation waveform is digitally synthesized and played on the output channels of the USB CODEC. Test signals are played prior to measurements and the gain adjustments are done by automatically controlling the volume (AGC). The excitation and sense signals are played and recorded simultaneously; this technique is generally referred to as full-duplex. Spectral transformations are done on the loop back and the sense signals for impedance calculations. These impedance values are further subjected to PCA for dimensionality reduction and Support Vector Machines (SVM) for pattern classification. Details of these processes are mentioned in following subsections. The software has been implemented in MATLAB using data acquisition toolbox.

2.2.1. Signal generation and acquisition

In practice, EIS measurements are usually performed with very small amplitude excitation signals in order to achieve linearity, due to this the measurements suffer from poor signal to noise ratio apart from losing the ability to describe the initial rapidly evolving stages of an electrochemical phenomenon. Broadband impedance measurements based on simultaneous measurement of multiple frequency components have been used since long. However, multisine excitation approaches for EIS have a rather limited number of experimental investigations [21] due to prenotions over the advantages in comparison to the conventional single sine excitation methods. In fact, multisine measurements not only facilitate reduction in experimental time, but also, offer a scope to detect and quantify the presence of time-variance and measurement of non-linear distortions when specific data treatment procedures are used. Hubin and co-workers [22–24] have demonstrated that by using an odd random phase multisine signal; noise, non-linear distortions and non-stationary behavior of an electrochemical system can be quantified. This issue has been duly addressed by Van Ingelgem et al. [21] and the advantages of odd random phase multisine electrochemical impedance measurements have been demonstrated in terms of reliability and accuracy in determination of the impedance parameters. The reliability and accuracy of the impedance parameters are of key importance, especially when the results from impedance measurements are to be incorporated in analytical studies. Although, this technique has been primarily demonstrated for corrosion studies [25,26], but has tremendous scope in other applications [27]. This encouraged us to use a multisine odd random phase signal. This periodic broadband signal consists of the sum of N harmonically related sine waves with amplitude U_k , frequency f_k and random phases φ_k (chosen between 0 and 2π such that $E[\exp(j\varphi_k)] = 0$ where $E[\cdot]$ is the expected value [28]). The time domain representation of this signal is as follows:

$$u(t) = \sum_{k=1}^N U_k \cos(2\pi f_k t + \varphi_k) \quad (1)$$

where, the *rms* value of $u(t)$ is independent of N and hence, U_k decreases as $1/\sqrt{N}$. The values of U_k are chosen so as to flatten the frequency response of PCM2900b and odd frequencies in the range of 20 Hz to 10 kHz with logarithmic spacing are chosen i.e. the spacing of 10 Hz between 20 Hz and 100 Hz, 100 Hz spacing between 100 Hz and 1000 Hz and so on. For the range between 10 kHz and 20 kHz frequencies with an interval of 1000 Hz are chosen. This yielded a total of 37 frequency points. The signal is generated for 4 s with a sampling rate of 96 000 samples/s and permanently stored for playback. This signal is played simultaneously on the left and right output channels of PCM2900B wherein, one of the channels is looped back to one of the input channels for automatic voltage control, whereas, the other channel excites the counter electrode.

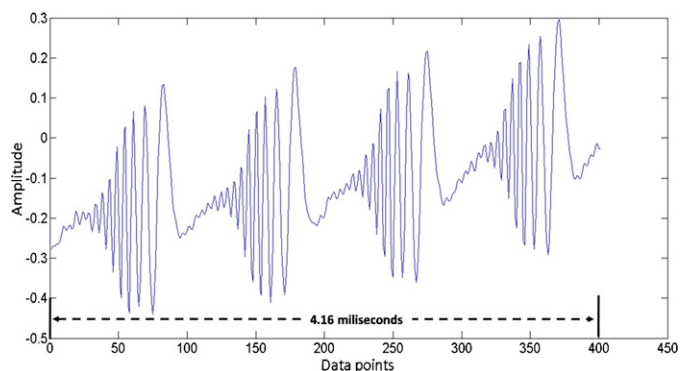


Fig. 3. Generated excitation waveform.

The signal from the working electrode is acquired for 1 s with the sampling rate of 96 000 samples/s through the other input channel. The acquired data is stored in ASCII format for further processing. Fig. 3 shows the generated excitation waveform.

2.2.2. Feature extraction

The signals are acquired from the working and counter (loop back signal) electrodes, and each acquisition frame is divided into 11 segments of 8192 data points after omitting 5888 initial data points. These segments are then transformed into frequency domain using 8192 point FFT and the Intra-Frame Averaging is performed. This yields a frequency resolution of 11.72 Hz, and the magnitudes of the frequencies nearest to the original 37 frequencies are retained. Further, the complex impedance of the electrochemical cell at each of the retained frequencies, $Z(f)$ is then determined by [29]

$$Z(f) = |Z(f)|e^{j\theta(f)} = \frac{|V(f)|}{I(f)} e^{j(\theta_v(f) - \theta_i(f))} \quad (2)$$

where $\theta(f)$ is the phase difference between the excitation signal and the current through the electrochemical cell. Fig. 4 shows the typical Bode plot for 0.1 M NaCl solution. However, it may be noted that the value of $Z(f)$ is not the true impedance, as the system dynamics of the hardware are not being taken into consideration by assuming that these dynamics are independent of sample variability.

Principal Component Analysis (PCA) is a statistical technique for the reduction of input data dimension and is largely used for feature extraction. It captures the relevant information in a set of input data providing a lower dimension. It sequentially creates a

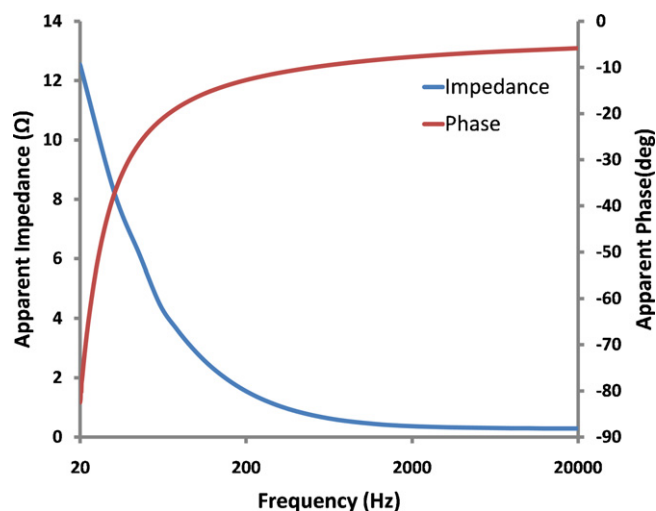


Fig. 4. Typical Bode plot of 0.1 M NaCl solution.

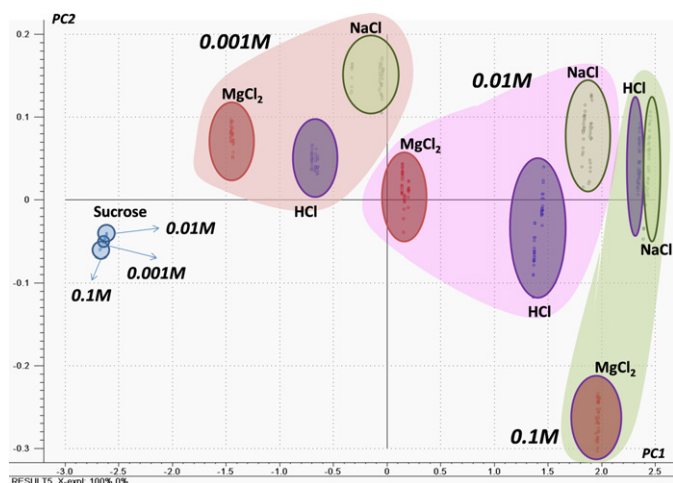


Fig. 5. PCA score plot of taste solutions at 0.001 M, 0.01 M and 0.1 M concentrations.

set of eigen vectors called as principal components from the original data. The first principal component (PC1) maps the maximum variance and information of the input data followed by the other principal components (PC2, PC3 and so on) in decreasing order of the variance. Generally a good discrimination is mapped by the first few principal components. Here, the 37 apparent impedance features are subjected to PCA and the first five PCs are used for pattern classification.

2.2.3. Pattern classification

Good generalization performance for a finite amount of training data, by striking a balance between goodness of fit on the training data and error free prediction on test data is the prime objective of any machine learning algorithm. Support Vector Machines (SVM) proposed by Vapnik [30] has been proven to achieve this with no prior knowledge of data. It has been successfully applied to a gamut of applications ranging from object recognition [31] and text characterization to electronic tongue and nose [15,32]. The SVM in general maps the input data onto a higher dimensional feature space which is non-linearly related to the input space and determines a separating hyperplane such that a maximum margin between two classes is achieved. A Support Vector Machine is a maximal margin hyperplane in the feature space which is determined by a kernel function. By using these kernel functions in the input space the optimal hyperplane can be determined without any computations in the higher dimension feature space. Commonly used kernels are linear, polynomial and radial basis function [33]. In its elementary form, SVM is primarily used for binary classification; however, it may be extended to multiclass problems using strategies such as one-against-rest or one-against-one. Concise overviews on SVM with respect to their applications in chemometrics can be found in [34,35]. In the present study, LIBSVM library [36] has been used for implementation of SVM. The LIBSVM library which uses one-against-one strategy has been used with linear kernel and cost parameter as 1.

2.2.4. Qualitative performance analysis

For qualitative methods, performance is usually expressed in terms of false response rates and detection capability, rather in terms of precision and trueness or measurement uncertainty. Usually, parameters such as FPR, FNR, SPR and SER are used for the performance evaluation of a qualitative method [37–40]. Since the analysis done by the proposed ET is qualitative in nature, these parameters have been chosen for its performance analysis and are defined as follows.

FPR is the probability that a test sample is a known negative, given that the test sample has been classified as positive by the method [40,41]. FPR is defined as:

$$\text{FPR} = \frac{\text{Number of False Positives}}{\text{Number of True Negatives} + \text{Number of False Positives}} \quad (3)$$

Similarly, FNR is the probability that a test sample is a known positive, given that the test sample has been classified as negative by the method [40,41]. FNR is defined as:

$$\text{FNR} = \frac{\text{Number of False Negatives}}{\text{Number of True Positives} + \text{Number of False Negatives}} \quad (4)$$

Further, specificity is defined as the ability of a method to detect truly negative samples as negative [42] and the SPR is the probability for a given concentration, that the method will classify the test sample as negative, given that the test sample is a known negative [40]. SPR is defined as:

$$\text{SPR} = \frac{\text{Number of True Negatives}}{\text{Number of False Positives} + \text{Number of True Negatives}} \quad (5)$$

Similarly, sensitivity is the ability of a method to detect truly positive samples as positive [42] and the SER is the probability for a given concentration that the method will classify test sample as positive, given that the test sample is a known positive [40]. SER is defined as:

$$\text{SER} = \frac{\text{Number of True Positives}}{\text{Number of False Negatives} + \text{Number of True Positives}} \quad (6)$$

The aforementioned parameters have been calculated for the qualitative performance analysis of the proposed ET.

3. Sample preparation and data collection

3.1. Basic taste samples

Four basic taste solutions, namely, sour, salty, bitter and sweet were prepared using HCl, NaCl, MgCl_2 and sucrose, respectively. Each taste solution was prepared for 0.1 M, 0.01 M and 0.001 M concentrations in distilled water. Impedance spectra of 40 samples for each taste at each concentration were acquired. The impedance spectra of each sample was acquired five times and averaged. Further, the 40 samples for each taste at each concentration were split into two equal batches randomly for training and testing. The experiments were performed at room temperature and the electrodes were duly polished before measurement.

3.2. Water samples

Six brands of packaged water, easily available in Indian markets i.e. Aquafina (AQU), Bisleri (BIS), Equal (EQU), Himalya (HIM), Kinley (KIN), Rail Neer (RAN) along with distilled (DIS) [Type 1 ultra pure, $18.2 \text{ M}\Omega \text{ cm}$ at 25°C] and tap (TAP) water were used for experimentation. Four bottles of different batches were collected for each of the brands and 10 samples each of 5 ml were obtained from each bottle. For each class, 40 samples were used to obtain the impedance spectra. The impedance spectra of each sample was measured five times and averaged. Further, these 40 samples per class were divided into 2 equal batches for training and test purpose. The experiments were performed at room temperature and the electrodes were duly polished before measurement.

3.3. Tea samples

Kangra orthodox black tea (*Camellia sinensis* (L.) O. Kuntze) samples were collected during single batch production at the "The Palampur Co-Operative Tea Factory Ltd.", Palampur, Himachal

Output Class	HCl	0.001M	20 8.3%	0	0	0	0	0	0	0	0	0	0	0	0	100%	0.0%
		0.01M	0	20 8.3%	0	0	0	0	0	0	0	0	0	0	0	100%	0.0%
		0.1M	0	0	20 8.3%	0	0	0	0	0	0	5	2.1%	0	0	80.0%	20.0%
	MgCl ₂	0.001M	0	0	0	20 8.3%	0	0	0	0	0	0	0	0	0	100%	0.0%
		0.01M	0	0	0	0	20 8.3%	0	0	0	0	0	0	0	0	100%	0.0%
		0.1M	0	0	0	0	0	20 8.3%	0	0	0	0	0	0	0	100%	0.0%
	NaCl	0.001M	0	0	0	0	0	0	20 8.3%	0	0	0	0	0	0	100%	0.0%
		0.01M	0	0	0	0	0	0	0	20 8.3%	0	0	0	0	0	100%	0.0%
		0.1M	0	0	0	0	0	0	0	0	15 6.3%	0	0	0	0	100%	0.0%
	Sucrose	0.001M	0	0	0	0	0	0	0	0	0	12 5.0%	0	0	0	100%	0.0%
		0.01M	0	0	0	0	0	0	0	0	0	8 3.3%	20 8.3%	0	0	71.4%	28.6%
		0.1M	0	0	0	0	0	0	0	0	0	0	0	20 8.3%	0	100%	0.0%
Target Class		0.001M	0.01M	0.1M	0.001M	0.01M	0.1M	0.001M	0.01M	0.1M	0.001M	0.01M	0.1M	0.001M	0.01M	0.1M	
		HCl			MgCl ₂			NaCl			Sucrose						
		100%	100%	100%	100%	100%	100%	100%	100%	75.0%	60.0%	100%	100%	94.6%	5.4%		
		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	25.0%	40.0%	0.0%	0.0%	0.0%	0.0%		

Fig. 6. Confusion matrix for taste classes at 0.001 M, 0.01 M and 0.1 M concentrations.

Pradesh, India. Three samples each were picked after withering (WIT), during fermentation (50, 60 and 70 min, abbreviated as 50M, 60M and 70M, respectively) and after mechanical grading. The samples were classified as “Super Fine Tippy Flowery Orange Pekoe” (SFTFOP-1, abbreviated as SFT), “Tippy Golden Broken Orange Pekoe” (TGBOP, abbreviated as TGB), “Golden Orange Fannings” (GOF) and “Flowery Orange Fannings” (FOF) according to internationally accepted black tea grading methodology based upon particle size and appearance [43–45]. Thus, 24 tea samples belonging to 8 classes were obtained.

0.4 g of sample was added to 25 ml of distilled boiling water in a flask and boiling was continued for 10 min in a water bath and subsequently allowed to cool down to room temperature. The infusions were then filtered using Whatman Filter No. 42. The infusions were prepared in triplicates yielding 45 infusions of 5 ml each, per class. Impedance spectra of each infusion was measured five times and averaged. Further, these 45 samples were divided into 2 batches of 25 and 20 samples for training and testing purpose, respectively. The experiments were performed at room temperature and the electrodes were duly polished before measurement.

4. Results and discussions

Fig. 5 shows the PCA score plot of all the tastes at each concentration level (0.1 M, 0.01 M, and 0.001 M). Both the Principal components account for a total of 100% of variance in the data. Distinct clusters and sub-clusters of concentrations and tastes can be observed in this figure except for sucrose, being a weak electrolyte as compared to others. These plots give us an intuition that clear decision boundaries could be obtained by the classification algorithms with an expected error for the sucrose samples which is evident from the confusion matrix (Fig. 6) derived after classification using SVM. Careful observation of this confusion matrix shows a maximum misclassification rate of 40% that too, amongst the sucrose samples only. A misclassification rate of 25% can also

be observed between HCl and NaCl at 0.1 M concentration. The overall classification rate obtained for these samples is 94.6%. Table 1 shows the percentage qualitative performance parameters for these samples. Wherein, sucrose 0.01 M has the highest FPR of 3.64% followed by HCl 0.1 M (2.27%), whereas, sucrose 0.001 M has the highest FNR of 40% followed by NaCl 0.1 M (25%). Similarly, 0.001 M has the lowest SER of 60% preceded by NaCl 0.1 M (75%), whereas, Sucrose 0.01 M has the lowest SPR of 96.36% preceded by HCl 0.1 M (97.73%).

Similarly, Fig. 7 shows the PCA score plot of first two principal components derived from 37 apparent impedance features of water samples. Both the Principal components account for a total of 99% of variance in the data. Distinct clusters of water samples can be observed in this plot. It is also interesting to observe the vicinities of these clusters on the basis of manufacturing information on the water packages. A major factor affecting the inter-class variance is

Table 1

Qualitative performance parameters for basic taste samples.

Samples	%FPR	%FNR	%SER	%SPR
Sour				
HCl 0.001 M	0	0	100	100
HCl 0.01 M	0	0	100	100
HCl 0.1 M	2.27	0	100	97.73
Bitter				
MgCl ₂ 0.001 M	0	0	100	100
MgCl ₂ 0.01 M	0	0	100	100
MgCl ₂ 0.1 M	0	0	100	100
Salty				
NaCl 0.001 M	0	0	100	100
NaCl 0.01 M	0	0	100	100
NaCl 0.1 M	0	25	75	100
Sweet				
Sucrose 0.001 M	0	40	60	100
Sucrose 0.01 M	3.64	0	100	96.36
Sucrose 0.1 M	0	0	100	100

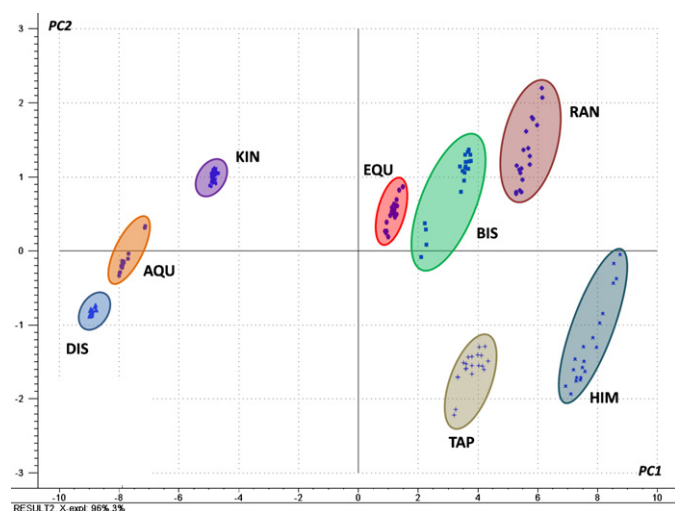


Fig. 7. PCA score plot of water samples.

the geographical location of manufacturing. The close placement of AQU-KIN and EQU-BIS-RAN clusters may be attributed to a common geographic region, Punjab and Delhi region of India respectively. Also, the vicinity of AQU-KIN clusters to DIS cluster suggests low mineral content as compared to others. However, the placement of TAP and HIM clusters does not explain a common geographical origin but ascertains a fact that HIM is natural spring water containing added minerals and TAP water contains more salts than the others. Further, AQU and KIN were manufactured using a similar process except for filtration and variations in added salts. Further, the data were subjected to classification and an overall classification rate of 97.5% has been achieved. Fig. 8 shows the confusion matrix for the classification results obtained, four BIS samples misclassified as EQU can be observed in the matrix. Table 2 shows the percentage qualitative performance parameters for the water samples. Wherein, EQU and BIS have the highest FPR (2.86%) and FNR (20%), respectively. Similarly, BIS and EQU has the lowest SER (80%) and SPR (97.14%), respectively.

Output Class									Target Class
	AQU	BIS	DIS	EQU	HIM	KIN	RAN	TAP	
	20	0	0	0	0	0	0	0	
	12.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	16	0	0	0	0	0	0	
	0.0%	10.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	0	20	0	0	0	0	0	
	0.0%	0.0%	12.5%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	4	0	20	0	0	0	0	
	0.0%	2.5%	0.0%	12.5%	0.0%	0.0%	0.0%	0.0%	
Output Class									Target Class
	AQU	BIS	DIS	EQU	HIM	KIN	RAN	TAP	
	0	0	0	20	0	0	0	0	
	0.0%	0.0%	0.0%	12.5%	0.0%	0.0%	0.0%	0.0%	
	0	0	0	0	20	0	0	0	
	0.0%	0.0%	0.0%	0.0%	12.5%	0.0%	0.0%	0.0%	
	0	0	0	0	0	20	0	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	12.5%	0.0%	0.0%	
	0	0	0	0	0	0	20	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	12.5%	0.0%	
Output Class									Target Class
	AQU	BIS	DIS	EQU	HIM	KIN	RAN	TAP	
	100	0	0	0	0	0	0	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	80	0	0	0	0	0	0	0	
	80.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	0	0	0	0	0	0	20	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	12.5%	
	0	0	0	0	0	0	0	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	

Fig. 8. Confusion matrix for water samples.

Table 2
Qualitative performance parameters for water samples.

Samples	%FPR	%FNR	%SER	%SPR
AQU	0	0	100	100
BIS	0	20	80	100
DIS	0	0	100	100
EQU	2.86	0	100	97.14
HIM	0	0	100	100
KIN	0	0	100	100
RAN	0	0	100	100
TAP	0	0	100	100

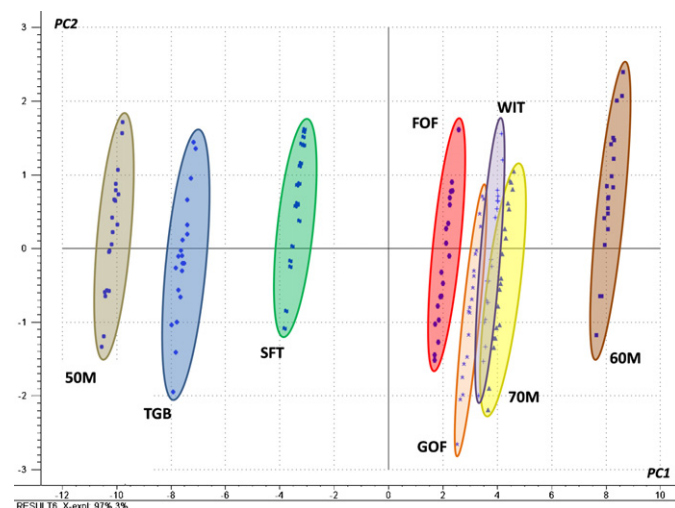


Fig. 9. PCA score plot of tea samples.

Further, Fig. 9 shows the PCA score plot of first two principal components derived from 37 apparent impedance features of tea samples. Both the Principal components account for a total of 100% of variance in the data. Overlapping clusters of GOF, WIT and 70M can be observed; hence finding classification boundaries for the classifier becomes difficult. Fig. 10 shows the confusion matrix obtained after classification using SVM and an overall

Output Class									Target Class
	50M	60M	70M	FOF	GOF	SFT	TGB	WIT	
	20	0	0	0	0	0	0	0	
	12.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	20	0	0	0	0	0	0	
	0.0%	12.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	0	13	0	0	0	0	8	
	0.0%	0.0%	8.1%	0.0%	0.0%	0.0%	0.0%	5.0%	
	0	0	0	20	0	0	0	0	
	0.0%	0.0%	0.0%	12.5%	0.0%	0.0%	0.0%	0.0%	
Output Class									Target Class
	50M	60M	70M	FOF	GOF	SFT	TGB	WIT	
	0	0	1	0	14	0	0	1	
	0.0%	0.0%	0.6%	0.0%	8.8%	0.0%	0.0%	0.6%	
	0	0	0	0	0	20	0	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	12.5%	0.0%	0.0%	
	0	0	0	0	0	0	20	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	12.5%	0.0%	
	0	0	6	0	6	0	0	11	
	0.0%	0.0%	3.8%	0.0%	3.8%	0.0%	0.0%	6.9%	
Output Class									Target Class
	50M	60M	70M	FOF	GOF	SFT	TGB	WIT	
	100	100	65.0%	100%	70.0%	100%	100%	55.0%	
	0.0%	0.0%	35.0%	0.0%	30.0%	0.0%	0.0%	45.0%	
	0	0	0	0	0	0	0	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	0	0	0	0	0	0	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
	0	0	0	0	0	0	0	0	
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	

Fig. 10. Confusion matrix for tea samples.

Table 3
Qualitative performance parameters for tea samples.

Samples	%FPR	%FNR	%SER	%SPR
50M	0	0	100	100
60M	0	0	100	100
70M	5.71	35	65	94.29
FOF	0	0	100	100
GOF	1.43	30	70	98.57
SFT	0	0	100	100
TGB	0	0	100	100
WIT	8.57	45	55	91.43

classification rate of 86.3% is obtained and a high confusion between 70M-GOF and WIT clusters is observable. However, a better classification rate may be obtained by optimal selection of impedance features and kernel as discussed in [14–16]. Table 3 shows the percentage qualitative performance parameters for the tea samples. Wherein, WIT has the highest FPR of 8.57% followed by 70M (5.71%) and GOF (1.43%) samples. Also, a similar trend can be seen for the FNR values i.e. WIT (45%), 70M (35%) and GOF (30%). Similarly, WIT has the lowest SER and SPR i.e. 55% and 91.43%, respectively, preceded by 70M (65% and 94.29%, respectively) and GOF (70% and 98.57%, respectively). Further, to find the correlations and reasons for the placement of clusters in Fig. 9, and for the ease of visualization, the biochemical data pertaining to major chemical constituents of tea, namely, gallic acid (GA), caffeine (CAF), (+)-catechin (C), (–)-epicatechin (EC), (–)-epicatechin gallate (ECG), (–)-epigallocatechin (EGC) and (–)-epigallocatechin gallate (EGCG), determined by HPLC, has been plotted on a PCA biplot (Fig. 11). This bi-plot shows the correlations between the relative concentrations of these biochemicals amongst the tea samples, the pattern of cluster formation and their vicinity are similar to Fig. 9 i.e. if we look along the PC2 axis of Fig. 11; the sequence of placement of tea clusters is almost similar to the sequence along PC1 axis of Fig. 9. In Fig. 9, two major cluster groups can be observed i.e. 50M-TGB-SFT and FOF-GOF-WIT-70M-60M on the left and right hand side of PC2, respectively. Similar cluster groups are observable in the upper and lower half of Fig. 11. The separation of these two cluster groups can mainly be attributed to the relative difference in concentrations of C, EGC and EC, as observed by the contributions of their loading vectors (red arrows) along the PC2 axis in Fig. 11. (For interpretation of the references to color in this text, the reader is referred to the web version of this article.) This suggests that, somehow the ET is majorly mapping the relative concentration of C, EGC and EC and is a matter of further investigations.

Further, it may be observed that, all the PCA plots mentioned above, pertaining to ET data, show samples of individual classes to be spread in a linear fashion. This effect may be confused with the drift associated with electrode polishing [46], which indeed is not the case here. This was confirmed by the facts that, the electrodes were duly polished, and the placement of samples in linear order is random (i.e. they do not line up in time order of repeated measurements). However, this effect could be attributed to system noise, system dynamics, and non-linearity. Indeed, this is a matter of further investigation and motivation for further design improvements.

The above results show that although, the apparent impedance features were used for pattern classification, a good classification accuracy has been achieved for single specie samples (taste samples), multi-species samples (water) and complex samples (tea). The proposed low cost instrument is capable of classifying the samples with concentration and manufacturing variability. However, calibrating the instrument for absolute impedances may take the instrument a step further for quantitative analysis as well. Above all, the instrument provides a rapid analysis with comparatively easy implementation. Since, the instrument uses an audio CODEC, any wave shape in the range of 20 Hz to 20 kHz can be programmed and analyzed for various electrochemical applications.

5. Conclusions

A novel electronic tongue (ET) using a simple audio CODEC chip has been demonstrated. This ET uses odd random phase multisine waveform, in the range of 20 Hz to 20 kHz, as an excitation signal. Fourier transform has been applied on the excitation and measured signal to calculate apparent impedance. SVM in conjunction with PCA has been used for pattern classification. A good classification accuracy has been achieved for single specie samples (taste samples), multi-species samples (water) and complex samples (tea). The proposed low cost instrument is capable of classifying the samples with concentration and manufacturing variability. The performance of proposed ET can be improved by optimally selecting the frequencies and classifier parameters. Calibrating the instrument for absolute impedances may take the proposed ET a step further for quantitative analysis as well. The instrument provides a rapid analysis with comparatively easy implementation.

Acknowledgements

The authors are thankful to the management of “The Palampur Co-Operative Tea Factory Ltd.”, Palampur, Himachal Pradesh, India for providing tea samples, Dr. Ashu Gulati, Scientist, IHBT, Palampur for her valuable guidance and Monika Singla for sample collection.

References

- [1] C. Di Natale, F. Davide, J.A.J. Brunink, A. D'Amico, Y. Vlasov, A. Legin, A. Rudnitskaya, Multicomponent analysis of heavy metal cations and inorganic anions in liquids by a non-selective chalcogenide glass sensor array, *Sensors and Actuators B: Chemical* 34 (1996) 539–542.
- [2] S. Iiyama, S. Ezaki, K. Toko, T. Matsuno, K. Yamafuji, Study of astringency and pungency with multichannel taste sensor made of lipid membranes, *Sensors and Actuators B: Chemical* 24 (1995) 75–79.
- [3] A. Legin, Y. Vlasov, A. Rudnitskaya, E. Bychkov, Cross-sensitivity of chalcogenide glass sensors in solutions of heavy metal ions, *Sensors and Actuators B: Chemical* 34 (1996) 456–461.
- [4] L. Lvova, A. Legin, Y. Vlasov, G.S. Cha, H. Nam, Multicomponent analysis of Korean green tea by means of disposable all-solid-state potentiometric electronic tongue microsystem, *Sensors and Actuators B: Chemical* 95 (2003) 391–399.
- [5] P. Ivarsson, C. Krantz-Rülcker, F. Winquist, I. Lundström, A voltammetric electronic tongue, *Chemical Senses* 30 (2005) i258–i259.
- [6] P. Ivarsson, S. Holmin, N.-E. Höjer, C. Krantz-Rülcker, F. Winquist, Discrimination of tea by means of a voltammetric electronic tongue and different applied waveforms, *Sensors and Actuators B: Chemical* 76 (2001) 449–454.

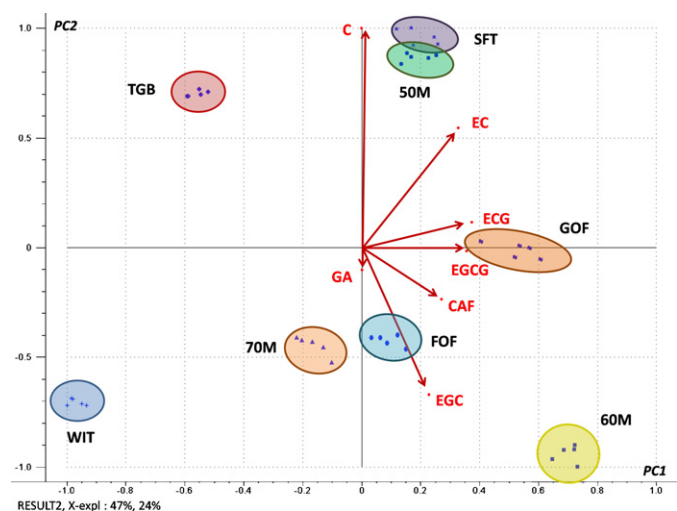


Fig. 11. PCA bi-plot of tea samples with chemical concentration loadings.

- [7] Z. Hai-Jun, D. Shao-Ping, T. Shi-Yi, A virtual smart tongue for multifrequency large amplitude pulse voltammetry, in: GCIS '09. WRI Global Congress on Intelligent Systems, 2009, pp. 8–13.
- [8] M. Scampicchio, S. Benedetti, B. Brunetti, S. Mannino, Amperometric electronic tongue for the evaluation of the tea astringency, *Electroanalysis* 18 (2006) 1643–1648.
- [9] M. Cortina-Puig, X. Muñoz-Berbel, M.A. Alonso-Lomillo, F.J. Muñoz-Pascual, M. del Valle, EIS multianalyte sensing with an automated SIA system – an electronic tongue employing the impedimetric signal, *Talanta* 72 (2007) 774–779.
- [10] G. Pioggia, F. Di Francesco, A. Marchetti, M. Ferro, A. Ahluwalia, A composite sensor array impedimetric electronic tongue: Part I. Characterization, *Biosensors and Bioelectronics* 22 (2007) 2618–2623.
- [11] A. Riul, R.R. Malmegrim, F.J. Fonseca, L.H.C. Mattoso, An artificial taste sensor based on conducting polymers, *Biosensors and Bioelectronics* 18 (2003) 1365–1369.
- [12] A. Riul, A.M. Gallardo Soto, S.V. Mello, S. Bone, D.M. Taylor, L.H.C. Mattoso, An electronic tongue using polypyrrole and polyaniline, *Synthetic Metals* 132 (2003) 109–116.
- [13] A. Riul, H.C. de Sousa, R.R. Malmegrim, D.S. dos Santos, A.C.P.L.F. Carvalho, F.J. Fonseca, O.N. Oliveira, L.H.C. Mattoso, Wine classification by taste sensors made from ultra-thin films and using neural networks, *Sensors and Actuators B: Chemical* 98 (2004) 77–82.
- [14] A.P. Bhondekar, R. Vig, A. Gulati, M. Singla, P. Kapur, Performance evaluation of a novel iTongue for Indian black tea discrimination, *IEEE Sensors Journal* 11 (2011) 3462–3468.
- [15] A.P. Bhondekar, R. Kaur, R. Kumar, R. Vig, P. Kapur, A novel approach using Dynamic Social Impact Theory for optimization of impedance-Tongue (iTongue), *Chemometrics and Intelligent Laboratory Systems* 109 (2011) 65–76.
- [16] A.P. Bhondekar, M. Dhiman, A. Sharma, A. Bhakta, A. Ganguli, S.S. Bari, R. Vig, P. Kapur, M.L. Singla, A novel iTongue for Indian black tea discrimination, *Sensors and Actuators B: Chemical* 148 (2010) 601–609.
- [17] F. Winquist, S. Holmin, C. Krantz-Rülcker, P. Wide, I. Lundström, A hybrid electronic tongue, *Analytica Chimica Acta* 406 (2000) 147–157.
- [18] M. Del Valle, Electronic tongues employing electrochemical sensors, *Electroanalysis* 22 (2010) 1539–1555.
- [19] A. Riul Jr., C.A.R. Dantas, C.M. Miyazaki, O.N. Oliveira Jr., Recent advances in electronic tongues, *Analyst* 135 (2010) 2481–2495.
- [20] T. Instruments, Stereo audio codec with usb interface, single-ended analog input/output, and s/pdif, <http://www.ti.com/lit/ds/symlink/pcm2900b.pdf> (accessed 05.03.12).
- [21] Y. Van Ingelgem, E. Tourwé, O. Blajiev, R. Pintelon, A. Hubin, Advantages of odd random phase multisine electrochemical impedance measurements, *Electroanalysis* 21 (2009) 730–739.
- [22] O.L. Blajiev, R. Pintelon, A. Hubin, Detection and evaluation of measurement noise and stochastic non-linear distortions in electrochemical impedance measurements by a model based on a broadband periodic excitation, *Journal of Electroanalytical Chemistry* 576 (2005) 65–72.
- [23] O.L. Blajiev, T. Breugelmans, R. Pintelon, A. Hubin, Improvement of the impedance measurement reliability by some new experimental and data treatment procedures applied to the behavior of copper in neutral chloride solutions containing small heterocycle molecules, *Electrochimica Acta* 51 (2006) 1403–1412.
- [24] E. Van Gheem, R. Pintelon, A. Hubin, J. Schoukens, P. Verboven, O. Blajiev, J. Vereecken, Electrochemical impedance spectroscopy in the presence of non-linear distortions and non-stationary behaviour: Part II. Application to crystallographic pitting corrosion of aluminium, *Electrochimica Acta* 51 (2006) 1443–1452.
- [25] T. Breugelmans, E. Tourwé, Y. Van Ingelgem, J. Wielant, T. Hauffman, R. Hausbrand, R. Pintelon, A. Hubin, Odd random phase multisine EIS as a detection method for the onset of corrosion of coated steel, *Electrochemistry Communications* 12 (2010) 2–5.
- [26] T. Breugelmans, E. Tourwé, J.-B. Jorcin, A. Alvarez-Pampliega, B. Geboes, H. Terryn, A. Hubin, Odd random phase multisine EIS for organic coating analysis, *Progress in Organic Coatings* 69 (2010) 215–218.
- [27] O. Olarte, W. Van Moer, K. Barbe, Y. Van Ingelgem, A. Hubin, Using random phase multisines to perform non-invasive glucose measurements, in: 2011 IEEE International Workshop on Medical Measurements and Applications Proceedings (MeMeA), 2011, pp. 300–304.
- [28] J. Schoukens, R. Pintelon, T. Dobrowiecki, Linear modeling in the presence of nonlinear distortions, *IEEE Transactions on Instrumentation and Measurement* 51 (2002) 786–792.
- [29] M. Nahvi, B.S. Hoyle, Electrical impedance spectroscopy sensing for industrial processes, *IEEE Sensors Journal* 9 (2009) 1808–1816.
- [30] V.N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, 1995.
- [31] M. Pontil, A. Verri, Support vector machines for 3D object recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (1998) 637–646.
- [32] M. Pardo, G. Sberveglieri, Classification of electronic nose data with support vector machines, *Sensors and Actuators B: Chemical* 107 (2005) 730–737.
- [33] J.A.K. Suykens, J. Vandewalle, Least squares support vector machine classifiers, *Neural Processing Letters* 9 (1999) 293–300.
- [34] R.G. Brereton, G.R. Lloyd, Support Vector Machines for classification and regression, *Analyst* 135 (2010) 230–267.
- [35] J. Luts, F. Ojeda, R. Van de Plas, B. De Moor, S. Van Huffel, J.A.K. Suykens, A tutorial on support vector machine-based methods for classification problems in chemometrics, *Analytica Chimica Acta* 665 (2010) 129–145.
- [36] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, *ACM Transactions on Intelligent Systems and Technology* 2 (2011), 27:21–27:27.
- [37] K. Higgins, J.H. Kalivas, E. Andries, Evaluation of target factor analysis and net analyte signal as processes for classification purposes with application to benchmark data sets and extra virgin olive oil adulterant identification, *Journal of Chemometrics* 26 (2012) 66–75.
- [38] S.L.R. Ellison, T. Fearn, Characterising the performance of qualitative analytical methods: statistics and terminology, *TrAC Trends in Analytical Chemistry* 24 (2005) 468–476.
- [39] A. Pulido, I. Ruisánchez, R. Boqué, F.X. Rius, Uncertainty of results in routine qualitative analysis, *TrAC Trends in Analytical Chemistry* 22 (2003) 647–654.
- [40] E. Trullols, I. Ruisánchez, F.X. Rius, Validation of qualitative analytical methods, *TrAC Trends in Analytical Chemistry* 23 (2004) 137–145.
- [41] P. Feldsine, C. Abeyta, W.H. Andrews, AOAC international methods committee guidelines for validation of qualitative and quantitative food microbiological official methods of analysis, *Journal of AOAC International* 85 (2002) 1187–1200.
- [42] J.J. O'Rangers, R.J. Condon, Promise and pitfalls of rapid contaminant analysis, in: J.F. Kay, J.D. MacNeil, J.J. O'Rangers (Eds.), *Current Issues in Regulatory Chemistry*, AOAC, International, MD, USA, 2000, p. p207.
- [43] The grading of tea, Wissotzky Tea, <http://www.wtea.com/about-tea-grading.aspx> (accessed 05.03.12).
- [44] Grading, sorting and packing, Tea Research Association, India, <http://www.tocklai.net/TeaManufacture/sorting.aspx> (accessed 05.03.12).
- [45] S. Borah, E.L. Hines, M. Bhuyan, Wavelet transform based image texture analysis for size estimation applied to the sorting of tea granules, *Journal of Food Engineering* 79 (2007) 629–639.
- [46] J. Olsson, F. Winquist, I. Lundström, A self polishing electronic tongue, *Sensors and Actuators B: Chemical* 118 (2006) 461–465.

Biographies

Ritesh Kumar is M.Tech in Advanced Instrumentation Engineering and currently working as a Scientist at CSIR-CSIO, Chandigarh. His research interests are evolutionary computational techniques, machine learning and artificial organoleptic systems.

Amol P. Bhondekar is M.Tech in Electronics and currently working as a Senior Scientist at CSIR-CSIO, Chandigarh. His research interests are in computational instrumentation and performance enhancement of artificial organoleptic systems.

Rishemjit Kaur is M.Tech in Advanced Instrumentation Engineering and currently working as a Scientist at CSIR-CSIO, Chandigarh. Her research interests are signal processing, machine learning and artificial organoleptic systems.

Saru Vig is an undergraduate trainee at CSIR-CSIO and pursuing her B.Tech in Electronics & Communications Engineering from Manipal Institute of Technology, Manipal.

Anupma Sharma is MSc in chemistry. She is presently working in the area of chromatography and tea chemistry.

Pawan Kapur is a Doctorate in Bio-medical engineering and is the Director of CSIR-CSIO. His research interests are in intelligent instrumentation, process automation and soft computing.