

Application of SVM in Gas Sensor: Quantitative Analysis of CO₂

HongYan Chen

College of Mechanical and Electrical Engineering
China Jiliang University
HangZhou,China
bbchy@163.com

WenZhen Liu

College of Mechanical and Electrical Engineering
China Jiliang University
HangZhou,China
bbchy@163.com

Jian Qu

College of Mechanical and Electrical Engineering
China Jiliang University
HangZhou,China
848990467@qq.com

ZhiBin Li

College of Mechanical and Electrical Engineering
China Jiliang University
HangZhou,China
1247139378@qq.com

Abstract—Objective:According to the difficult in selecting parameter of SVM when modeling on the gas quantitative analysis, and existing methods need long time to finish, SVM optimized by improved grid search method was proposed to build a model to quantitatively analyse infrared spectrum of CO₂ gas. **Methods:**We analyze 15 samples of CO₂ gas at concentrations ranging from 500 ppm to 18% based on SVM. **According to this method,**the spectrum data of CO₂ is optimized. **The kernel function leads SVM and calculate the concentration.**By using improved grid search, quantitatively analyzed 15 different concentrations of CO₂ in the range between 500 ppm-18%. **Results:**The experiment results show that this method gets $c=1.412, g=0.25$. **And the prediction error is less than 5%.****Conclusion:**And method of grid search combined with SVM has a certain potential for development and mining space in gas quantitative analysis modeling in the infrared spectrum of CO₂ gas.

Keywords- SVM; grid search; gas sensor; infrared spectrum.

I. INTRODUCTION

Vehicle exhaust emissions have become an important factor in environmental pollution, emission of polluting gases are mainly CO, NO, CO₂, HC compound. CO₂ is a greenhouse gas, excess emissions will seriously affect our living environment. In the quantitative analysis of CO₂ gas, the main method is infrared absorption method[1-3], but due to the sensor components aging, temperature changes, supply voltage fluctuations and other factors, resulting in the accuracy of measurement is not very high.

We have artificial neural networks[4-5](ANN) method and support vector machine[6-8] method to improve the measurement accuracy of the gas. The convergence of neural network algorithm and its prediction accuracy by a great effect of the selection of initial weights and thresholds, so the output has inconsistencies, easily to fall into local minimum, and support vector machine algorithm can effectively avoid the above problems.

In this paper, we analyze 15 samples of CO₂ gas at concentrations ranging from 500 ppm to 18% based on SVM[9-11], and at the problem of selection of SVM parameters, in order to further reduce the mean absolute

error in the quantitative analysis, reduce the optimization time and obtain the best combination of parameters, an improved grid search optimization method is proposed to establish the optimal prediction model.

This gas analytical sensor developed in this paper is achieved using Non-Dispersive Infrared (NDIR) techniques. It could rapidly measure the concentration of carbon dioxide, carbon monoxide and hydrocarbon in the emission exhaust at the same time.

Based on the analysis of status quo of vehicle exhaust testing and non-dispersive infrared detection technique, this article establishes a full system test scenario and designs a detection system structure. Choose the right infrared lamp, infrared detectors, gas chamber, window films, etc. as the optical, pneumatic system components. According to the characteristics of the infrared light source, we design a steady-voltage light source driven modulation circuit. And according to the IR thermopile detector output characteristics, a weak signal amplification filter circuit is designed in this paper. Then we designed the following A/D conversion circuit, MCU control circuit and peripheral circuit. By developing the program of hypogynous machine and epigynous machine, complete the whole design of sensor system platform.

In this paper, the calibration experiment of sensor system platform is studied. By measuring a series of standard gases of each component, we can get the relationship between each channel voltage signal data and corresponding gas concentration. By analyzing the experimental data and processing, fitting out of the three components CO, HC, CO₂ concentration curves using Origin software. And through detailed analysis of the factors influencing the system, compensate to the main affecting factors like the environment temperature and pressure, and finally the sensor system get more accuracy.

II. THEORY OF SVM REGRESSION MODEL

In this paper, we take a known concentration of CO₂ gas samples as the training set of, $T = \{(x_1, y_1), \dots, (x_N, y_N)\} \in (R^n \times R)^N$ and $x_i \in R^n$ i

s i-th sample of tested CO₂ gas, $x_i = (x_1, x_2, \dots, x_L)$ is L spectral datas within the scanning range of the wavelength, $y_i \in R$ is the concentration of tested CO₂ gas, $i=1,2,\dots,N$. In the highdimensional space, the regression function $f(x_i)$ of spectrum needed to solve and the concentration of CO₂ gas can be expressed as follows:

$$f(x_i) = \omega \cdot \phi(x_i) + b \quad (1)$$

Among them, $\omega \cdot \phi(x)$ is the inner product of ω and $\phi(x)$; ω is regression coefficients; b is threshold.

According to SRM, the formula (1) is converted to a convex quadratic programming problem:

$$\min_{\omega, \xi} \frac{1}{2} \omega^T \omega + C \sum_{i=1}^N (\xi + \xi^*) \quad (2)$$

Lagrange function is introduced to solve the equation (2), by the kernel function $k(x_i, x_j)$, the inner product of the conversion of highdimensional space is calculated in the original two-dimensional space, we get

$$\omega = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \phi(x_i) \quad (3)$$

$$s.t. \pm [y_i - \omega \cdot \phi(x_i) - b] \leq \varepsilon + \xi_i, i = 1, 2, \dots, N \quad (4)$$

Then the regression function of SVM regression model is:

$$f(x) = \omega \cdot \phi(x) + b = \sum_{i=1}^N (\alpha_i - \alpha_i^*) k(x_i, x_j) + b \quad (5)$$

In formula (5), if α_i is not zero or α_i^* is not zero, indicates that this sample is the support vector.

Commonly used kernel functions have linear kernel, polynomial kernel function, radial basis (RBF) kernel function, Sigmoid Kernel, we select Sigmoid kernel function in this paper.

III. EXPERIMENT AND ANALYSIS

In this study, CO₂ concentration is detected by the application of nondispersive infrared method (NDIR). The infrared light which is modulated by MCU go through a certain length of the gas chamber, which is filled with CO₂ gas concentration measured.

The infrared light source emits a particular wavelength spectrum, after a specific wavelength light source through the gas, due to the CO₂'s absorption of infrared energy in the appointed wavelength of infrared band, after the light passes through, in the corresponding line of the light intensity attenuation will occur, the infrared energy will be reduced, the probe detects the remaining light intensity and converts into electrical signals, amplified and filtered as model input.

Principle of the test system is shown in Figure 1.

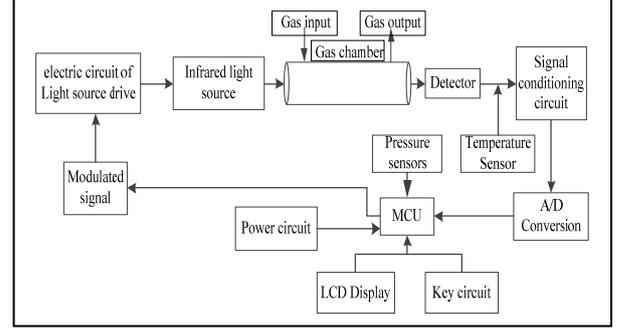


Figure 1. The NDIR test system schematic

Standard concentration between 500ppm ~ 18% of the 15 groups of CO₂ gas as the experimental sample gas, the concentration of the standard were normalized, that is the desired output of the model. Model needs to find the optimal penalty parameter c and RBF kernel parameter g . In this paper, we apply a method of grid search[12] to optimize parameter. The basic principle of grid search method is to make c and g meshing and through all points in the grid in a certain range and get values, for given c and g , calculate the accuracy of classification of training set taking advantage of K-CV under this group c and g , and ultimately taking c and g that make the highest classification accuracy of the training set as the best parameters. The results are shown in Figure 2.

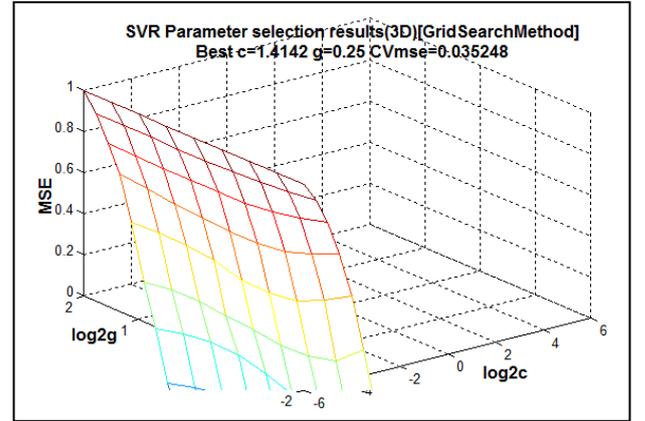


Figure 2. Parameter selection results

Well selected optimal parameter combination (c, g), the SVM regression model can be established for standard sample data, so that you can treat the measured gas sample regression analysis. Because of the standard sample values were normalized, so the output prediction model is normalized concentration too, to get the standard value, the actual output of the model need anti-normalization.

The comparison chart of actual output and expected output of the model, as well as the absolute error are shown in Fig.3、 Fig.4 and TABLE I:

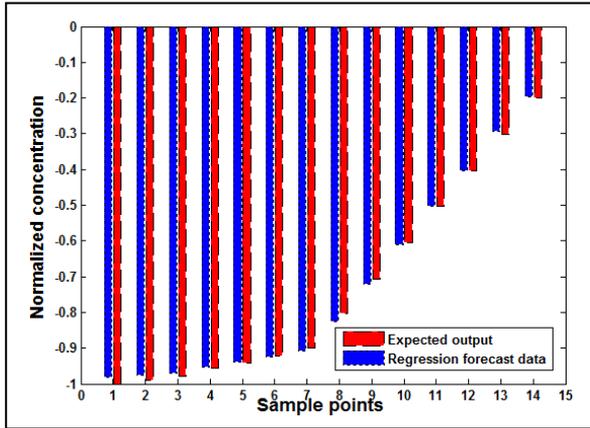


Figure 3. The comparison chart

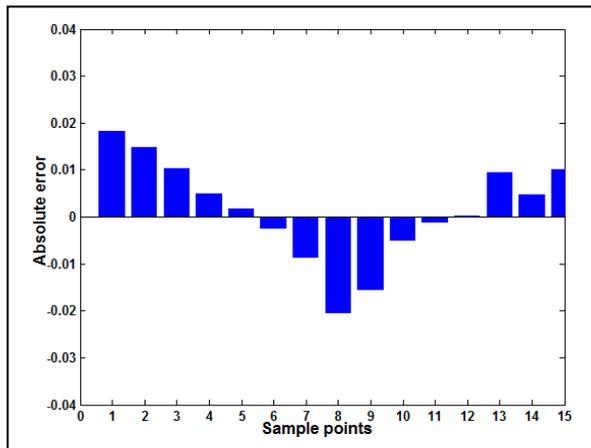


Figure 4. The absolute error

And the data of simulation is shown in TABLE I:

TABLE I. SIMULATION DATAS

Table Head	Values Need to Be Compared/%	
	Calibration Values	Improved Grid Search
S	0.025	0.028
I	0.02	0.068
M	0.5	1.249
U	1.02	1.575
L	2.10	2.307
A	2.99	3.002
T	4.10	3.952
I	5.02	4.805
O	6.04	5.837
N	7.05	6.837
	8.12	8.032
D	10.01	10.109
A	12.03	12.330
T	14.02	14.338
A	16.20	16.380
S	18.10	17.897
	MAE	0.0417

IV. CONCLUSION

As we all know, to control the automobile vehicle emissions, it must be started with the sources of pollution and enhance the detection and control of motor vehicle exhaust emissions. The demand of stable and accurate gas analytical systems has become very urgent.

Based on the analysis of status quo of vehicle exhaust testing and non-dispersive infrared detection technique, this article establishes a full system test scenario and designs a detection system structure. Choose the right infrared lamp, infrared detectors, gas chamber, window films, etc. as the optical, pneumatic system components. According to the characteristics of the infrared light source, we design a steady-voltage light source driven modulation circuit. And according to the IR thermopile detector output characteristics, a weak signal amplification filter circuit is designed in this paper. Then we designed the following A/D conversion circuit, MCU control circuit and peripheral circuit. By developing the program of hypogynous machine and epigynous machine, complete the whole design of sensor system platform.

We can find in Fig.3 and Fig.4:the regression forecast results are very close to the expected datas.And the absolute error is very small,between -0.02 to 0.02.And from table 1,two groups of datas tell us:the first few datas and the final few datas have larger difference ,but the middle is fine.And the total MAE is about 0.0417,less than 0.05.So the model established by SVM can finely analyse and forecast the concentration of CO₂ gas.

In this paper, the method of grid search is applied to optimize SVM parameters, and quantitatively analyze the concentration range of 500ppm ~ 18% of the standard CO₂ gas, with advantages of fast convergence and high model prediction accuracy. Absolute error of simulation models is less than 5%, in line with national testing standards on emissions and completely meet the actual demand of gas measurement. Therefore, method of grid search combined with SVM has a certain potential for development and mining space in gas quantitative analysis modeling in the infrared spectrum of CO₂ gas.

We can foresee,with the guidance of right policies and in the detection of high-precision monitoring equipment, the future of automotive exhaust emissions will be controlled effectively, our living environment will be better.

REFERENCES

- [1] Yi-Feng Wang.The Infrared Technology of China in 2013[J].Infrared Technology, 2014,36(1):10-21.
- [2] Yi-Feng Wang.The Infrared Technology of China in 2013 II [J].Infrared Technology, 2014,36(1):10-21.
- [3] Yi-Feng Wang. The Infrared Technology of China in 2013III[J].Infrared Technology, 2014,36(1):10-21.
- [4] Min-Ming Tong,Yu Zhang,Mei-Xing Qi.The Mixed Inflammable Gas Analys- is Based on BP Neural Network.ActaMetrologica Sinica,2006,(2):169.
- [5] He Wang , Dong-Xiang Zhou, Wei-Yuan Song,et al. Research of Recognition of Gas Mixture Based on Genetic Algorithms.Journal of Huazhong University of Science and Technology•Nature Science Edition),2007,35(9):118.
- [6] Xin Bao ,Lian-kui Dai .Spectral Quantitative Analysis Based on Local Least Square Support Vector Machine Regression.Chinese Journal of Analytical Chemistry,2008,36(1):75.

- [7] Ji-Peng Lin, Jun-Hua Liu. Multi-Component Gas Analyzing Based on Support Vector Machine. *Journal of Xi'an Jiaotong University*, 2005, 06: 586-589.
- [8] Qi-Lin Zhang. Infrared Spectrum Mixed Gas Composition Analysis Based on Support Vector Machine. *Computer Era*, 2010, 01: 9-11.
- [9] Peng Bai, Jun-Hua Liu. Algorithm of Mixed Gas Concentration Analysis Based on Support Vector Machine and Multidimensional Spectrum. *Control and Instruments in Chemical Industry*, 2005, 05: 43-47.
- [10] Yuan-Yuan Chen, Ji-Long Zhang, Xiao Li, et al. Research on Concentration of Multi-Component Pollution Gas Based on SVM with Kernel Optimized by Rough Set[J]. *Spectroscopy and Spectral Analysis*, 2010, 30(12): 3384-3387.
- [11] Manouchehrian Amin, Sharifzadeh Mostafa, Hamidzadeh, et al. Selection of regression models for predicting strength and deformability properties of rocks using GA[J]. *International Journal of Mining Science and Technology*, 2013, (4): 492-498.
- [12] Xing-Ling Wang, Zhan-Bing Li. Identifying the Parameters of the Kernel Function in Support Vector Machines Based on the Grid-Search Method. *Journal of Ocean University of Qingdao*, 2005, 35(5): 859-862.