



Robust Sparse Learning Based Sensor Array Optimization for Multi-feature Fusion Classification

Leilei Zhao¹, Fengchun Tian² (✉), Junhui Qian¹, Ran Liu³,
and Anyan Jiang¹

¹ School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China

{leileizhao, junhuiq, 20161202003}@cqu.edu.cn

² Chongqing Key Laboratory of Bio-perception and Intelligent Information Processing, Chongqing University, Chongqing 400044, China

fengchuntian@cqu.edu.cn

³ College of Computer Science, Chongqing University, Chongqing 400044, China
rliu@cqu.edu.cn

Abstract. In this paper, we propose a robust sensor array optimization method based on sparse learning for multi-feature fusion data classification. The proposed approach contains three key characteristics. First, it considers the intrinsic group structure among features by combining an $\ell_{F,1}$ norm regularizer design and least squares regression framework. Second, in sensor selection, insignificant feature groups can be eliminated by grouped row sparse coefficients generated by the model, while the ϵ -dragging trick is introduced to improve the classification ability. Third, an efficient alternating iteration algorithm is presented to optimize the convex objective function. The results compared with the other classical methods on gas sensor array data sets demonstrate that the proposed method can effectively reduce the number of sensors with higher classification accuracy.

Keywords: Sensor array optimization · Sparse learning · Multiple feature fusion

1 Introduction

Gas sensor array is an important part of classical electronic nose systems, which converts different gases into different electrical signals thus enabling pattern recognition [5]. The time series signals of the sensor response are usually represented in a low-dimensional way by feature extraction [13]. However, in practice, some sensors in the array are useless for gas detection because they do not respond to the target gas or are heavily disturbed by noise, then the corresponding extracted features will be redundant for pattern recognition task. The

This work is funded by the Natural Science Foundation of China (No. 62171066).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

E. Pimenidis et al. (Eds.): ICANN 2022, LNCS 13532, pp. 176–186, 2022.

https://doi.org/10.1007/978-3-031-15937-4_15

optimization of the sensor array, i.e., the selection of the optimal sensor combination, will improve the accuracy of pattern recognition [8]. In addition, it will also reduce the complexity and cost of the subsequent system design.

Sensor array optimization is usually performed by combining feature selection methods due to their similar purpose [7, 9, 11]. In [7], three types of feature selection methods, t-statistics, Fisher's criterion and minimum redundancy maximum relevance, were used to select the most informative features, experimental results show that performance of electronic nose system was improved by 6–10%. In [9], five methods were used to optimize sensor array, then linear discriminant analysis (LDA) achieved the best experimental results, with a 9.6% increase in recognition rate while reducing the number of sensors by 10. Most of the above methods obtain the sensor importance ranking by evaluating the importance of individual features, but the best combination of features is not necessarily composed of the best features. In many scenarios, especially where multiple features are extracted for each sensor, considering the correlation among features is more valuable for selecting the optimal feature set.

Recently, sparse learning based feature selection methods have received considerable attention due to its good performance and interpretability [2–4]. In these methods, the ℓ_p norm or $\ell_{p,q}$ norm regularization terms is often used, which forces important features to have large coefficients and unimportant features to have small or zero coefficients, thus completing the selection of important features. Specially, in [4], the joint $\ell_{2,1}$ norms minimization was designed for multi-class classification problems, it selects features across all data points with joint sparsity. In [3], sparse group LASSO was used to select most informative features and thus improve the accuracy of the binary classification problem. The group structure between features is considered in group LASSO-based approaches, but the ℓ_1 norm in them is commonly used to constrain variables in vector form. The $\ell_{2,1}$ norm-based methods eliminate unimportant features by generating row sparse solutions, but does not consider the correlation between features.

Motivated by the previous works, in this paper, we propose a robust sensor array optimization method (RSAO) for multi-feature fusion data classification by combining the least squares regression framework and an $\ell_{F,1}$ norm regularizer design. When each sensor is characterized by multiple features, there are clear group structures divided by sensor category among the features. Therefore, the $\ell_{F,1}$ norm regularizer is designed to enforce unimportant feature groups to have small or zero coefficients, and then the corresponding sensors can be removed. Compared with traditional methods, the proposed method takes into account the intrinsic relevance of features and selects important sensors by directly scoring the feature groups. Besides, to improve the discriminative ability of the model, we further introduce the ε -dragging technique proposed in [12] to increase the inter-class distance. Meanwhile, an efficient alternating iteration algorithm is presented to solve the convex optimization problem. Experimental results on the gas sensor data sets show that the sensor combinations selected by the proposed method have better classification accuracy than the other conventional methods.

The rest of the paper is organized as follows. In Sect. 2, we present the proposed method and its optimization algorithm. In Sect. 3, we report experimental results. Finally, conclusions are offered in Sect. 4.

2 The Proposed Method

In this section, we propose a robust sensor array optimization method by combining an $\ell_{F,1}$ norm regularizer design and least squares regression framework. Meanwhile, an efficient iterative algorithm is presented to optimize the convex objective function.

2.1 $\ell_{F,1}$ Norm Regularization Term

Here we first summarize the common norms for vectors and matrices. For any vector $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, its ℓ_1 norm and ℓ_2 norm are

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \quad (1)$$

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \quad (2)$$

For any matrix $\mathbf{X} \in \mathbb{R}^{n \times d}$, its i -th row and j -th column element is denoted as X_{ij} , then its Frobenius norm and $\ell_{2,1}$ norm can be calculated by

$$\|\mathbf{X}\|_F = \left(\sum_{i=1}^n \sum_{j=1}^d X_{ij}^2 \right)^{1/2} \quad (3)$$

$$\|\mathbf{X}\|_{2,1} = \sum_{i=1}^n \sqrt{\sum_{j=1}^d X_{ij}^2}. \quad (4)$$

Models based on ℓ_1 norm regularizer can usually generate sparse solutions, and based on similar principles, row sparse solutions can also be generated for models based on $\ell_{2,1}$ norm regularizer. To enable a model to produce grouped row sparse solutions, we design the $\ell_{F,1}$ norm as follows.

Suppose $\mathbf{X} \in \mathbb{R}^{n \times d}$ is the feature fusion matrix extracted from the sensor array response values, and d features can be divided into m groups by sensor category, that is, \mathbf{X} is divided into m block matrices by column $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m]$, where n and m are the number of samples and sensors, respectively. Correspondingly, the transformation matrix \mathbf{W} can also be divided into m block matrices by row $\mathbf{W} = [\mathbf{W}_1^T, \mathbf{W}_2^T, \dots, \mathbf{W}_m^T]^T$, and then we define $\ell_{F,1}$ norm of matrix as

$$\|\mathbf{W}\|_{F,1} = \sum_{s=1}^m \|\mathbf{W}_s\|_F \quad (5)$$

Obviously, $\|\cdot\|_{F,1}$ is a norm due to satisfying the positive definiteness, absolute homogeneity and triangle inequality.

2.2 Sensor Selection Model

For binary classification problems, the ℓ_1 norm-based models are often used to select important features, such as LASSO model

$$\min_{\mathbf{w}} \|\mathbf{X}\mathbf{w} + \mathbf{1}b - \mathbf{y}\|_2^2 + \lambda\|\mathbf{w}\|_1 \quad (6)$$

where $\mathbf{1}$ is a vector with all elements one. In multi-class classification problems, class label vector \mathbf{y}_i is usually transformed into matrix consisting of “0/1” element by one-hot coding. At this point, feature selection can be accomplished by using $\ell_{2,1}$ norm-based models, such as

$$\min_{\mathbf{W}} \|\mathbf{X}\mathbf{W} + \mathbf{1}\mathbf{b}^T - \mathbf{Y}\|_F^2 + \lambda\|\mathbf{W}\|_{2,1}. \quad (7)$$

Usually, traditional feature selection methods obtain the importance ranking of sensors by scoring individual features, in which the intrinsic structure between features is not considered. However, in many sensor array optimization tasks, each sensor is represented by multiple features, it is valuable to consider the group structure among features divided by sensor type. Therefore, following the least squares regression framework, we propose a novel sensor array optimization method for multi-feature fusion data classification by using the $\ell_{F,1}$ norm regularizer as follows:

$$\min_{\mathbf{W}} \|\mathbf{X}\mathbf{W} + \mathbf{1}\mathbf{b}^T - \mathbf{Y}\|_F^2 + \lambda\|\mathbf{W}\|_{F,1}. \quad (8)$$

In Eq. (8), the intrinsic group structure of the features is considered and the unimportant feature groups are forced to have small or zero coefficients, so that the corresponding sensors are removed. In addition, to increase the robustness of the model, the Frobenius norm is replaced by the $\ell_{2,1}$ norm, while the ε -dragging trick is introduced to improve the discriminative performance of the model, and finally we obtain a robust sensor array optimization method (RSAO) as follows:

$$\min_{\mathbf{M}, \mathbf{W}, \mathbf{b}} \|\mathbf{X}\mathbf{W} + \mathbf{1}\mathbf{b}^T - \mathbf{Y} - \mathbf{E} \odot \mathbf{M}\|_{2,1} + \lambda\|\mathbf{W}\|_{F,1} \quad (9)$$

where symbol \odot represents Hadamard product of matrix, matrix \mathbf{M} consists of positive elements, λ is a positive regularized parameter, and matrix \mathbf{E} is defined as

$$E_{ij} = \begin{cases} +1, & \text{if } Y_{ij} = 1 \\ -1, & \text{if } Y_{ij} = 0. \end{cases} \quad (10)$$

In sensor selection process, the proposed method uses $\|\mathbf{W}_i\|_F$ as a score for the feature subset \mathbf{X}_i to obtain the importance ranking of the sensors, which is more efficient and has better interpretability.

2.3 Model Optimization

Clearly, there are three unknown variables to learn in Eq. (9), so we present an alternating iteration algorithm to solve it. First, given \mathbf{M} , let $\tilde{\mathbf{Y}} = \mathbf{Y} + \mathbf{E} \odot \mathbf{M}$, we will solve the following problem

$$\min_{\mathbf{W}, \mathbf{b}} \|\mathbf{XW} + \mathbf{1b}^T - \tilde{\mathbf{Y}}\|_{2,1} + \lambda \|\mathbf{W}\|_{F,1}. \quad (11)$$

Since problem (11) has no analytical solution, we present an efficient iterative algorithm to solve for \mathbf{W} and \mathbf{b} .

Let $J(\mathbf{W}, \mathbf{b})$ be objective function of problem (11), taking the derivative of the function $J(\mathbf{W}, \mathbf{b})$ with respect to \mathbf{W} and \mathbf{b} , we have

$$\frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{W}} = \mathbf{X}^T \mathbf{U}_1 (\mathbf{XW} + \mathbf{1b}^T - \mathbf{Y}) + \lambda \mathbf{U}_2 \mathbf{W} \quad (12)$$

$$\frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{b}^T} = \mathbf{1}^T \mathbf{U}_1 (\mathbf{XW} + \mathbf{1b}^T - \mathbf{Y}) \quad (13)$$

where \mathbf{U}_1 and \mathbf{U}_2 are diagonal matrix, and their diagonal elements are

$$U_1^{tt} = \frac{1}{\|\mathbf{e}_t\|_2}, \quad t = 1, 2, \dots, n \quad (14)$$

$$U_2^{ii} = \frac{1}{\|\mathbf{W}_j\|_F}, \quad \mathbf{w}_i \in \mathbf{W}_j, j = 1, 2, \dots, m \quad (15)$$

where \mathbf{e}_t is t -th row of matrix $\mathbf{XW} + \mathbf{1b}^T - \tilde{\mathbf{Y}}$ and \mathbf{w}_i is i -th row of matrix \mathbf{W} . Setting Eq. (13) equal to zero, we can get

$$\mathbf{b}^T = c \mathbf{1}^T \mathbf{U}_1 (\mathbf{Y} - \mathbf{XW}) \quad (16)$$

where c is equal to $(\mathbf{1}^T \mathbf{U}_1 \mathbf{1})^{-1}$. Then setting Eq. (12) equal to zero and using Eq. (16), we can obtain

$$\mathbf{W} = (\mathbf{X}^T \mathbf{LX} + \lambda \mathbf{U}_2)^{-1} \mathbf{X}^T \mathbf{LY} \quad (17)$$

where $\mathbf{L} = \mathbf{U}_1 - c \mathbf{U}_1 \mathbf{1} \mathbf{1}^T \mathbf{U}_1$. Note that the computation of \mathbf{U}_1 and \mathbf{U}_2 depends on \mathbf{W} and \mathbf{b} , so \mathbf{W} and \mathbf{b} can be iteratively updated by using \mathbf{U}_1 and \mathbf{U}_2 from the previous step.

Second, we perform the optimization of \mathbf{M} . Given \mathbf{W} and \mathbf{b} , let $\mathbf{T} = \mathbf{XW} + \mathbf{1b}^T - \mathbf{Y}$, then we need to solve the following problem

$$\min_{\mathbf{M}} \|\mathbf{T} - \mathbf{E} \odot \mathbf{M}\|_{2,1} \quad (18)$$

this problem can be decomposed into subproblems by row

$$\min_{\mathbf{m}_i} \|\mathbf{t}_i - \mathbf{e}_i \odot \mathbf{m}_i\|_2, \quad i = 1, 2, \dots, n \quad (19)$$

where \mathbf{t}_i , \mathbf{e}_i , and \mathbf{m}_i , are i -th row of the matrix \mathbf{T} , \mathbf{E} , and \mathbf{M} , respectively.

Let $J(\mathbf{m}_i)$ be objective function of problem (19), taking the derivative of $J(\mathbf{m}_i)$ with respect to M_{ij} , we have

$$\frac{\partial J(\mathbf{m}_i)}{\partial M_{ij}} = \frac{E_{ij}T_{ij} - M_{ij}}{\|\mathbf{t}_i - \mathbf{e}_i \odot \mathbf{m}_i\|_2} \quad (20)$$

and set it equal to zero, we obtain

$$M_{ij} = \max\{0, E_{ij}T_{ij}\} \quad (21)$$

where M_{ij} , E_{ij} , and T_{ij} are i -th row and j -th column of the matrix \mathbf{M} , \mathbf{E} , and \mathbf{T} , respectively. So the solution of problem (18) can be write as

$$\mathbf{M} = \max\{\mathbf{0}, \mathbf{E} \odot \mathbf{T}\} \quad (22)$$

In short, we present an alternating iterative method to solve problem (9), and it mainly includes two steps. First, optimize vector \mathbf{b} and matrix \mathbf{W} with fixed matrix \mathbf{M} according to Eq. (16) and Eq. (17); Second, optimize matrix \mathbf{M} with fixed vector \mathbf{b} and matrix \mathbf{W} according to Eq. (22). The proposed iterative algorithm is summarized as Algorithm 1.

Algorithm 1. RSS for Sensor Selection

Input: feature matrix $\mathbf{X} \in \mathbb{R}^{n \times d}$ with divided into m groups by column, label matrix $\mathbf{Y} \in \mathbb{R}^{n \times k}$, parameters λ .
Output: indices of the first s sensors
1: initialize \mathbf{M} , \mathbf{W} and \mathbf{b} for solving Eq. (9)
2: **while not converged do**
3: $\hat{\mathbf{Y}} = \mathbf{Y} + \mathbf{E} \odot \mathbf{M}$
4: **while not converged do**
5: calculate \mathbf{U}_1 and \mathbf{U}_2 according to Eq. (14) and (15)
6: calculate $\mathbf{b}^T = \mathbf{c}1^T \mathbf{U}_1 (\mathbf{Y} - \mathbf{X}\mathbf{W})$
7: calculate $\mathbf{W} = (\mathbf{X}^T \mathbf{L}\mathbf{X} + \lambda \mathbf{U}_2)^{-1} \mathbf{X}^T \mathbf{L}\mathbf{Y}$
8: **end while**
9: $\mathbf{T} = \mathbf{X}\mathbf{W} + \mathbf{1}\mathbf{b}^T - \hat{\mathbf{Y}}$
10: $\mathbf{M} = \max\{\mathbf{0}, \mathbf{E} \odot \mathbf{T}\}$
11: **end while**
12: calculate $\|\mathbf{W}_i\|_F$ as an importance score of the i -th sensor
13: sort the scores in descending order and output the first s indices

2.4 Complexity Analysis

The computational cost of Algorithm 1 is mainly concentrated in three parts due to matrix inverse and matrix product calculation. In step 6, the complexity of calculating \mathbf{b} is $O(n^2k + ndk)$, where n and k are the number of samples and classes, d is the dimension of features. In step 7, the complexity of calculating

\mathbf{W} is $O(n^2d + nd^2 + ndk + d^3 + d^2k)$. The sum of the complexity of computing \mathbf{W} and \mathbf{b} is $O(n^2d + nd^2 + ndk + d^3 + d^2k + n^2k)$. In step 9, the complexity of calculating \mathbf{XW} is $O(ndk)$. Since the number of classes is much smaller than the number of samples and the feature dimension, neglecting the lower order quantities, the total computational complexity of algorithm 1 is $O(\tau(n^2d + d^3))$, where τ is the number of iterations.

3 Experiment

In this section, we evaluate the proposed method on gas sensor array data sets while comparing other classical methods.

3.1 Data Sets

We provide a brief description of all the data sets used in the experiments as follows. Note that three data sets are from the UCI Machine Learning Repository and the corresponding papers are cited.

Gas sensor array under flow modulation data set (GSAFM) [15]: this data set was collected from an array of 16 metal-oxide gas sensors under gas flow modulation conditions. It contains four categories and a total of 58 samples. Each sample included 16 time series (one time series per sensor), and then 27 features contain one maximum features, 13 high-frequency features and 13 low-frequency features extracted from each time series as corresponding sensor features, so each sample has 432 features.

Gas sensor array drift data set (GSAD) [6, 10]: this data set was collected from an array of 16 chemical sensors exposed six gases. Its first batch contains six classes and a total of 445 samples. Each sample contains 16 time series, four steady-state features and four dynamic features are extracted from each time series, thus each sample is characterized by 128 features.

Gas sensors for home activity monitoring data set (GSHAM) [1]: this data set has recorded time series signals of eight gas sensors in response to wine, banana and background activity. It contains three classes and a total of 100 samples. After the time series signals are filtered by FIR low-pass filter, each sensor signal is represented by three features, minimum, average, and minimum slope, i.e., each sample contains 24 features.

3.2 Experiments Settings

We will compare our method, RSAO, with the classical T-test [7], LDA [9], MI [14] and ERFS [4]. All data are normalized by Z-score method. Classification accuracy is used to evaluate the performance of the selected sensors. A linear

SVM classifier is trained on the training set, and then its classification accuracy on the test set is calculated by the function `fitsvm` in Matlab. For all data sets, ten times five-fold cross-validation is performed, i.e., 50 classification accuracies are obtained, and finally we present the average accuracy and standard deviation of the different methods to compare.

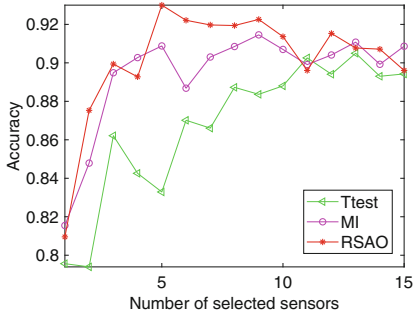
In ERFS and RSAO, the regularization parameter λ needs to be tuned. In addition, the regularization parameter C in SVM needs to be tuned for all methods. In each experiment, the optimal parameters are selected by the grid search method with three-fold cross validation as the evaluation criterion based on the training set. The candidate set for log value of parameter λ is $\{-2, -1, 0, 1, 2, 3, 4\}$ and the candidate set for log value of parameter C is $\{-3, -2, -1, 0, 1, 2, 3\}$.

3.3 Comparison of Classification Accuracy

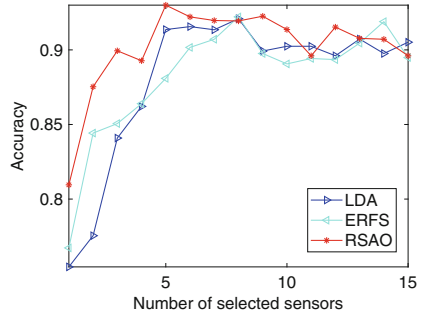
Figure 1 presents the relationship between the classification accuracy and the number of sensors selected by the five methods. It can be seen that the proposed method surpasses the other methods in most points. Table 1 presents the classification accuracies of the raw sensor array features and the optimized sensor array features, as well as the corresponding number of selected sensors. It can be seen that the classification accuracy has been significantly improved after sensor array optimization, and the proposed method achieves better accuracy than other methods while fewer sensors are selected. In addition, Fig. 2 presents the variation of the values of the objective function (11) and the objective function (9) with the number of iterations in the proposed optimization algorithm. It can be seen that the objective function values are monotonically decreasing and are convergent after fewer iteration steps.

Table 1. Optimal classification accuracy and corresponding number of selected sensors after executing sensor array optimization.

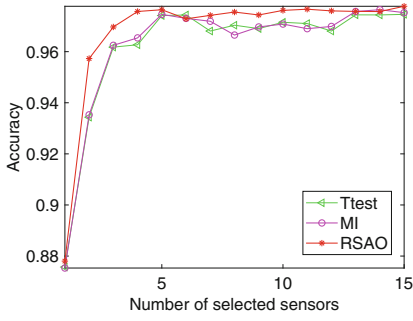
Method	Gas sensor array data sets					
	GSAFM		GSAD		GSHAM	
	Number of sensor	Accuracy	Number of sensor	Accuracy	Number of sensor	Accuracy
Raw	16	0.8986 \pm 0.0845	16	0.9755 \pm 0.0130	8	0.7464 \pm 0.0972
Ttest	13	0.9050 \pm 0.0934	6	0.9744 \pm 0.0174	7	0.7499 \pm 0.0848
MI	9	0.9145 \pm 0.0846	5	0.9746 \pm 0.0184	6	0.7546 \pm 0.0933
LDA	8	0.9209 \pm 0.0693	7	0.9733 \pm 0.0150	5	0.7646 \pm 0.0842
ERFS	8	0.9221 \pm 0.0821	7	0.9760 \pm 0.0137	7	0.7479 \pm 0.07781
RSAO	5	0.9300 \pm 0.0675	5	0.9764 \pm 0.0156	3	0.7650 \pm 0.1000



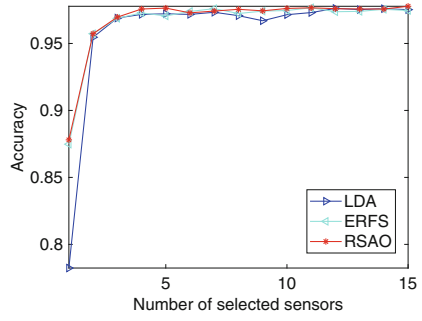
(a) GSAFM



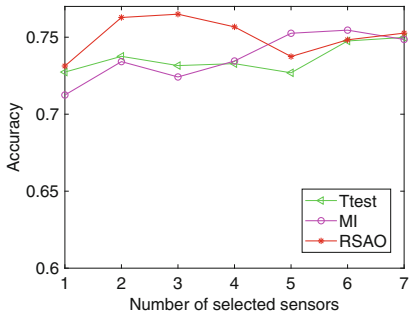
(b) GSAFM



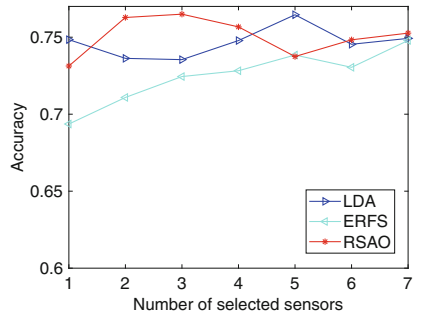
(c) GSAD



(d) GSAD



(e) GSHAM



(f) GSHAM

Fig. 1. Relationship between classification accuracy and the number of selected sensors on the three gas sensor array data sets.

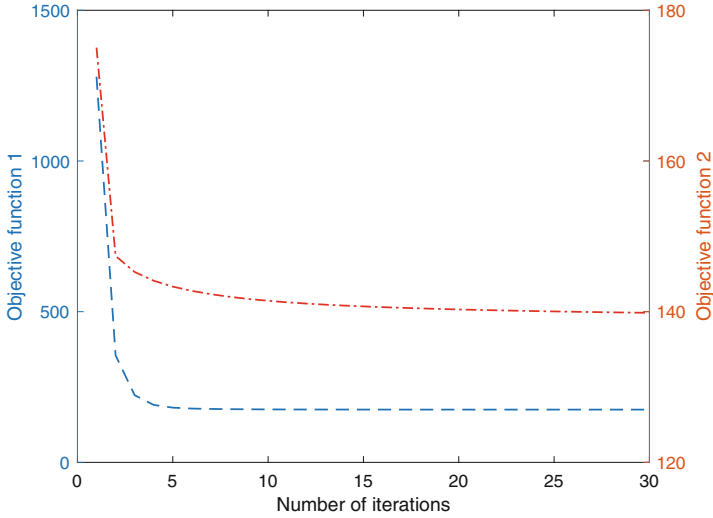


Fig. 2. Relationship between the objective function value and the number of algorithm iterations.

4 Conclusion and Future Work

In this paper, we propose a novel sensor array optimization method for multi-feature fusion data classification. The intrinsic group structure of the sensor features is considered by combining a least squares regression framework and an $\ell_{F,1}$ norm regularization design. Experimental results on the gas sensor array data sets demonstrate that the proposed method can effectively improve the classification accuracy while reducing the number of sensors compared to other classical methods.

However, the work in this paper has a limitation that it is only evaluated on gas sensor array data sets. In the future, the generalization of the proposed method will be validated for more feature fusion tasks, such as EEG signals, machine fault detection, radar array signals, etc.

References

1. Huerta, R., Mosqueiro, T., Fonollosa, J., Rulkov, N.F., Rodriguez-Lujan, I.: Online decorrelation of humidity and temperature in chemical sensors for continuous monitoring. *Chemometr. Intell. Lab. Syst.* **157**, 169–176 (2016)
2. Li, J., et al.: Feature selection: a data perspective. *ACM Comput. Surv. (CSUR)* **50**(6), 1–45 (2017)
3. Liu, B., et al.: Lung cancer detection via breath by electronic nose enhanced with a sparse group feature selection approach. *Sens. Actuators B Chem.* **339**, 129896 (2021)

4. Nie, F., Huang, H., Cai, X., Ding, C.: Efficient and robust feature selection via joint ℓ_2 , 1-norms minimization. *Adv. Neural Information Process. Syst.* **23**, 1813–1821 (2010)
5. Röck, F., Barsan, N., Weimar, U.: Electronic nose: current status and future trends. *Chem. Rev.* **108**(2), 705–725 (2008)
6. Rodriguez-Lujan, I., Fonollosa, J., Vergara, A., Homer, M., Huerta, R.: On the calibration of sensor arrays for pattern recognition using the minimal number of experiments. *Chemometr. Intell. Lab. Syst.* **130**, 123–134 (2014)
7. Saha, P., Ghorai, S., Tudu, B., Bandyopadhyay, R., Bhattacharyya, N.: Optimization of sensor array in electronic nose by combinational feature selection method. In: Mason, A., Mukhopadhyay, S.C., Jayasundera, K.P., Bhattacharyya, N. (eds.) *Sensing Technology: Current Status and Future Trends II. SSMI*, vol. 8, pp. 189–205. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-02315-1_9
8. Scott, S.M., James, D., Ali, Z.: Data analysis for electronic nose systems. *Microchim. Acta* **156**(3), 183–207 (2006). <https://doi.org/10.1007/s00604-006-0623-9>
9. Sun, H., et al.: Sensor array optimization of electronic nose for detection of bacteria in wound infection. *IEEE Trans. Ind. Electron.* **64**(9), 7350–7358 (2017)
10. Vergara, A., Vembu, S., Ayhan, T., Ryan, M.A., Homer, M.L., Huerta, R.: Chemical gas sensor drift compensation using classifier ensembles. *Sens. Actuators B Chem.* **166**, 320–329 (2012)
11. Wei, G., Zhao, J., Yu, Z., Feng, Y., Li, G., Sun, X.: An effective gas sensor array optimization method based on random forest. In: 2018 IEEE SENSORS, pp. 1–4. IEEE (2018)
12. Xiang, S., Nie, F., Meng, G., Pan, C., Zhang, C.: Discriminative least squares regression for multiclass classification and feature selection. *IEEE Trans. Neural Netw. Learn. Syst.* **23**(11), 1738–1754 (2012)
13. Yan, J., et al.: Electronic nose feature extraction methods: A review. *Sensors* **15**(11), 27804–27831 (2015)
14. Zhou, J., Welling, C.M., Kawadiya, S., Deshusses, M.A., Grego, S., Chakrabarty, K.: Sensor-array optimization based on mutual information for sanitation-related malodor alerts. In: 2019 IEEE Biomedical Circuits and Systems Conference (BioCAS), pp. 1–4. IEEE (2019)
15. Ziyatdinov, A., Fonollosa, J., Fernandez, L., Gutierrez-Galvez, A., Marco, S., Perera, A.: Bioinspired early detection through gas flow modulation in chemo-sensory systems. *Sens. Actuators B Chem.* **206**, 538–547 (2015)