Digital Signal Processing 60 (2017) 350–359

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Compressive sensing based electronic nose platform

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

KINDI Center for Computing Research, Qatar University, Doha, P.O. Box: 2713, Qatar

ARTICLE INFO

Article history:
Available online 17 October 2016

Keywords:
Compressive sensing
Distributed compressive sensing
Reconstruction algorithms
Classification
Gas sensors

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.

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. Thus, 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 is an electronic nose (EN) system. 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/odour mixtures leakage. EN systems were first 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 overcome 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.

* Corresponding author.
E-mail address: djelouat.hmz@gmail.com (H. Djelouat).

http://dx.doi.org/10.1016/j.dsp.2016.10.006

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. The 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 Fig. 1. The gas sensors are exposed to different gases each with different concentrations. The data acquisition process is performed as follows: first, the chamber is flushed with air for 750 s, then, the new concentration of gas is established in the chamber for the next 750 s, resulting in measurement cycle of 1500 s to provide a single pattern. 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 analyzed in [24].

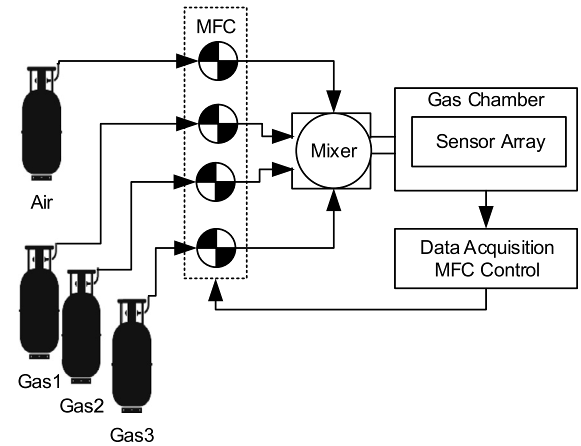

Fig. 1. Data acquisition system.

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

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

The proposed CS based EN system is shown in Fig. 2. 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. The 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 $\mathbf{X} \in \mathbb{R}^{N \times J}$ 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 jth sensor.