Compressive Sensing based Electronic Nose Platform

Hamza Djelouat*, Amine Ait Si Ali, Abbes Amira, Faycal Bensaali

College of Engineering
Qatar University
Doha, Qatar, P. O. Box: 2713

Abstract

Electronic nose (EN) systems play a significant role for gas monitoring and identification in gas plants. Using an EN system which consists of an array of sensors provides a high performance. Nevertheless, this performance is bottlenecked by the high system complexity incorporated with the high number of sensors. In this paper a new EN system is proposed using data sets collected from an in-house fabricated 4×4 tin-oxide gas array sensor. The system exploits the theory of compressive sensing (CS) and distributed compressive sensing (DCS) to reduce the storage capacity and power consumption. The obtained results have shown that compressing the transmitted data to 20 % of its original size will preserve the information by achieving a high reconstruction quality. Moreover, exploiting DCS will maintain the same reconstruction quality for just 15 % of the original size. This high quality of reconstruction is explored for classification using several classifiers such as decision tree (DT), K-nearest neighbour (KNN) and extended nearest neighbour (ENN) along with linear discrimination analysis (LDA) as feature reduction technique. CS-based reconstructed data has achieved a 95% classification accuracy. Furthermore, DCS-based reconstructed data achieved a 98.33% classification accuracy which is the same as using original data without compression.

Keywords: Compressive sensing, Distributed compressive sensing, Reconstruction algorithms, Classification, Gas Sensors

*Corresponding author
Email address: djelouat.hmz@gmail.com (Hamza Djelouat)

1. Introduction

The main breakthrough in compressive sensing (CS) paradigm was introduced by Donoho in [1] and Candès, Romberg and Tao in [2], in which they show that any signal that has a sparse representation in some basis can be recovered exactly from a small set of linear, non-adaptive measurements. This result suggests that it may be possible to sense sparse signals by taking far fewer measurements than what the famous Shannon-Nyquist theorem states [3], hence the name compressed sensing. This fewer number of measurements contain the pertinent information from the original data, after this measurements are collected, processed and transmitted, The original data can be recovered efficiently at the receiver under certain conditions.

Since its first introduction, several emerging fields have witnessed the exploitation of CS theory such as computer science, applied mathematics and medicine. Furthermore, in [4] Baron et al. introduce the theory of distributed compressive sensing (DCS) to enable new distributed coding algorithms that exploit both intra- and inter-signal correlation structures. In a typical DCS scenario, a number of sensors measure signals that are each individually sparse in some basis and also correlated from sensor to sensor. Following these results, CS and DCS has gained a lot of attention in wireless sensors network systems[5].

One of the typical wireless sensor systems are electronic nose (EN) systems. A typical EN system consists of a multi-sensor array, an information-processing unit such as an artificial neural network (ANN) and software with digital pattern-recognition algorithms. EN systems have been used in diverse applications such as spoilage detection of foodstuffs [6], disease diagnosis [7] and in the current gas industry to detect any gas/odor mixtures leakage. EN systems were firstly introduced in [8]. In [9], Victor et al. proposed an EN with tin-oxide based microarray, which can discriminate between various gases in air. Using EN for gas identification based on fingerprints obtained from gas sensors responses has been presented in [10, 11]. The problem with such EN systems is that the exposure to reactive gases for long period of time can result in a change of the gas sensor properties, which is known as the drift problem [12] and non selectivity of the sensors [13] which relates with the reactivity of a chemical sensor to so called interference gases which are different from the nominal gas towards which the sensor is targeted. The problem of the non selectivity can be overcame by using more than one sensor at a time such that each sensor shows different sensitivity or response to each gas. Guo et al. in [14], proposed a 4×4 array gas sensor in which each sensor provides a different response for the same gas. This approach can help to provide a time efficient data acquisition system as all sensors are acquiring data at the same time.

The collected data is exploited to improve the gas identification process, however, dealing with big data will increase the computational complexity [14]. Therefore an appropriate feature reduction technique is required to extract the most useful information from the data and rearrange the data for improved classification. Different feature reduction techniques have already been proposed such as multidimensional scaling [15], independent component analysis [16], principal component analysis (PCA) [17] and linear discriminant analysis (LDA) [18]. A performance evaluation and hardware implementation for PCA and LDA for gas identification using data from two different type of gas sensors was presented in [19]. The data was collected from seven commercial Figaro sensors and in-house fabricated 4×4 tin-oxide gas sensor.

In addition to feature reduction techniques, several classifiers used for pattern recognition application have been adopted for gas identification [20]. The most simplified classifiers for pattern recognition applications which can also be easily adopted on hardware are based on binary decision tree (BDT) and K-nearest neighbours (KNN), extended nearest neighbour (ENN) and committee machine (CM) which combines more than one classifier in order to improve the classification. In [13], a gas identification ensemble machine (GIEM) is presented, where five different classifiers have been used to implement the CM.

Exploiting CS theory for gas identification application has not been widely considered in the literature. However, there have been some research that are quite relevant to gas identification problems. In[21], Razzaque et al. provided a quantitative analysis of the main operational energy costs of popular sensors. where they clearly show that temperature, seismic and CO_2 signals are sparsely representable and, so, compressible, allowing CS to be effectively applied. A comparative study between CS and transform coding for wireless sensor networks based gas emission monitoring system is presented in [22], the obtained results show that CS outperform transform coding in terms of overall energy costs. Moreover, in order to optimize the power consumption in EN systems, De Vito et al. proposed in [23] an on-board processing method that allows the transmission of only the informative data packet in order to send data with the significant concentrations. This proposed system is expected to reduce the power consumption and to have a one year lifespan.

In this paper we propose a new framework for developing a CS-based EN system for gas monitoring and identification exploiting the sparsity of the different gases responses. The paper quantifies the quality of the gas data reconstruction using CS recovery algorithm as well as the usefulness of the these reconstructed data for gas identification. Distributed compressive sensing will also be investigated in order to exploit the collaboration between the sensors as all of them are measuring the same data in order to maintain the same reconstruction quality while transmitting a much fewer samples than conventional CS.

The remainder of this paper is organized as follows. Section 2 presents the EN system with detailed description of the experimental setup, data collection and the proposed CS-based EN system. Section 3 provides a mathematical background of CS, DCS and their associated reconstruction as well as a description of feature extraction, dimensionality reduction using LDA and classification using DT, KNN and ENN. In section 4, simulation results for the software implementation are presented and discussed. Section 5 concludes the paper.

2. Data Acquisition System

The experiment is conducted in a controlled lab environment containing gas chamber, cylinders of the target gases, mass flow controllers (MFCs). The 4×4 gas sensor array is installed in a gas chamber as shown in Figure 1. The gas sensors are exposed to different gases each with different concentrations. The data acquisition process is performed as follow, first, the chamber is flushed with air for 750 sec, then, the new concentration of gas is established in the chamber for the next 750 sec, resulting in measurement cycle of 1500 sec to provide a single pattern.

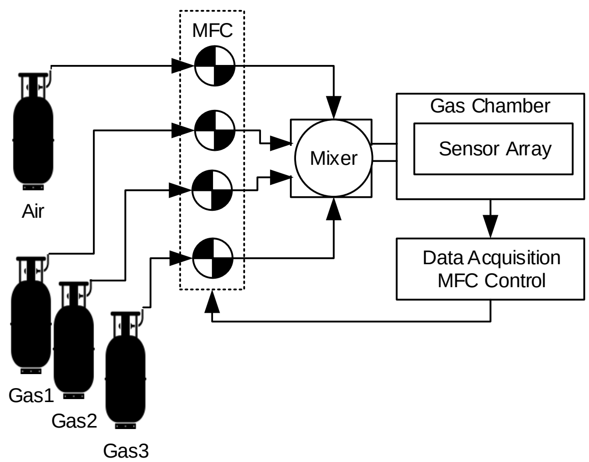


Figure 1: Data Acquisition System

Furthermore, in order to examine the behaviour of the gas sensor for different operating temperatures, the data acquisition for five most hazardous gases (C_6H_6, CH_2O, CO, NO_2 and SO_2) is performed at three different temperatures (200°C , 300°C and 400°C) where the optimal operating temperature (OT) for gas sensor is analysed in [24].

In this experiment four different concentrations for each gas have been used, Table 1 lists the ranges of different gas concentrations used:

Table 1: Gases And Their Concentration Ranges

Gas	**Concentration Range (ppm)**
C_6H_6	0.25-5
CH_2O	0.25-5
CO	5-200
NO_2	1-10
SO_2	1-25

The proposed CS based-EN system is shown in Figure2. At the data acquisition stage, a selected data is used for training and transmitted directly to the processing unit. After that, the remaining collected data will be used for testing. This latter will be processed through two main stages, compression stage and identification stages.

At the compression stage, the testing data are compressed following the theory of CS and DCS. Next, the compressed gas data are transmitted from the sensors to the processing unit.

At the processing unit, the received data are reconstructed using several recovery algorithms associated with CS and DCS. After reconstruction, several combination of feature reduction techniques and classification algorithms are used to quantify the performance of the proposed EN system in terms of classification accuracy.

3. Mathematical Overview

3.1. Compressive sensing (CS)

The data extracted by the sensors can be modelled by a matrix $\mathbf{X} \in \mathbb{R}^{N \times J}$ Matrix such that $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_j, \cdots, \mathbf{x}_J]$, where N denotes the number of samples extracted from each sensor, J is the total number of sensors used to acquire the gas data and $\mathbf{x}_j$ represents the data acquired by the j^{th} sensor.

CS allows the data to be acquired and effectively reconstructed with significantly fewer samples than its original dimension. CS relies on the sparsity

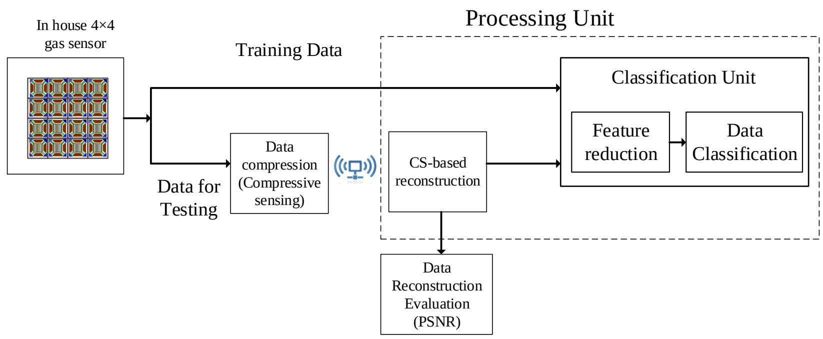


Figure 2: The proposed CS-based gas monitoring platform

and/or compressibility of the data in an appropriate transform domain. CS is generally performed by multiplying the input signal by a measurement matrix. The reconstruction of original signal from compressively sampled signal consists of finding the best solution to an underdetermined system of linear equations given by $\mathbf{y} = \mathbf{\Phi x}$. No a priori information about the original signal $\mathbf{x}$ is required for reconstruction except that it is sparse or compressible in a specific domain.

Now, Given a basis $\{\Psi_i\}_{i=1}^N$ for $\mathbb{R}^N$, we can represent every signal $\mathbf{x}_j \in \mathbb{R}^N$ in terms of N coefficient $\{s_i\}_{i=1}^N$ as $\mathbf{x}_j = \sum_{i=1}^N \Psi_i s_i$. The measurement signal $\mathbf{x}_j$ is said to be K-sparse in the basis or frame $\{\Psi_i\}$ if there exists a vector $\mathbf{S} \in \mathbb{R}^N$ with only $K \ll N$ nonzero entries such that $\mathbf{x}_j = \mathbf{\Psi S}$. The set of the indices corresponding to of the nonzero entries of $\mathbf{S}$ is called the support of $\mathbf{S}$. In general Wavelet basis provide a good sparse representation for several natural signals.

CS model for the proposed system can be written as follow:

$$\mathbf{y}_j = \mathbf{\Phi}_j \mathbf{x}_j \qquad j = 1, \cdots, J \tag{1}$$

$\mathbf{y}_j$ represents the compressed gas data to be transmitted for the j^{th} sensor such that $\mathbf{y}_j \in \mathbb{R}^M$ with $M < N$ and $\mathbf{\Phi}_j \in \mathbb{R}^{M \times N}$ is the sensing matrix that compresses the acquired N-length gas data $\mathbf{x}_j$ to the M-length compressed gas data $\mathbf{y}_j$.

Furthermore, to ensure that the data can be reconstructed efficiently with a minimum number of samples, the sensing matrix should satisfy the conditions on the restricted isometry property (RIP) and the coherence [25, 26]. In order to satisfy the RIP and the coherence conditions, several type of matrices have been well considered. The most used ones are the matrices whose entries follow a sub-Gaussian distribution[27, 28].

At the receiver, the data has to be reconstructed where the data of each sensor is processed and recovered individually. Several CS reconstruction algorithms have been proposed in the literature such as convex relaxation approach, Bayesian approach [29] and greedy algorithms. Convex relaxation approaches are based on $\ell 1$ -minimization known as basis pursuit [30] and considers the solution

$$\hat{\mathbf{x}}_j = arg\min \|\mathbf{x}_j\|_1 \quad subject \quad to \quad \mathbf{y}_j = \mathbf{\Phi}_j \mathbf{x}_j, \quad j = 1, \cdots, J \tag{2}$$

In the case when some prior knowledge about the distribution of the sparse vector is available, it would make sense to incorporate that prior knowledge into the recovery process. Bayesian methods provide a systematic framework for doing that. By making use of Bayes rule, these methods update the prior knowledge about the sparse vector in accordance with the new evidence or observations. At the same time, they suffer higher degradation in performance when prior assumptions about the signal distribution do not hold. In [31], Zayyani et al. proposed a new recovery algorithm for sparse data whose elements follow a Bernoulli-Gaussian distribution using an iterative method for to estimate the maximum a posteriori (MAP) of the sparse vector.

Greedy algorithms solve the reconstruction problem by greedily optimizing a metric that minimize the norm of the difference of the measurement vector and a residual. Matching pursuit [32], one class of CS greedy algorithms, attempts to find the columns of the measurement matrix $\mathbf{\Phi}_j$ that contribute the most in the measurement $\mathbf{y}_j$. The column(s) with the strongest correlation with the residual is (are) added to the support vector. Afterwards, their contribution is subtracted from the current residual. Several well considered greedy algorithm are of orthogonal matching pursuit OMP [33],[34], compressed sampling matching pursuit (CoSaMP) [35] and subspace pursuit algorithm (SP) [36]. The OMP pseudo code to recover the data of each sensor individually is given in Algorithm 1.

Algorithm 1: OMP

Input:

$\mathbf{y} \in \mathbb{R}^M$: measurement vector

$\mathbf{\Phi} \in \mathbb{R}^{M \times N}$: sensing matrix

ϵ: stopping criterion

Output:

$\hat{\mathbf{x}} \in \mathbb{R}^N$: reconstructed signal

Initialization:

$\Omega = \{\emptyset\}$

$\mathbf{r}^{[\mathbf{0}]} = \mathbf{y}$

current iteration i= 0

Procedure: **While** $\|\mathbf{r}^{[i]}$-$\mathbf{r}^{[i-1]}\|_2 \geq \epsilon$

1. i=i+1
2. $G^{[i]} = \Phi^* r^{[i]}$
3. $\Omega^{[i]}$= { index corresponding to the largest absolute value of $G^{[i]}$ }
4. $\Omega = \Omega \cup \Omega^{[i]}$
5. $\hat{\mathbf{x}}_T^{[i]} = \mathbf{\Phi}_\Omega^\dagger \mathbf{y}$
6. $\mathbf{r}^{[\mathbf{i}]} = \mathbf{y} - \mathbf{\Phi}_\Omega \mathbf{x}^{[i]}$

Where $\mathbf{\Phi}_\Omega^\dagger$ denotes the pseudo-inverse of $\mathbf{\Phi}_\Omega$ which is calculated as follow:

$$\mathbf{\Phi}_\Omega^\dagger = (\mathbf{\Phi}_\Omega^* \mathbf{\Phi}_\Omega)^{-1} \mathbf{\Phi}_\Omega^*$$

3.2. Distributed Compressive Sensing

CS theory as presented previously is designed mainly to exploit intra-signal structures (sparsity and compressibility) at a single sensor. However, if the system consists of a multi-sensor platform where several sensors acquire information about a physical or environmental phenomena and all the sensors measure the same data, than the different signals acquired by the different sensors are likely to share certain structures, like sparsity (in case the signal of interest is sparse).

In a multi-sensor setting, an intuitive approach is to acquire each signal and recover it separately. However, exploiting the collaboration between the sensors in the recovery process i.e., combining all of the sensors measurements to reconstruct all of their data simultaneously can result in a better reconstruction quality. This process is called joint measurement setting.

Distributed compressive sensing (DCS) presents a new distributed coding algorithm that exploits both intra- and inter-signal correlation structures of the signals. In addition, DCS requires no collaboration between the sensors during signal acquisition. Nevertheless, DCS permits to exploit the inter-signal correlation by using all of the obtained measurements to recover all the signals simultaneously, under the right conditions.

However, since multi-sensor measurement architecture described by 1 are different in real-world scenarios, different forms of correlation within an ensemble of sparse signals can occur. Thus, different setting, namely, joint sparsity models (JSMs) were introduced in [4]. Three different JSMs have been assigned to three different scenarios. In the first and the second model each of the measured signal is itself sparse whereas the third model, no signal is itself sparse, yet there still exists a joint sparsity among the signals.

In JSM-1, all signals share a common sparse component while each individual signal contains a sparse innovations component:

$$\mathbf{x}_j = \mathbf{z}_C + \mathbf{z}_j \qquad j = j, \cdots, J \tag{3}$$

Thus, the signal $\mathbf{z}_C$ is common to all of the $\mathbf{x}_j$ and the signals $\mathbf{z}_j$ are the unique portions of the $\mathbf{x}_j$. A practical situation well-modelled by JSM-1 is a group of the same sensors measuring temperatures at a number of locations throughout the day. The temperature readings $\mathbf{x}_j$ have both temporal (intra-signal) and spatial (inter-signal) correlations.

JSM-2 presents a model where all signals are constructed from the same sparse index set of basis vectors, but with different coefficients. A practical situation well-modelled by JSM-2 is where multiple sensors acquire the same signal but with different attenuation coefficients due to the unique characteristics of each sensor.

The proposed system is well modelled by JSM-2 where each of the sensors in the 4×4 gas array sensor measure the response of each of the different gases used through the experiment.

Recovery algorithms associated with DCS are variants of their CS counterparts, such as Multichannel-BPDN [37], simultaneous OMP (SOMP) and distributed compressive sensing OMP (DCS-SOMP)[4] which is the general form of SOMP. The DCS-SOMP pseudo code is provided in Algorithm 2.

3.3. Dimensionality Reduction

After data collection, the most significant features of the data have to be extracted. The most common used features are the steady states (SSs) values. SSs corresponding to all gasses and concentrations are extracted manually form the data set by taking the values corresponding to the end of each gas injection period. The extracted features can be used directly to train and test the system or can be injected to various feature reduction algorithms such as LDA.

3.3.1. Linear Discrimination Analysis

LDA is most commonly used as dimensionality reduction technique in the pre-processing step for pattern recognition and machine learning applications. The goal is to project a dataset onto a lower-dimensional space with good class-separability in order to avoid overfitting and also reduce computational costs. In addition to finding the component axes that maximize the variance between inter classes of the data and simultaneously reducing inner classes variances. LDA is interested in the axes that maximize the separation between multiple classes.

Algorithm 2: DCS-SOMP

Input:

$\mathbf{y}_j \in \mathbb{R}^M$: measurement vectors, $j = 1, \cdots, J$

$\mathbf{\Phi}_j \in \mathbb{R}^{M \times N}$: Measurement matrices, $j = 1, \cdots, J$

ϵ: stopping criterion

Output:

$\hat{\mathbf{x}}_j \in \mathbb{R}^N$: reconstructed signals foe the j^{th} sensor.

Initialization:

$\Omega = \{\emptyset\}$, i= 0

$\mathbf{r}_j^{[0]} = \mathbf{y}_j$

Procedure While $\|\mathbf{r}^{[i]} - \mathbf{r}^{[i-1]}\|_2 \geq \epsilon$

1. $i \leftarrow i + 1$
2. $\mathbf{G_j} = \Phi_j^T \mathbf{r}_j$, $j = 1, \cdots, J$
3. $\mathbf{G}_j = \sum_{j=1}^{J} |\mathbf{G}_j|$
4. $\Omega^{[i]}$= { index corresponding to the largest absolute value of $\mathbf{G}_j$}
5. $\Omega = \Omega \cup \Omega^{[i]}$
6. **for** $j = 1, \cdots, J$

   $\hat{\mathbf{x}}_{j|\Omega}^{[i]} = \mathbf{\Phi}_{j|\Omega}^{\dagger} \mathbf{y}_j$

   $\mathbf{r}_j^{[i]} = \mathbf{y}_j - \mathbf{\Phi}_j \hat{\mathbf{x}}_j^{[i]}$

   end for

The extracted training data set can be modelled by a matrix $\mathbf{T} \in \mathbb{R}^{L \times J}$ where L is the number of samples in the training data. Moreover, each row vector of $\mathbf{T}$ is assigned with a class label $\mathbf{C}_x$ forming a class label matrix $\mathbf{C} \in \mathbb{R}^L$. To perform LDA-based feature reduction on the training data, each training gas sample is assigned to a specific class. The training samples belonging to the same class form a sub-matrix $\mathbf{T}_{C_i}$ with $i = [1, \cdots, S]$ and S denotes the total number of classes.

In addition, LDA technique is also applied to the reconstructed testing data

$\hat{\mathbf{X}}$ in order to evaluate the classification performance. The pseudo code for LDA for both training and testing data is provided in Algorithm 3.

Algorithm 3: LDA Training and Testing

Input:

$\mathbf{T}$: Training data set

$\mathbf{T}_{C_i}$; training data for the i^{th} class, $i = 1, \cdots, S$

j_i :number of samples of i^{th} gas

Procedure

1. $\mu_{T_i} = [\mu_{T_1} \quad \mu_{T_2} \quad \cdots \quad \mu_{T_S}]$ $\rightarrow$ μ_{T_i}: the mean for the i^{th} class
2. $\mu = \sum_{i=1}^{S} \frac{\mu_{T_i}}{S}$ $\rightarrow$ overall average mean
3. $\mathbf{COV}_i = cov(\mathbf{T}_{C_i})$ $\rightarrow$ The within-class-variance
4. $COV = \sum_{i=1}^{S} \frac{\mathbf{COV}_i}{S}$ $\rightarrow$ average Covariance
5. $\mu\mathbf{B}_i = \mu\mathbf{T}_{C_i} - \mu$
6. $COV_{\mu\mathbf{B}_i} = j_i(\mu\mathbf{B}_i)^T(\mu\mathbf{B}_i)$
7. $\mu\mathbf{B} = \sum_{i=1}^{S} \frac{COV_{\mu\mathbf{B}_i}}{S}$
8. [Ev,Eval]= $Eig(\mu\mathbf{B}, COV)$ $\rightarrow$ Ev: Eigen vector

Output

$LDA_{training}$=$Ev \times \mathbf{T}$

$LDA_{testing}$=$Ev \times \hat{\mathbf{X}}$

3.4. Classification algorithm overview

After feature extraction and dimensionality reduction phase, the data are set to classification. The training data and the label class matrix are used as an input for several identification algorithm in order to classify the reconstructed data $\hat{\mathbf{X}}$.

3.4.1. Binary Decision Tree

BDT-based classifier is selected because of its simplicity in terms of software and hardware implementation [38]. BDT is a supervised learning technique with a set of class labelled data as the input of the learning algorithm and a binary tree as its output. The generated tree is used for the classification of a the testing data $\hat{\mathbf{X}}$. BDT training algorithm requires two inputs, the training data matrix $\mathbf{T}$ and the class label matrix $\mathbf{C}$.

3.4.2. K-Nearest Neighbour (KNN)

KNN is a non-parametric technique widely used in pattern recognition and statistical estimation to classify the unobserved data on the basis of similarity measures. KNN classifiers are based on learning from the corresponding neighbours by comparing a given test case with training samples that are similar to it [39]. KNN algorithm depends on the parameter K, this coefficient determines how many neighbours influence the classification. A sample is classified by a majority vote of its neighbours, with the sample being assigned to the class most common amongst its K nearest neighbours measured by a distance function.

Algorithm 4: KNN

Input:

$\mathbf{T} \in \mathbb{R}^{L \times J}$: Training data

$\hat{\mathbf{x}} \in \mathbb{R}^{1 \times J}$: Testing data sample

$\mathbf{C} \in \mathbb{R}^{L}$: Class label for $\mathbf{T}$

K: selected number of neighbours

for $i = 1$ **to** L **do** $\rightarrow L$: number of samples in the training data.

Compute distance $d([T]_i, \hat{\mathbf{x}})$

end for

Compute set of $\mathcal{I}$ containing indices for the K smallest distance $d(T, \hat{\mathbf{x}}_i)$

Output: Majority label for $\{\mathbf{C}_i \text{ where } i \in \mathcal{I}\}$

$$d([T]_i, \hat{\mathbf{x}}) = \|[T]_i - \hat{\mathbf{x}}_i\|_2 \tag{4}$$

Where $[\mathbf{T}]_i$ is the i^{th} row vector of $\mathbf{T}$.

3.4.3. Extended Nearest Neighbour(ENN)

ENN is a novel classification technique [40]. ENN is considered as an improved version of KNN, the key point of ENN is that it makes a prediction for a new test sample based on a 'two-way communication' style, unlike KNN where only the K-neighbours of the test sample are taken into consideration for the prediction. ENN considers the entire training data to find not only the K-nearest neighbours for the test sample, but also who are the samples from the training data set that consider the test sample as one of their K-nearest neighbours. Generalized class-wise statistics is used to achieve this.

4. Software implementation

In this section, the performance of the proposed CS-based gas monitoring presented in Figure. 2 is investigated in terms of both data reconstruction quality and classification accuracy using the reconstructed data. All the simulation presented herein are carried out using MATLAB software. The performance metrics used for the evaluation of the proposed algorithms are:

Compression Ratio

In the following the term compression ratio (CR) is used which is defined as the percentage of the number of samples in the compressed data M over the number of samples in the original data N before compression. i.e,

$$CR\ (\%) = \frac{M}{N} \times 100$$

Peak Signal-to-Noise Ratio (PSNR)

The term peak Signal-to-Noise Ratio (PSNR) is an expression for the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation [41]. The relation of PSNR can be calculated as follow:

$$PSNR = \frac{1}{J}\sum_{i=1}^{J} 20\log\frac{max\|\mathbf{x}_i\|_2}{\|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_2} \tag{5}$$

where J denotes the total number of sensors, $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$ represent the original data and reconstructed data at the i^{th} sensor, respectively.

4.1. Data Reconstruction

In this section, the performance of the CS-based gas monitoring is quantified in terms of the quality of the reconstructed data compared to the original data collected directly from the gas sensor array. Two different approaches have been adopted, the first one uses conventional CS, with OMP algorithm to be the recovery algorithm. The second approach exploits DCS theory, where DCS-SOMP is adopted to be the recovery algorithm.

The Simulations are conducted on a data set that consists of 16 sensors, each with 1800 samples, The sensing matrix is random matrix such that $\mathbf{\Phi} \sim (0, \frac{1}{M})$. The experiments results are averaged on 100 trials.

Figure 3 shows the results for data reconstruction accuracy in terms of PSNR versus the CR using OMP algorithm. Intuitively, increasing the number of samples in the compressed data improves the reconstruction quality as indicated by the increase of PSNR values (up to 60 dB using 25 % of the total samples). However, we notice a difference in reconstruction performance from gas to another, which can be explained by the fact that the responses of the sensors to each gas r exhibit a different level of sparsity. Figure 4 shows a comparison between the original CO data acquired from sensor 1 to sensor 8 with its reconstructed ones using different CR values. The data is reconstructed almost perfectly using a compression ratio of 20%.

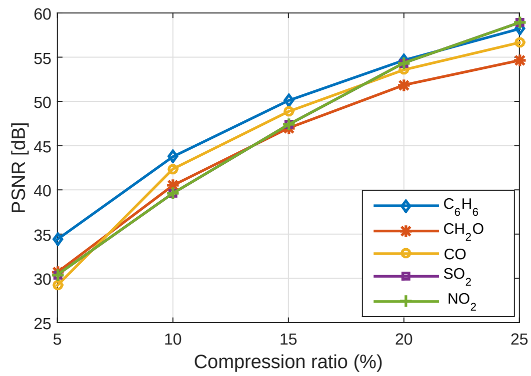


Figure 3: Reconstruction accuracy in terms of different compression ratio using OMP

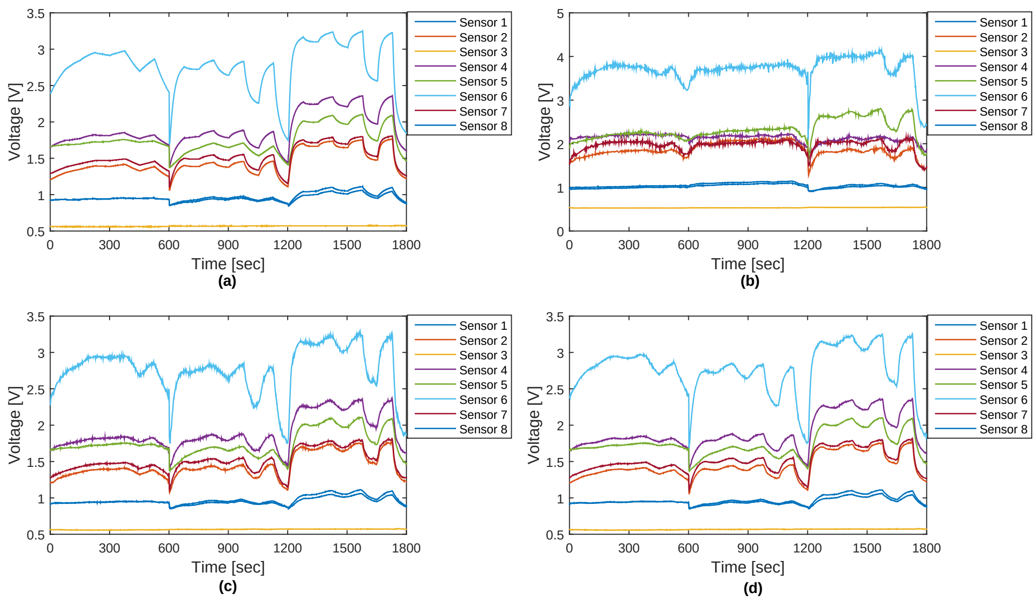


Figure 4: Comparison between the: (a) original CO data and the reconstructed ones using OMP with (b) CR=10%, (c) CR=15%, (d) CR=20%.

Now, since the signals are measurement of the responses of the same gas and as they smoothly varying in time, this causes the sensor readings to be close in value to each other, a situation well captured by the JSM-2 models.

Therefore using the same analogy as before, Figure 5 shows the attainable PSNR values versus CR exploiting DCS. The obtained results using DCS-SOMP as reconstruction algorithm consolidate the previous results regarding the quality of the reconstruction (up to 60 dB using 15% of the total samples). Comparison between original C_6H_6 data with the its reconstructed one with different CR values is shown in Figure 6. The obtained results reveals that the JSM2 models provide a good approximation for the joint sparsity structure for the proposed system and that DCS offers a promising approach for such sensing environments.

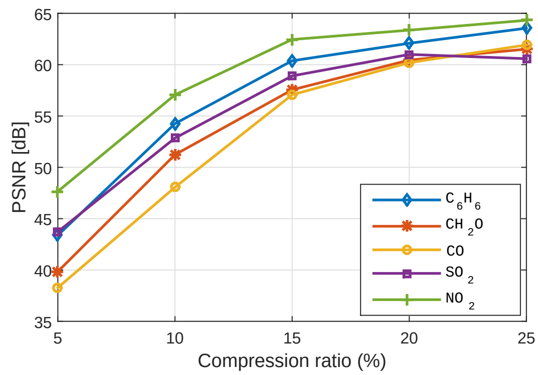


Figure 5: Reconstruction accuracy in terms of different compression ratio using DCS-SOMP

Furthermore, comparing the different obtained results, we note that exploiting the intra-signal correlation between the different sensors represented by DCS-SOMP results brings a notable enhancement on the reconstruction accuracy compared to the one using OMP to recover each sensor data individually. To clearly illustrate that, we refer to Figure 7 which presents the average PSNR values over all the used gases. The DCS-SOMP reconstruction approach outperforms the OMP whatever the CR used. Figure 8 shows a comparison between CH_2O data acquired from sensor 10 to sensor 12 with its reconstructed ones using both OMP and DCS-SOMP for CR value of 15%. The DCS recovery algorithm identifies the common structure emphasized by JSM-2, recovering salient common features for all signals. present

Moreover, the performance of standard compression techniques is quantified

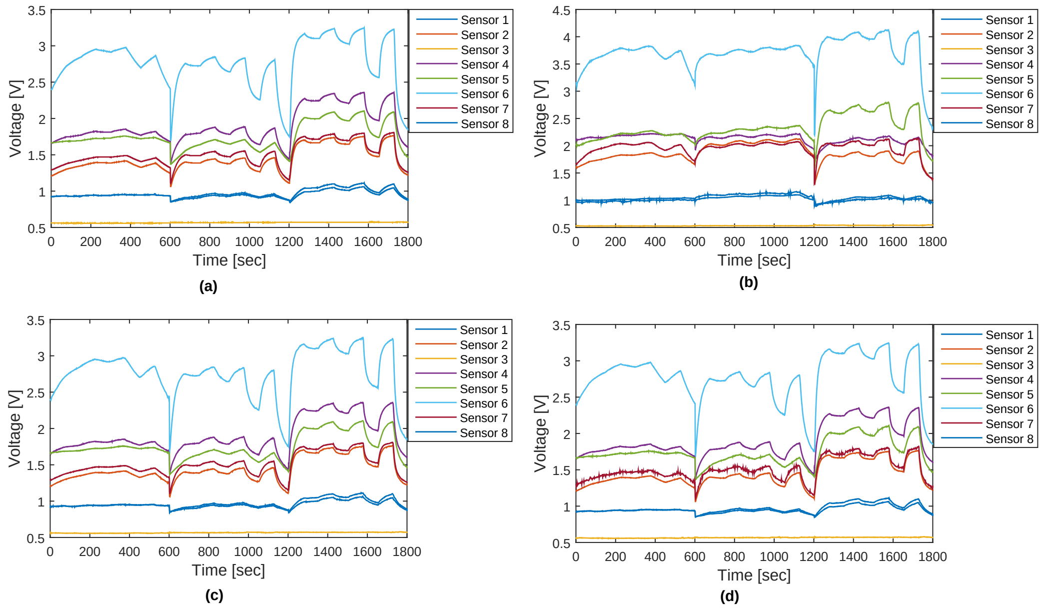


Figure 6: Comparison between the: (a) original C_6H_6 data and the reconstructed ones using DCS-SOMP with (b) CR=10, (c) CR=15, (d) CR=20

for all the gas data used in experiment. The investigated data compression techniques are based on the discrete cosine transform (DCT) and the discrete wavelet transform (DWT), it is worth mentioning that for DWT compression technique, the Daubechies wavelet family has been adopted. To compare the performance of these standard compression technique to the CS-based EN system, the same set of coefficients number to be transmitted have been selected. Figure 7 shows the performance of the different techniques in terms of the average PSNR attained for all the gases responses. The result reveals that DCT and DWT outperform OMP for a CR values in the interval $[0, 15]$. Nevertheless, increasing the number of transmitted data represented by increasing the CR values, OMP tends to achieve a remarkable performance over DCT and DWT. Moreover, introducing DCS approach for EN system presented herein, provides with no doubt the best performance in term of high quality reconstruction for the lowest number of transmitted data.

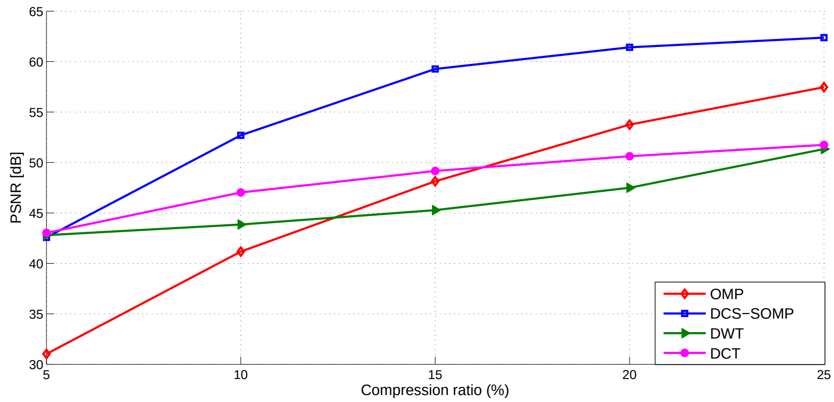


Figure 7: Comparison between the average PSNR attained for all the gases.

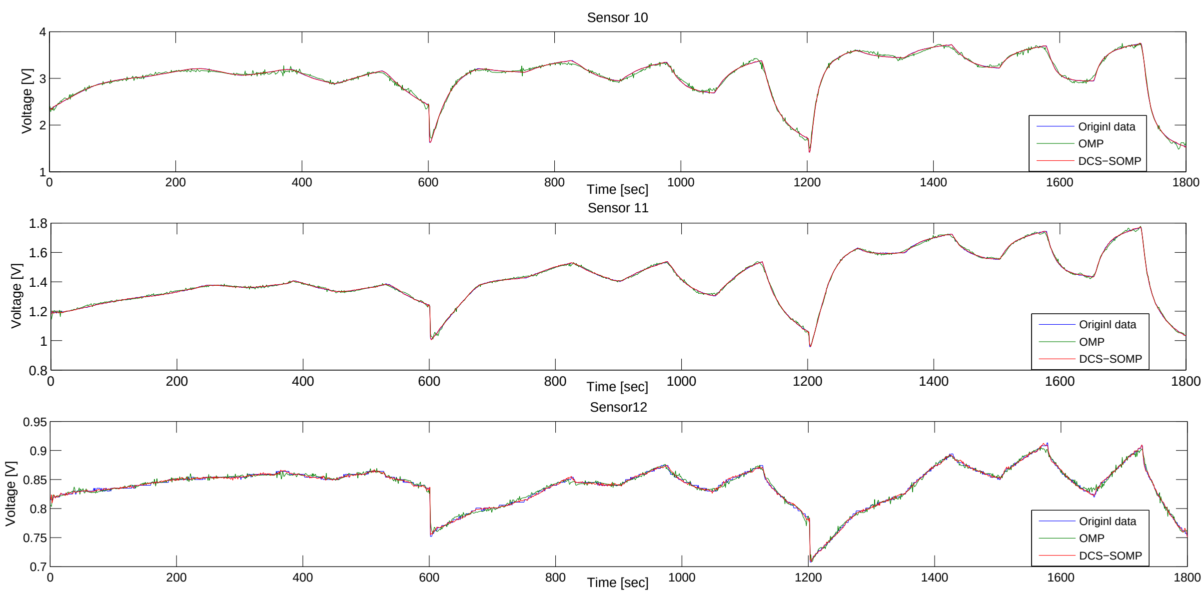


Figure 8: Comparison between the original CH_2O data and the reconstructed ones with 15% using OMP and DCS-SOMP

4.2. Data identification

The performance of the reconstructed data for classification is quantified in this section. The training data are taken directly from the output of the sensor array without performing any compression on them. However, for the testing data, they are first compressed, transmitted and then reconstructed using CS and DCS recovery algorithms.

Table 2 and Table 3 present the classification accuracy of both data reconstructed by OMP and DCS-SOMP, using KNN, ENN and BDT as classification algorithms. The data is tested after performing feature reduction technique using LDA(Table 3) and without feature reduction technique using SSs values (Table 2). the results shown in column labelled original data represents the classification accuracy obtained by using original testing data without compression. It is worth mentioning that for KNN and ENN the algorithm parameter is set to be $K = 1$ and the number of LDAs components is set to 4 (4-LDA).

The results reveal that KNN and ENN provide almost the same performance that outperforms always BDT performance regardless of the algorithm used for data reconstruction and feature selection. Moreover, the obtained results show an unacceptable classification performance when the number of samples in the compressed data is less than 10% of the original one whatever the reconstruction algorithm used which can be explained by the fact the values the recovered signal at the points where the features have to be extracted are completely different than the original values and that they hold no significant for the classification process. Using OMP and increasing the number of sample will improve the classification, attaining a maximum accuracy of 95%, yet with small error compared to the classification obtained by the original data, this improvement is due to the enhancement in the quality of the reconstructed signals as the number of transmitted samples increases. Moreover, exploring DCS theory by adopting DCS-SOMP, a much higher classification accuracy up to 98.33 % for a CR of 20% is achieved and it is equal to the one achieved by original data. Moreover, The classification accuracy obtained by applying LDA preprocessing approach to the extracted data is better than the one using the SSs values whatever the used"recovery algorithm- classifier" combination.

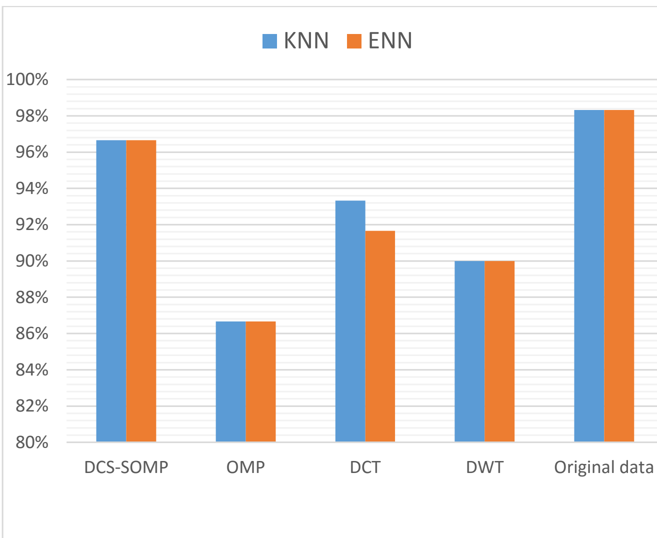
Table 2: Classification accuracy (%) results using the SSs values in terms of different compression ratio values

Classifiers technique	**Recovery approach**	Compression Ratio %				original data
		5	10	15	20	
BDT	OMP	75	75	76.33	76.66	76.66
	DCS-OMP	75	75	76.33	76.66	
KNN	OMP	71.33	86.66	91.66	91.66	95
	DCS-OMP	83.33	91.66	95	95	
ENN	OMP	85	90	91.66	91.66	95
	DCS-OMP	86.66	91.66	95	95	

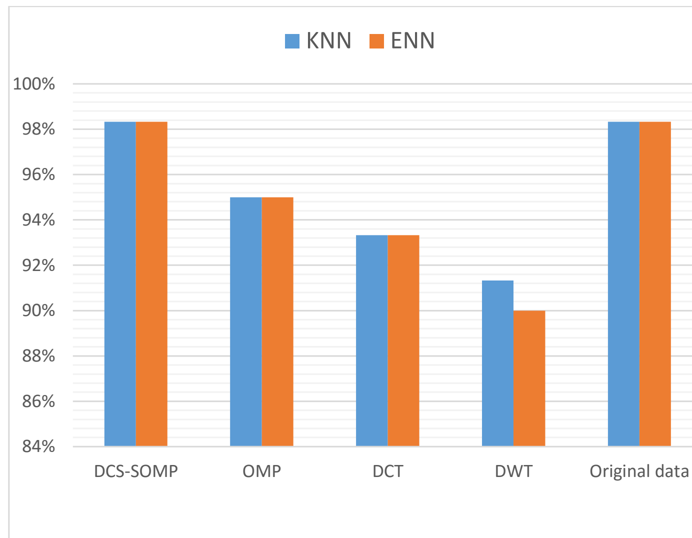
Table 3: Classification accuracy (%) results using the 4-LDA values in terms of different compression ratio values

Classifiers technique	**Recovery approach**	Compression Ratio %				original data
		5	10	15	20	
BDT	OMP	61.66	83.33	91.66	91.66	93.33
	DCS-OMP	61.66	83.33	88.33	90	
KNN	OMP	71.33	86.66	91.66	95	98.33
	DCS-OMP	71.66	96.66	96.66	98.33	
ENN	OMP	71.33	86.66	91.66	95	98.33
	DCS-OMP	76.66	96.66	96.66	98.33	

Figure 9 (a) and Figure 9 (b) present the comparison between the performance of CS-based EN system and the two compression technique used before using KNN and ENN as classification method and 4-LDA as feature reduction technique for two different values of $CR = 10\%$ and $CR = 20\%$, respectively. The results show that the identification performance exhibits a similar pattern as the reconstruction quality i.e., Using DCS-SOMP provides the best performance whatever the value of CR. In addition , using conventional compression, the classification accuracy attained outperform the one of the OMP algorithm for $CR = 10\%$. However, increasing the CR up to 20% a remarkable enhancement of the classification accuracy is achieved using OMP compared to conventional compression techniques.



(a)



(b)

Figure 9: comparison of identification performance of different approaches using:(a) $CR = 10\%$, (b) $CR = 20\%$.

5. Conclusion

This paper proposes a novel CS-based EN system for gas monitoring and identification. The performance of the proposed system has been investigated in terms of data reconstruction quality and classification accuracy. The proposed EN system exploits the theory of compressed sensing in order to provide an efficient compression scheme while maintaining good quality for the data after reconstruction. Moreover, using the multi-sensor architecture of our system, sensors collaboration has been exploited using DCS for simultaneous reconstruction. The algorithms adopted for CS and DCS are OMP and DCS-SOMP, respectively. The results show a remarkable reconstruction quality for a CR of 15% and higher. Furthermore, using DCS-SOMP shows to render a much higher reconstruction quality compared to OMP.

Furthermore, regarding classification accuracy, using the reconstructed data using DCS-SOMP outperforms the one using OMP up to 3.33% of classification accuracy which consolidates the results of data reconstruction. Moreover, transmitting just 20% of the samples while using DCS-SOMP provides a classification accuracy of 98.33% which is the the same using the original acquired data without compression. In addition the performances of the different CS approaches has been compared to the performance of standard compression techniques, the obtained results clearly shows the superiority and the performance improvement achieved using a CS-based EN system.

Acknowledgement

This paper was made possible by National Priorities Research Program (NPRP) grant No. 5-080-2-028 from the Qatar National Research Fund (a member of Qatar Foundation). Also, we would like to thank Prof. Amine Bermak (Hamad bin Khalifa University, Qatar) for providing the data used in this paper. The statements made herein are solely the responsibility of the authors.

References

[1] D. L. Donoho, Compressed sensing, Information Theory, IEEE Transactions on 52 (4) (2006) 1289–1306.

[2] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, Information Theory, IEEE Transactions on 52 (2) (2006) 489–509.

[3] M. Unser, Sampling-50 years after shannon, Proceedings of the IEEE 88 (4) (2000) 569–587.

[4] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, R. G. Baraniuk, Distributed compressive sensing, arXiv preprint arXiv:0901.3403.

[5] M. F. Duarte, V. Cevher, R. G. Baraniuk, Model-based compressive sensing for signal ensembles, in: Proceedings of the Allerton Conference on Communication, Control, and Computing, 2009, pp. 585–600.

[6] S. Marin, M. Vinaixa, J. Brezmes, E. Llobet, X. Vilanova, X. Correig, A. Ramos, V. Sanchis, Use of a ms-electronic nose for prediction of early fungal spoilage of bakery products, International journal of food microbiology 114 (1) (2007) 10–16.

[7] M. Shykhon, D. Morgan, R. Dutta, E. Hines, J. Gardner, Clinical evaluation of the electronic nose in the diagnosis of ear, nose and throat infection: a preliminary study, The Journal of Laryngology & Otology 118 (09) (2004) 706–709.

[8] V. V. Sysoev, B. K. Button, K. Wepsiec, S. Dmitriev, A. Kolmakov, Toward the nanoscopic 'electronic nose': Hydrogen vs carbon monoxide discrimination with an array of individual metal oxide nano-and mesowire sensors, Nano letters 6 (8) (2006) 1584–1588.

[9] V. V. Sysoev, J. Goschnick, T. Schneider, E. Strelcov, A. Kolmakov, A gradient microarray electronic nose based on percolating sno2 nanowire sensing elements, Nano letters 7 (10) (2007) 3182–3188.

[10] S. Bedoui, R. Faleh, H. Samet, A. Kachouri, Electronic nose system and principal component analysis technique for gases identification, in: Systems, Signals & Devices (SSD), 2013 10th International Multi-Conference on, IEEE, 2013, pp. 1–6.

[11] S. C. Das, B. Tudu, N. Bhattacharyya, R. Bandyopadhyay, P. Pramanik, Doped zno nanostructured sensor in electronic nose for detection of am-

monia, hydrogen and liquefied petroleum gas, in: Advanced Nanomaterials and Nanotechnology, Springer, 2013, pp. 475–484.

[12] A. Vergara, R. Calavia, R. M. Vázquez, A. Mozalev, A. Abdelghani, R. Huerta, E. H. Hines, E. Llobet, Multifrequency interrogation of nanostructured gas sensor arrays: a tool for analyzing response kinetics, Analytical chemistry 84 (17) (2012) 7502–7510.

[13] M. Shi, A. Bermak, S. B. Belhouari, P. C. Chan, Gas identification based on committee machine for microelectronic gas sensor, Instrumentation and Measurement, IEEE Transactions on 55 (5) (2006) 1786–1793.

[14] B. Guo, A. Bermak, P. C. Chan, G.-Z. Yan, An integrated surface micromachined convex microhotplate structure for tin oxide gas sensor array, Sensors Journal, IEEE Transactions on 7 (12) (2007) 1720–1726.

[15] N. P. Chandrasiri, M. C. Park, T. Naemura, H. Harashima, Personal facial expression space based on multidimensional scaling for the recognition improvement, in: Signal Processing and Its Applications, 1999. ISSPA'99. Proceedings of the Fifth International Symposium on, Vol. 2, IEEE, 1999, pp. 943–946.

[16] S. Z. Li, X. Lu, X. Hou, X. Peng, Q. Cheng, Learning multiview face subspaces and facial pose estimation using independent component analysis, Image Processing, IEEE Transactions on 14 (6) (2005) 705–712.

[17] P. Honeine, Online kernel principal component analysis: A reduced-order model, Pattern Analysis and Machine Intelligence, IEEE Transactions on 34 (9) (2012) 1814–1826.

[18] E. K. Tang, P. N. Suganthan, X. Yao, A. K. Qin, Linear dimensionality reduction using relevance weighted lda, Pattern recognition 38 (4) (2005) 485–493.

[19] M. A. Akbar, A. Ait Si Ali, A. Amira, F. Bensaali, M. Benammar, M. Hassan, A. Bermak, An empirical study for pca- and lda-based feature reduc-

tion for gas identification, Sensors Journal, IEEE Transactions on 16 (14) (2016) 5734–5746. doi:10.1109/JSEN.2016.2565721.

[20] S. Brahim-Belhouari, A. Bermak, Gas identification using density models, Pattern recognition letters 26 (6) (2005) 699–706.

[21] M. A. Razzaque, S. Dobson, Energy-efficient sensing in wireless sensor networks using compressed sensing, Sensors 14 (2) (2014) 2822–2859.

[22] M. A. Razzaque, M. Akhtaruzzaman Adnan, A. H. Abdullah, Energy efficient gas emission monitoring systems using wireless sensor networks, in: Ubiquitous and Future Networks (ICUFN), 2013 Fifth International Conference on, IEEE, 2013, pp. 588–593.

[23] S. De Vito, P. Di Palma, C. Ambrosino, E. Massera, G. Burrasca, M. L. Miglietta, G. Di Francia, Wireless sensor networks for distributed chemical sensing: Addressing power consumption limits with on-board intelligence, IEEE Sensors journal 11 (4) (2011) 947–955.

[24] A. Akbar, A. Ait Si Ali, A. Amira, F. Bensaali, M. Benammar, M. Hassan, A. Bermak, Impact of feature reduction and operating temperature on gas identification, ARPN Journal of Engineering and Applied Sciences 10 (19) (2015) 1–8.

[25] E. J. Candès, M. B. Wakin, An introduction to compressive sampling, Signal Processing, IEEE Transactions on 25 (2) (2008) 21–30.

[26] D. L. Donoho, X. Huo, Uncertainty principles and ideal atomic decomposition, Information Theory, IEEE Transactions on 47 (7) (2001) 2845–2862.

[27] R. Baraniuk, M. Davenport, R. DeVore, M. Wakin, A simple proof of the restricted isometry property for random matrices, Constructive Approximation 28 (3) (2008) 253–263.

[28] M. P. Pereira, L. Lovisolo, E. A. Da Silva, M. L. De Campos, On the design of maximally incoherent sensing matrices for compressed sensing

using orthogonal bases and its extension for biorthogonal bases case, Digital Signal Processing 27 (2014) 12–22.

[29] S. Ji, Y. Xue, L. Carin, Bayesian compressive sensing, Signal Processing, IEEE Transactions on 56 (6) (2008) 2346–2356.

[30] S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit, SIAM journal on scientific computing 20 (1) (1998) 33–61.

[31] H. Zayyani, M. Babaie-Zadeh, C. Jutten, An iterative bayesian algorithm for sparse component analysis in presence of noise, Signal Processing, IEEE Transactions on 57 (11) (2009) 4378–4390.

[32] S. G. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, Signal Processing, IEEE Transactions on 41 (12) (1993) 3397–3415.

[33] J. A. Tropp, A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, Information Theory, IEEE Transactions on 53 (12) (2007) 4655–4666.

[34] J. Tropp, A. C. Gilbert, Signal recovery from partial information via orthogonal matching pursuit (2005).

[35] D. Needell, J. A. Tropp, Cosamp: Iterative signal recovery from incomplete and inaccurate samples, Applied and Computational Harmonic Analysis 26 (3) (2009) 301–321.

[36] W. Dai, O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, Information Theory, IEEE Transactions on 55 (5) (2009) 2230–2249.

[37] J. A. Tropp, Algorithms for simultaneous sparse approximation. part ii: Convex relaxation, Signal Processing, IEEE Transactions on 86 (3) (2006) 589–602.

[38] Q. Li, A. Bermak, A low-power hardware-friendly binary decision tree classifier for gas identification, Journal of Low Power Electronics and Applications 1 (1) (2011) 45–58.

[39] Y.-C. Liaw, C.-M. Wu, M.-L. Leou, Fast k-nearest neighbors search using modified principal axis search tree, Digital Signal Processing 20 (5) (2010) 1494–1501.

[40] B. Tang, H. He, Enn: Extended nearest neighbor method for pattern recognition [research frontier], Computational Intelligence Magazine, IEEE Transactions on 10 (3) (2015) 52–60.

[41] Q. Huynh-Thu, M. Ghanbari, Scope of validity of psnr in image/video quality assessment, Electronics letters 44 (13) (2008) 800–801.