Article

A Novel Sparse Representation Classification Method for Gas Identification Using Self-Adapted Temperature Modulated Gas Sensors

Aixiang He [1],* , **Guangfen Wei [1]**, **Jun Yu [2]**, **Meihua Li [1]**, **Zhongzhou Li [2]** and **Zhenan Tang [2],***

[1] School of Information & Electronic Engineering, Shandong Technology and Business University,
    Yantai 264005, China; guangfen.wei@sdtbu.edu.cn (G.W.); limeihua@sdtbu.edu.cn (M.L.)
[2] The Key Laboratory of Liaoning for Integrated Circuits Technology, Faculty of Electronic Information and
    Electrical Engineering, Dalian University of Technology, Dalian 116023, China; junyu@dlut.edu.cn (J.Y.);
    lizhongzhou127@163.com (Z.L.)
* Correspondence: sdgshax@126.com (A.H.); tangza@dlut.edu.cn (Z.T.);
    Tel.: +86-136-4645-4139 (A.H.); +86-138-9863-3170 (Z.T.)

check for
updates

Abstract: A novel sparse representation classification method (SRC), namly SRC based on Method of Optimal Directions (SRC_MOD), is proposed for electronic nose system in this paper. By finding both a synthesis dictionary and a corresponding coefficient vector, the *i*-th class training samples are approximated as a linear combination of a few of the dictionary atoms. The optimal solutions of the synthesis dictionary and coefficient vector are found by MOD. Finally, testing samples are identified by evaluating which class causes the least reconstruction error. The proposed algorithm is evaluated on the analysis of hydrogen, methane, carbon monoxide, and benzene at self-adapted modulated operating temperature. Experimental results show that the proposed method is quite efficient and computationally inexpensive to obtain excellent identification for the target gases.

---

1. Introduction

Electronic noses are technical devices that contain a gas sensor array and pattern recognition system [1]. However, the pattern recognition of electronic noses, in many cases, is plagued with problems. It is quite usual to encounter drift, scattering due to concentration effects, highly correlated features, or non-Gaussian data distributions [2,3]. In addition, due to high calibration costs and complex experimental conditions, the number of training samples is limited. Hence, the performance of classifier is very important to electronic noses, as they can improve the robustness to the problems mentioned above.

At present, there are many classification methods for gas sensor data [4–9] such as deep learning and support vector machine (SVM). Since the concept of deep learning was put forward, it has attracted the attention of many scholars [10–12]. Peng et al. proposed a novel Deep Convolutional Neural Network (DCNN) tailored for gas classification [13]. Wei et al. also proposed a new improved LeNet-5 gas identification convolutional neural network structure for electronic noses [14].

Because support vector machine (SVM) has good generalization properties and robustness against the curse of dimensionality [15], SVM has been widely applied to gas identification [16–18]. Vergara et al. used Inhibitory Support Vector Machine (ISVM) to detect and identify odor under complex environmental conditions [17]. Sakumura et al. also used SVM to detect respiratory samples, and achieved high detection accuracy [18].

Other explored classification methods are artificial neural network [19], decision tree [20], and Bayesian Networks [21], etc. For example, Aleixandre used probabilistic neural network (PNN) and multilayer perceptrons (MLP) to discriminate four different pollutant gases [22]. Cho et al. used classification and regression tree to identify acetone, nitrobenzene, octane, and nitrotoluene, and obtained 94% classification accuracy [23].

Sparse representation classification (SRC) is proved to be robust to outliers, noise, and even incomplete measurements [24–26], and some scholars have successfully used SRC to solve the problem of gas classification, such as, Guo et al. [27] who used sparse representation-based classification to identify breath samples. However, SRC is time consuming, which limits its application.

The contribution of this paper is to propose a novel sparse representation classification method, namely SRC based on Method of Optimal Directions (SRC_MOD), to improve the classification performance of electronic nose. In order to improve the learning speed, the training set is divided into some subsets according to the label of samples, and the optimal solution of synthesis dictionary submatrix and coefficient submatrix are solved by MOD. Moreover, the testing phase is separated from the training phase. The structure of the article is as follows: Section 2 describes briefly introduces the experimental setup and data collection. Section 3 analyzes the proposed SRC_MOD method. Section 4 discusses comparison results with other classifiers. Section 5 presents the conclusions.

2. Experimental Set-Up and Data Collection

2.1. The Measurement Circuit

The selectivity and sensitivity of the metal oxide gas sensors can be improved by optimizing the operation temperature of the sensors [28,29]. Martinelli designed a self-adapted temperature modulation circuit and achieved high detection accuracy. This method implements the concept of self-adapted temperature modulation, and it is based on the evidence that the sensitivity to the gas of the sensor resistance depends on the operating temperature, and, conversely, the sensitivity to the temperature depends on the gas [28].

In this paper, we proposed an improved self-adapted temperature modulated measurement circuit to improve the performance of the electronic nose. The measurement circuit is shown in Figure 1. It mainly contains a multivibrator circuit, three gas sensors and a comparator C_1. The resistances of sensor 1 and sensor 2 are part of the multivibrator circuit. V_{REFL} and V_{REFH} represent low and high reference voltage of heating voltage, respectively. In this paper, V_{REFL} = 2 V and V_{REFH} = 5 V.
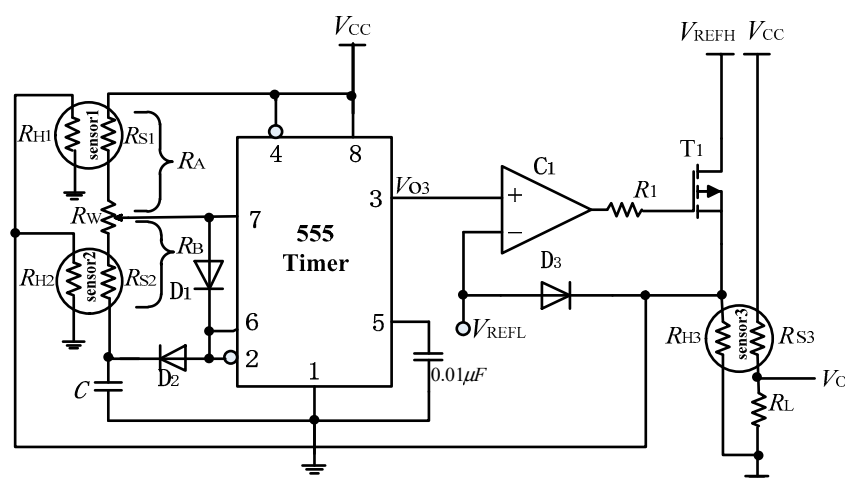
Figure 1. The measurement circuit. Sensor 1 is TGS2610, sensor 2 is TGS2610, and sensor 3 is TGS2620. V_{REFL} and V_{REFH} represent low and high reference voltage of heating voltage, respectively.

We can see from Figure 1 that if the output voltage V_{O3} of the 555 timer is low ($V_{O3} \approx 0$ V), $V_{O3} < V_{REFL}$, the output voltage of comparator C_1 is low, and the transistor T_1 will be cut off. When this happens, the heating voltage across sensor 3 or the three sensors? will be close to V_{REFL}. On the other hand, if the output voltage V_{O3} is high ($V_{O3} \approx V_{CC}$), $V_{O3} > V_{REFL}$, the output voltage of comparator C_1 is high, the transistor T_1 will be turned on. When this occurs, the heating voltage across sensor 3 or the three sensors? will be close to V_{REFH}.

V_{O3} is a square wave signal whose period depends on the capacitor C and the sensor resistances of sensor 1 and sensor 2. The charging time of capacitor C is given by:

$$T_1 = (R_A + R_B)C \ln \frac{V_{CC} - V_{T-}}{V_{CC} - V_{T+}} \tag{1}$$

where, $V_{T+} = \frac{2}{3}V_{CC}$, $V_{T-} = \frac{1}{3}V_{CC}$.

And the discharge time of the capacitor C is given by:

$$T_2 = R_B C \ln \frac{0 - V_{T+}}{0 - V_{T-}} \tag{2}$$

hence, the period of square wave signal is:

$$T = T_1 + T_2 = (R_A + 2R_B)C \ln 2 \tag{3}$$

and the duty cycle of the square wave signal V_{O3} is:

$$q = \frac{R_A}{R_A + R_B} \tag{4}$$

The output voltage of the third sensor 3 is given by:

$$V_O = \frac{R_L}{R_{S3} + R_L} V_{CC} \tag{5}$$

where, R_{S3} is the sensor resistor of the third TGS2620 sensor.

2.2. The Experimental Set-Up

Figure 2 shows the experimental set-up. The testing system uses two computer-controlled, digital mass flow controllers (MFCs). The testing gas at the desired concentration is conveyed to a 300 mL volume testing chamber by MFCs with highly reproducibility and higher accuracy. We keep the total flow constant for each test. In this paper, the total flow rate is set to 500 sccm.
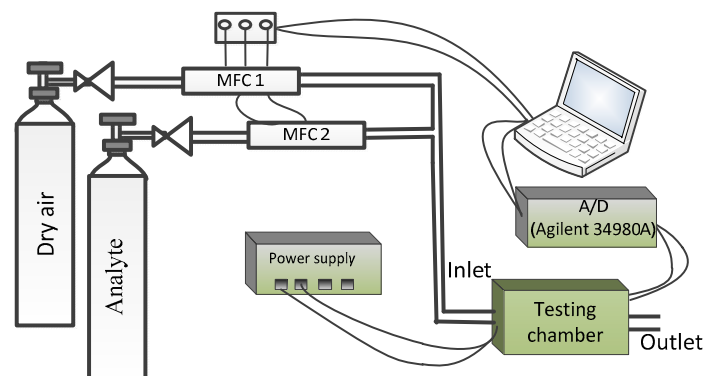

Figure 2. The experimental set-up.

The above-mentioned measurement circuit that contains three gas sensors (TGS2610, TGS2610, and TGS2620, Figaro, Inc. Japan) is placed into the testing chamber. In order to collect all experimental samples, a LabVIEW environment program running on a PC platform, and the sample frequency is set to 1 Hz.

The measurement procedure is as follows:

(1)   Clean testing chamber with dry air for 50 s.
(2)   The testing gas at the desired concentration is conveyed to the testing chamber by MFCs for 100 s.
(3)   Clean the testing chamber with dry air for 100 s.

When the three sensors return to baseline steady-state response, repeat step 1 to 3 for the next test until all the experiments are completed.

2.3. Data Collection

Four chemical analytes with different concentrations are tested by the electronic nose system. As shown in Table 1, the tested gases are hydrogen, methane, carbon monoxide, and benzene. Each test is repeated 20 times, and finally 400 samples are collected.

Table 1. Dataset in detail.

Label	Analyte	Concentration (ppm)	Number
1	hydrogen	1000, 2000, 3000, 4000, 5000	100
2	methane	1000, 2000, 3000, 4000, 5000	100
3	carbon monoxide	100, 200, 300, 400, 500	100
4	benzene	10, 15, 20, 25, 30	100
Total			400

Figure 3 shows the heating voltage of 30 ppm benzene and the output voltages of four analytes. We can see from Figure 3a, the frequency of heating waveform in the middle is higher than that on both sides. The reason for this phenomenon is related to the change of resistances of sensor 1 and sensor 2. The reducing gas is injected into the testing chamber from the time of 51 to 150 s, which leads to the decrease of sensor resistances and the increase of waveform frequency. In this paper, setting $C = 100$ μF, the periods T of the heating voltage range from 2 to 25 s and the frequencies range from 40 to 500 mHz.

(a) (b)
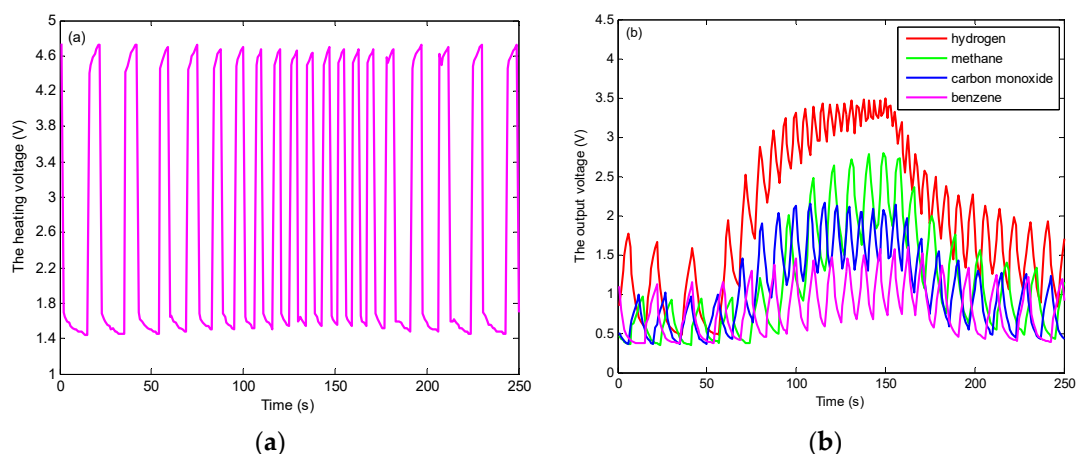
Figure 3. (**a**) The heating voltage of 30 ppm benzene; (**b**) The output voltages of 3000 ppm hydrogen, 3000 ppm methane, 300 ppm carbon monoxide, and 30 ppm benzene.

In this paper, 40 samples are chosen as a testing set, and the other samples for training and validation. In order to improve robustness of the algorithm, the classifiers use 10-fold cross-validation

method. 360 samples are randomly divided into 10 subsets with equal size. A single subset is retained as the validation set, and the remaining nine subsets are used as training set. Hence, the number of validation samples is 36 and the number of training samples is 324. The program runs 10 times, with each of the 10 subsets used exactly once as the validation set. The prediction model with the highest recognition rate is used as the final model to identify the testing gas.

3. The Proposed SRC_MOD Algorithm for Gas Identification

Suppose $V_O \in R^{250 \times 1}$ denotes a sensor signal that is a time-based variable and in total has 250 points. Firstly, the sensor siganl is removed additive noise or drift by $x = V_O - V_{ref}$, where, V_{ref} is baseline steady-state output voltage in dry air. Then, the sensor sample x is normalized by $x = (x - \min(x))/(\max(x) - \min(x))$, where, min() and max() denote the sample minimum and maximum value.

The normalized training sample dataset is represented by a matrix $X = [X_1, X_2, \cdots, X_n] \in R^{250 \times 324}$, where, $X_i, i = 1, 2, \cdots, n$, is a submatrix of training sample corresponding to class i and each column is a sensor sample, n is total number of categories.

The ith class training samples X_i is approximated as a linear combination of some few of the dictionary atoms. The approximation X_i^* can be written as:

$$X_i^* = D_i W_i \quad s.t. \quad \min\|W_i\|_0 \tag{6}$$

where, $\|\cdot\|_0$ is l_0-norm, W_i is the coefficients of the ith class training samples and most of the entries in W_i are zero, D_i is a synthesis dictionary corresponding to class i. Equation (6) describes each given sensor signal as the sparsest representation W_i over the synthesis dictionary D_i, and aims to jointly find the proper representations and the dictionary. If a solution has been found such that every representation has fewer non-zero entries, a candidate feasible model has been found. In this paper, the synthesis dictionary D_i is initialized as a random matrix with the size of 250×250.

Equation (6) can be formulated as an optimization problem with respect to W_i and D_i. With γ, we may put it as:

$$\{D_i, W_i\} = \arg\min_{D_i, W_i} (\|X_i - D_i W_i\|_2^2 + \gamma\|W_i\|_0) \tag{7}$$

As γ increases, the solution is getting more dense. Solutions of Equation (7) can be found by the Method of Optimal Directions (MOD). MOD is a dictionary learning algorithm [30]. It's aim is to find both a dictionary D_i and a corresponding coefficient matrix W_i such that the representation error $R = X_i - D_i W_i$ is minimized and W_i fulfill some sparseness criterion. The procedures of obtaining the optimal solution of D_i and W_i are summarized in Algorithm 1.

Algorithm 1. Obtain the optimal solution of D_i and W_i

Input: The i-th class training samples X_i, maximum error ε, $k = 1$.

  1. Initialize dictionary $D_i^{(0)}$ and $W_i^{(0)}$ with two random matrices.

  2. *while* (*error* > ε)

    update $D_i^{(k+1)} = X_i (W_i^{(k)})^T (W_i^{(k)} (W_i^{(k)})^T)^{-1}$;

    update $W_i^{(k+1)} = \arg\min_{W_i} (\|X_i - D_i^{(k)} W_i^{(k)}\|_2^2)$ by pursuit algorithm [31];

    error = $\|X_i - D_i^{(k)} W_i^{(k)}\|_F^2$;

    $k \leftarrow k + 1$;

    end while

Output: The i-th class synthesis dictionary D_i and coefficient matrix W_i.

For $\forall i$, the training samples X can be projected onto a coding coefficient space via $P_i X$, where P_i is an analysis dictionary corresponding to class i. The coding coefficient matrix W_i is given by:

$$W_i = P_i X_i \tag{8}$$

where, P_i is a full-rank matrix. If most of large coefficients generated by $P_i X$ are concentrated in W_i, while the coding coefficients of the other class training samples over P_i is as small as possible, the discrimination power of P_i can be promoted. Hence, we may improve the discrimination power of P_i by min $\|P_i \overline{X}_i\|_F^2$, where, $\overline{X}_i$ is the complementary data matrix of X_i in the whole training set X and $\| \cdot \|_F$ is the Hilbert-Schmidt norm or the Frobenius norm.

We evaluate the error using a Frobenius norm. The i-th class analysis dictionary P_i can be approximated by

$$P_i = \underset{P_i}{\operatorname{argmin}}(\|W_i - P_i X_i\|_F^2 + \alpha\|P_i \overline{X}_i\|_F^2 + \beta\|P_i\|_F^2) \tag{9}$$

where, the first term minimizes the error of the ith class coding coefficients, the second term is used to improve the recognition performance of the analysis dictionary P_i, and α is a scalar constant. The third term is to avoid a high risk of overfitting to training samples, and β is a regularization parameter.

We then solve for P_i by least-squares. Differentiating Formula (9) with respect to P_i, such a differentiation results in:

$$\begin{aligned} &\frac{\partial}{\partial P_i}(\|W_i - P_i X_i\|_F^2 + \alpha\|P_i \overline{X}_i\|_F^2 + \beta\|P_i\|_F^2) \\ &= 2(P_i X_i - W_i)X_i^T + 2\alpha P_i \overline{X}_i \overline{X}_i^T + 2\beta P_i \end{aligned} \tag{10}$$

Setting Equation (10) equal to zero gives the optimum P_i as:

$$P_i = (W_i X_i^T)(X_i X_i^T + \alpha \overline{X}_i \overline{X}_i^T + \beta I)^{-1} \tag{11}$$

Since $W_i = P_i X_i$, Equation (6) can also be given by

$$X_i^* = D_i P_i X_i \tag{12}$$

Define $\Phi_i = D_i P_i$ as a projection matrix, and the approximation X_i^* is then rewritten as

$$X_i^* = \Phi_i X_i \tag{13}$$

Using Equation (13), an arbitrary testing sample x_{test} can be reconstructed as:

$$x_i^* = \Phi_i x_{test} \tag{14}$$

We obtain n approximations $x_1^*, x_2^*, \cdots, x_n^*$, and calculate the residual between x_{test} and x_i^* by

$$Label(x_{test}) = \underset{i}{\operatorname{argmin}}\|x_i^* - x_{test}\|_2^2 \tag{15}$$

where, $\| \cdot \|_2$ is l_2-norm. The label corresponding to the minimum residual is the class of the testing sample.

Figure 4 shows the original testing sample x_{test} and four reconstrcted samples x_i^*, $(i = 1, 2, 3, 4)$. x_i^* is obtained by Formula (14). From Figure 4, we can see that x_2^* is closest to the original testing sample x_{test} The residuals between x_{test} and x_i^*, $(i = 1, 2, 3, 4)$ are represented in Figure 5. From Figure 5, we can see that the second residual is the smallest. Hence, the testing sample x_{test} is the 2nd class, namely methane.
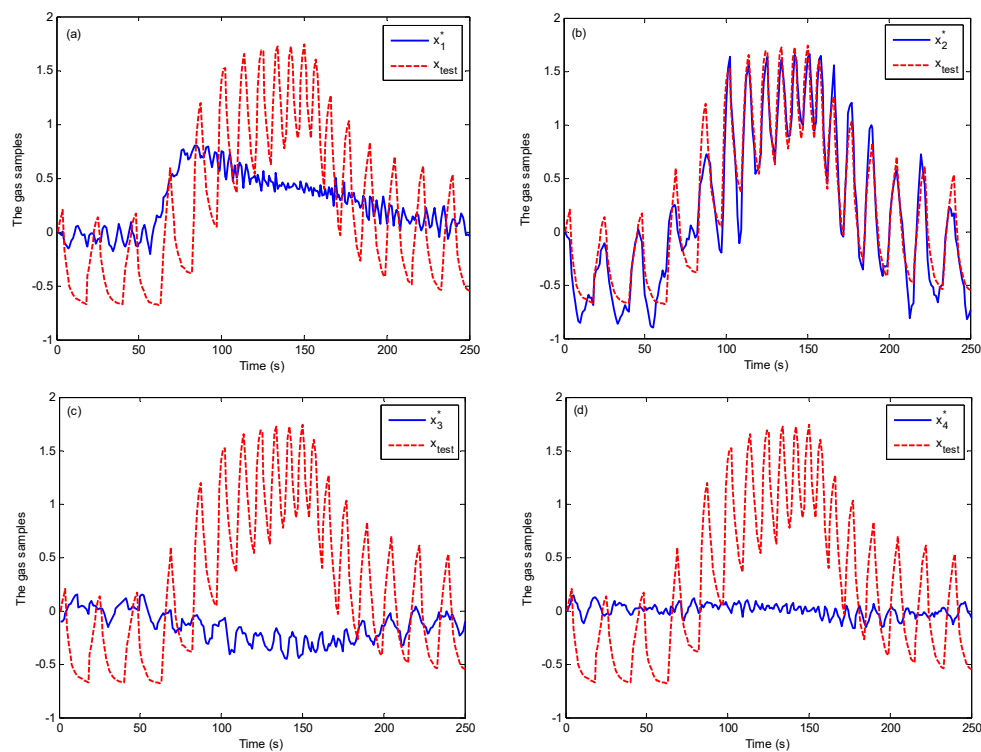
Figure 4. The original testing sample and four reconstrcted samples for 1000 ppm methane (**a**) The original testing sample x_{test} and the 1th class reconstrcted sample x_1^*; (**b**) The original testing sample x_{test} and the 2th class reconstrcted sample x_2^*; (**c**) The original testing sample x_{test} and the 3th class reconstrcted sample x_3^*; (**d**) The original testing sample x_{test} and the 4th class reconstrcted sample x_4^*.
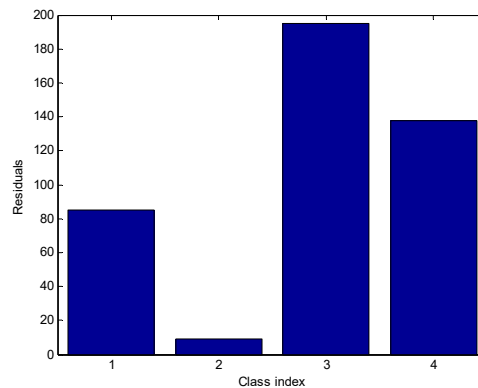
Figure 5. Residuals between the original sample and reconstructed samples.

The proposed SRC_MOD algorithm is summarized in Algorithm 2.

Algorithm 2. The proposed SRC_MOD algorithm

Input: The training samples for *n* classes, testing samples.
    1. *for i = 1:n*
        obtain D_i and W_i by Algorithm 1;
        obtain P_i by Equation (11);
    end for
    2. $\Phi_i = D_i P_i$
    3. Reconstructing the testing sample by Equation (14)
    4. To identify the testing sample by Equation (15)
Output: The label of the testing sample.

---

4. Comparisons with Other Classifiers

In order to evaluate the performance of the proposed algorithm, we compare it with other algorithms, such as SRC (used in [27]), the dictionary learning (DL) classifier (proposed in [32]), deep learning (used in [14]) and BP artificial neural network. All experiments in this paper run on a dual-core processor with a CPU main frequency of 2.4 GHz. Python is applied for deep learning and MATLAB for the other algorithms. Accuracy and processing time are the average values of all testing samples.

For SRC, the algorithm is the same as that in [27]. At first, the testing sample is approximated as a linear combination of all training samples

$$x_{test} = WX_{train} \tag{16}$$

where, W is a sparse coefficient vector, X_{train} is a matrix of all training samples.

The sparsest solution of Equation (16) is defined as the following an l_1-minimization problem:

$$W = \underset{W}{\operatorname{argmin}}(\|WX_{train} - x_{test}\|^2 + \lambda\|W\|_1) \tag{17}$$

where, $\|\cdot\|_1$ is l_1-norm, λ is the regularization parameter. The solution to the l_1-minimization problem can also be obtained by using the MATLAB package provided by Reference [33]. Keep the ith class coefficients and set the other coefficients equal to zero. We have

$$W_i^* = \{0, 0, W_i, \cdots, 0\} \tag{18}$$

Reconstruct the testing sample by $x_i^* = W_i^* X_{train}$, $(i = 1, 2, 3, 4)$ and use Equation (15) to obtain the class of the testing sample.

For DL classifier, the i-th class analysis dictionary, coding coefficients, and synthesis dictionary are trained together to generate the prediction model. It is a simple and effective dictionary learning algorithm. This is our previous work and more details of the algorithm are shown in [32].

For deep learning, as shown in Figure 6, a LeNet-5 convolutional neural network structure is built. C1 and C3 are convolutional layers with kernel size of 3×3 and 2×2, respectively. C1 computes 20 filters over its input. The first convolutional layer C1 takes a matrix with size of $25 \times 10 \times 1$ and outputs a matrix with size of $25 \times 10 \times 20$. Pooling layer P2 and P4 are all done with 2×2 windows. Max pooling consists of extracting windows from the input features and outputting the max value of each channel. Before the first max-pooling layer P2, the feature map is 25×10, but the pooling operation halves it to 12×5. The numbers of fully connected layer F5 and F6 are 120 and 84, respectively. Finally, the label of gas sample is obtained. This is also our previous work and more details are shown in [14].
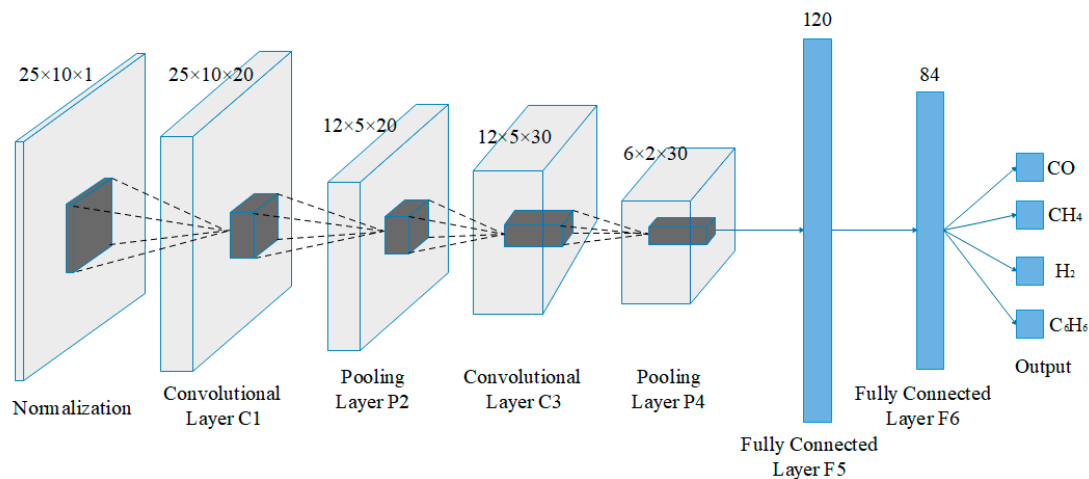

Figure 6. The LeNet-5 structure.

For the BP artificial neural network, the structure of the BP network is set to 250-501-4. The transfer function 'tansig' is applied for hidden layer, and 'purelin' for output layer.

The experimental results of the proposed SRC_MOD method are shown in Table 2. From Table 2, we can see that the average accuracy of SRC_MOD is about 98.44%, the average training time is nearly 0.2061 s and the average testing time is nearly 3.1 ms. SRC_MOD classifier obtains high accuracy in short testing time. Hence, the performance of the SRC_MOD classifier is perfect.

Table 2. Performance comparison with other algorithms.

Algorithm	Accuracy (%)	Training Time (s)	Testing Time (ms)
SRC_MOD	98.44	0.2061	3.1
SRC	98.52	no need	1987.9
DL	96.88	0.8061	6.5
Deep learning	91.87	12.84	23.5
BP	84.51	6.399	12.7

The experimental results of other classifiers are also shown in Table 2. The average accuracy of SRC is about 98.52% and the average testing time is nearly 1987.9 ms. The testing time is 641 times longer than that of SRC_MOD. The main reason is that training and testing of SRC are conducted simultaneously, solving l_1-minimization problem is time consuming, repeatedly training for each test leads to longer processing time.

The average accuracy of DL classifier is about 96.88% and the testing time is nearly 6.5 ms. We find that SRC_MOD method is superior to the DL classifier in both recognition accuracy and testing time. Since alternating direction method of multipliers(ADMM) is more complex than MOD when obtaining the optimal solution of D_i.

The average accuracy of deep learning is about 91.87%, the average training time is nearly 12.84 s and testing time is nearly 23.5 ms. The performance of deep learning significantly worse than SRC_MOD method. The main reason is that deep learning is more suitable for large training sets. In this experiment, the size of training samples is too small to show its advantage. For BP artificial neural network, the performance significantly worse than the other classifiers.

In a word, comparison results show that the proposed SRC_MOD method is quite efficient and computationally inexpensive to obtain excellent identification for the target gases.

5. Conclusions

This paper presents a SRC_MOD gas recognition algorithm. First the *i*-th class training samples are approximated as a linear combination of the synthesis dictionary atoms. Next, MOD is applied to solve the optimal solution of the synthesis dictionary and coefficient vector. Finally, we obtain the analysis dictionary and establish the prediction model. Compared with other classical classifiers (such as SRC, dictionary learning classifier, deep learning and BP artificial neural network), the experimental results show that SRC_MOD has better performance, not only in recognition rate but also in testing speed.

References

1.　Persaud, K.; Dodd, G. Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. *Nature* **1982**, *299*, 352. [CrossRef]

2.　Hierlemann, A.; Gutierrez-Osuna, R. Higher-order chemical sensing. *Chem. Rev.* **2008**, *108*, 563–613. [CrossRef]

3.　Marco, S.; Gutierrez-Galvez, A. Signal and data processing for machine olfaction and chemical sensing: A review. *IEEE Sens. J.* **2012**, *12*, 3189–3214. [CrossRef]

4.　Pardo, M.; Sberveglieri, G. Comparing the performance of different features in sensor arrays. *Sens. Actuators B Chem.* **2007**, *123*, 437–443. [CrossRef]

5.　Boeker, P. On 'electronic nose'methodology. *Sens. Actuators B Chem.* **2014**, *204*, 2–17. [CrossRef]

6.　Alizadeh, T. Chemiresistor sensors array optimization by using the method of coupled statistical techniques and its application as an electronic nose for some organic vapors recognition. *Sens. Actuators B Chem.* **2010**, *143*, 740–749. [CrossRef]

7.　Llobet, E.; Brezmes, J.; Gualdron, O.; Vilanova, X.; Correig, X. Building parsimonious fuzzy ARTMAP models by variable selection with a cascaded genetic algorithm: Application to multisensor systems for gas analysis. *Sens. Actuators B Chem.* **2004**, *99*, 267–272. [CrossRef]

8.　Gualdrón, O.; Brezmes, J.; Llobet, E.; Amari, A.; Vilanova, X.; Bouchikhi, B.; Correig, X. Variable selection for support vector machine based multisensor systems. *Sens. Actuators B Chem.* **2007**, *122*, 259–268. [CrossRef]

9.　Yang, Y.S.; Ha, S.-C.; Kim, Y.S. A matched-profile method for simple and robust vapor recognition in electronic nose (E-nose) system. *Sens. Actuators B Chem.* **2005**, *106*, 263–270. [CrossRef]

10.　LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [CrossRef]

11.　Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]

12.　Erhan, D.; Bengio, Y.; Courville, A.; Manzagol, P.A.; Vincent, P.; Bengio, S. Why does unsupervised pre-training help deep learning? *J. Mach. Learn. Res.* **2010**, *11*, 625–660.

13.　Peng, P.; Xiaojin, Z.; Xiaofang, P. Gas Classification Using Deep Convolutional Neural Networks. *Sensors* **2018**, *18*, 157. [CrossRef]

14.　Wei, G.; Li, G.; Zhao, J.; He, A. Development of a LeNet-5 Gas Identification CNN Structure for Electronic Noses. *Sensors* **2019**, *19*, 217. [CrossRef]

15.　Cortes, C.; Vapnik, V. Support vector machine. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

16.　Brudzewski, K.; Osowski, S.; Markiewicz, T. Classification of milk by means of an electronic nose and SVM neural network. *Sens. Actuators B Chem.* **2004**, *98*, 291–298. [CrossRef]

17.　Vergara, A.; Fonollosa, J.; Mahiques, J.; Trincavelli, M.; Rulkov, N.; Huerta, R. On the performance of gas sensor arrays in open sampling systems using Inhibitory Support Vector Machines. *Sens. Actuators B Chem.* **2013**, *185*, 462–477. [CrossRef]

18.　Sakumura, Y.; Koyama, Y.; Tokutake, H.; Hida, T.; Sato, K.; Itoh, T.; Akamatsu, T.; Shin, W.; Seitz, W.R. Diagnosis by volatile organic compounds in exhaled breath from lung cancer patients using support vector machine algorithm. *Sensors* **2017**, *17*, 287. [CrossRef]

19.　Shahid, A.; Choi, J.-H.; Rana, A.U.H.S.; Kim, H.-S. Least Squares Neural Network-Based Wireless E-Nose System Using an SnO2 Sensor Array. *Sensors* **2018**, *18*, 1446. [CrossRef]

20.　He, A.; Yu, J.; Wei, G.; Chen, Y.; Wu, H.; Tang, Z. Short-Time Fourier Transform and Decision Tree-Based Pattern Recognition for Gas Identification Using Temperature Modulated Microhotplate Gas Sensors. *J. Sens.* **2016**, *2016*, 7603931. [CrossRef]

21.　Li, C.; Heinemann, P.; Sherry, R. Neural network and Bayesian network fusion models to fuse electronic nose and surface acoustic wave sensor data for apple defect detection. *Sens. Actuators B Chem.* **2007**, *125*, 301–310. [CrossRef]

22.　Aleixandre, M.; Sayago, I.; Horrillo, M.; Fernández, M.; Arés, L.; García, M.; Santos, J.; Gutierrez, J. Analysis of neural networks and analysis of feature selection with genetic algorithm to discriminate among pollutant gas. *Sens. Actuators B Chem.* **2004**, *103*, 122–128. [CrossRef]

23.　Cho, J.; Li, X.; Gu, Z.; Kurup, P.U. Recognition of explosive precursors using nanowire sensor array and decision tree learning. *IEEE Sens. J.* **2012**, *12*, 2384–2391. [CrossRef]

24.　Yang, M.; Zhang, L.; Feng, X.; Zhang, D. Sparse representation based fisher discrimination dictionary learning for image classification. *Int. J. Comput. Vis.* **2014**, *109*, 209–232. [CrossRef]

25. Zheng, C.-H.; Zhang, L.; Ng, T.-Y.; Shiu, C.K.S.; Huang, D.-S. Metasample-based sparse representation for tumor classification. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2011**, *8*, 1273–1282. [CrossRef] [PubMed]

26. Zhang, L.; Yang, M.; Feng, X. Sparse representation or collaborative representation: Which helps face recognition? In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 471–478.

27. Guo, D.; Zhang, D.; Zhang, L. Sparse representation-based classification for breath sample identification. *Sens. Actuators B Chem.* **2011**, *158*, 43–53. [CrossRef]

28. Martinelli, E.; Polese, D.; Catini, A.; D'Amico, A.; Di Natale, C. Self-adapted temperature modulation in metal-oxide semiconductor gas sensors. *Sens. Actuators B Chem.* **2012**, *161*, 534–541. [CrossRef]

29. Gosangi, R.; Gutierrez-Osuna, R. Active temperature programming for metal-oxide chemoresistors. *IEEE Sens. J.* **2010**, *10*, 1075–1082. [CrossRef]

30. Engan, K.; Aase, S.O.; Husøy, J.H. Multi-frame compression: Theory and design. *Signal Process.* **2000**, *80*, 2121–2140. [CrossRef]

31. Tropp, J.A.; Gilbert, A.C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inf. Theory* **2007**, *53*, 4655–4666. [CrossRef]

32. He, A.; Wei, G.; Yu, J.; Tang, Z.; Lin, Z.; Wang, P. A novel dictionary learning method for gas identification with a gas sensor array. *IEEE Trans. Ind. Electron.* **2017**, *64*, 9709–9715. [CrossRef]

33. Koh, K.; Kim, S.; Boyd, S. *l*1 ls: A Matlab Solver for large-scale *l*1-Regularized Least Squares Problems. Available online: https://web.stanford.edu/~{}boyd/l1_ls/l1_ls_usrguide.pdf (accessed on 9 May 2019).