Data in Brief 3 (2015) 85–89

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Chemical gas sensor array dataset

Jordi Fonollosa *, Irene Rodríguez-Luján, Ramón Huerta

BioCircuits Institute, University of California, San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Received 31 December 2014
Accepted 5 January 2015
Available online 17 February 2015

Keywords:
Chemometrics
Machine olfaction
Electronic nose
Chemical sensing
Machine learning

ABSTRACT

To address drift in chemical sensing, an extensive dataset was collected over a period of three years. An array of 16 metal-oxide gas sensors was exposed to six different volatile organic compounds at different concentration levels under tightly-controlled operating conditions. Moreover, the generated dataset is suitable to tackle a variety of challenges in chemical sensing such as sensor drift, sensor failure or system calibration. The data is related to "Chemical gas sensor drift compensation using classifier ensembles", by Vergara et al. [1], and "On the calibration of sensor arrays for pattern recognition using the minimal number of experiments", by Rodriguez-Lujan et al. [2]

The dataset can be accessed publicly at the UCI repository upon citation of: http://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations

* Corresponding author.
E-mail address: fonollosa@ucsd.edu (J. Fonollosa).

http://dx.doi.org/10.1016/j.dib.2015.01.003

Specifications table

Subject area	*Chemistry*
More specific subject area	*Chemometrics, Machine Olfaction, Electronic Nose, Chemical Sensing, Machine Learning*
Type of data	*Text Files*
How data was acquired	*Metal Oxide (MOX) gas sensors provided by Figaro Inc. (TGS2600, TGS2602, TGS2610, TGS2620; four of each type) exposed to different gas conditions over a period of 36 months.*
Data format	*Processed*
Experimental factors	*For each measurement a 128-component vector is processed from the sensors' responses to extract steady-state and transient features.*
Experimental features	*Sensors were exposed to clean air before and after sample presentation.*
Data source location	*San Diego, California, US.*
Data accessibility	*Data in public repository:* http://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations *Citation of [1,2] is required.*

Value of the data

- Response of the same chemical sensor array measured consistently over a period of 36 months. Drift in sensors' sensitivity can be evaluated over time.
- Extensive dataset (13,910 measurements) generated from chemical sensors exposed to six different volatiles, each volatile presented at different concentration levels. The problem can be formulated either as a classification problem to determine which gas is present or as a regression task to determine the gas concentration levels.
- The dataset can be utilized to address sensor drift, sensor failure, system calibration, sensor poisoning, among other common challenges in chemical sensing [1–4].
- It can also be applied to concept drift, active learning, and pattern recognition in Machine Learning.
- Dataset suitable for the benchmark of different Machine Learning techniques designed for chemical sensing.

1. Experimental design, materials and methods

1.1. Experimental setup

The chemical detection platform included 16 commercially available metal-oxide gas sensors manufactured and commercialized by Figaro Inc. The sensor array had four types of sensors (four of each type) tagged as TGS2600, TGS2602, TGS2610, TGS2620. Hence, the detection platform generates a multivariate response upon exposure to different volatiles.

The operating temperature of the sensors is controlled by the voltage applied to the built-in sensors' heaters. The voltage on the heaters was kept constant at 5 V.

We placed the sensor array into a 60 ml air-tight chamber where the volatiles of interest in gaseous form were injected in random order. The test chamber was attached in series to a vapor delivery system that provided the selected concentrations of the chemical substances by means of three digital mass flow controllers and calibrated gas cylinders. The total flow rate across the sensing chamber was set to 200 ml/min and kept constant for the whole measurement process. The entire measurement system setup was fully operated by a computerized environment and provided versatility for setting the concentrations with high accuracy and in a highly reproducible manner (see Fig. 1).