Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

# Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes

Zuyu Yin [a,*], Jian Hou [b]

[a] Harbin Institute of Technology, Harbin, China
[b] Bohai University, Jinzhou, China

ARTICLE INFO

*Article history:*
Received 9 June 2015
Received in revised form
19 August 2015
Accepted 24 September 2015
Communicated by Rongni Yang

*Keywords:*
Support vector machine
Fault monitoring and diagnosis
Survey

ABSTRACT

With the advancement of industrial systems, fault monitoring and diagnosis methods based on the data-driven attract much attention in recent years. This kind of methods are widely used in engineering projects, especially in those big and complicated machines, whose conditions are difficult to obtain from straight view. They can provide the administrator with effective fault information in initial phase and therefore reduce the loss caused by faults. This paper reviews the research and development of fault diagnosis and monitoring approach based on support vector machine (SVM). While many other methods, such as expert system and artificial neural network, have been used in fault monitoring and diagnosis, SVM shows its advantage in generalization performance and in case of small sample. Therefore, it should attract more attention.



## 1. Introduction

Fault monitoring and diagnosis methods originated from America and their development satisfied the crying needs in aerospace and military, etc. In 1964, there were a series of serious equipment failures in the Apollo Project of The United States, which urged American Naval Research Laboratory (ONR) to built the Machinery Faults Predict Group (MFPG) from the avocations of NASA in 1967. This created a new field of equipment fault diagnosis and made an important contribution to the fault diagnostic theory and technology.

Many industrial developed countries, such as Japan and some European countries, spent a large amount of manpower, financial resources and material resources to develop equipment diagnosis technology from the end of the 1960s to the early of 1970s. In United Kingdom, U.K. Mechanical Health Monitoring Center was the first organization to start the research and got decent results in many aspects, such as publicity, training, consulting, planning, fault analysis and the development of diagnostic technology.

China started the equipment fault diagnosis technology researches in the mid-1980s. In general, China followed the pace of these industrial advanced countries and obtained some decent results in some specific fields. Colleges, research institutions and factories are responsible for research, development and application, respectively.

The key point to fault diagnosis is not the faults itself, but the diagnosis method. Because of the complexity of fault diagnosis process, it is necessary to use not a single method but a variety of methods to deal with the problem. At present, existing fault diagnosis methods can be roughly divided into three categories, i.e., traditional diagnostic methods, the methods based on the signal and intelligent diagnosis methods. Correspondingly, the development of fault diagnosis technology can also be roughly divided into the following three stages.

The first stage is dominated the traditional methods which rely heavily on humans' senses to obtain fault information.

In this stage, experts directly use the information, such as vibration, sound, light, heat, electricity, magnetic and chemical, to diagnose the faults. They combine these information with the fault phenomenon and observe their changing rules and characteristics to dispose the faults directly. This kind of methods are rapid and easy to understand, but only part of faults can be detected with it.

In the second stage, the majority of efforts are devoted to signal based methods, whose foundations are sensor technology, dynamic testing technology, signal analysis and processing technology.

Experts have been working on the collection, selection, processing and analysis process of fault signal. Equipment running signals, such as vibration, rotary speed, can be collected by sensors and the equipment conditions or faults can be identified and evaluated though processing in time-domain and frequency-domain. The fault diagnostic technique in this stage obtained a number of applications in practice and generate huge economic

* Corresponding author.
*E-mail address:* yinzuyuhit@gmail.com (Z. Yin).

http://dx.doi.org/10.1016/j.neucom.2015.09.081
0925-2312/

benefits, which further promoted the development of this technology.

The third stage is represented by intelligent diagnosis methods, whose foundation is traditional fault diagnosis methods and the core is artificial intelligence technology.

The main tools used in methods of this stage are knowledge application, knowledge processing and knowledge reasoning instead of model building and information processing. There are big differences between these methods and the traditional ones in research contents and implement methods. New ideas and methods constantly emerge and fault diagnosis technology develops rapidly in this field.

The characteristic of intelligent fault diagnosis is that it can effectively obtain, transmit, process and regenerate the diagnosis information and can diagnose and forecast the information accurately. Fault intelligent diagnosis technology is closely related to the development of artificial intelligence technology. A variety of new theories and methods, especially the expert system, provide the possibility for the intelligentialize of fault diagnosis.

Expert system is proposed to use the experts' domain knowledge to deal with fault diagnosis issues. Its application relies heavily on the experts' domain knowledge which is hard to acquire. Knowledge acquisition is the bottleneck problem in expert system. Besides, expert system has its limitations in the adaptive capacity, learning ability and real-time performance. Machine learning is another important research field in the technology of artificial intelligence and it has been used in the diagnostic techniques since the 1980s.

With machine learning techniques, diagnosis system does not rely on the experience knowledge and can learn new knowledge from changes of the environment. These abilities make machine learning become a main way to improve the performance of intelligent fault diagnosis. In machine learning, artificial neural network is an efficient approach to dispose the problem of complicated non-linear issues. It has a rapid development over the past decade and has been applied in a range of fields of the fault intelligent diagnosis.

The neural network algorithm is based on traditional statistics which plays the basis role in machine learning. The content of traditional statistical research is the gradual theory, namely the ultimate properties when the sample tends to infinity. However, this precondition cannot be guaranteed in practical applications, as samples are usually rare.

In the case of a limited number of samples, the neural network algorithm often shows a poor generalization ability, namely the over-fitting problem. It is based on empirical risk minimization, but not expected risk minimization, and this leads to poor minimization performance. With the development of neural network, people realize the neural network algorithm is lack of quantitative analysis and mechanism of the complex theory and its research encounters difficulties.

An ideal machine learning approach should have strong generalization ability and be able to find out as much information as possible from the limited samples. Statistical learning theory is good at looking for the optimal solution in small samples and does not need to use the progressive condition when sample tends to infinity. In addition, it has strong promotion value and it is suitable for fault diagnosis of the small sample situation in many practical engineering problems.

Structural risk minimization principle is an effective tool for practical applications of statistical learning theory, and it is the basis of the SVM. The greatest advantage of SVM in fault diagnosis is that it is suitable for small sample based decision, as it can maximally excavate the implicit classification knowledge in the data.

This paper is inspired by the work in [1], which reviewed the articles from 1996 to 2006 in the field of machine condition monitoring and machine fault diagnosis using SVM. In this paper we follow [1] and make a survey of works in the same field from 2007 to 2015.

## 2. Support vector machine

SVM was firstly proposed by Vapnik et al. in [2], and it has become a popular approach to deal with the problem of classification. SVM is based on Structural Risk Minimization (SRM) principle in the statistical learning theory, and it has outstanding generalization performance [3]. Structural Risk Minimization means maximizing the margin between different classes. That is why SVM is not only a useful statistical theory but also a method that can be used to deal with engineering problems [4]. In Fig. 1,

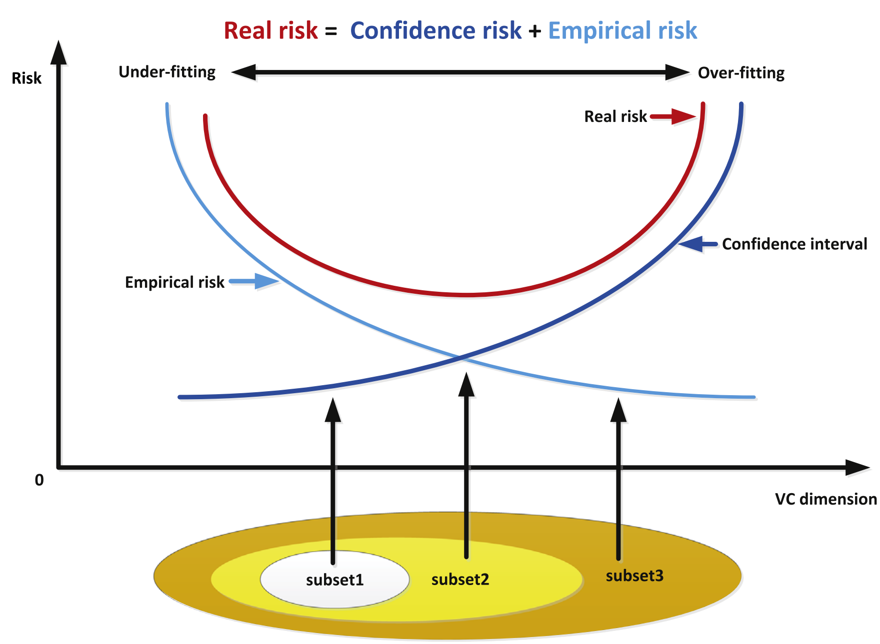


**Fig. 1.** Sketch map of SRM.

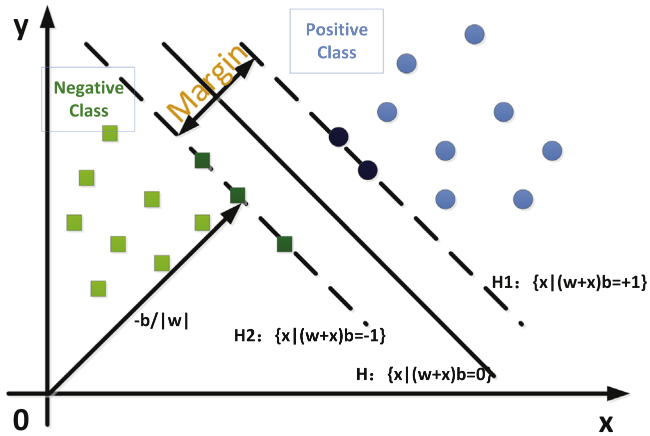


**Fig. 2.** Two classes classification using SVM.

$s_1 \subset s_2 \subset s_3$ and the magnitude of Vapnik–Chervonenkis dimension is $h_1 \le h_2 \le h_3$.

### 2.1. Introduction to svm

The initial idea of SVM is to use a linear separating hyperplane to divide the training samples into two classes. Generally, there are two types of methods suitable to accomplish this task [5]. The first one is finding out the optimal decision hyperplane which bisects the two closest samples into two convex hulls. The second one is to research the optimal decision hyperplane which makes the margin between the two parallel supporting planes maximum. Both these two methods can produce the optimal decision hyperplane and the support points (support vectors).

### 2.2. Linear classification

The linear classification function of SVM is used to separate the training data into two classes with a linear separating hyperplane.

Fig. 2 shows data of two classes, where the squares belong to the negative class and the circles are of the positive class. The SVM tries to put a linear boundary in the middle of the two classes. The maximum margin is represented as the distance between two imaginary lines. The points, located on the imaginary lines, are called support vectors. Support vectors are the most significant samples as they contain all the information used to design the classifier [1]. The training data represented by

$$(x_1, y_1), \ldots, (x_l, y_l), x \in R^n, y \in \{+1, -1\} \tag{1}$$

can be divided by a hyperplane equation:

$$(w \cdot x) + b = 0 \tag{2}$$

If hyperplane can make sure the distance between different classes of samples is maximum, this hyperplane is the optimal hyperplane. We can explain the separating hyperplane using the form below:

$$(w \cdot x_i) + b \ge 1 \quad \text{if } y_i = 1 \tag{3}$$

$$(w \cdot x_i) + b \le 1 \quad \text{if } y_i = -1 \tag{4}$$

Or equivalently

$$y_i[(w \cdot x_i + b)] \ge 1 \quad i = 1, 2, \ldots, l \tag{5}$$

Then the problem of searching for the optimal hyperplane could be transferred into the problem of obtaining the minimum of the function below subject to the constraints in Eq. (6).

$$\Phi(w) = \frac{1}{2}\|w\|^2 = \frac{1}{2}(w \cdot w) \tag{6}$$

Therefore, the optimal hyperplane can be obtained by solving a simple quadratic programming problem.

According to the optimization theory of convex quadratic programming method, this problem can be solved by converting it into a Wolfe dual problem.

Lagrange function is constructed in Eq. (7).

$$L(w, b, \alpha) = \frac{1}{2} w^T \cdot w - \sum_{i=1}^{n} \alpha_i [y_i (w^T \cdot x_i - b) - 1] \tag{7}$$

$\alpha_i$ is Lagrange multiplier, which should satisfy the constraint $\alpha_i \ge 0, i = 1, \ldots, l$. When condition (6) reaches its extremum, the corresponding points should satisfy Eqs. (8) and (9).

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0 \tag{8}$$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{n} \alpha_i y_i x_i \tag{9}$$

Substitute Eqs. (8) and (9) into Lagrange function and eliminate $w$ and $b$.

$$w(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j}^{l} \alpha_i \alpha_j y_i y_j (x_i, x_j) \tag{10}$$

Eq. (10) should under the constraint $\alpha_i \ge 0, i = 1, \ldots, l$ and $\sum_{i=1}^{l} \alpha_i y_i = 0$. Those $x_i \cdot s$ with $\alpha_i > 0$ are termed support vectors (SVs). The label of a testing data $x_i$ can then be obtained by

$$f(x) = sign\left( \sum_{x_i \in X_{SVM}} \alpha_i y_i (x \cdot x_i) + b \right) \tag{11}$$

where

$$w = \sum_{x_i \in X_{SVM}} \alpha_i y_i x_i$$

### 2.3. Non-linear classification

Non-linear is always caused by nonlinearities or noise, which is the source of non-linear data. In the case of the training samples is not linearly separable, Cortes introduced nonnegative variables $\xi_i$ and penalty function $F(\xi) = \sum \xi_i$ in order to promote the optimal hyperplane to general situation. A slack variable is introduced to condition (5).

$$y_i(w^T \cdot \varphi(x_i) + b) \ge 1 - \xi_i \tag{12}$$

The general classification hyperplane is minimum value of Eq. (13) under the constrain (12).

$$\min \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{l} \xi_i \tag{13}$$

The condition $y_i(w \cdot x_i + b) \ge 1 - \xi_i$ can decide the value of the penalty term(c). The formula used in the Non-linear SVM is the same as the one used in the linear SVM, but the constraint $\alpha_i$ is different, which should satisfy the constrain $0 \le \alpha_i \le C$. Based on the theory of functional, as long as a kernel function $K(x_i, y_i)$ satisfy the Mercer condition, it is the inner product of one transformation space. Therefore, using an appropriate inner product function $K(x_i, y_i)$ in the optimal classification plane can achieve linear classification after a nonlinear transformation, and the computational

**Table 1**
Examples of kernel functions.

| Kernel functions | Formula |
|---|---|
| Linear | $k(x_i, y_i) = x_i \cdot x_j$ |
| Polynomial | $k(x_i, y_i) = (\gamma x_i \cdot x_j + r)^d, \gamma > 0$ |
| Radian basis function | $k(x_i, y_i) = \exp(-\gamma \Vert x_i - x_j \Vert^2), \gamma > 0$ |
| Sigmoid | $k(x_i, y_i) = tanh(\gamma x_i \cdot x_j + r)$ |

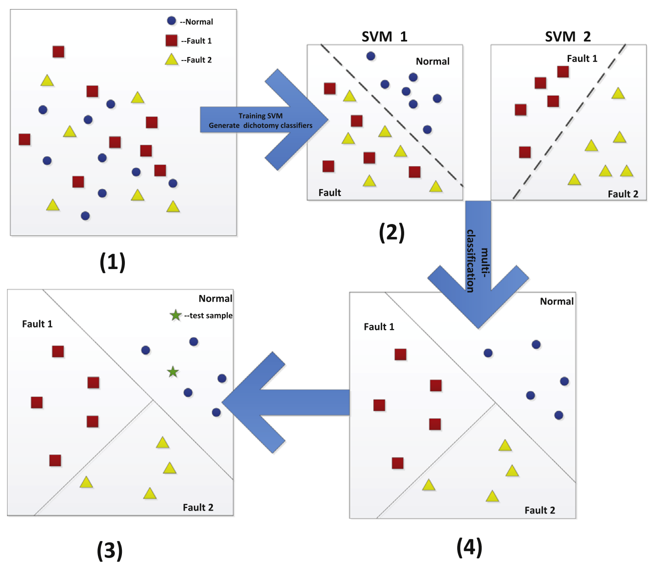


**Fig. 3.** Diagnosis principle using SVM.

complexity is not increased. The objective function is represented below.

$$W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \tag{14}$$

The kernel is always used in input space because it can map the input samples into feature space and apply dot product in that space. There are many different kinds of kernel, such as Polynomial Kernel Function, Linear Kernel Function, Sigmoid Kernel Function and Radian Basis Function (RBF), which are commonly used and showed in Table 1.

**Table 2**
Conclusion of research starus.

| Objects | References | SVM-combined |
|---|---|---|
| Rolling element bearings | Widodo and Kim [8] | RVM, envelope analysis, SMO, PCA |
| | Konar and Chattopadhyay [9] | CWT, ANN |
| | Tyagi [11] | DWT, ANN |
| | Sugumaran et al. [12] | Decision tree, PSVM |
| | Yang et al. [13] | Fractal dimension, 11 time-domain statistical features |
| | Liu et al. [14] | EMD, WSVM, PSO |
| Induction motors | Armaki and Roshanfekr [20] | FFT |
| | Keskes et al. [22] | SWPT, MWSVM |
| | Das et al. [23] | CWT |
| | Widodo et al. [24] | Independent component analysis |
| | Widodo and Yang [25] | WSVM, KPCA |
| | Ghate and Dudul [26] | ANN-SVM, optimal neural network model, PCA |
| Pump | Wang et al. [31,32] | LSVM |
| | Azadeh et al. [33] | ANN, GA, SVC-PSO |
| | Tian et al. [34] | GA |
| Compressors | Verma et al. [35] | DDAG |
| | Qin et al. [36] | Wave matching method |
| | Wang et al. [37] | Indicator diagram |
| | Cui et al. [38] | Information entropy |
| | Moreira et al. [39] | in/a |
| Turbines | Lee et al. [40] | ANN |
| | Salahshoor et al. [41] | ANFIS, OWA |
| | Liu et al. [42] | Diagonal spectrum, cluster binary tree |
| HAVC machines | Guo et al. [44] | On-line incremental SVM, online soft-margin SVM |
| | Dehestani et al. [45,46] | Online incremental SVM |
| | Dehestani et al. [47] | ANN, online SVM |
| | Liang and Du [48] | Residual analysis method |
| | Beghi et al. [49] | PCA |
| | Tanaka et al. [50] | Building management system |
| | Li et al. [51] | Rough set theory |
| | Li et al. [52] | KPCA |
| Gear and gear box | Chen et al. [53] | AR model, IMFs |
| | Saravanan et al. [55] | Morlet wavelet, PSVM |
| | Saravanan et al. [56] | PSVM, ANN |
| Power systems | Bacha et al. [57] | Dissolved gas analysis |
| | Fei and Zhang [59] | SVMG |
| | Ni et al. [60] | KPCA |

### 2.4. Fault monitoring and diagnosis

Fault diagnosis methods based on SVM regard the diagnosis problem as a classification one. They uses the historical data to train the classifiers and divides the data space into various areas, where each area corresponds to a running state. Different areas corresponds to different running states. The test data will be mapped into the data space and the corresponding running status is predicted based on the position of the test data in the data space. Fig. 3 shows an example of the SVM classification process with objects of three classes (Table 2).

In Fig. 3(1), data of different running states are firstly selected to form the training set. Then, the training data are then used by SVM to learn a set of dichotomy classifiers as illustrated in Fig. 3 (2). Since fault diagnosis is usually a multi-class classification problem, these dichotomy classifiers will be combined into a multi-class classifier, as shown in Fig. 3(3). Finally, we can isolate the fault though mapping the test sample and identify the corresponding area in the data space, showed in Fig. 3(4).

In a broad sense, fault diagnosis is composed of three main aspects. Fault detection determines whether a system is malfunctioned, fault isolation position the fault and fault identification estimates the violence level of the fault. Narrowly speaking, fault diagnosis means fault detection and isolation (FDI).

Some reviews of the works on FDI can be found in [1,6,7].

## 3. The research status of fault diagnosis and condition monitoring

Fault diagnosis and condition monitoring technology has been widely applied in engineering fields, especially in large-scale and complex mechanical equipment. Fault diagnosis technology can provide effective fault information in the initial stage, therefore reducing the loss caused by failures.

Fault diagnosis is able to improve the reliability of equipment and avoid the economic loss caused by the damage of machines. It is also helpful to accurately predict the equipment conditions and maximize the equipment's function and production efficiency.

Please cite this article as: Z. Yin, J. Hou, Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes, Neurocomputing (2015), http://dx.doi.org/10.1016/j.neucom.2015.09.081

### 3.1. Rolling element bearings

Widodo and Kim [8] employed relevance vector machine (RVM) and SVM to diagnose the faults in low speed bearing. A low speed test rig was construed to simulate different kinds of bearing defects. Accelerometer sensors and Acoustic emission (AE) were employed to acquire the data. In order to isolate the unwanted signals, envelope analysis was employed to extract features and process signal. From the perspective of the methods used in SVM and RVM, one-against-one approach and sequential minimal optimization (SMO) approach were employed in SVM and multinomial logistic regression and the Newtons method were employed in RVM. The dimensional of original features could be reduced though component analysis, and the acoustic emission signal generates better classification accuracy than the vibration signal. It was found that RVM performs better than SVM in low speed bearing.

The combination of continuous wavelet transform (CWT) and SVM has become a research focus in recent years.

Konar and Chattopadhyay introduced a CWT-SVM approach in [9]. They employed machinery fault simulator to simulate various types of faults and used computers with four-channel data acquisition (DAQ) system with accelerometer probe to collect data. Continuous wavelet transform was the main approach employed in their paper and the authors pointed the limitations of Fourier analysis, i.e., it is not suitable for non-stationary or transitory signals. Besides, the limitation of short time Fourier transform (STFT) lies in the choice of time window. SVM and artificial neural networks (ANN) were employed in post processing and diagnosis of faults process. Experimental results indicate that SVM performs much better than ANN. Finally, the authors regarded CWT-SVM as a promising approach as it is more efficient than Discrete Wavelet Transform (DWT)/ANN based fault classification approach. The approach was quite similar to in Abbasion and Rafsanjani's work in [10].

Tyagi [11] compared the effectiveness of SVM and ANN though diagnosing the rolling element bearings' faults. In experiments 102,400 samples were collected and divided into 40 bins, and 10 features were then extracted from these 40 bins. From the result the author concluded that SVM outperforms ANN. In addition, DWT was beneficial to both approaches as it could reduce the number of iterations performed in training.

Sugumaran et al. [12] employed decision tree to select good features which were important in pattern recognition. Moreover, proximal support vector machines (PSVM) and SVM were utilized to classify the faults and the result showed PSVM was better than SVM in classification efficiency.

Yang et al. [13] found that fractal dimension could be employed to describe non-linear vibration signal. They applied the capacity dimension, information dimension and correlation dimension to classify various types of faults and found that these fractal dimensions describe different kinds of fault information. Therefore, these three fractal dimensions should be combined. Finally, 11 time-domain statistical features and three fractal dimensions were combined to improve the classification performance of the SVM.

Liu et al. [14] proposed an approach which was based on wavelet support vector machine (WSVM) with empirical model decomposition (EMD), particle swarm optimization (PSO) and distance evaluation technique. This approach was employed to detect the bearing fault of electric locomotives and experimental result indicated that WSVM was better than SVM in accuracy.

There are many other approaches proposed to detect faults of rolling element bearings, such as physical model training based method [15], pattern spectrum entropy method [16], neighborhood score multi-class SVM [17], ensemble empirical mode decomposition and optimized support vector machines [18] and features extracted from time-domain and multi-class SVM [19].

### 3.2. Induction motors

Armaki and Roshanfekr [20] proposed a new method to diagnose the broken rotor bar fault in induction motors. They employed Fast Fourier Transform (FFT) to extract new features from power spectral density (PSD) and found harmonic curve area and harmonic crest angle were suitable for fault detection in induction motors. SVM was employed to classify these features and this method had been used in 1500W three-phase induction motor. The experimental result was good. Kurek and Osowski [21] also researched the broken rotor bar fault in squirrel-cage induction motor. The spectral information of shaft field of one phase registered, motor current and voltage were selected as the features.

Wavelet SVM is widely used to diagnose faults in induction motors. Keskes et al. [22] employed the combination of multi-class wavelet support vector machines (MWSVM) and stationary wavelet packet transform (SWPT) to diagnose the broken rotor bar's faults and found wavelet kernel function had a higher accuracy among the various kernel functions. Das et al. [23] detected faults in stator winding of induction motors. The faults can be detected though minor short circuit by Parks transformation and CWT. SVM was the classifier. Widodo et al. [24] proposed the method of the combination of independent component analysis and SVM. One year later, Widodo and Yang [25] not only employed WSVM but also introduced principal component analysis (PCA) and kernel principal component analysis (KPCA) to decrease the features' dimension. It was also a decent approach to extract features.

Ghate and Dudul [26] proposed ANN-SVM model to detect faults in induction motors. Optimal neural network model and PCA were also introduced to decrease the dimension of the features.

Besides what we discussed above, there are still many other methods which is employed to work with SVM, such as the motor current signal analysis (MCSA) [27], artificial immune system [28], advanced Hilbert-Park transform [29] and DQ0 VOLTAGE COMPONENTS [30]. These approaches improve the performance of SVM.

### 3.3. Pumps, compressors, valves and turbines

(1) *Pump*: Wang et al. [31,32] proposed to use Lagrangian support vector machine (LSVM) to detect the ventricular suction faults in rotary blood pumps. This method had been tested though vivo experimental data and the experiments showed that this method performs better than the other three existed methods and original SVM in terms of classification accuracy, learning speed, stability and robustness.

Azadeh et al. [33] firstly combined artificial neural networks (ANNs), support vector classification and particle swarm optimization (SVC-PSO) based on genetic algorithm to detect faults in a centrifugal pump. Their experimental results showed that the robustness and the efficiency were improved compared with existing algorithms.

Tian et al. introduced SVM to detect faults of oil pump. It was usually difficult to detect faults in oil pump because it is a complicated and nonlinear system. In [34] Tian et al. proposed to use SVM to detect faults of oil pump, and they employed the genetic algorithm (GA) to optimize SVM parameters. From the experimental results, they concluded that this algorithm was feasible and effective.

(2) *Compressors*: Verma et al. [35] proposed an optimized fault diagnosis method based on RBF kernel and adaptive SVM parameters and decision directed acyclic graph (DDAG). They applied this method in the fault diagnosis of reciprocating air compressors and found it performs better than the original SVM.

The approach based on the wave matching, basis pursuit (BP) and SVM was presented by Qin et al. [36], and employed to detect faults in reciprocating compressor valves. Wave matching was adopted as it can extract a small number of features with clear physical meanings. This approach was shown to be accurate and reliable in experiments.

Wang et al. [37] proposed to use indicator diagram and SVM to detect faults of reciprocating air compressor. The indicator diagram was based on image processing methods and moment theory.

(3) *Valves*: Cui et al. [38] employed SVM and information entropy to detect faults in reciprocating compressor valve. The reason for this selection is that information entropy is a decent approach to deal with non-linearity and SVM is good at dealing with small sample issues.

Moreira et al. [39] found that SVM was an efficient classifier which can be employed to detect aircraft bleed valves faults. This approach was meaningful because it could reduce the high cost of maintenance in aircraft systems.

(4) *Turbines*: Lee et al. [40] presented an method based on SVM and ANN, where SVM is used to help ANN overcome its disadvantages, such as low classification accuracy and local minima. This method was employed to defect faults in gas turbine engine and was shown to accurate and efficient.

SVM with adaptive neuro-fuzzy inference system (ANFIS) was proposed by Salahshoor et al. [41]. Ordered weighted averaging (OWA) operators was utilized to weighting the multi-attribute data, which can be added into the total value. Industrial steam turbines' faults could be diagnosed and detected by this approach.

Liu et al. [42] detected faults in wind turbine by SVM, diagonal spectrum and cluster binary tree. Diagonal spectrum was employed to extract features in wind turbine and the features were clustered to build the cluster binary tree. This method was shown to be effective by experimental result of gear-box.

Yin et al. [43] overcame the difficulties such as nonlinearity, disturbances and measurement noise in wind turbine diagnosis and proposed a performance index and an optimization criterion to achieve the robustness of the residual signals to disturbances. The simulation on the wind turbine benchmark model valuated the effectiveness of this approach.

### 3.4. HAVC machines

Incremental on-line SVM has been employed in HAVC machines. Guo et al. [44] applied on-line incremental SVM to deal with big datasets and proposed on-line soft-margin SVM to measure the misclassification level. Dehestani et al. [45,46] proposed an on-line fault detection and isolation (FDI) system for HAVC, based on-line incremental SVM. They also presented a novel approach using the previously unknown faults to update the classifier. Dehestani et al. [47] proposed a black box ANN model and an on-line SVM classifier to detect the faults of air fan and dampers. Moreover, the robustness of SVM helped to enhance the probability of detection in their research.

Liang and Du [48] proposed a model-based approach for HAVC. They built a lumped-parameter model for HAVC and employed residual analysis method to detect faults. Finally, they used SVM to classify the faults.

Beghi et al. [49] employed one-class SVM and PCA to deal with the faults of the chiller systems in HAVC system. Unsupervised One-Class SVM was used to detect the unforeseen phenomena and abnormalities and PCA was utilized to reduce feature dimension.

Tanaka et al. [50] proposed an approach based on SVM and built management system (BMS) data. A physical model of air handling unit (AHU) was built and employed to generate normal and fault data. This approach was validated by the data from artificial faults of a real building.

Li et al. [51] presented a method with SVM and rough set theory (RST), based on the observation that SVM has decent generalization performance and RST has the ability to dispose the uncertainty information. This method was shown to reduce the maintain cost and save the energy.

Li et al. [52] proposed an integrated method combining SVM and KPCA. As an improvement of PCA, KPCA adopts a nonlinear kernel function which is employed to extract optimal features. This approach was applied to detect faults in fan machinery system of HAVC and experimental result showed that KPCA-SVM generates higher recognition rate than KPCA.

### 3.5. Other aspects of the application

In the following we mainly introduce the application of SVM based fault diagnosis methods in gear and gear box power systems. These methods are also applied in some other systems, e.g., three tank system [61] and industrial hot strip mill [62].

Chen et al. [53] proposed an approach, where accurate autoregressive (AR) model, SVM and intrinsic mode functions (IMFs) are combined. This approach was employed to detect faults of gears. Wang et al. [54] proposed an LWPR based data-driven fault detection approach and applied it to CSTH (continuous stirred tank heater) process. Saravanan et al. [55] introduced morel wavelet to extract features and employed PSVM and SVM to classify these features in bevel gear box. Experimental results indicate that PSVM performs better than SVM in classification accuracy. Later, they [56] employed PSVM and ANN to detect fault of spur bevel gear box. They found that PSVM and ANN has almost the same classification capability but ANN is computationally more expensive in training and classification. In addition, Yin et al. [63] employed a fuzzy positivistic C-means clustering in vehicle suspension system.

SVM is also widely employed in power systems. Bacha et al. [57] proposed a method, where dissolved gas analysis (DGA) and SVM were combined. The proposed method was used to classify the faults in power transformer. Yin et al. [58] proposed data-based techniques focused on modern industry which can be applied to new power systems. Fei and Zhang [59] employed GA to select suitable SVM parameters and presented the so-called SVMG algorithm. This algorithm was applied to classify data from Chinese electric power companies and it was shown to perform better than SVM, IEC three ratios and artificial neural network. Ni et al. [60] proposed to combine SVM and an adaptive KPCA method. This method could delete the redundant data and reduce the number of samples. It was employed to detect high-voltage circuit breakers (HVCBs). Dash et al. [64] employed three SVMs to identify and classify faults in an advanced series-compensated transmission line.

In addition to what we discussed above, there are many fault diagnosis aspects, which could be disposed by SVM, such as an industrial benchmark of Tennessee Eastman process [65,66], a mine hoist [67], wireless networked control systems [68], a thermal power plant [69] and even passive systems [70]and TE (Tennessee Eastman) process [71].

## 4. Discussion and conclusion

### 4.1. Discussion

Table 2 shows the objects, authors and SVM-combined approaches we reviewed in this paper, where n/a means that the content is not available. From the table, we arrive at the conclusion that SVM and related approaches can be widely employed in many objects. Moreover,the combination of SVM and other approaches usually performs better than SVM alone. Therefore, it is a trend to develop SVM-combined approaches. As there are a large amount of related papers in the literature, this paper only introduces those closely related to the keywords we survey and written in English.

### 4.2. Conclusion

This paper surveys the papers on SVM-based fault diagnosis and process monitoring in complicated industrial processes from 2007 to 2015. Many approaches based on SVM have been proposed, and they promote the advance of fault diagnosis and monitoring techniques significantly. While the algorithms such as PCA, SMO and ANN has been used in combination with SVM, few knowledge based algorithms have been proposed. It is also necessary to develop new methods to deal with very complicated industrial systems.

## References

[1] A. Widodo, B.-S. Yang, Support vector machine in machine condition monitoring and fault diagnosis, Mech. Syst. Signal Process. 21 (6) (2007) 2560–2574.
[2] V. Cherkassky, The nature of statistical learning theory, IEEE Trans. Neural Netw. 8 (November (6)) (1997) 1564.
[3] V.N. Vapnik, V. Vapnik, Statistical Learning Theory, vol. 2, Wiley, New York, 1998.
[4] V. Vapnik, An overview of statistical learning theory, IEEE Trans. Neural Netw. 10 (September (5)) (1999) 988–999.
[5] K.P. Bennett, C. Campbell, Support vector machines: hype or hallelujah?, ACM SIGKDD Explor. Newsl. 2 (2) (2000) 1–13.
[6] S. Yin, S.X. Ding, X. Xie, H. Luo, A review on basic data-driven approaches for industrial process monitoring, IEEE Trans. Ind. Electron. 61 (11) (2014) 6418–6428.
[7] S. Yin, O. Kaynak, Big data for modern industry: challenges and trends, Proceedings of the IEEE, 102(3):143-146, 2015.
[8] A. Widodo, E.Y. Kim, J.-D. Son, B.-S. Yang, A.C. Tan, D.-S. Gu, B.-K. Choi, J. Mathew, Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine, Expert Syst. Appl. 36 (3) (2009) 7252–7261.
[9] P. Konar, P. Chattopadhyay, Bearing fault detection of induction motor using wavelet and support vector machines (svms), Appl. Soft Comput. 11 (6) (2011) 4203–4211.
[10] S. Abbasion, A. Rafsanjani, A. Farshidianfar, N. Irani, Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine, Mech. Syst. Signal Process. 21 (7) (2007) 2933–2945.
[11] C.S. Tyagi, A comparative study of svm classifiers and artificial neural networks application for rolling element bearing fault diagnosis using wavelet transform preprocessing, Neuron 1 (2008) 309–317.
[12] V. Sugumaran, V. Muralidharan, K. Ramachandran, Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing, Mech. Syst. Signal Process. 21 (2) (2007) 930–942.
[13] J. Yang, Y. Zhang, Y. Zhu, Intelligent fault diagnosis of rolling element bearing based on svms and fractal dimension, Mech. Syst. Signal Process. 21 (5) (2007) 2012–2024.
[14] Z. Liu, H. Cao, X. Chen, Z. He, Z. Shen, Multi-fault classification based on wavelet svm with pso algorithm to analyze vibration signals from rolling element bearings, Neurocomputing 99 (2013) 399–410.
[15] K.C. Gryllias, I.A. Antoniadis, A support vector machine approach based on physical model training for rolling element bearing fault detection in industrial environments, Eng. Appl. Artif. Intell. 25 (2) (2012) 326–344.
[16] R. Hao, Z. Peng, Z. Feng, F. Chu, Application of support vector machine based on pattern spectrum entropy in fault diagnostics of rolling element bearings, Meas. Sci. Technol. 22 (4) (2011) 045708.
[17] V. Sugumaran, G. Sabareesh, K. Ramachandran, Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine, Expert Syst. Appl. 34 (4) (2008) 3090–3098.
[18] X. Zhang, J. Zhou, Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines, Mech. Syst. Signal Process. 41 (1) (2013) 127–140.
[19] O. Seryasat, M.A. Shoorehdeli, F. Honarvar, A. Rahmani, Multi-fault diagnosis of ball bearing based on features extracted from time-domain and multi-class support vector machine (msvm), in: 2010 IEEE International Conference on Systems Man and Cybernetics (SMC), IEEE, 2010, pp. 4300–4303.
[20] M.G. Armaki, R. Roshanfekr, A new approach for fault detection of broken rotor bars in induction motor based on support vector machine, in: 2010 18th Iranian Conference on Electrical Engineering (ICEE), IEEE, 2010, pp. 732–738.
[21] J. Kurek, S. Osowski, Support vector machine for fault diagnosis of the broken rotor bars of squirrel-cage induction motor, Neural Comput. Appl. 19 (4) (2010) 557–564.
[22] H. Keskes, A. Braham, Z. Lachiri, Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet svm, Electr. Power Syst. Res. 97 (2013) 151–157.
[23] S. Das, C. Koley, P. Purkait, S. Chakravorti, Wavelet aided svm classifier for stator inter-turn fault monitoring in induction motors in: Power and Energy Society General Meeting, 2010 IEEE, 2010, 1–6.
[24] A. Widodo, B.-S. Yang, T. Han, Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors, Expert Syst. Appl. 32 (2) (2007) 299–312.
[25] A. Widodo, B.-S. Yang, Wavelet support vector machine for induction machine fault diagnosis based on transient current signal, Expert Syst. Appl. 35 (1) (2008) 307–316.
[26] V. Ghate, S. Dudul, Induction machine fault detection using support vector machine based classifier, WSEAS Trans. Syst. 8 (5) (2009) 591–603.
[27] R. Fang, H. Ma, Application of mcsa and svm to induction machine rotor fault diagnosis, in: The Sixth World Congress on Intelligent Control and Automation, 2006, WCICA, 2006, vo. 2, 2006, pp. 5543–5547.
[28] I. Aydin, M. Karakose, E. Akin, Artificial immune based support vector machine algorithm for fault diagnosis of induction motors, in: International Aegean Conference on Electrical Machines and Power Electronics, 2007. ACEMP'07. IEEE, 2007, pp. 217–221.
[29] S.B. Salem, K. Bacha, A. Chaari, Support vector machine based decision for mechanical fault condition monitoring in induction motor using an advanced hilbert-park transform, ISA Trans. 51 (5) (2012) 566–572.
[30] C. Manop, V. Kinnares, S. Chunwiphat, Svm-based stator fault diagnosis for induction motors using dq0 voltage components, Int. J. Innovat. Comput. Inf. Control 6 (10) (2010) 4539–4553.
[31] Y. Wang, M.A. Simaan, A suction detection system for rotary blood pumps based on the Lagrangian support vector machine algorithm, IEEE J. Biomed. Health Inform. 17 (3) (2013) 654–663.
[32] Y. Wang, G. Faragallah, E. Divo, M.A. Simaan, Detection of ventricular suction in an implantable rotary blood pump using support vector machines, in: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, IEEE, 2011, pp. 3318–3321.
[33] A. Azadeh, M. Saberi, A. Kazem, V. Ebrahimipour, A. Nourmohammadzadeh, Z. Saberi, A flexible algorithm for fault diagnosis in a centrifugal pump with corrupted data and noise based on ann and support vector machine with hyper-parameters optimization, Appl. Soft Comput. 13 (3) (2013) 1478–1485.
[34] J. Tian, M. Gao, K. Li, H. Zhou, Fault detection of oil pump based on classify support vector machine, in: IEEE International Conference on Control and Automation, 2007, ICCA 2007, IEEE, 2007, pp. 549–553.
[35] N.K. Verma, A. Roy, A. Salour, An optimized fault diagnosis method for reciprocating air compressors based on svm, in: 2011 IEEE International Conference on System Engineering and Technology (ICSET), IEEE, 2011, pp. 65–69.
[36] Q. Qin, Z.-N. Jiang, K. Feng, W. He, A novel scheme for fault detection of reciprocating compressor valves based on basis pursuit, wave matching and support vector machine, Measurement 45 (5) (2012) 897–908.
[37] F. Wang, L. Song, L. Zhang, H. Li, Fault diagnosis for reciprocating air compressor valve using pv indicator diagram and svm, in: 2010 International Symposium on Information Science and Engineering (ISISE), IEEE, 2010, 2010, 255–258.
[38] H. Cui, L. Zhang, R. Kang, X. Lan, Research on fault diagnosis for reciprocating compressor valve using information entropy and svm method, J. Loss Prev. Process Ind. 22 (6) (2009) 864–867.
[39] R. de Pádua Moreira, C. Nascimento, Prognostics of aircraft bleed valves using a svm classification algorithm, in: Aerospace Conference, 2012 IEEE, IEEE, 2012, 1–8.
[40] I. Nam-Gu, S. KOREA, Defect diagnostics of gas turbine engine using hybrid svm-artificial neural network method, 2007.
[41] K. Salahshoor, M. Kordestani, M.S. Khoshro, Fault detection and diagnosis of an industrial steam turbine using fusion of svm (support vector machine) and anfis (adaptive neuro-fuzzy inference system) classifiers, Energy 35 (12) (2010) 5472–5482.
[42] L. Wenyi, W. Zhenfeng, H. Jiguang, W. Guangfeng, Wind turbine fault diagnosis method based on diagonal spectrum and clustering binary tree svm, Renew. Energy 50 (2013) 1–6.
[43] S. Yin, G. Wang, H.R. Karimi, Data-driven design of robust fault detection system for wind turbines, Mechatronics 24 (4) (2014) 298–306.
[44] Y. Guo, D. Dehestani, J. Li, J. Wall, S. West, S. Su, Intelligent outlier detection for hvac system fault detection, in: 10th International Conference on Healthy Buildings, July, 2012, pp. 8–12.
[45] D. Dehestani, Hvac model based fault detection by incremental online support vector machine, Int. J. Comput. Theory Eng. 5 (2013) 3.
[46] D. Dehestani, F. Eftekhari, Y. Guo, S. Ling, S. Su, H. Nguyen, Online support vector machine application for model based fault detection and isolation of hvac system, Int. J. Mach. Learn. Comput, vol. 1 (2011) 1–7.
[47] D. Dehestani, S. Su, H. Nguyen, Y. Guo, Robust fault tolerant application for hvac system based on combination of online svm and ann black box model, in: 2013 European Control Conference (ECC), IEEE, 2013, pp. 2976–2981.

Please cite this article as: Z. Yin, J. Hou, Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes, Neurocomputing (2015), http://dx.doi.org/10.1016/j.neucom.2015.09.081

[48] J. Liang, R. Du, Model-based fault detection and diagnosis of hvac systems using support vector machine method, Int. J. Refrigeration 30 (6) (2007) 1104–1114.
[49] A. Beghi, L. Cecchinato, C. Corazzol, M. Rampazzo, F. Simmini, G.A. Susto, A one-class svm based tool for machine learning novelty detection in hvac chiller systems, in: World Congress, 19 (1) (2014) 1953–1958.
[50] T. Tanaka, S.-i. Tanabe, E. Togashi, K. Yokoyama, T. Watanabe, Development of fault diagnosis method for hvac equipment using support vector machine.
[51] L. Xuemei, S. Ming, D. Lixing, Hvac fault diagnosis system using rough set theory and support vector machine, in: Second International Workshop on Knowledge Discovery and Data Mining, 2009, WKDD 2009, IEEE, 2009, pp. 895–899.
[52] L. Xuemei, S. Ming, D. Lixing, X. Gang, L. Jibin, Novel hvac fan machinery fault diagnosis method based on kpca and svm, in: International Conference on Industrial Mechatronics and Automation, 2009, ICIMA 2009, IEEE, 2009, pp. 492–496.
[53] J. Cheng, D. Yu, Y. Yang, A fault diagnosis approach for gears based on imf ar model and svm, EURASIP J. Adv. Signal Process. 1 (2008) 647135.
[54] G. Wang, S. Yin, O. Kaynak, A LWPR based data-driven fault detection approach for nonlinear process monitoring, IEEE Transactions on Industrial Electronics 10 (4) (2014) 2016–2023.
[55] N. Saravanan, V.K. Siddabattuni, K. Ramachandran, A comparative study on classification of features by svm and psvm extracted using morlet wavelet for fault diagnosis of spur bevel gear box, Expert Syst. Appl. 35 (3) (2008) 1351–1366.
[56] N. Saravanan, V.K. Siddabattuni, K. Ramachandran, Fault diagnosis of spur bevel gear box using artificial neural network (ann), and proximal support vector machine (psvm), Appl. Soft Comput., 10 (1) (2010) 344–360.
[57] K. Bacha, S. Souahlia, M. Gossa, Power transformer fault diagnosis based on dissolved gas analysis by support vector machine, Electr. Power Syst. Res. 83 (1) (2012) 73–79.
[58] S. Yin, X. Li, H. Gao, O. Kaynak, Data-based techniques focused on modern industry, An overview, IEEE Transactions on Industrial Electronics 62 (1) (2015) 657–667.
[59] S.-w. Fei, X.-b. Zhang, Fault diagnosis of power transformer based on support vector machine with genetic algorithm, Expert Syst. Appl. 36 (8) (2009) 11352–11357.
[60] J. Ni, C. Zhang, S.X. Yang, An adaptive approach based on kpca and svm for real-time fault diagnosis of hvcbs, IEEE Trans. Power Deliv. 26 (3) (2011) 1960–1971.
[61] S. Yin, X. Zhu, Intelligent particle filter and its application on fault detection of nonlinear system, IEEE Transactions on Industrial Electronics 62 (6) (2015) 3852–3861.
[62] S.X. Ding, S. Yin, K. Peng, H. Hao, B. Shen, A novel scheme for key performance indicator prediction and diagnosis with application to an industrial hot strip mill, IEEE Trans. Ind. Inform. 9 (4) (2013) 2239–2247.
[63] S. Yin, Z. Huang, Performance monitoring for vehicle suspension system via fuzzy positivistic C-means clustering based on accelerometer measurements, IEEE/ASME Transactions on Mechatronics, 20(5):2613-2620, 2015.
[64] P. Dash, S. Samantaray, G. Panda, Fault classification and section identification of an advanced series-compensated transmission line using support vector machine, IEEE Trans. Power Deliv. 22 (1) (2007) 67–73.
[65] S. Yin, G. Wang, X. Yang, Robust pls approach for kpi-related prediction and diagnosis against outliers and missing data, Int. J. Syst. Sci. 45 (7) (2014) 1375–1382.
[66] S. Yin, X. Zhu, O. Kaynak, Improved pls focused on key-performance-indicator-related fault diagnosis, IEEE Trans. Ind. Electron. 62 (3) (2015) 1651–1658.
[67] Y.-w. Chang, Y.-c. Wang, L. Tao, Z.-J. Wang, Fault diagnosis of a mine hoist using pca and svm techniques, J. China Univ. Min. Technol. 18 (3) (2008) 327–331.
[68] S.X. Ding, P. Zhang, S. Yin, E.L. Ding, An integrated design framework of fault-tolerant wireless networked control systems for industrial automatic control applications, IEEE Trans. Ind. Inform. 1 (9) (2013) 462–471.
[69] K.-Y. Chen, L.-S. Chen, M.-C. Chen, C.-L. Lee, Using svm based method for equipment fault detection in a thermal power plant, Comput. Ind. 62 (1) (2011) 42–50.
[70] X. Li, S. Yin, H. Gao, Passivity-preserving model reduction with finite frequency $h$ approximation performance, Automatica 50 (9) (2014) 2294–2303.
[71] G. Wang, S. Yin, Quality-related fault detection approach based on orthogonalsignal correction and modified PLS, IEEE Transactions on Industrial Informatics, 11(2):398-405.

**Zuyu Yin** received his B.E. degree in automation from Harbin Institute of Technology, Harbin, China. He is currently working toward the master degree in control science and engineering with the Key Laboratory of Robotics and Systems in Harbin Institute of Technology. His research interests include fault detection and diagnosis, and machine learning.

**Jian Hou** received the Ph.D. degree in 2007 from Harbin Institute of Technology, Harbin, China. He is currently an associate professor with the Research Institute of Mechatronics and Control, Bohai University, Jinzhou, China. His research interests include pattern recognition, machine learning, computer vision and image processing.