ATLANTIS PRESS
Advances in Engineering Research, volume 139
2nd International Conference on Electrical Engineering and Automation (ICEEA 2018)

Gas Composition Recognition Based on Analyzing Acoustic Relaxation Absorption Spectra: Wavelet Decomposition and Support Vector Machine Classifier

Yaqiong Jia[1], Bin Yu[2], Mingdi Du[1] and Xiaoli Wang[1]
[1]Department of Electronics and Information Engineering, Key Laboratory of Signal &Information Processing, Hunan Institute of Technology, Hengyang City of Hunan Province 421002, China
[2]Department of Computer and Information Science, Key Laboratory of Signal & Information Processing, Hunan Institute of Technology, Hengyang 421002, China

***Abstract*—Gas acoustic spectrum represents properties of acoustic propagation, which can distinguish gas compositions. However in few existing methods of gas-composition-unknown recognition, the approaches of programmatically processing the acoustic spectrum curves have not yet been presented. We propose a method for gas-composition-unknown recognition by analyzing gas acoustic relaxation absorption spectrum (GARAS) based on wavelet multi-resolution analysis (MRA) and multi-class support vector machine (SVM). Features of GARAS are extracted by wavelet MRA, and then selected to obtain a few feature coefficients which are utilized to train and test multi-class SVM. Simulation results show that the proposed method completely classifies four examples of gas mixtures, and that it recognizes the mixtures with the same and similar concentration or temperature. This method realizes the numerically extracting and programmatically processing the information of GARAS, and implements gas-composition-unknown sensing based on acoustic spectrums.**

Keywords—gas compositions recognition; gas acoustic relaxation absorption spectrum; wavelet multi-resolution analysis; multi-class support vector machine

I. INTRODUCTION

The acoustic-based gas sensing is a promising gas detection technology advantageous in without gas preprocessing or sampling, low cost, robustness and real-time response[1]. As a key of acoustic gas sensing, acoustic properties of different gas mixtures should be acquired comprehensively and accurately[2,3]. Lueptow et al inherited and accomplished gas probing by measuring sound velocity, which has been widely adopted for gas detection [3]. And these gases were treated as pseudo-binary mixtures. In Phillips' work, the acoustic attenuation as the second property, was joined into sound speed to determine the concentration of the three-component mixtures [1]. By adding effective relaxation frequency as the third property, Zhu et al developed the method to detect the concentration of gas mixtures with four components [4]. Both Petculescu[5] and Yan[6] pointed out that the gas acoustic spectrum, which mainly represents dependence of acoustic absorption on frequency, is the future research trend to develop multi-component mixtures sensing. Gas acoustic absorption relaxation spectrum (GARAS), which is the core of acoustic absorption spectrum, contains various gas properties such as effective relaxation frequency and relaxation absorption coefficient, and represents great potential for acoustic gas sensing [7].

Petculescu developed a method of gas-composition-unknown sensing based on analyzing acoustic spectrum, which was achieved by comparing the spectrum position of candidate gases with that of the contaminant gases [8]. In this method, the comparison was accomplished by intuitively observing, not by programming, and the compositions of the candidate gas and contaminant gas should be known before. In practical gas detection applications, it is desirable to develop a gas-composition-unknown sensing technology since gas compositions are always unknown before detection. To establish a gas-composition-unknown sensing technology, the GARAS database should be constructed firstly and then the gas properties are learned, thus the classes of unknown gas can be recognized.

In this paper, a gas-composition-unknown recognition method is proposed to recognize mixtures by analyzing GARAS. The method utilizes wavelet MRA to extract the detail information of the GARAS and employs multi-class SVM to study the information and recognize the corresponding gas compositions. Validated by the simulation results, the proposed recognition algorithm can programmatically extract gaseous characteristics from acoustic spectrum for gas-composition-unknown sensing. The paper is organized as follows. The related theories are presented in Section 2. The proposed method based on wavelet MRA and multi-class SVM is described in Section 3. The simulation results are presented and analyzed in Section 4. The conclusion is presented in the final section.

II. GARAS FROM THEORETICAL MODELS AND EXPERIMENTAL DATA

GARAS is a curve which presents the gas acoustic dimensionless relaxation absorption coefficient ($\alpha\lambda$) depending on acoustic frequency (f) [2, 3]. Figure I displays the experimental data from [10] (circles and squares), the GARAS curves calculated from Dain's theory [2, 3] (the dash lines) and our previous work [11,12] (the solid lines). The experimental data shows the gas acoustic relaxation absorption coefficient varying with frequency, and it can be represented by GARAS curves obtained from several acoustic absorption theories as shown in Figure 1.

The solid GARAS curves are obtained from our previous model developed from the relationship of effective heat capacity and relaxation time in gas acoustic relaxation process [11]. Our proposed model matches with the experimental data better than Dain's theory, especially on the peaks of the corresponding experimental data as shown in Figure I. Therefore, the calculated curves of our previous model are used to construct the database for gas recognition in this paper.

According to Petculescu's theory in [7] and [11], GARAS varied with the changes of gas compositions, proved by Figure I. Therefore, the characteristics of GARAS can be utilized to recognize the compositions. The differences of the characteristics roughly demonstrate the coordinate location of the spectrum curves and the sharp degrees of the peaks. To recognize the acoustic spectrum, the first step is extracting the gas properties from curves; the second is classing the curves from these properties.

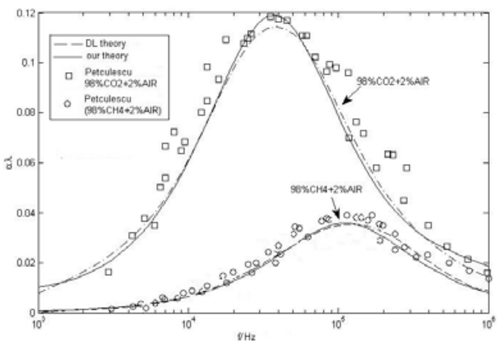

FIGURE I. THE GAS ACOUSTIC RELAXATION ABSORPTION SPECTRUMS

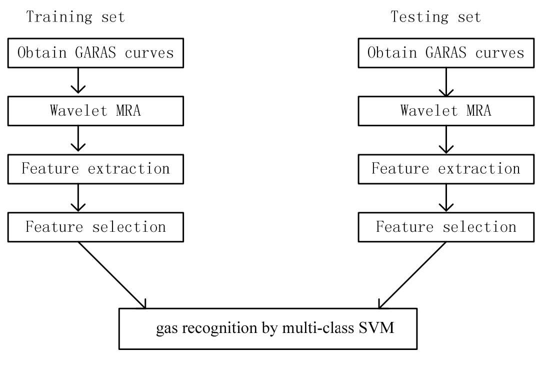

FIGURE II. THE PROCEDURE BLOCK DIAGRAM OF THE GAS RECOGNITION BASED ON WAVELET MRA AND MULTI-CLASS SVM

III. PROCEDURE OF THE RECOGNITION METHOD

The procedure of gas-composition-unknown recognition method by analyzing GARAS based on wavelet MRA and multi-class SVM contains the following five steps, which are illustrated in Figure IV. In this method, the first step is obtainment of GARAS curves, the second step is decomposition of GARAS curves by wavelet MRA, the third step is extraction from original decomposed results, the forth step is selection in the extracted features, and the final step is recognition of gas compositions corresponding the statistical features by multi-class SVM. The method learns a training set with the given gas compositions, and then processes a testing set with the unknown compositions to accomplish the gas-composition-unknown recognition. In this paper, four sorts of gas mixtures, air-CO2, air-CO, air-CH4 and air are taken as examples to prove the proposed method.

Step 1: The obtainment of the GARAS curves

The GARAS curves are obtained from our theoretical model [11,12]. In this paper, four sorts of gas mixtures, air-CO2, air-CO, air-CH4 and air are taken as examples to validate the proposed method, in which the composition proportion of air is given as $N_2 : O_2 : H_2O : CO_2 = 77.632 : 21.734 : 0.604 : 0.03$. For air, the spectrum curves are acquired by changing the temperature from 293K to 302K with step 1K. Other three groups of mixture curves with additive gas CO2, CO and CH4 in air at 293K are acquired by increasing the concentrations of the three gases from 1% to 10% with step 1%. As a result of step 1, four groups of GARAS curves corresponding to the four gas mixtures are obtained, and each group includes 10 curves.

Step 2: Wavelet MRA of the GARAS curves

Wavelet MRA is utilized for analyzing the GARAS curves to extract their detailed information. , Daubechies(db) wavelets are widely used in analyzing signals, because the decomposition order can be controlled for specific requirements[13]. Among the different dbN wavelets, db4 is the most widely adopted with its advantage of practicality [14]. 6 levels of decomposition are utilized on the ground that it can provide high classification accuracy of the SVM as well as small computation [14].

As a result, the general db4 wavelet with 6 levels of decomposition is selected in this study. For each GARAS curve obtained from step 1, 88 coefficients including 6 levels detail coefficients and 6th level approximation coefficients are derived by wavelet MRA. However, 88 coefficients are too many for the multi-class SVM classification system and a feature extraction approach is required.

Step 3: Feature extraction

88 coefficients altogether can represent the corresponding curve, while only one of them can not. However, the statistical characteristics of data on each decomposition level represent the information of every level data. Therefore, the feature extraction approach is necessary before the feature selection. There are many statistical characteristics of discrete data, such as mean, standard deviation, energy, entropy, kurtosis and form factor [13]. They are utilized to extract the statistical characteristics from the results of wavelet MRA, which not only simplify data from wavelet MRA but also enrich and sum up the information amount of every decomposition level.

In this paper, the mean, standard deviation, energy and entropy are chosen to utilize to extract the four types of features for the GARAS curves. According to Equations (3) to (6), 24 feature coefficients are obtained from the six decomposition levels. And 4 feature coefficients of the approximation coefficients at 6th decomposition level are calculated by Equation (7). Finally, 88 coefficients are reduced to 28 feature coefficients for each curve. However, in order to further reduce the computation, a feature selection approach is still needed.

Step 4: Feature selection

Features selection can reduce computation, while if some inappropriate features were selected, the recognition accuracy would decrease. Consequently, the feature coefficients with the best classification accuracy will be selected by comparing respective recognition accuracy of the extracted ones, in order to reach the balance between the computation and recognition accuracy. The recognition accuracy for one type of parameter is calculated with the following formula: $k_{cla} = \frac{T_{right}}{T_{all}} \times 100\%$. k_{cla} is the recognition accuracy, T_{right} and T_{all} represent the number of right testing samples and the number of all testing samples respectively.

In this work, feature selection is accomplished by the preprocessing based on multi-class SVM of the next step which is to obtain the respective recognition accuracy of 28 feature coefficients. In the step, the process is based on the assumption that the coefficients from the first-five curves of each gas mixture are regarded as the training samples and the coefficients from the rest-five curves as the testing sample. The training and testing process are repeated for 28 times to select several most appropriate feature coefficients, majority of which have the highest classification accuracy.

Step 5: gas recognition by multi-class SVM

To recognize four different classes in multi-class SVM, FOUR feature coefficients are usually not enough. The aim of this step is to obtain the minimum number of feature coefficients via adding one by one into the FOUR until the recognition accuracy in the multi-class SVM for each gas class reaches 100%. Consequently, the multi-class SVM based on the feature coefficients with minimum number could realize the method of gas-composition-unknown recognition with the low computation and complete classification.

Finally, through the five steps, each GARAS curve would be transformed into a few feature coefficients, which could be utilized to programmatically process the acoustic spectrum. Therefore, the gas-composition-unknown recognition framework based on GARAS analysis with wavelet MRA and multi-class SVM is established.

IV. RESULT AND DISCUSSION

A. *The Four Sorts of Garas Curves*

Four classes of gas mixtures, including air, air-CO2, air-CO, and air-CH4 are considered for simulation, which are numbered from class 1 to class 4 respectively. The GARAS curves of the four classes are put in the same coordinate, shown in Figure III. For each curve of different compositions in Figure 3, two characteristics can be observed intuitively: 1) the numbers of peaks are different; 2) the positions of each peak are also different. The difference for the number of peaks between Figure I and Figure V results from different ranges of coordinate and the range in Figure I is limited by the experimental condition. These characteristics, which are appropriate to be extracted by wavelet MRA, can be used for distinguishing different curves belonging to different gas classes.

The different characteristics of above curves are resulted from gas acoustic relaxation process. The peaks of GARAS curve, namely the largest sound absorption shown in Figure V, occur when the acoustic period is consistent with gas molecular relaxation time [10]. The number of peaks, especially for the gas mixtures consisting of different compositions, depends on the number of their molecular relaxation time. Both the position and the sharp degrees of each peak depend on the molecular relaxation time and absorption values which are related with the concentration of an additional gas [15]. Among all these characteristics, the number of the peaks is the most essential nature, because the number of the molecular relaxation time is the basic parameter to describe the relaxation process. This number of the peaks carries the intrinsic frequency information of a curve which can be extracted by wavelet MRA.

B. *The Results of Wavelet Mra*

According to the method mentioned in Step 2 Section 3, db4 wavelet with 6 levels of decomposition is adopted to analyze the GARAS curves. By the example of the curve from air at 293K in Figure III, the results of its wavelet MRA are shown in Figure VI, where the circles represent the discrete data.

The curve of air can be decomposed as discrete coefficients, including $D1$, $D2$, $D3$, $D4$, $D5$, $D6$ and $A6$, as shown in Figure IV.The span of the horizontal axis in Figure 6 is the number of wavelet coefficients. The results of summing all curves constructed respectively by the discrete results are consistent with the curve of air. For example, the location of the arrow in Figure III is basically determined by the sum of the seventh coefficient in $D5$, the sixth in $D6$, and the sixth in $A6$ as shown in Figure IV. Consequently, the discrete data of wavelet MRA can directly reflect the characteristics of GARAS curves, and can be utilized to multi-class SVM.

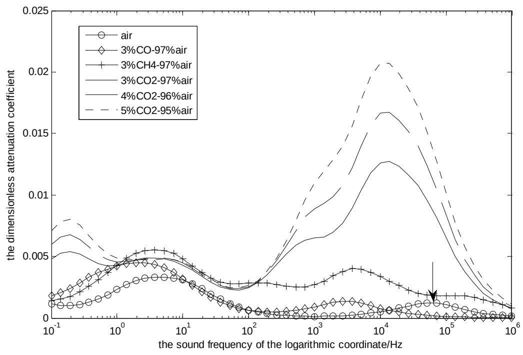

FIGURE III. COMPARISON OF THE GARAS CURVES FOR THE FOUR SORTS OF GAS MIXTURES AT 293K

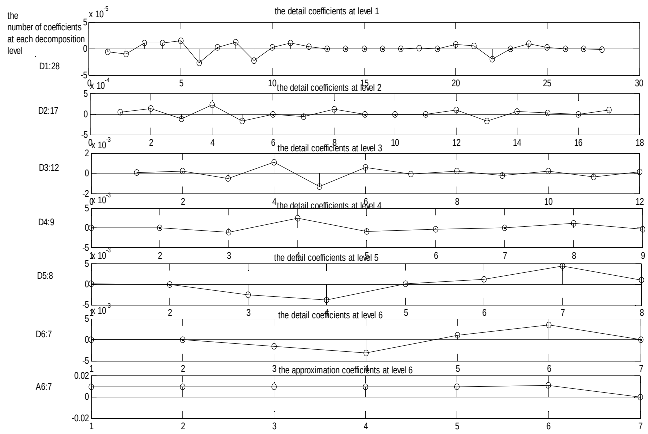

FIGURE IV. THE RESULTS OF WAVELET MRA OF THE AIR ARAS CURVE

The number of coefficients at each decomposition level is given in the left part of Figure 4, such as the number of $D1$ is 28. For the whole curve, there are 81 detail coefficients and 7 approximation coefficients. After the wavelet MRA, 88 coefficients are obtained from each GARAS curve. Obviously, these coefficients are too many for recognizing program, and a feature extraction approach is required.

C. *The Results of Feature Extraction*

The mean, standard deviation, energy and entropy of the coefficients at each decomposition level are calculated as the results of feature extraction. And 28 feature coefficients are extracted from 88 decomposition coefficients for each curve. The order of magnitude for each decomposition level is unified in order to compute conveniently.

So the values of feature coefficient at each decomposition level are different among four classes of gas mixtures. Therefore comparing with the direct results from the wavelet MRA, the feature coefficients by extraction indirectly reflect the characteristics of GARAS curves, and can also be used for numerically discriminating the curves.

D. *The Results of Feature Selection*

According to step 4, the recognition accuracies of 28 feature coefficients for each class of gas mixture are calculated. The data is horizontally divided into four parts corresponding to the four types of feature coefficients, including the mean, standard deviation, energy and entropy. Vertically the columns named class 1, 2, 3, and 4 are corresponding to the 28 feature coefficients of four classes of curves respectively.

The recognition accuracies of each feature coefficient are calculated. Because the total number of testing samples is 5, the accuracy of each feature coefficient can vary from 0% to 100% with step 20%. The average accuracy of each feature is evaluated by $average\ accuracy = \frac{\sum_{i=1}^{4} k_{cla_i}}{4} \times 100\%$, where k_{cla_i} represents the recognition accuracies of each feature for the class i.

Since the above feature coefficients focus on the energy and the entropy, other feature coefficients should be concerned with the mean and standard deviation. Thus, the recognition accuracies of all feature coefficients need to be vertically compared. Among 28 feature coefficients of air, air-CO2 and air-CO, the majority of them have 100% recognition accuracy. For air-CH4, feature coefficients with 100% recognition accuracy are less than other gas mixtures', and thereby the ones with 80% recognition accuracy should be taken into consideration. Thus, a compromise is made between air-CH4 and the other classes.

E. *The Results of Multi-Class Svm*

Next, we attempt to recognize four classes of gas mixtures by utilizing four feature coefficients. All the four types of statistical data are chosen, because different types of statistical characteristics can represent different features of the curves. Firstly for entropy, all of the three feature coefficients --NO.19, 20, and 22, have the highest average accuracy of 95%, and each of them can be used to distinguish the mixtures. Secondly, NO.16 is utilized because of its highest average accuracy among the type of energy. The recognition accuracies of entropy and energy for air, air-CO and air-CO2 reach 100%, but the highest accuracy for air-CH4 is only 80%. Thus, the choice of standard deviation and mean should focus on the recognition of air-CH4. Thirdly, the feature NO.10 has the highest average accuracy in the features of the standard deviation and is able to recognize the mixtures of air-CO2 and air-CH4. Finally, the feature NO.5 is chosen as mean data, because NO.4 fails to completely distinguish the mixtures of air-CH4 and air-CO2, although NO.4 has the highest average accuracy of 75% among the same type features. Consequently, NO.5, 10, 16 and 19 are proposed as the basic combination. Thus, the features of NO.5, 6, 10, 16, and 19 are finally chosen to identify the four mixtures.

Finally, the gas recognition of four mixtures is accomplished by multi-class SVM. Each group of GARAS curves corresponding to the four gas mixtures includes 10 curves by varying concentration and temperature. The training samples of the multi-class SVM are 5×5 results, which are five feature coefficients and first-five GARAS curves of each gas class. The testing sample utilizes all the 10 curves of each group, and consists of 40×5 feature coefficients from the ten GARAS curves. Thus, the testing sample includes not only the same ones but also the similar ones to the training data. In practical gas detection applications, the feature coefficients of unknown gas mixtures are regarded as the testing samples.

The kernel function of multi-class SVM is a polynomial kernel, and the kernel parameter q is taken as 2, and the penalty parameter C is taken as 1000. After the training samples are put into the multi-class SVM, the testing samples are entered into the multi-class SVM. Finally the overall accuracy is 100%. Thus, the method of gas-composition-unknown recognition can classify the same ones and the similar ones to the prior gas properties.

V. CONCLUSIONS

This paper proposes a new approach to recognize the gas mixtures by analyzing GARAS based on wavelet MRA and multi-class SVM. Here four sorts of gas mixtures, air-CO2, air-CO, air-CH4 and air are taken as examples to validate the method. By comparing respective recognition accuracy, 5 feature coefficients with best classification accuracy are selected and they are utilized to train and test multi-class SVM to finish gas-composition-unknown recognition. The simulation results show that the final classification accuracy achieves 100%. Consequently, the proposed method has an effective classification capability to recognize the four kinds of gas mixtures. In this method, the acoustic spectrum could be programmatically processed because each spectrum curve is transformed into a few feature coefficients. And the gas recognition method through GARAS curves is accomplished to detect an unknown gas. This paper validates the feasibility and shows the practicality of the proposed method, which offers a promising approach of utilizing GARAS for gas recognition. In future, this method could be popularized in detecting gas mixtures by obtaining sufficient classes of GARAS curves.

ROLE OF THE FUNDING SOURCES

This work is supported by a grant from the Natural Science Foundation of Hunan Province of China (Grant No. 14JJ6046), the Key Project of Technology Department of Hunan Province of China (Grant No. 2017GK2204), the scientific research fund of Hunan provincial education department (Grant No. 17A050), Hengyang Key Laboratory of Chaotic Secure Communication based on Fractional-order Chaos System (Grant No. 2017KJ186, and 2017KJ257), Science & Technology fund of Hengyang City (Grant No. 2017KJ257), and Industry-University-Research Collaboration of Hengyang City (Grant No. 2017KJ158). This work is also supported by the construct program of the key discipline in Hunan Institute of Technology.

REFERENCE

[1] S Phillips , Y Dain , and R M Lueptow, Theory for a gas composition sensor based on acoustic properties, Meas. Sci. Technol. 14 (2003) 70-75

[2] Y Dain and R M Lueptow, Acoustic attenuation in three-component gas mixtures-Theory, J. Acoust. Soc. Am. 109 (2001) 1955-1964

[3] R M Lueptow and S Phillips, Acoustic sensor for determining combustion properties of natural gas, Meas. Sci. Technol. 11 (1994)1375-1381

[4] Ming Zhu, Shu Wang, Shutao Wang, Donghai Xia, Developed acoustic multi-relaxation algorithm for humidity measurement in gas mixtures, 3th Int. Conf. on Natural Computation, (2007)

[5] A Petculescu, Future trends in acoustic gas monitoring and sensing, J Optoelectron ADV. M. 8 (2006) 217-221

[6] S Yan and S Wang, Reconstruction algorithm of relaxation attenuation spectrum in polyatomic gas, Acta. Phys. Sin. 57 (2008) 4282-4291

[7] K S Zhang , S Wang, M Zhu , Y Hu and Y Q Jia, Analytical model for acoustic multi-relaxation spectrum in gas mixtures, Acta. Phys. Sin. 61 (2012) 174301

[8] A Petculescu and R M Lueptow, Quantitative acoustic relaxational spectroscopy for real-time monitoring of natural gas: A perspective on its potential, Sensor Actuat B 169 (2012) 121-127

[9] C Cortes and V N Vapnik, Support-vector networks, Mach. Learn. 20 (1995)273-297

[10] A Petculescu, B Hall, R Fraenzle , S Phillips, and R M Lueptow, A prototype acoustic gas sensor based on attenuation(L), J. Acoust. Soc. Am.120 (2006) 1779-1782

[11] Y Q Jia, S Wang, M Zhu, K S Zhang, and F G Yuan, The analytic model between effective heat capacity and relaxation time in gas acoustic relaxation process, Acta Phys. Sin. 61 (2012) 095101

[12] Y Q Jia, S Wang, M Zhu ,and F G Yuan, Multi-relaxation time model and its application in construction of gas acoustic attenuation spectrum, Indian J. Phys. 87 (2013) 419-424

[13] H Eristia, A Ucarb and Y Demirb, Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines, Electr. Pow. Syst. Res. 80 (2010) 743-752

[14] V S Murthy, K Tarakanath, D K Mohanta, and S Gupta, Insulator condition analysis for overhead distribution lines using combined wavelet support vector machine (SVM), IEEE Trans. Dielect. El. In. 17 (2010) 89-99

[15] S G Ejakov, S Phillips, Y Dain, R M Lueptow, and J H Visser, Acoustic attenuation in gas mixtures with nitrogen: experimental data and calculations, J. Acoust. Soc. Am. 113 (2003)1871-1879