IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 23, 2019, accepted September 6, 2019, date of publication September 16, 2019,
date of current version September 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2941497

Fault Diagnosis of Gas Pressure Regulators Based on CEEMDAN and Feature Clustering

SHEN TIAN[1], XIAOYU BIAN[2], ZHIPENG TANG[2], KUAN YANG[2], AND LEI LI[2]
[1]School of Business, Zhengzhou University, Zhengzhou 450001, China
[2]School of Physics and Engineering, Zhengzhou University, Zhengzhou 450001, China

Corresponding author: Lei Li (lilei@zzu.edu.cn)

ABSTRACT Gas pressure regulators are widely applied in natural gas pipeline networks, correspondingly, establishing an efficient fault diagnosis approach of regulators plays a critical role in optimizing the safety and reliability of pipeline network systems. In our paper, considering that the outlet pressure signals of gas regulators are nonstationary and nonlinear, we propose a fault diagnosis approach combining a complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and fuzzy *c*-means (FCM) clustering to classify three typical faults of gas regulators. First, we propose to apply the CEEMDAN approach for decomposing intrinsic mode functions (IMFs). Then feature vectors of the typical faults are established by Hilbert marginal spectrum (HMS) of IMFs. Finally, we adopt cluster centers and feature clustering algorithm to distinguish the types of faults. The experimental results indicate the high performance of the present fault diagnosis approach. The membership degrees of test samples obtained from the CEEMDAN algorithm are optimized to be within 0.9 to 1.

INDEX TERMS Gas pressure regulators, fault diagnosis, CEEMDAN, feature extraction, spectral analysis, fuzzy *c*-means clustering.

I. INTRODUCTION

Pressure regulators are designed to maintain constant output pressure regardless of the variations in the upstream pressure or the downstream flow [1]. These control valves are widely applied in the fields of industries and household, such as aircraft [2], aerospace [3], vehicle [4], mining [5], etc. In cooking and heating fields [6], the compressed natural gas, regulated by a series of pressure regulators, eventually goes to household supply systems at a lower pressure. Any fault of a gas pressure regulator in this chain may lead to the leakage of explosive gas, causing economic losses and residential casualties [7]–[8]. Therefore, an effective fault detection and identification approach for gas pressure regulators has become an urgent problem to be solved [9]–[14].

In low-pressure gas pipeline networks, the stability of outlet pressure is one of the most important parameters that reflects the performance of a gas regulator. Depending on its maximum allowable operating pressure (MAOP), we can define three types of networks: high, middle and low pressure gas networks [15]. The pressure value of low-pressure networks which eventually goes to household supply systems ranges from 20 kPa down to 2 kPa [16]. When a faulty regulator is operating, its outlet pressure signal, which fluctuates abnormally, is rich in state information. Most common faults in a pilot-operated gas pressure regulator shown in Fig. 1 include three types: high frequency surge, low outlet pressure at peak hours, and high closing pressure at night [17]. The schematic outlet pressure signals of a healthy and three faulty regulators are shown in Fig. 2, and the reasons of typical faults are summarized in Table 1. In the stage of signal processing, the signals gathered from gas regulators usually show nonlinearity and nonstationarity due to the variety of interference factors and unstable conditions in a pipeline [14]. Hence, a thorough solution for outlet pressure signal processing is crucial to fault diagnosis of gas regulators.

In the literature, fault diagnosis has been shown to be performed through various algorithms, i.e., fast Fourier transform (FFT) [18]–[20], short-time Fourier transform (STFT) [21]–[23], wavelet transform (WT) and wavelet packet decomposition (WPD) [24]–[26], empirical mode decomposition (EMD) [27]–[29], and Hilbert-Huang transform (HHT) [30]–[32]. These strategies can be successfully

The associate editor coordinating the review of this manuscript and approving it for publication was Youqing Wang.

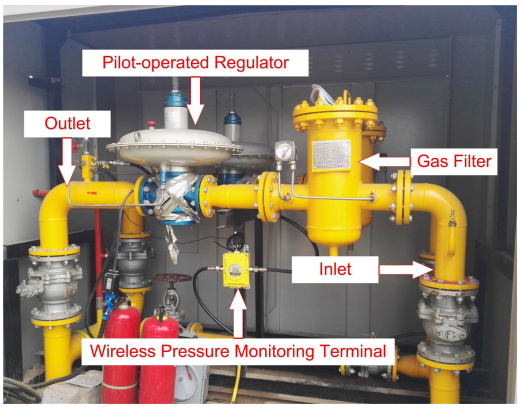


FIGURE 1. Experimental platform: A pilot-operated gas regulator in regulating box.

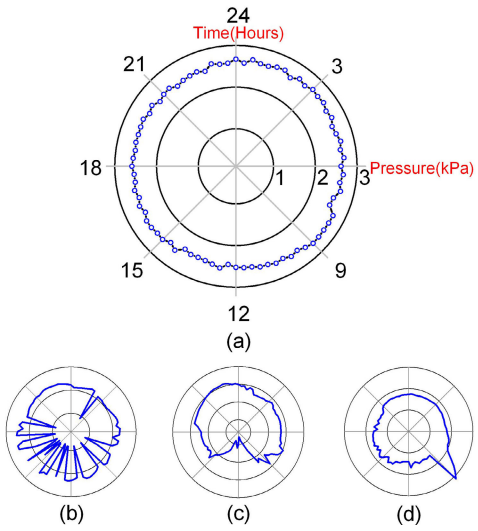


FIGURE 2. The outlet pressure signal from (a) a healthy regulator, (b) a regulator with surge, (c) a regulator with low outlet pressure, and (d) a regulator with high closing pressure. (The features of the original time-varying waveforms above are schematic and optimized.)

applied to the fault diagnosis in certain scenarios, but with limitations. The FFT has an advantage of a higher extraction efficiency than all other methods, but it performs poorly in nonlinear and nonstationary conditions. The STFT has adopted its decomposition method from global to local, which is used on nonstationary signal processing; however, it shows inefficiency in multi-resolution analysis and nonlinearity conditions. The wavelet-based techniques have improved the loss of information and the resolution limitations introduced by Fourier analysis; however, desired wavelet requires a strict design of filters for the diverse fault types, and thus any improper selection of the basis function can affect the analysis results [33, 47].

As an ideal time-frequency analysis approach, EMD exhibits a better performance in analyzing amplitude-frequency modulated (AM-FM) and multicomponent signals. Compared with traditional FFT and WT, the data driven method of the EMD does not require designated assumptions

TABLE 1. Typical faults and reasons.

Types	Reasons
Surge	1) the stem of regulator rubs against its guide hole 2) the support bracket of regulator rubs against other parts
Low outlet pressure at midday peak	1) the port of pilot is clogged 2) the spring stiffness of regulator is reduced (material fatigue) 3) the gas filter contains impurities (upstream problem) 4) the actual flow rate of regulator is higher than the rated (improper selection of regulator) 5) inpropriate inlet pressure (upstream problem)
High closing pressure at night	1) the sealing part is aged or damaged 2) the valve gasket contains impurities which lead to gas leakage 3) the pliot is not working

behind a fundamental model and is suitable for both nonlinear and nonstationary signals. Meanwhile, the EMD decomposes a signal into approximative monocomponents called instinct mode functions (IMFs). Some of these IMFs are sensitive and relevant to specific faults, thus making the EMD ideal for fault feature extraction of a system. Moreover, the HHT method [34] is an advanced method which combines the self-adaptive EMD algorithm and the Hilbert transform (HT) in order to produce a time-frequency distribution of amplitude called Hilbert spectrum (HS). So far, the EMD has been applied in many cases, but its application in gas pressure regulation is rarely reported.

Although the EMD displays decent performance in diverse fields, its drawbacks in separating close modes cannot be ignored. As a dyadic filter bank [35], the EMD suffers from sampling rate issues and insufficient performance in noisy industrial environments, which have hindered its application in fault diagnosis. A vital drawback of the EMD method is the presence of oscillations of very different amplitudes in a mode, or very similar ones in different modes, termed ''mode mixing'' [36]. To alleviate this problem, several methods were proposed [37]–[40]. These methods, however, can only reduce mode mixing to some extent due to different attempts of signal mixed with noise generating different decomposition results. Some of the methods lead to a significant residual noise while others increase the decomposition level [41]. Therefore, these algorithms can hardly be applied in diagnostic schemes for gas regulators.

In this paper, we proposed an automatic diagnostic framework for gas pressure regulators, which is for the most part manual and inaccurate nowadays. Utilizing a complete ensemble EMD with adaptive noise (CEEMDAN) to extract weak features from pressure signals, a holistic diagnosis

approach is applied based on CEEMDAN and feature clustering. First, to alleviate mode mixing and improve diagnostic accuracy of results, the CEEMDAN algorithm is utilized; then IMFs are computed based on a unique residue [42]. Second, we collect a finite number of IMFs, which contain the fault features of a gas regulator, to reconstruct a feature vector based on Hilbert marginal spectrum (HMS). Third, we use feature vectors to construct faulty sample centers by a fuzzy *c*-means (FCM) clustering method [43]. Finally, the fault diagnosis method is verified by collected data from residential gas regulators.

The remaining parts of our paper are organized as follows: Section II introduces the CEEMDAN algorithm and FCM clustering. Section III gives a full account of experiment results and analysis aiming at one healthy pressure regulator and three faulty pressure regulators as we adopt the traditional and CEEMDAN algorithms to diagnose the outlet pressure of gas regulator respectively. Finally, the FCM clustering is used to classify the fault types. Conclusions are drawn in the last section.

II. METHOD

A. EMPIRICAL MODE DECOMPOSITION AND MODE MIXING

Empirical mode decomposition is a self-adaptive algorithm that was first proposed by Huang *et al*. in [34]. It can decompose the original signals into a series of IMFs and a residue function. Suppose the original signal is $x(n)$, in EMD $x(n)$ can be written as:

$$x(n) = \sum_{k=1}^{K} IMF_k(n) + r(n) \tag{1}$$

where K denotes the number of decomposed IMFs; IMF_k denotes the kth IMF; and $r(n)$ is the residue of the signal $x(n)$.

The realization of EMD algorithm consists of four steps: extract local maximum and minimum of a signal, interpolate extremum points to generate lower and upper envelopes, calculate the mean of the upper and lower envelopes, judge whether the difference between the signal *x(n)* and the mean of envelope is an IMF, and sift and iterate. Moreover, each IMF must satisfy two conditions:

1) The numbers of local extremum points and zero crossings must be equal or differ by one.
2) The local mean of the upper and lower envelopes is zero.

Although the EMD method has proven to be efficient for signal processing, the drawbacks are obvious, such as end effect, sampling rate issue, mode mixing, etc. [44]. If mode mixing occurs, an IMF component no longer has a physical significance by itself, suggesting falsely that there may be different physical processes represented in a mode [45]. An improved method which can decompose IMFs with a narrower frequency spectrum is thus in need.

B. COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH ADAPTIVE NOISE

Aiming at the mode mixing phenomenon, Torres *et al*. [42] recently proposed a method called CEEMDAN. Given a signal *x(n)*, this algorithm defines an operator $E_j(\cdot)$ which produces the *j*th mode decomposed by the EMD. Because adding the white Gaussian noise at each stage of the decomposition directly will lead to incomplete decomposition with residual noise, in this approach, a particular noise $E_j(\omega^i(n))$ is added to extract the *j*th IMF component. The steps of the CEEMDAN algorithm are described as follows:

1) Add $\varepsilon_0\omega^i(n)$ to the original signal $x(n)$ to obtain realization $X(n) = x(n) + \varepsilon_0\omega^i(n)$, where $\omega^i(n)$ denotes the ith added white noise with $N(0,1)$, and ε_k is the kth signal-to-noise ratio (SNR) coefficient. Decompose by EMD I realizations $X(n)$ to obtain the first IMFs and calculate the mean value by

$$\widetilde{IMF}_1(n) = \frac{1}{I}\sum_{i=1}^{I} IMF_1^i(n) = \overline{IMF_1}(n)$$

2) Then calculate the first residue: $r_1(n) = x(n) - \widetilde{IMF}_1(n)$. By adding $\varepsilon_1 E_1[\omega^i(n)]$ to the first residue $r_1(n)$, we can obtain the realizations $r_1(n) + \varepsilon_1 E_1[\omega^i(n)]$, $i = 1, 2...I$.
3) Decompose realizations $r_1(n) + \varepsilon_1 E_1[\omega^i(n)]$ I times, compute the mean value and then obtain the second IMF:

$$\widetilde{IMF}_2(n) = \frac{1}{I}\sum_{i=1}^{I} E_1\{r_1(n) + \varepsilon_1 E_1[\omega^i(n)]\}$$

4) For $k = 2, 3 \ldots K$, calculate the kth residue: $r_k(n) = r_{k-1} - \widetilde{IMF}_k(n)$, decompose the realizations $r_k(n) + \varepsilon_k E_k[\omega^i(n)]$, $i = 1, 2 \ldots I$, and then the $(k+1)$th IMF can be defined as:

$$\widetilde{IMF}_{k+1}(n) = \frac{1}{I}\sum_{i=1}^{I} E_1\{r_k(n) + \varepsilon_k E_k[\omega^i(n)]\}$$

5) Repeat Step 4 for next k until $r_k(n)$ cannot be decomposed. The final residue is:

$$R(n) = x(n) - \sum_{k=1}^{K} \widetilde{IMF}_k$$

Therefore, the original signal is written as:

$$x(n) = \sum_{k=1}^{K} \widetilde{IMF}_k + R(n)$$

Although the single experiment may certainly produce very noisy results, the added $E_j(\omega^i(n))$ can cancel out each other in the ensemble mean of enough experiments. The ensemble mean is regarded as the true IMF.

The ε_k coefficients enable us to choose the SNR at each phase. To determine the parameter settings of the CEEMDAN, the noise amplitude needs to be reduced when the

gas pressure signal is dominated by high-frequency components. When the gas pressure signal is dominated by low-frequency components, the amplitude of added noise should be increased [46]. To alleviate mode mixing and reduce errors, we can increase the ensemble number to a few hundred and the error caused by the added white noise can be reduced to a very small extent or even negligible. Therefore, we used the CEEMDAN with an ensemble number of 500, and the amplitude of noise was 0.2 times standard deviation of the pressure signal.

C. MODIFIED HILBERT–HUANG TRANSFORM AND FEATURE VECTOR

The traditional HHT consists of two parts: the EMD and Hilbert spectrum analysis (HSA). The almost monocomponent IMFs decomposed by EMD provide a proper method for the instantaneous frequency analysis of complex signals. Applying the Hilbert transform to each IMF, we obtain a time-frequency-energy distribution, called Hilbert spectrum. Nevertheless, an IMF can cease to have physical meaning by itself due to the mode mixing originated from the EMD. Thus obtained Hilbert spectrum cannot reveal the signal features accurately. To solve this problem, we used CEEMDAN instead of EMD to decompose the original signal.

After applying the CEEMDAN algorithm on a signal, a set of IMFs with different feature scales is obtained. Suppose the ith IMF is c_i, the Hilbert transform is defined in the following equation:

$$d_i(t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{c_i(\tau)}{t-\tau}d\tau \tag{2}$$

Combining c_i and d_i, an analytic signal can be expressed as:

$$\begin{cases} z_i(t) = c_i(t) + jd_i(t) = a_i(t)e^{j\theta_i(t)} \\ a_i(t) = \sqrt{c_i^2(t) + d_i^2(t)} \\ \theta_i(t) = \arctan\dfrac{d_i(t)}{c_i(t)} \end{cases} \tag{3}$$

where $a_i(t)$ denotes the instantaneous amplitude (IA) and $\theta_i(t)$ is the instantaneous phase.

If the signal $x(t)$ is monocomponent, the instantaneous frequency (IF) can be expressed as:

$$\omega_i(t) = \frac{d\theta_i(t)}{dt} \tag{4}$$

With definitions above, the original signal $x(t)$ can be written as:

$$x(t) = R_e\sum_{i=1}^{n}a_i(t)e^{j\int\omega_i(t)dt} \tag{5}$$

where R_e denotes the real part. According to the Equation (5), the Hilbert spectrum can be defined as:

$$H(\omega, t) = R_e\sum_{i=1}^{n}a_i(t)e^{j\int\omega_i(t)dt} \tag{6}$$

According to Equation (6), the Hilbert marginal spectrum is written as:

$$h(\omega) = \int_0^T H(\omega, t)dt \tag{7}$$

where T denotes the sampling time period. Then according to Equation (7), the total energy of the original signal is defined as:

$$E = \int_0^{\infty} h(\omega)d\omega \tag{8}$$

The Hilbert marginal spectrum can describe the frequency-energy distribution of a pressure signal. In order to analyze these fault features precisely, we divided the HMS further into five frequency bands. The energy of each frequency band is defined as:

$$E_j = \int_{a_j}^{b_j}|h(\omega)|d\omega \qquad (j = 1, 2\cdots 5) \tag{9}$$

where E_j denotes the energy of jth frequency band, a_j and b_j are the lower and upper limits of jth frequency band respectively. With these definitions, we can establish a feature vector $\mathbf{V}$ of an outlet pressure signal as:

$$\mathbf{V} = [E_1/E, \quad E_2/E \quad \cdots \quad E_5/E] \tag{10}$$

$$E = \sum_{i=1}^{5}E_i \tag{11}$$

D. FUZZY C-MEANS AND CLUSTER CENTER

The FCM algorithm, which was proposed by Bezdek *et al*. [43], is a fuzzy analysis method focused on evaluating the centroids of multiple clusters and the activation levels of data modes. Based on a large number of experiments, we found that a gas regulator may have multiple faults in its life cycle, and each fault has a unique clustering structure. Thus, we used the FCM algorithm to identify the types of faults in diagnostic schemes.

The FCM algorithm is based on an objective function, which can be written as:

$$\begin{cases} J(U, Z) = \sum_{i=1}^{C}\sum_{j=1}^{n}(\mu_{ij})^m(d_{ij})^2 \\ i \in (1\ldots t\ldots C) \\ j \in (1\ldots k\ldots n) \end{cases} \tag{12}$$

where U, Z, n, and m denote the membership matrix, cluster center, the number of samples, and the weight index (also known as smoothing factor which is 2 in most cases), respectively. In this paper, we chose random values to initialize a prototype (cluster center) of samples.

The goal of fuzzy c-means clustering is to compute a group of center vectors to minimize the objective function, which can be defined as:

$$\min\{J(U, Z)\} = \sum_{j=1}^{n}\min\{\sum_{i=1}^{C}(\mu_{ij})^m(d_{ij})^2\} \tag{13}$$

In this definition, the membership degree μ_{ij} satisfies the following relationship:

$$\sum_{i=1}^{C} \mu_{ij} = 1 \tag{14}$$

where C denotes the number of the categories. d_{ij}, the degree of distortion between the jth sample point and the ith cluster center, is generally expressed by the distance between two vectors:

$$d_{ij}^2 = \| x_j - z_i \|_2 \tag{15}$$

where x_j denotes the sample point and z_i denotes cluster center.

With these definitions and according to Equation (14), we can use Lagrange method to calculate membership degrees. Suppose the Lagrange function is:

$$F = \sum_{i=1}^{C} (\mu_{ij})^m (d_{ij})^2 + \lambda (\sum_{i=1}^{C} \mu_{ij} - 1) \tag{16}$$

where λ is a parameter. Suppose $\frac{\partial F}{\partial \lambda} = \sum_{i=1}^{C} \mu_{ij} - 1 = 0$, $\frac{\partial F}{\partial \mu_{tk}} = [m(\mu_{tk})^{m-1}(d_{tk})^2 - \lambda] = 0$. We can obtain the following equation:

$$\mu_{tk} = \left[\frac{\lambda}{m(d_{tk})^2}\right]^{\frac{1}{m-1}} = \left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} \left(\frac{1}{(d_{tk})^2}\right)^{\frac{1}{m-1}} \tag{17}$$

Bring Equation (17) into $\sum_{i=1}^{C} \mu_{ik} = 1$, the following equation can be obtained:

$$\sum_{i=1}^{C} \mu_{ik} = \left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} \left\{\sum_{i=1}^{C} [\frac{1}{(d_{ik})^2}]\right\}^{\frac{1}{m-1}} = 1 \tag{18}$$

and the equation in (18) can be written as:

$$\left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} = \frac{1}{\left\{\sum_{i=1}^{C} [\frac{1}{(d_{ik})^2}]\right\}^{\frac{1}{m-1}}} \tag{19}$$

Then bring Equation (19) into Equation (17), the μ_{tk} can be obtained:

$$\mu_{tk} = \frac{1}{\sum_{i=1}^{C} [\frac{d_{tk}}{d_{ik}}]^{-\frac{2}{m-1}}} \tag{20}$$

According to the membership degree μ_{tk}, we can obtain the membership matrix U. For the iteration of updating the prototype, suppose $\frac{\partial J(U,Z)}{\partial z_i} = 0$, we can obtain the following equation:

$$\begin{aligned}
\frac{\partial J(U,Z)}{\partial z_i} &= \sum_{k=1}^{n} (\mu_{ik})^m \frac{2\partial[(x_k - z_i)]^T (x_k - z_i)}{\partial z_i} \\
&= \sum_{k=1}^{n} (\mu_{ik})^m [-4(x_k - z_i)] \\
&= -4[\sum_{k=1}^{n} (\mu_{ik})^m (x_k - z_i)] \\
&= -4[\sum_{k=1}^{n} (\mu_i k)^m x_k - \sum_{k=1}^{n} (\mu_i k)^m z_i] = 0
\end{aligned} \tag{21}$$

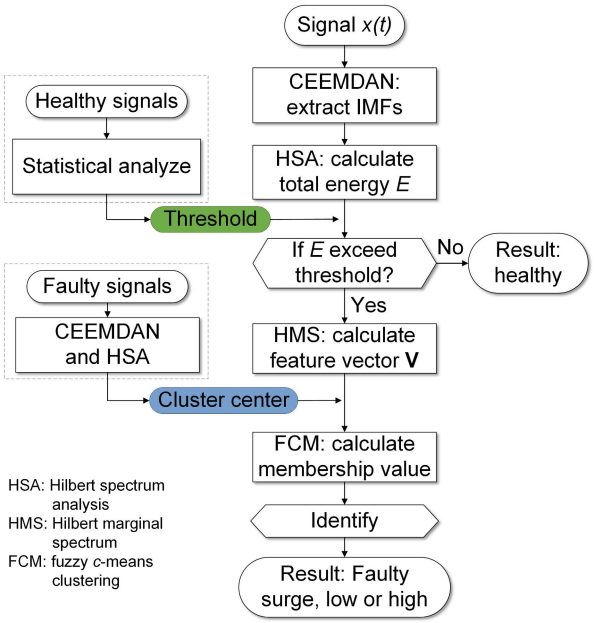


FIGURE 3. Flowchart of the diagnostic scheme for gas regulators.

Then, the prototype (cluster center) can be updated as:

$$z_i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m x_k}{\sum_{k=1}^{n} (\mu_{ik})^m} \tag{22}$$

Repeat Equations (13)–(22) until the z_i coincides with the centroid of a cluster. The final z_i is regarded as the standard cluster center, and we can thus obtain the corresponding membership degrees of samples.

E. DIAGNOSTIC SCHEME BASED ON CEEMDAN AND FCM

The diagnostic scheme for gas regulators consists of two parts: building parameters and diagnostic verification, which are as follows:

1) Calculate total energy E of healthy regulators by CEEMDAN, HSA and statistical analyze, and then set a healthy threshold.
2) Analyze faulty samples to establish a cluster center for each of the three typical faults by our method.
3) Input testing signal $x(t)$. First, calculate its energy and confirm if value exceeds the threshold. If not, this regulator is healthy.
4) If yes, this regulator is faulty. Then extract its IMFs by CEEMDAN, calculate feature vector $\mathbf{V}$ by HMS and identify membership degree by FCM.

The diagnostic scheme for gas regulators based on CEEMDAN and FCM is shown in Fig. 3.

III. EXPERIMENT RESULTS AND ANALYSIS

A. EXPERIMENTAL PLATFORM

The approach proposed above was verified in a gas regulating box as shown in Fig. 1. This platform consisted of a pilot-operated regulator (RTZ-/0.6 AQ, made by Hebei Anxin Co., Ltd.), a gas filter, an inlet pressure

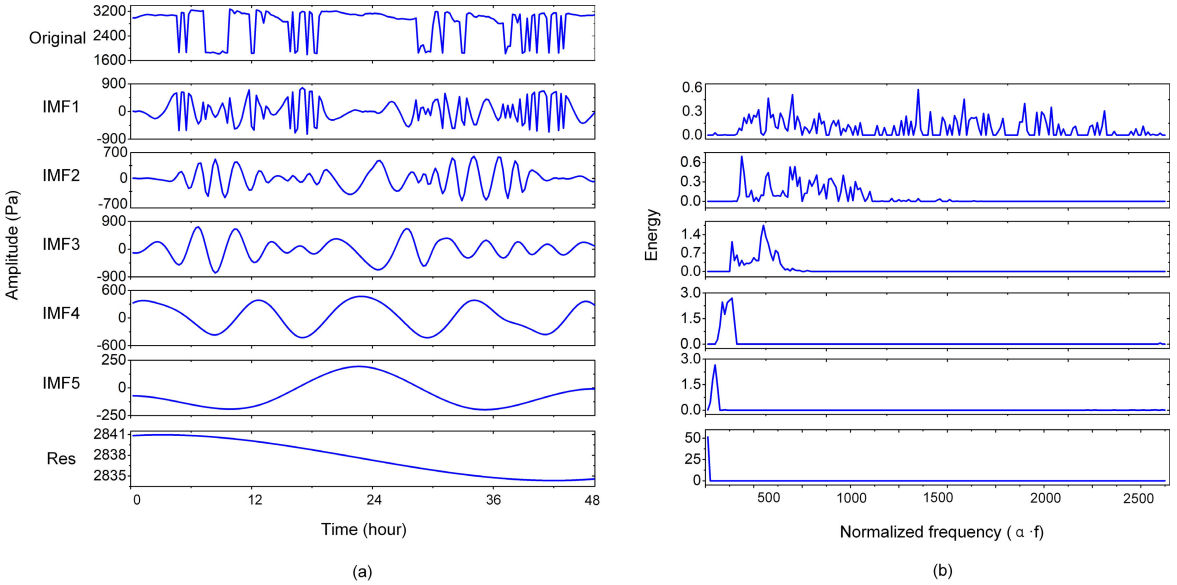


FIGURE 4. The results from a regulator with surge: (a) IMFs based on the EMD, and (b) Hilbert marginal spectra of IMFs.

valve, an outlet pressure valve, and other auxiliary devices.

In identical running conditions, we mainly researched four gas regulators: a regulator with a deformed stem (surge simulation), a regulator with clogged ports (low outlet pressure simulation), a regulator with damaged sealing rings (high closing pressure simulation), and a healthy regulator. We replaced the tested regulator after each group of experiments. The rated running pressure of tested regulators was 2500 Pa, and the inlet pressure was maintained at 20 kPa by upstream regulators. In the regulating box, we installed a wireless pressure monitoring terminal which enabled us to derive sufficient outlet pressure data. For each regulator, similar experiments were conducted and each set of data collection lasted for 2 days. All pressure data from regulators were recorded at 15-minute intervals, totalling 96 pressure values in one day. In order to display experimental results clearly, all frequency domains of figures in next sections ranging from 1 to 2500 were normalized ($\alpha \cdot f, \alpha = 4.6 \times 10^6$).

B. FEATURE EXTRACTION BASED ON CEEMDAN

After the pressure signal was collected, the EMD was applied to decompose the signal into a set of IMFs. The CEEMDAN with a noise amplitude of 0.2 and an ensemble size of 500 was applied on the same signal, respectively. Figs. 4(a) and 5(a) illustrate the decomposed IMF components from a regulator with surge. Meanwhile, Figs. 4(b) and 5(b) describe the corresponding frequency-amplitude (energy) distribution of these IMFs, which is also known as the Hilbert marginal spectrum. By comparing Fig. 4(a) and Fig. 5(a), it is obvious that the CEEMDAN decomposed this faulty signal thoroughly into more monocomponents. Moreover, in Fig. 5(b), each IMF occupied almost a unique frequency band in HMS. Hence, the CEEMDAN algorithm was verified effective for fault feature extraction.

In order to observe the influence of mode mixing, the Hilbert marginal spectrum of IMFs 1 to 5 were obtained by EMD and CEEMDAN, as shown in Fig. 6. We discovered that the lines of spectra based on CEEMDAN were less overlapped than those by EMD. Fig. 5(b) displays a clearer separation of frequency distribution between each IMF, which indicates that CEEMDAN can alleviate mode mixing well enough for further analysis.

C. FEATURE VECTOR BASED ON HILBERT SPECTRUM ANALYSIS

After the IMFs were obtained, the system energy of a regulator was calculated to identify the severity of faults, and necessary parameters including healthy threshold, feature vectors were computed.

First, according to Equations (2)–(7), Fig. 7 shows the obtained Hilbert spectrum of a signal from a healthy regulator and three defective regulators. The x axis of 4 corresponding marginal spectra ranges from 0 to 3.398, which corresponds to the logarithmic value of normalized frequency from 1 to 2500. Any fluctuation of the original signal will produce a local energy peak in time domain of the Hilbert spectrum. By Fig. 7, we found the features of time-frequency-energy distribution as follows:

1) The energy distribution of the signals from healthy regulators is more stable than those from the faulty ones.
2) In high frequency bands, the energy of a regulator with surge is even larger than the other two types of faults.
3) The energy of regulators with the other two types of faults mainly distributed in lower frequency bands, but the energy is much distinct in Fig. 7(c).

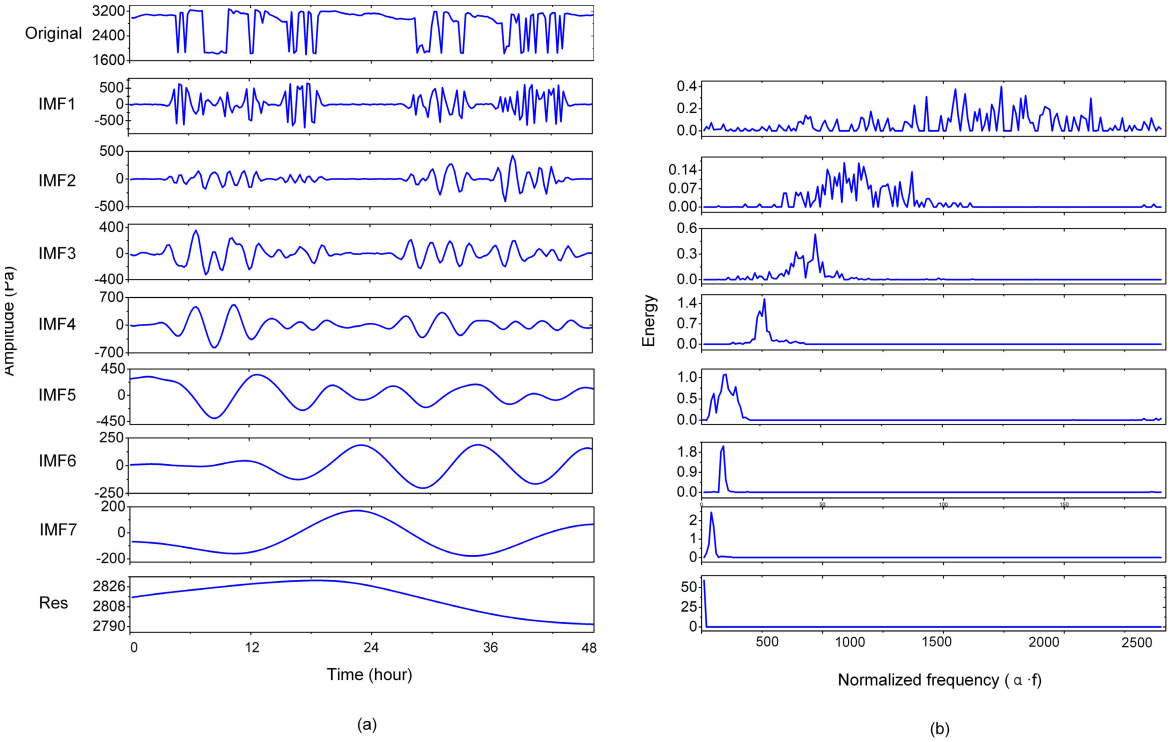


FIGURE 5. The results from a regulator with surge: (a) IMFs based on the CEEMDAN, and (b) Hilbert marginal spectra of IMFs.

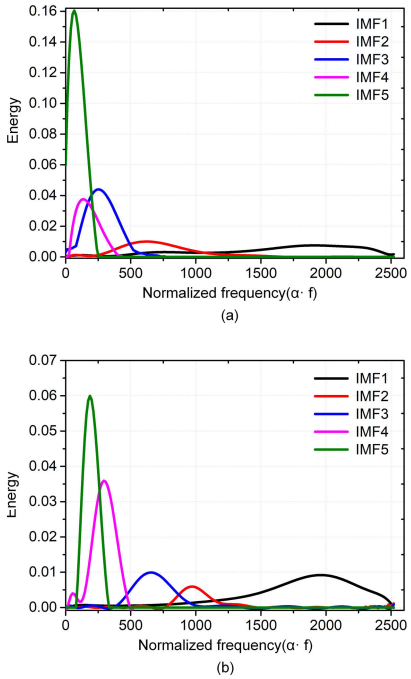


FIGURE 6. Spectra of IMFs 1 to 5 obtained by (a) the EMD, and (b) the CEEMDAN.

Second, according to Equation (8), the total energy of each operating state can be calculated based on their marginal spectra. As shown in Fig. 7, we found that the energy value of a healthy system is far below other defective ones and this is consistent with the energy distribution in Hilbert spectrum. Therefore, there should be a healthy threshold which can identify the severity of a fault. In this paper, we computed this threshold by 51 data sets from a healthy pressure regulator. The method to acquire each data set is the same as mentioned in Experimental Platform section. The normalized threshold 7.7646 in Fig. 8 is the average of data plus three times standard deviation. By this threshold, we could pick healthy signals out of an unknown signal pool.

For the defective signals, the key parameter to distinguish the fault types was its feature vector. The y axis of the HMS was divided into five frequency bands to calculate the local energy component E_i. As shown in Fig. 7, the normalized frequency band from 1 to 500 in every Hilbert spectra included massive energy, so it was divided into two bands. As a result, these normalized frequency bands vary in the range of 1–253, 254–508, 509–1021, 1022–1533, and 1534–2500. Then according to Equations (8)–(10), the energy ratios were calculated to build the feature vectors. For gas regulators monitored in this experiment, the feature vectors of three typical faults are shown in Fig. 9. It can be seen that the energy ratios of high frequencies E_1/E, E_2/E, E_3/E, E_4/E are higher for a regulator with surge (see Fig. 9). In comparison, the energy ratio of low frequency E_5/E is significantly greater for regulators with other two faults. Based on the above analysis, the distinct differences among vectors indicate that they are able to represent corresponding faults.

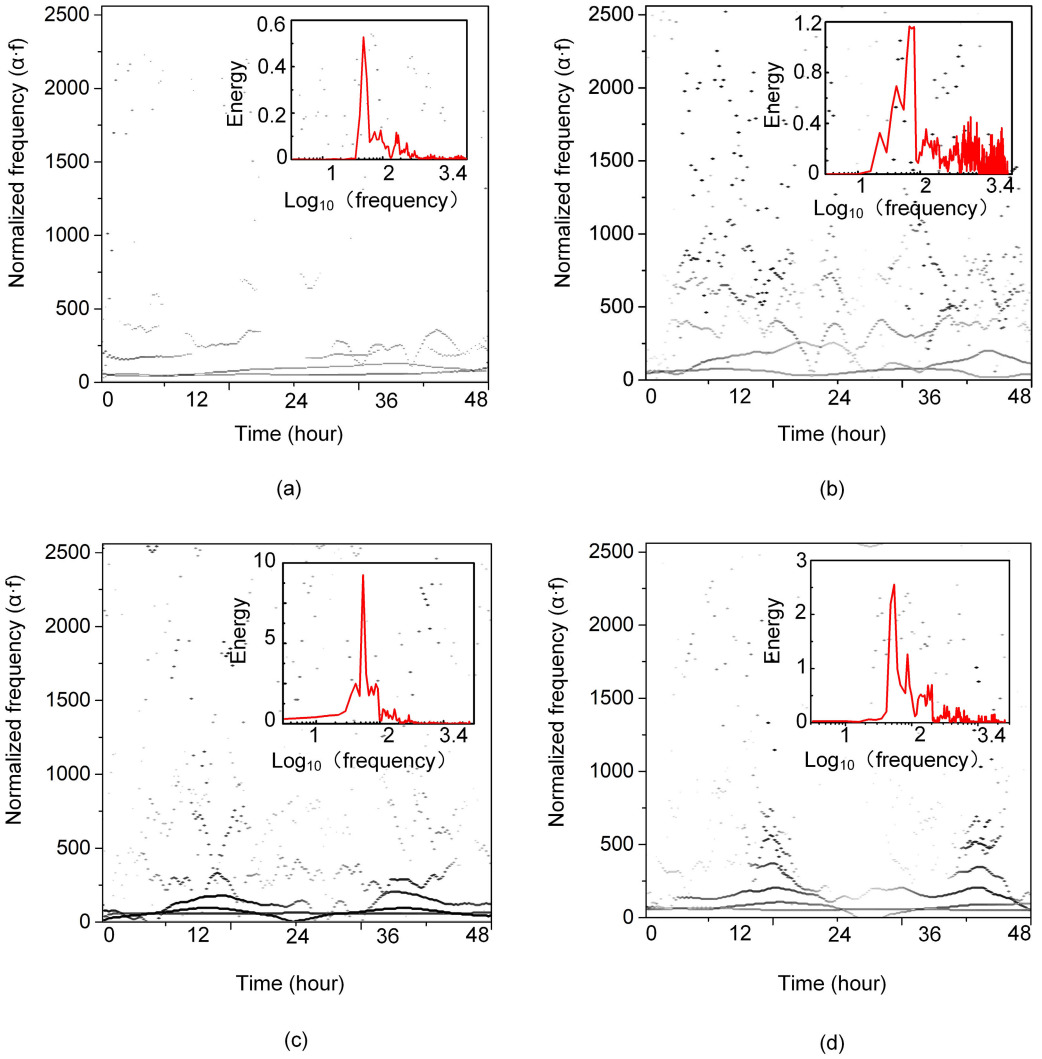


FIGURE 7. The HS and corresponding HMS of a signal from (a) healthy, (b) surge, (c) low outlet pressure, and (d) high closing pressure regulator.

TABLE 2. The standard cluster center of FCM algorithm based on the EMD.

	E_1	E_2	E_3	E_4	E_5
Surge	0.0867	0.0661	0.1461	0.2480	0.4530
Low Outlet Pressure	0.0273	0.0141	0.0393	0.0780	0.8412
High Closing Pressure	0.0382	0.0230	0.0659	0.1905	0.6823

D. FAULT DIAGNOSIS BASED ON FCM CLUSTERING

As described in Experimental Platform section, three groups of experiments – a regulator with surge fault, a regulator with low outlet pressure, and a regulator with high closing pressure – were conducted to classify the fault types. The data sampling interval for each experimental group was 15 minutes and a data collection procedure lasted for two days. Each experimental group consisted of two phases: building standard cluster centers of a typical fault, and computing the membership degrees of test samples.

The goal of Phase 1 was to build 3 standard cluster centers of typical faults. The sampling procedure was repeated 8 times to obtain a total of 8 data sets for each experimental group. Thus for 8 data sets, the number of feature vectors is 24. Then, the FCM clustering method was used to calculate a cluster center for each cluster of feature vectors, and these cluster centers were set as the standard cluster centers. To verify the stability of our method, we also applied the EMD on the same data sets as comparison groups. Tables 2 and 3 present the standard cluster

TABLE 3. The standard cluster center of FCM algorithm based on the CEEMDAN.

	E_1	E_2	E_3	E_4	E_5
Surge	0.1685	0.0840	0.1829	0.1480	0.4166
Low Outlet Pressure	0.0390	0.0156	0.0393	0.0718	0.8343
High Closing Pressure	0.0595	0.0284	0.0852	0.1579	0.6689

TABLE 4. The membership matrix of the test samples from the EMD after FCM.

Diagnosis Results		Test Samples											
		Surge				Low Outlet Pressure				High Closing Pressure			
Cluster Center	Surge	0.882	0.771	0.921	0.843	0.020	0.042	0.023	0.042	0.554	0.039	0.038	0.035
	Low	0.025	0.052	0.020	0.037	0.834	0.309	0.894	0.309	0.072	0.351	0.557	0.197
	High	0.093	0.177	0.059	0.120	0.147	0.649	0.083	0.649	0.347	0.610	0.405	0.768

TABLE 5. The membership matrix of the test samples from the CEEMDAN after FCM.

Diagnosis Results		Test Samples											
		Surge				Low Outlet Pressure				High Closing Pressure			
Cluster Center	Surge	0.955	0.977	0.967	0.965	0.006	0.001	0.010	0.001	0.009	0.008	0.004	0.002
	Low	0.012	0.007	0.010	0.010	0.957	0.993	0.945	0.992	0.019	0.026	0.012	0.005
	High	0.033	0.016	0.023	0.026	0.038	0.006	0.045	0.007	0.972	0.966	0.984	0.994

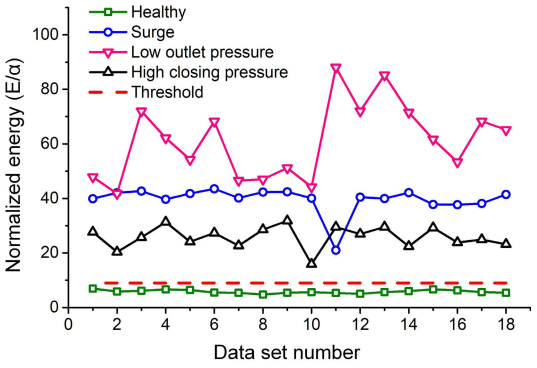


FIGURE 8. Energy of signals from a healthy and three typical faulty gas regulators. (Normalized energy $= E/\alpha$, $\alpha = 4.6 \times 10^6$.)

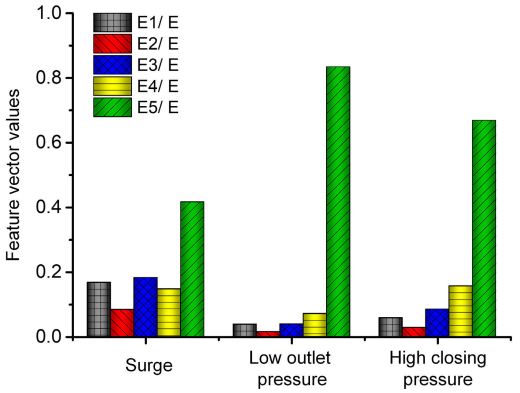


FIGURE 9. The feature vector of a signal from (a) a regulator with surge, (b) a regulator with low outlet pressure, and (c) a regulator with high closing pressure.

centers of FCM algorithm based on the EMD and the CEEMDAN.

The goal of Phase 2 was to compute the membership degrees of test samples. To classify the fault type of an unknown signal, we repeated the sampling procedure to collect another 18 more data sets. Then, we calculated the feature vectors of these data sets, and obtained their membership degrees with standard cluster centers. A part of the diagnosis results are shown in Tables 4 and 5. In Table 4, false classification results are presented. Two samples with low outlet pressure were classified as ones with high closing pressure. Another two samples with high closing pressure were classified as one with surge, and one with low outlet pressure. In Table 5, by contrast, we can observe that the diagnostic results based on the CEEMDAN algorithm is relatively ideal.

To validate the diagnostic accuracy of proposed approach, we used identical data sets to conduct two more groups of experiments based on STFT with Hamming window function, and WT using a db4 mother wavelet. The whole diagnosis results of 18 data sets (54 test samples) based on STFT, WT, EMD, and CEEMDAN are shown in Fig. 10. It can be seen that the diagnostic accuracy calculated by the STFT, WT and EMD algorithms are unstable. Some membership degrees

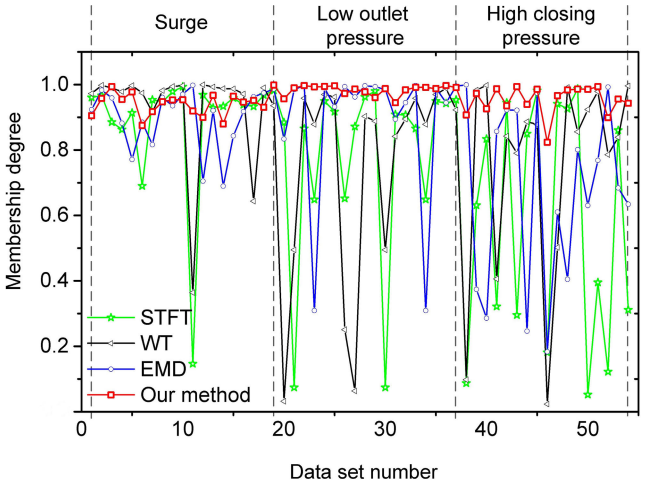


FIGURE 10. **The stability of diagnosis results based on STFT, WT, EMD, and CEEMDAN.**

are significantly low, and false diagnosis results often occur: eleven errors in STFT, nine errors in WT, and seven errors in EMD. Hence, the fault diagnosis accuracy of STFT, WT, and EMD is 79.6%, 83.3%, and 87.0% respectively. By contrast, for the CEEMDAN algorithm, all the samples are classified into three categories correctly, and the membership degree of test samples ranges between 0.9 to 1. Overall, the proposed diagnostic method has been proved to be effective by experimental results.

IV. CONCLUSION

In this paper, a novel fault diagnosis approach for gas pressure regulators based on CEEMDAN and feature clustering is proposed. The CEEMDAN algorithm was used to extract fault features of a regulator. In comparison with the EMD, the CEEMDAN was able to alleviate the mode mixing significantly. Hence, we can obtain a more accurate Hilbert spectrum to reflect fault information distibution. Then, FCM was used to classify diverse types of faults. Experimental results indicate that the proposed method can achieve diagnostic results with a higher degree of accuracy.

The currently proposed fault diagnosis approach for gas pressure regulators is competent for further application. When the parameters or structures of gas regulator are changed, accurate diagnosis results may not be obtained. Even so, the proposed method is capable of diagnosing more fault types effectively if requisite parameters mentioned in this paper are known. However, compound faults in single regulators are not taken into account in this paper. In future research, an improved approach which can identify compound faults on single gas pressure regulators will be considered.

REFERENCES

[1] N. Zafer and G. R. Luecke, "Stability of gas pressure regulators," *Appl. Math. Model.*, vol. 32, no. 1, pp. 61–82, 2008.

[2] C. Werner, G. Preiß, F. Gores, M. Griebenow, and S. Heitmann, "A comparison of low-pressure and supercharged operation of polymer electrolyte membrane fuel cell systems for aircraft applications," *Prog. Aerosp. Sci.*, vol. 85, pp. 51–64, Aug. 2016.

[3] B. Sun, Q. Xu, and C. Yang, "Dynamic modeling and simulation of a pressurized system used in flight vehicle," *Chin. J. Aeronaut.*, vol. 31, no. 6, pp. 1232–1248, 2018.

[4] L. Gao, L. Jin, F. Wang, Y. Zheng, and K. Li, "Genetic algorithm–based varying parameter linear quadratic regulator control for four-wheel independent steering vehicle," *Adv. Mech. Eng.*, vol. 7, no. 11, 2015, Art. no. 1687814015618632.

[5] S. V. Serdyukov, A. V. Azarov, P. A. Dergach, and A. A. Duchkov, "Equipment for microseismic monitoring of geodynamic processes in underground hard mineral mining," *J. Mining Sci.*, vol. 51, no. 3, pp. 634–640, 2015.

[6] J. Szoplik, "Forecasting of natural gas consumption with artificial neural networks," *Energy*, vol. 85, pp. 208–220, Jun. 2015.

[7] G.-W. Bang, B.-S. Kim, and N.-G. Gwangju, "A development of automotive solenoid regulator," *Int. J. Pure Appl. Math.*, vol. 116, no. 23, pp. 733–741, 2017.

[8] N. Bariha, I. M. Mishra, and V. C. Srivastava, "Hazard analysis of failure of natural gas and petroleum gas pipelines," *J. Loss Prevention Process Ind.*, vol. 40, pp. 217–226, Mar. 2016.

[9] W. Yahui, Z. Fucheng, and C. Fangwen, "The gas regulator fault diagnosis based on PCA-RBF neural network," in *Proc. 20th Int. Conf. Elect. Mach. Syst. (ICEMS)*, Aug. 2017, pp. 1–6.

[10] H. V. Krechkovs'ka, A. B. Mytsyk, O. Z. Student, and H. M. Nykyforchyn, "Diagnostic indications of the in-service degradation of the pressure regulator of a gas-transportation system," *Mater. Sci.*, vol. 52, no. 2, pp. 233–239, 2016.

[11] Z. Lou, D. Shen, and Y. Wang, "Preliminary-summation-based principal component analysis for non-Gaussian processes," *Chemometrics Intell. Lab. Syst.*, vol. 146, pp. 270–289, Aug. 2015.

[12] Y. Wang, Y. Si, B. Huang, and Z. Lou, "Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008–2017," *Can. J. Chem. Eng.*, vol. 96, no. 10, pp. 2073–2085, 2018.

[13] Y. Dong, H. Wang, Y. Wang, and F. Chen, "Research on fault diagnosis algorithm of gas pressure regulator based on compressed sensing theory," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, May 2017, pp. 2702–2707.

[14] X. Hao, X. An, B. Wu, and S. He, "Application of a support vector machine algorithm to the safety precaution technique of medium-low pressure gas regulators," *J. Therm. Sci.*, vol. 27, no. 1, pp. 74–77, 2018.

[15] J. Szoplik, "The gas transportation in a pipeline network," in *Advances in Natural Gas Technology*. Rijeka, Croatia: InTechOpen, 2012.

[16] H. Mishra, P. Karkulali, A. Ukil, and J. Dauwels, "Testbed for real-time monitoring of leak in low pressure gas pipeline," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2016, pp. 459–462.

[17] X. Hao, Q. Liu, G. Yang, and Y. Du, "Using the EMD method to determine fault criterion for medium-low pressure gas regulators," *J. Therm. Sci.*, vol. 24, no. 6, pp. 557–561, 2015.

[18] T. Wang, J. Qi, H. Xu, Y. Wang, L. Liu, and D. Gao, "Fault diagnosis method based on FFT-RPCA-SVM for cascaded-multilevel inverter," *ISA Trans.*, vol. 60, pp. 156–163, Jan. 2016.

[19] A. Sapena-Bañó, M. Pineda-Sanchez, R. Puche-Panadero, J. Martinez-Roman, and D. Matic, "Fault diagnosis of rotating electrical machines in transient regime using a single stator current's FFT," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 11, pp. 3137–3146, 2015.

[20] C. Gan, J. Wu, S. Yang, Y. Hu, W. Cao, and J. Si, "Fault diagnosis scheme for open-circuit faults in switched reluctance motor drives using fast Fourier transform algorithm with bus current detection," *IET Power Electron.*, vol. 9, pp. 20–30, Jan. 2016.

[21] H. Liu, L. Li, and J. Ma, "Rolling bearing fault diagnosis based on STFT-deep learning and sound signals," *Shock Vib.*, vol. 2016, Jul. 2016, Art. no. 6127479.

[22] L.-H. Wang, X.-P. Zhao, J.-X. Wu, Y.-Y. Xie, and Y.-H. Zhang, "Motor fault diagnosis based on short-time Fourier transform and convolutional neural network," *Chin. J. Mech. Eng.*, vol. 30, no. 6, pp. 1357–1368, 2017.

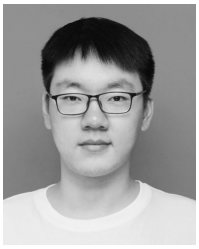
[23] A. F. Aimer, A. H. Boudinar, N. Benouzza, and A. Bendiabdellah, "Bearing fault diagnosis of a PWM inverter fed-induction motor using an improved short time Fourier transform," *J. Elect. Eng. Technol.*, vol. 14, no. 3, pp. 1201–1210, 2019.

[24] L. Y. Zhao, L. Wang, and R. Q. Yan, "Rolling bearing fault diagnosis based on wavelet packet decomposition and multi-scale permutation entropy," *Entropy*, vol. 17, no. 9, pp. 6447–6461, 2015.
[25] M. Sharma, R. B. Pachori, and U. R. Acharya, "A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension," *Pattern Recognit. Lett.*, vol. 94, pp. 172–179, Jul. 2017.
[26] J. Chen, Z. Li, J. Pan, G. Chen, Y. Zi, J. Yuan, B. Chen, and Z. He, "Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review," *Mech. Syst. Signal Process.*, vol. 70, pp. 1–35, Mar. 2016.
[27] X. Yu, F. Dong, E. J. Ding, S. P. Wu, and C. Y. Fan, "Rolling bearing fault diagnosis using modified LFDA and EMD with sensitive feature selection," *IEEE Access*, vol. 6, pp. 3715–3730, 2018.
[28] Y. Li, M. Xu, Y. Wei, and W. Huang, "An improvement EMD method based on the optimized rational Hermite interpolation approach and its application to gear fault diagnosis," *Measurement*, vol. 63, pp. 330–345, Mar. 2015.
[29] A. Mejia-Barron, M. Valtierra-Rodriguez, D. Granados-Lieberman, J. C. Olivares-Galvan, and R. Escarela-Perez, "The application of EMD-based methods for diagnosis of winding faults in a transformer using transient and steady state currents," *Measurement*, vol. 117, pp. 371–379, Mar. 2018.
[30] Y. Wang, H. Li, and P. Ye, "Fault feature extraction of hydraulic pump based on CNC de-noising and HHT," *J. Failure Anal. Prevention*, vol. 15, no. 1, pp. 139–151, 2015.
[31] Y. Ji and H. Wang, "A revised Hilbert–Huang transform and its application to fault diagnosis in a rotor system," *Sensors*, vol. 18, no. 12, p. 4329, 2018.
[32] Y. Jia, L. Qingmao, and Y. Xuyi, "The diaphragm pump spindle fault diagnosis using HHT based on EMD," *Open Autom. Control Syst. J.*, vol. 7, pp. 640–645, Jun. 2015. doi: 10.2174/1874444320150610E001.
[33] J. Yao, B. Tang, and J. Zhao, "Improved discrete Fourier transform algorithm for harmonic analysis of rotor system," *Measurement*, vol. 83, pp. 57–71, Apr. 2016.
[34] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
[35] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 112–114, Feb. 2004.
[36] S. Al-Baddai, K. Al-Subari, E. Lang, and B. Ludwig, "Optimizing approach for sifting process to solve a common type of empirical mode decomposition mode mixing," *Int. J. Elect. Comput. Eng.*, vol. 11, no. 6, pp. 678–681, 2017.
[37] C. Yi, D. Wang, W. Fan, K. L. Tsui, and J. Lin, "EEMD-based steady-state indexes and their applications to condition monitoring and fault diagnosis of railway axle bearings," *Sensors*, vol. 18, no. 3, p. 704, 2018.
[38] M. Z. S. Žvokelj and I. Prebil, "EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis," *J. Sound Vib.*, vol. 370, pp. 394–423, May 2016.
[39] L. Zhao, W. Yu, and R. Yan, "Gearbox fault diagnosis using complementary ensemble empirical mode decomposition and permutation entropy," *Shock Vib.*, vol. 2016, Oct. 2016, Art. no. 3891429.
[40] J. Chen, D. Zhou, C. Lyu, and C. Lu, "An integrated method based on CEEMD-SampEn and the correlation analysis algorithm for the fault diagnosis of a gearbox under different working conditions," *Mech. Syst. Signal Process.*, vol. 113, pp. 102–111, Dec. 2018.
[41] Y. Li, Y. Li, X. Chen, J. Yu, H. Yang, and L. Wang, "A new underwater acoustic signal denoising technique based on CEEMDAN, mutual information, permutation entropy, and wavelet threshold denoising," *Entropy*, vol. 20, no. 8, p. 563, 2018.
[42] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.
[43] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, 1984.
[44] R. Sharma, L. Vignolo, G. Schlotthauer, M. A. Colominas, H. L. Rufiner, and S. R. M. Prasanna, "Empirical mode decomposition for adaptive AM-FM analysis of speech: A review," *Speech Commun.*, vol. 88, pp. 39–64, Apr. 2017.
[45] Y. Lei, "Fault diagnosis of rotating machinery based on empirical mode decomposition," in *Structural Health Monitoring*. Cham, Switzerland: Springer, 2017, pp. 259–292.
[46] M. Kedadouche, M. Thomas, and A. Tahan, "A comparative study between empirical wavelet transforms and empirical mode decomposition methods: Application to bearing defect diagnosis," *Mech. Syst. Signal Process.*, vol. 81, pp. 88–107, Dec. 2016.
[47] P. Y. Dibal, E. N. Onwuka, J. Agajo, and C. O. Alenoghena, "Application of wavelet transform in spectrum sensing for cognitive radio: A survey," *Phys. Commun.*, vol. 28, pp. 45–57, 2018.

SHEN TIAN received the B.E. degree from North China Electric Power University, Beijing, China, in 2012, and the master's degree in business administration from Zhengzhou University, in 2018. In 2013, he was a Project Manager with the Research and Development Center, Hanwei Electronics Group Corporation, Zhengzhou, China. His current research interests include sensor application and signal processing.

XIAOYU BIAN is currently pursuing the bachelor's degree in electrical engineering with Zhengzhou University, Zhengzhou, China. His current research interests include array signal processing, machine learning, and computer vision.

ZHIPENG TANG received the B.E. degree in electronic engineering from Zhengzhou University, Zhengzhou, China, in 2018, where he is currently pursuing the master's degree in electrical engineering. His current research interests include array signal processing, machine learning, and computer vision.



KUAN YANG received the B.S. degree in electrical engineering from Zhengzhou University, Zhengzhou, China, in 2017, where he is currently pursuing the master's degree in instrumentation engineering. His current research interests include array signal processing, machine learning, holographic imaging, and computational imaging.

LEI LI received the B.S. degree in electronic engineering from Zhengzhou University, Zhengzhou, China, in 2004, and the Ph.D. degree in electronic engineering from the Institute of Acoustics, Chinese Academy of Sciences, in 2009. He is currently an Associate Professor with the School of Physics and Engineering, Zhengzhou University. His current research interests include array signal processing, machine learning, and computer vision.