Journal of Computers Vol. 28, No. 5, 2017, pp. 39-50
doi:10.3966/199115992017102805004

Study of the Fault Diagnosis Method of Control Systems Based on MCCSAPSO-SVM

Ruicheng Guo[1], Pu Yang[1*], Xu Pan[1] and Jianwei Liu[1]

[1] College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, Jiangsu, China
{ruichengguo, ppyang, phxyy, ljw301}@nuaa.edu.cn

Received 18 June 2016; Revised 10 January 2017; Accepted 7 February 2017

Abstract. In order to improve the accuracy of actuator fault diagnosis of control system, a new method based on Multi-swarm cooperative chaos simulated annealing particle swarm optimization－support vector machine (MCCSAPSO-SVM) is proposed in this paper. Firstly, the noise reduction and feature extraction for the output signal are taken by the joint noise reduction and improved empirical mode decomposition (EMD) method. Secondly, structure parameters of SVM are optimized by MCCSAPSO, which not only can effectively avoid the premature convergence of particle swarm, but also can overcome the misjudgment problem caused by single particle information exchange. This algorithm can accelerate the convergence velocity and improve the accuracy of traditional PSO. Thirdly, the use of the mixture kernel function (MKF) can guarantee the good generalization and learning ability of SVM. Finally, a partial binary tree SVM is constructed by using the training data. This structure transforms a complex multi-classification problem into a number of two classification problems, which reduces the computational complexity and improves the real-time performance of the diagnosis. The experimental results of quad-rotor semi-physical simulation platform verify the feasibility and effectiveness of the proposed method.

Keywords: fault diagnosis, MCCSAPSO, MKF, partial binary tree SVM

1 Introduction

1.1 Research Background

Since the composition of modern control system is more and more complex and the system often works under different environmental conditions in high load for a long time, it will cause kinds of faults inevitably. Especially in the aerospace, medical, large-scale machinery production and other fields, subtle faults may cause serious economic loss and personnel injury. Faults can be removed quickly via a fast real-time diagnosis technology. Therefore, a good fault diagnosis method can avoid economic losses and accident casualties, which is more and more significant for today's industrial practical systems. Fault diagnosis technology has been developed rapidly since it was put forward [1]. The goal of fault diagnosis is mainly to determine whether systems remain a healthy state and to identify incipient problems before they lead to catastrophic failure [2]. Effective fault diagnosis approaches are warranted to detect and analyze faults of modern control systems and those approaches could eliminate the negative fault impacts to the lowest possible level. Hence, the studies of fault diagnosis methods have aroused the interest of scholars in related fields.

Fault diagnosis methods mainly include that methods based on control model, methods based on pattern recognition and methods based on artificial intelligence. The establishment of the corresponding control model is the foundation of the fault diagnosis based on control model method [3]. This method is mainly related to modeling, parameter- state estimations, and observer applications. However, this

* Corresponding Author

method also has many shortcomings and disadvantages. System faults may cause uncertainties to the form of system model structure and parameter variation. In essence, the fault diagnosis based on pattern recognition is to make use of the state information and the prior knowledge of systems to deal with the fault detection. And then the comprehensive evaluation process about the real time operation state of system can be obtained [4]. When system models are unknown or very complex, this method provides a simple and effective way to finish fault diagnosis. However, it is very complicatedly to train models. And other disadvantages includes that a bulk of computation, sensitive parameter selection and so on. Fault diagnosis based on artificial intelligence includes that based on expert system [5], neural network [6] and support vector machine [7]. Fault diagnosis based on expert system can use domain experts' experience and knowledge. But this method has great limitations, which include that knowledge set is not complete because this approach is overwhelmingly dependent on domain experts. Neural network, as a new pattern recognition technology, has shown great potential in the field of fault diagnosis. Fault diagnosis based on neural network has new information expression, highly parallel distributed processing, association, self-learning and self-organizing ability. However, there are a lot of deficiencies such that it is difficult to obtain a large number of training samples and can only get the local optimal solution but not the global optimal solution. Fault diagnosis based on support vector machine is becoming a new research focus in the field of machine learning and fault diagnosis, and will promote the further development of the fault diagnosis technology.

1.2 Related Works

The appearance and improvement of fault diagnosis technology are dependent on the development of signal acquisition technology, data processing technology and digital computer technology. Fault diagnosis can be divided into fault detection and fault isolation. These two technologies have different purposes and different processing methods. The purpose of fault detection is to determine whether the current running state is normal. In addition, the purpose of fault isolation is to determine the type and location of a specific fault. The keys to build a fast, effective and reliable fault diagnosis system include next two points. The first one is to accurate access to the characteristics of the current operating state of control systems. The second one is to establish a good generalization ability fault diagnosis algorithm to deal with the small sample problem.

In actual control system, fault signals are often affected by noise interference, and this would usually cause relatively low signal-to-noise ratio. Especially in the early stage of faults, the weak fault information is almost completely submerged in noise signals. Therefore, it is necessary to reduce noises and to take the feature extraction in the process of the fault signal acquisition. Fault feature extraction has been widely used in practical applications [8-9]. Empirical mode decomposition (EMD) is a kind of very effective feature extraction method, which is suitable for the nonlinear and non-stationary signal processing [10-12]. A signal is decomposed into a series of different frequency components of Intrinsic Mode Function (IMF) through the empirical mode decomposition. Compared with traditional Fourier transform, EMD makes the instantaneous frequency of decomposed signals have more practical significance instead of simple sine/cosine functions. Compared with other signal processing methods, EMD is more intuitive, direct and adaptive. However, there are lots of strong random noises mixing in acquired signals. Therefore, appropriate and necessary noise reduction-processing methods must be taken before the application of EMD. Zhang, Shu, Qin and Zhang propose dimensionless index processing algorithms based on EMD, which is used to decompose the collected vibration signals [13]. And then fault diagnosis can be achieved by calculating dimensionless parameter values to the IMF components. However, the effects of other noise signal are not removed at the same time very well. Although EMD could deal with some noises to a certain extent, this method has obvious shortcomings for the pulse noise and Gauss noise. The noise reduction method based on median filter and wavelet transform can indicate local information of signals in the time and frequency domain effectively and has the multi-resolution characteristic.

An efficient method for classification of fault of a transmission line is proposed [14], which deals with fault diagnosis by using EMD and probabilistic neural network. This way could have high classifying accuracy but neural network need a large number of fault samples and need a long time to train classifying network. Fault diagnosis is a typical small sample problem [15-16], because actuator faults of control systems are sudden and difficult to reproduce. However, the theory basis of traditional classifiers such as neural network is classical statistics. Classical statistics research is under the condition of the large sample capacity of statistical properties. Therefore, the performance of traditional classifiers can be guaranteed only when the sample size tends to infinity. Statistical learning theory is devoted to provide a unified framework for the learning problem of small sample size. This makes classifiers have a theoretical basis in the case of small sample size. Support vector machine (SVM) is a representative classifier based on statistical learning theory, which has a very strong generalization ability to deal with pattern recognition problem of small sample, nonlinear and high dimensional systems [17]. Thereby, SVM is more and more common to construct classifiers.

Jain, Bajaj and Kumar propose an automatic ECG beat least squares support vector machine classifier based on artificial bee colony and radial basis kernel function [17]. Although this simulation results illustrate that proposed classifier gives efficient classification, single kernel function still causes some mistaken classification results inevitably. In the actual application, to select effective kernel functions suitable for the sample data can enhance the efficiency and robustness of the decision function of SVM [18]. The mixture kernel function (MKF) is made up by a linear combination of multiple single kernel functions, which can compensate for the shortcomings of the single kernel function for the multi source complex sample data. The advantage of the MKF is that the new kernel functions can combine with the learning and generalization performance of SVM. Zhang and Cao propose a kind of fault diagnosis method for multivariable dynamic systems based on nonlinear spectrum and MKF-SVM [19]. The polynomial function and radial basis function are chosen as the mixture kernel function in paper [19], but there are still disadvantages in that paper. The main problem is that the choice of the weight of each single kernel function is not very suitable in many practical control systems, and the performance of the MKF is greatly affected by this weight. The parameters of the kernel function and the penalty factor are collectively referred to as the structural parameters of SVM. At present, optimization methods of structural parameters of SVM mainly include cut-and-trial, cross-validation, grid search method, genetic algorithm, particle swarm optimization (PSO), artificial immune algorithm, and ant colony algorithm. In the above methods, PSO only needs to adjust less parameter and has the characteristic of simple structure. However, traditional PSO is very easy to fall into local extrema, so it cannot find the global optimal solution in optimization process [20]. The idea of chaos is introduced into PSO [21], which not only improves the ability of particle swarm optimization algorithm to jump out of the local extrema, but also improves the convergence velocity and precision to a certain extent. However, due to the chaotic PSO cannot overcome the misjudgment defects caused by individual particle, the information exchange of particles can easily lead to false. Simulated Annealing (SA) can avoid the local optimal solution in the search process, and has a good global optimization performance [22]. But this algorithm has some limitations because the initial value determines the performance to find the global optimal solution of the problem.

With the rapid development of computer technology and the actual needs of industrial automation and other fields, the analysis and design of discrete control system have become an important part of control theory. In engineering practice, there are some errors in the process of modeling, and the physical structure of the system will be influenced by the working conditions. Therefore, it has become an urgent problem for engineering application to investigate and analyze whether there are actuator faults in control system.

Based on the above analysis, a fault diagnosis method for actuator based on MCCSAPSO-SVM is proposed in this paper. This fault diagnosis method is divided into fault detection and fault isolation. In fault detection part, the noise the output signal of the actuator is reduced and then the energy vector and the energy entropy are acquired by the improved EMD method firstly. And then the current operating state of the actuator is whether in normal state can be directly judged by the energy entropy. In fault isolation part, the partial binary tree multi kernel SVM based on MCCSAPSO is designed. The optimization of structure parameter of the SVM can effectively overcome the problem of false error caused by single particle information exchange. The idea of simulated annealing is introduced in the renewal process of velocity and position of each particle, which can make the PSO jump out of the local extrema and adjust the temperature of SA. At the same time, the inertia weight of the standard PSO algorithm is adjusted, and the convergence factor is introduced, which can accelerate the convergence rate of the particle swarm and improve the accuracy of PSO. This SVM adopts the linear combination of the Gauss kernel function and the polynomial kernel function, which has a good generalization ability of the global kernel function, and also has a good learning ability of local kernel function. The experimental test results of quad-rotor helicopter semi physical simulation platform show that the fault diagnosis method proposed in this paper not only can effectively determine if the fault happens, but also can accurately determine the fault type.

2 Real Time Fault Detection

The state signal of control systems, especially the fault state signal is a typical non-stationary signal. The anisotropic and multi-source noise pollution of the media of control systems increases the difficulty of state signal feature extraction. Yang, Guo, Pan and Liu propose a method of fault feature extraction based on joint noise reduction and empirical mode decomposition [23]. In this method, the large disturbances caused by impulse noise are removed via the median filter by the following formula:

$$\tilde{f}(k)=med\left(f(k)\right)=\begin{cases} f\left(\dfrac{(n+1)}{2}\right), & \text{n is the odd number} \\ \dfrac{1}{2}\left(f\left(\dfrac{n}{2}\right)+f\left(\dfrac{(n+2)}{2}\right)\right), & \text{n is the even number} \end{cases} \quad (1)$$

Where, $f(k)(k=1,2,...,N)$ is the signal with noises, $\tilde{f}(k)$ is the signal filtered out impulse noises by extracting medians in accordance with the center k and length n.

And then the effect of Gaussian white noise on active signal is inhibited by the improved wavelet threshold denoising method, where the improved threshold is determined by the following formula:

$$W_{\delta j}=\begin{cases} \operatorname{sgn}(W_j)\left(|W_j|-\dfrac{2p}{1+e^{q(|W_j|-\delta)}}\delta\right) & ,|W_j|\geq\delta \\ W_j\cdot\varphi^{-(\mu|W_j|-\delta)^2} & ,|W_j|<\delta \end{cases} \quad (2)$$

Among them, W_j is the wavelet coefficient of $\tilde{f}(k)$, δ is the threshold, $W_{\delta j}$ is the wavelet coefficient after noise reduction. $\mu(\mu>0), p(0\leq p\leq 1), q(q\geq 0)$ are all adjustable parameters and φ is a constant more than 1.

Each intrinsic mode function (IMF) component can be obtained by the improved EMD method that can effectively eliminate the end effect and modal aliasing phenomenon and then pseudo components are eliminated. The energy entropy is calculated by the following formula:

$$\begin{cases} E_i=\sum_{k=1}^{N}c_i^2(k) \quad (i=1,2,...,\eta) \\ H=-\sum_{i=1}^{\eta}\dfrac{E_i}{\sum_{i=1}^{\eta}E_i}\cdot\ln\left(\dfrac{E_i}{\sum_{i=1}^{\eta}E_i}\right) \end{cases} \quad (3)$$

Where, E_i is the energy of each IMF $c_i(k)$, and the feature vector is $\left[E_1\Big/\sum_{i=1}^{\eta}E_i, E_2\Big/\sum_{i=1}^{\eta}E_i, \cdots, E_\eta\Big/\sum_{i=1}^{\eta}E_i\right]$, N is the length of sampling signal, η is the number of IMF energy vectors after eliminating pseudo components, H is the energy entropy.

When the system is in normal state, the energy distribution is relatively uniform, so the entropy is larger. When the system is in fault state, the energy distribution is in the fault frequency range, so the entropy is smaller. Therefore, the current operating state of the actuator is whether in normal state can be directly judged by the energy entropy in real time.

3 Design of SVM Based on MKF

3.1 Structure of SVM Based on MKF

Mixture kernel function. The classification performance and generalization ability of SVM are affected directly by the selection of kernel functions, so different kernel functions are needed for different classification problems. Kernel functions can be divided into global kernel functions and local kernel functions. Local kernel functions have strong learning ability and global kernel functions have strong generalization ability [24]. Among them, the Gauss kernel function is a typical local kernel function, and the polynomial kernel function is a typical global kernel function. Combined with the two advantages, the construction of MKF is as follows:

$$K = \rho K_{poly} + (1-\rho) K_{RBF} = \rho (x x_i + 1)^d + (1-\rho) \exp\left(-\|x - x_i\|^2 / \sigma^2\right). \quad \textbf{(4)}$$

Where, d is the order of polynomial kernel function, σ is the Gauss kernel radius, ρ is the weight. According to Mercer theorem, the linear combination of kernel functions is still a kernel function [24]. Therefore, the construction of the MKF satisfies the Mercer condition.

Theory of SVM based on MKF. Decision function of single kernel function SVM is as follows:

$$f(x) = \operatorname{sgn}\left(\sum_{i=0}^{l} \alpha_i y_i K(x_i, x_j) + b \right). \quad \textbf{(5)}$$

Where, $\{x_i, y_i\}$ is the training sample, l is the number of training sample, $K(x_i, x_j)$ is the given kernel function and α_i, b are parameters derived from the training sample. According to the formula of the MKF, the decision function of the SVM can be expressed as follows:

$$f(x) = \operatorname{sgn}\left(\sum_{i=0}^{l} \alpha_i y_i \left(\rho K_{poly}(x_i, x_j) + (1-\rho) K_{RBF}(x_i, x_j) \right) + b \right). \quad \textbf{(6)}$$

The optimization model based on the MKF-SVM can be expressed as follows:

$$\begin{aligned} &\min_{\rho,(1-\rho),f_{poly},f_{RBF},b,\xi_i} \left(\frac{1}{\rho} \|f_{poly}\|^2 \right) + \left(\frac{1}{(1-\rho)} \|f_{RBF}\|^2 \right) + C \sum_{i}^{l} \xi_i \\ &s.t. \quad y_i \left(f_{poly} + f_{RBF} + b \right) \ge 1 - \xi_i \\ &\qquad\ \ \xi_i \ge 0, 0 \le \rho \le 1 \end{aligned} . \quad \textbf{(7)}$$

Where, C is the penalty factor, ξ_i is the slack variable. The above optimization problem can be transformed into the following equation:

$$\begin{aligned} &\min_{\rho,(1-\rho),f_{poly},f_{RBF},b,\xi_i} \frac{1}{2} \left[\left(\frac{1}{\rho} \|f_{poly}\|^2 \right) + \left(\frac{1}{(1-\rho)} \|f_{RBF}\|^2 \right) \right] + C \sum_{i}^{l} \xi_i \\ &s.t. \quad y_i \left(f_{poly} + f_{RBF} + b \right) \ge 1 - \xi_i \\ &\qquad\ \ \xi_i \ge 0, 0 \le \rho \le 1 \end{aligned} . \quad \textbf{(8)}$$

The following equation:

$$\begin{aligned} &\max_{\alpha} \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j \left(\rho K_{poly}(x_i, x_j) + (1-\rho) K_{RBF}(x_i, x_j) \right) \\ &s.t. \quad \sum_{i=1}^{l} \alpha_i y_i = 0, 0 \le \alpha_i \le C, i = 1, \cdots, l \\ &\qquad\ 0 \le \rho \le 1 \end{aligned} , \quad \textbf{(9)}$$

which is a dual problem by using Lagrange multiplier method and KKT condition.

The optimal solution $\alpha_i^*, b^*, \xi_i^*, \rho^*, C^*$ is obtained, and the optimal classification function is as following:

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{l} \alpha_i^* y_i \left(\rho^* K_{poly}\left(x_i, x_j\right) + \left(1-\rho^*\right) K_{RBF}\left(x_i, x_j\right) \right) + b^* \right). \tag{10}$$

3.2 Structure Parameter Optimization Method for SVM Based on MKF

The structural parameters have important influence on the classification accuracy and generalization ability for support vector machine. Although the advantages of PSO in the application of parameter optimization are obvious, it is easy to fall into local extreme value. MCCSAPSO that the chaos theory and SA are introduced into PSO not only makes PSO can jump out of the local extreme points, but also improves the convergence velocity and the precision of the algorithm to some extent. A multi swarm cooperative method with master-slave structure is adopted to avoid the false of the single particle information exchange effectively, which divides each generation group into a main group and many sub-groups. The main group focuses on local development and sub-groups focus on global exploration. The iterative formula of the standard PSO is adjusted in this paper, so that it can converge to the optimal search interval more quickly.

Adjustment of standard PSO. The iterative formula of standard PSO is as follows:

$$\begin{cases} v_i^{t+1} = w v_i^t + c_1 r_1 \left(p_{ibest} - x_i^t\right) + c_2 r_2 \left(g_{best} - x_i^t\right) \\ x_i^{t+1} = x_i^t + v_i^{t+1} \end{cases}. \tag{11}$$

Where, c_1, c_2 are learning factors, $r_1, r_2 \in [0,1]$ are random numbers, $p_{ibest} = \left(p_{i1}, p_{i2}, \cdots, p_{iD}\right)$ is the particle history optimal fitness value, $g_{best} = \left(g_1, g_2, \cdots, g_D\right)$ is the global optimal fitness value and w is the inertia weight. The inertia weight w is adjusted and the convergence factor δ is introduced in order to improve the optimization ability of PSO as follows:

$$\begin{cases} v_i^{t+1} = \delta\left(w v_i^t + c_1 r_1 \left(p_{ibest} - x_i^t\right) + c_2 r_2 \left(g_{best} - x_i^t\right)\right) \\ x_i^{t+1} = x_i^t + v_i^{t+1} \\ w = w_{\min} + \left[\left(t_{\max} - t\right)\left(w_{\max} - w_{\min}\right)\right] / t_{\max} \\ \delta = 2 \Big/ \left| 2 - \left(c_1 + c_2\right) - \sqrt{\left(c_1 + c_2\right)^2 - 4\left(c_1 + c_2\right)} \right| \end{cases}. \tag{12}$$

Where, $t_{\max}$ is the Maximum iteration number, $w_{\min}, w_{\max}$ are the minimum and maximum. With the iteration, the linear reduction of inertia weight w make the particle swarm have a strong search ability to search to the larger solution space and continue to search new areas. In the late iteration, the algorithm converges to the global optimum. By adjusting the convergence factor δ, the convergence rate of the particle swarm can be accelerated.

Steps of MCCSAPSO.

(Step 1) Initialization. Initialize the particle swarm size L, particle swarm dimension D, particle swarm position $x_i = \left(x_{i1}, x_{i2}, \cdots, x_{iD}\right)$, particle swarm velocity $v_i = \left(v_{i1}, v_{i2}, \cdots, v_{iD}\right)$, inertia weight w, the Maximum iteration number $t_{\max}$, learning factors c_1, c_2 of PSO. Initialize the current temperature T, end temperature T_0, annealing speed K_T.

(Step 2) Update sub-group. Update p_{ibest} and g_{best} by calculating the current fitness value Ψ_i of each particle in the sub-group and comparing with self-optimal fitness value p_{ibest}. The velocity and position of each particle is updated according to the equation (12).

(Step 3) Chaos optimization. Let g_{best} map to the definition domain of Logistic equation belonging to $[0,1]$, as follows:

$$R = \left(g_{best} - R_{\min}\right) / \left(R_{\max} - R_{\min}\right). \tag{13}$$

Where, $R_{\min}, R_{\max}$ are the minimum and maximum of g_{best}. The chaos sequence $R = \left(R_1, R_2, \cdots, R_P\right)$ is generated by the optimal particle position and velocity iterations. By equation (14), the optimal particle sequence shown in the formula (15) can be obtained.

$$g_{best} = R_{\min} + R_n\left(R_{\max} - R_{\min}\right), n = 1, 2, \cdots, P \tag{14}$$

$$g_{best}^* = \left(g_{best,1}^*, g_{best,2}^*, \cdots, g_{best,P}^*\right) \tag{15}$$

(Step 4) Update main group. Update the state and choose the best particle from sub-groups according to experiences of sub-groups. The velocity and position update equation of main group are as follows:

$$\begin{cases} v_{Mi}^{t+1} = \delta\left(wv_{Mi}^t + c_1 r_1\left(p_{Mibest} - x_{Mi}^t\right) + \phi c_2 r_2\left(g_{Mbest} - x_{Mi}^t\right) + \left(1 - \phi\right) c_3 r_3\left(g_{Sbest} - x_{Si}^t\right)\right) \\ x_{Mi}^{t+1} = x_{Mi}^t + v_{Mi}^{t+1} \end{cases}. \tag{16}$$

Where, M is the man group, S is sub-group, c_3 is the learning factor, r_3 is the random number, ϕ is the migration factor satisfied formula (17), g_{Mbest}, g_{Sbest} are the global optimal fitness values of main and sub group.

$$\phi = \begin{cases} 0, g_{Sbest} < g_{Mbest} \\ 0.5, g_{Sbest} = g_{Mbest} \\ 1, g_{Sbest} > g_{Mbest} \end{cases} \tag{17}$$

(Step 5) Annealing optimization. Calculate the fitness value Ψ_i' and its change rate $\Delta\Psi_i = \Psi_i' - \Psi_i$ of each particle after update. If $\Delta\Psi_i < 0$ or $\Delta\Psi_i > 0$, $\exp\left(-\Delta\Psi/T\right)$ is in $\left[0,1\right]$, and then take cooling operation $T \leftarrow K_T T$.

(Step 6) End condition. When the temperature drops to T_0 or the maximum number of iterations $t_{\max}$ is met, stopping iteration. Otherwise, return Step2.

4 The Design of Fault Diagnosis System

4.1 The Partial Binary Tree Support Vector Classifier

Control system actuator faults mainly include lock in place, hard over fault, constant gain, constant deviation, loss of effectiveness and so on. Obviously, if the fault type of the actuator is to be determined, a multi classification problem is there. The problem of multi classification solved by SVM mainly has two ways. One is to establish the objective function of multi class support vector machine. This method has a high computational complexity and bad real time performance. The other is the partial binary tree support vector classifier. This method transforms a multi-classification problem into a number of two classification problems.

$\left(Q-1\right)$ support vector classifiers should be trained when there are Q fault types. The first classifier SVM_1 classifies Fault 1 with other fault samples. The output of SVM_1 is +1 when system is in normal state and the output is -1 when there are the rest of the fault samples. The second classifier SVM_2 classifies Fault 2 with the rest fault samples. The output of SVM_2 is +1 when system has Fault 2 and the output is -1 when there are the rest of the fault samples and so on. The $\left(Q-1\right)^{\text{th}}$ classifier $SVM_{(Q-1)}$ classifies Fault$\left(Q-1\right)$ and Fault Q samples. The output of SVM_2 is +1 when system has Fault 2 and the output is -1 when there are the rest of the fault samples. The test data are used to train well $\left(Q-1\right)$ support vector machine classifier. According to the output value of each SVM, it can determine the specific actuator fault type.

4.2 Steps of Fault Diagnosis Mentioned in this Paper

(Step 1) Data preprocessing. By using the improved EMD method combined with the joint noise reduction, the collected actuator signals are preprocessed to eliminate the influence of noise and extract the energy feature vectors and energy entropy.
(Step 2) System real time fault detection. The running state of control system is judged whether it is in normal condition by the energy entropy.
(Step 3) Parameter optimization of SVM. Structural parameters of SVM are optimized by the MCCSAPSO algorithm after the adjustment of the iterative formula. Among them, the fitness function Ψ is selected as follows:

$$\Psi = 100r \qquad 0.1 \le r \le 1 \ . \tag{18}$$

When the MCCSAPSO algorithm is used to optimize the structure parameters of SVM. The fitness function is a multiple of the classifier accuracy to optimize the classifier in the maximum form as follows:

$$\max \Psi\left(C,\sigma,d,\rho\right) = 100r = 100 \times \sum_{j=1}^{Q} \frac{N_j^c}{N_j^c + N_j^w} . \tag{19}$$

Where, N_j^c is the number of SVM_j classifies correctly, N_j^w is the number of SVM_j classifies wrong.
(Step 4) Training and testing. The data sample is divided into training data and testing data. The training data is used to train the partial binary tree support vector classifiers and through the testing data to test them.

5 The Simulation of Semi-physical Platform

In order to verify the effectiveness of the designed actuator fault diagnosis method, the quad-rotor helicopter flight control system based on Qball-X4 provided by Quanser Company in Canada is used as the research object. As shown in Fig. 1, there are six dimension variables in Qball-X4. As shown in Fig. 1, there are six dimension variables $\left(X,Y,Z,\psi,\theta,\phi\right)$ in Qball-X4. Among them, X,Y,Z are the location variables, ψ is the yaw angle, θ is the pitch angle, and ϕ is the roll angle. In this paper, the speed of the motor is used as the test object in the helicopter to the X-axis under the condition of normal state, lock in place, hard over fault, constant gain, and constant deviation. Sampling frequency is 1 kHz and the number of sample points is 1000. Normal state and fault conditions sample data are shown in Fig. 2.

Firstly, the energy feature vectors and energy entropy of the collected actuator signals are extracted by using the improved EMD method combined with the joint noise reduction, as shown in Table 1. Set $\pm 10\%$ range of the normal state energy entropy as the actuator normal operating interval, which can quickly determine whether the actuator is in a fault state.

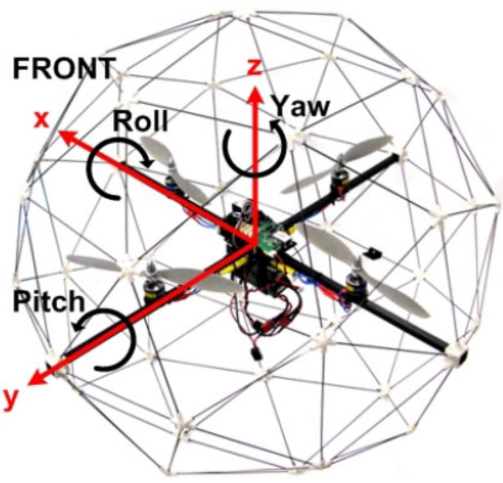

Fig. 1. The main body of Qball-X4

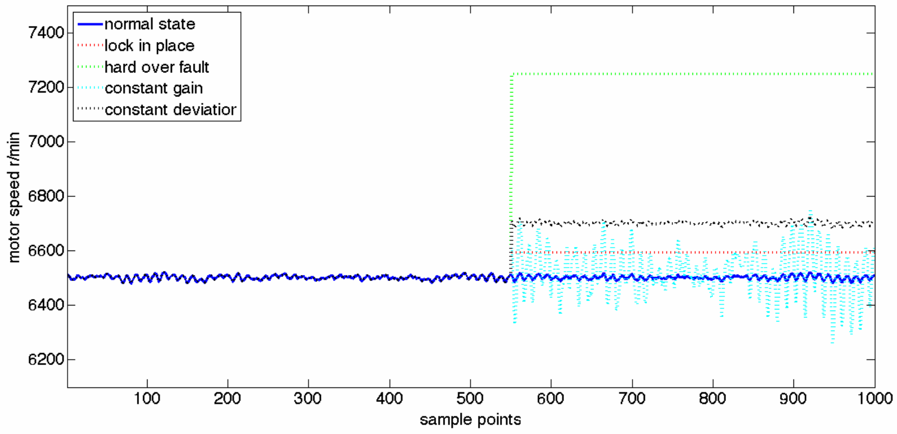

Fig. 2. Sample data of normal state and fault states

Table 1. The energy feature vectors and energy entropy under each state

Actuator state	Energy feature vectors	Energy entropy
Normal state	[0.1776,0.1722,0.1469,0.1102,0.1017, 0.0783,0.0877,0.1252]	1.7668
Lock in place	[0.1914,0.1802,0.1519,0.1195,0.1019, 0.0822,0.0677,0.1052]	1.3036
Hard over fault	[0.0847,0.0738,0.0667,0.0701,0.0821, 0.1179,0.2349,0.2707]	1.2256
Constant gain	[0.0827,0.0752,0.0699,0.0689,0.0817, 0.1122,0.2225,0.2898]	1.2161
Constant deviation	[0.0854,0.0777,0.0663,0.0679,0.0882, 0.1255,0.2149,0.2741]	1.2458

Secondly, the fault identification is carried out by using 3 SVM classifiers. The first classifier SVM_1 classifies lock in place state with other fault samples. The output of SVM_1 is +1 when system is in lock in place state and the output is -1 when there are the rest of the fault samples. The second classifier SVM_2 classifies hard over fault with the rest fault samples. The output of SVM_2 is +1 when system has fault hard over fault and the output is -1 when there are the rest of the fault samples. The third classifier SVM_3 classifies constant deviation with the constant gain. The output of SVM_3 is +1 when system has constant deviation fault and the output is -1 when there are the constant gain fault samples. Thirdly, the structure parameters of SVM are optimized by MCCSAPSO. Where, the number of main group and sub-group is respectively 1 and 3. The size of particle swarm is $L = 30$, maximum iteration number is $t_{\max} = 500$, $c_1 = 2.05$, $c_2 = 2.05$, $c_3 = 0.8$, $w_{\min} = 0.2$, $w_{\max} = 0.9$, $T = 100$, $T_0 = 0.01$, $K_T = 0.9$. Finally, the fault diagnosis is carried out by using the trained fault diagnosis SVM. Because of space limitation, only the lock in place and constant deviation fault identification for example.

The lock in place fault happens at 550th sample point and continued until the end of the sampling. The diagnosis result is shown in Fig. 3 to Fig. 4. From Fig. 4, SVM_1 can give accurate diagnostic results immediately when test data is used to SVM_1 - SVM_4. The constant deviation fault happens at 550th sample point and the diagnosis result is shown in Fig. 5 to Fig. 6. From Fig. 6, SVM_3 can give accurate diagnostic results immediately and last 10 sampling points. The results can be obtained by comprehending 4 classifiers.

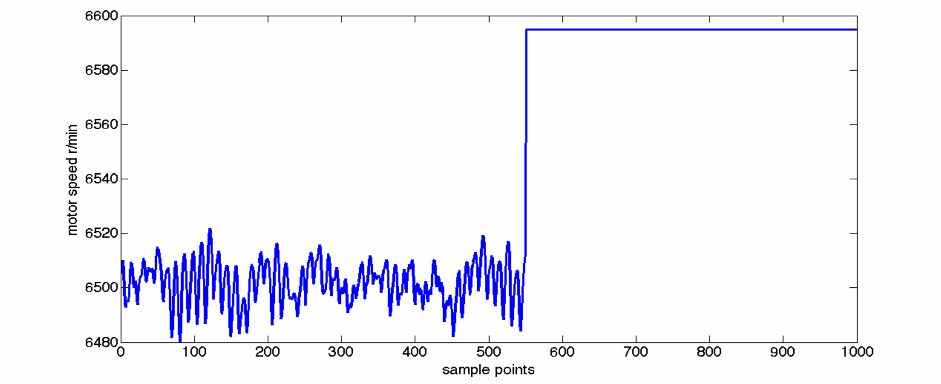

Fig. 3. Motor speed of lock in place

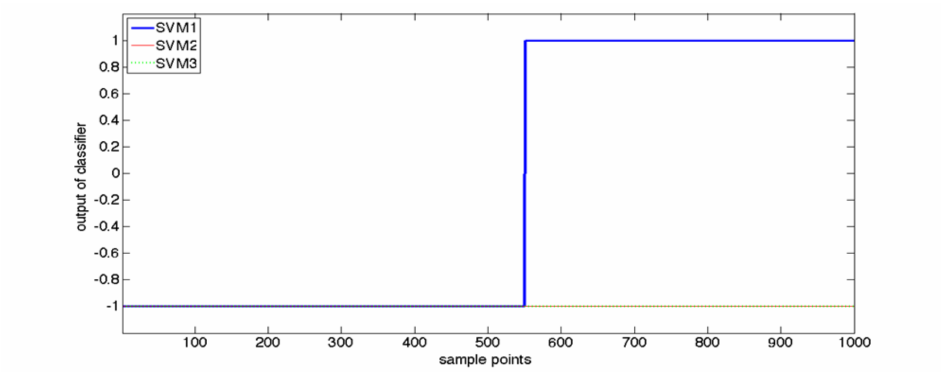

Fig. 4. Classification results of SVM

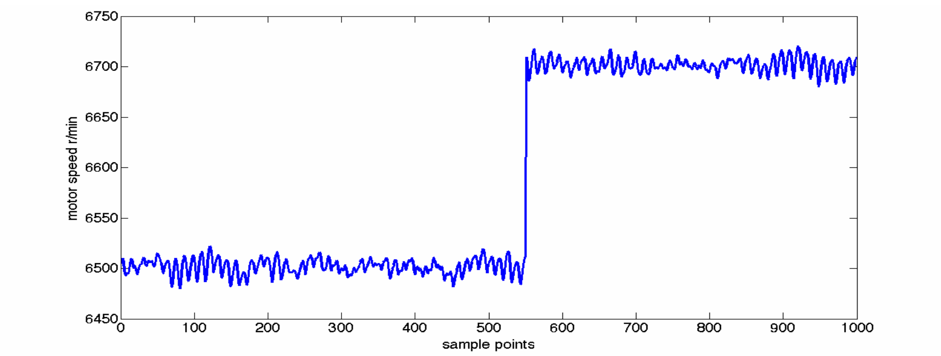

Fig. 5. Motor speed of constant deviation

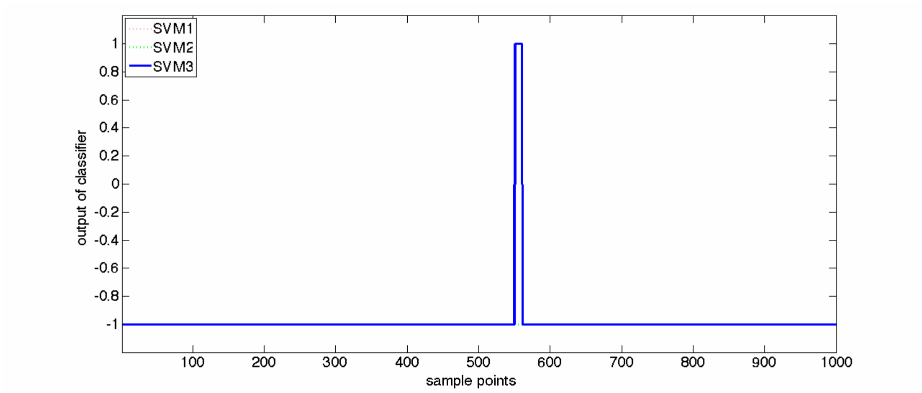

Fig. 6. Classification results of SVM

From the above simulation results of fault diagnosis, the support vector classifier can be used to diagnose the actuator faults, and it has the characteristics of high speed and accuracy.

6 Conclusion

In this paper, the fault diagnosis of the control system is studied based on the method of MCCSAPSO-SVM. Firstly, the preprocessing and feature extraction of the signal are extracted by using the method of joint noise reduction and EMD, which not only remove the pulse noise and Gauss noise affection but also finish preliminary fault detection. Secondly, the structure parameters of SVM are optimized by MCCSAPSO based on improved iterative formula. The problem of false error caused by single particle information exchange can be overcome via introducing chaos and multi-swarm cooperation. Simulated annealing can make the PSO jump out of the local extrema. Thirdly, the accuracy and real-time performance of the fault diagnosis are ensured by the partial binary tree SVM based on MKF. Finally, the simulation results show that the fault diagnosis method proposed in this paper has the advantages of accurate judgment and simple structure, so the effectiveness of this method has been verified. Further studies will focus on accelerating the computation velocity of this method through sampling some corresponding steps in order to guarantee control system real-time.

Acknowledgements

This work is supported by National Nature Science Foundation of China under Grant No. 61203090 and No. 61374130, the Fundamental Research Funds for the Central Universities No. NJ20160025 and No. NS2017018. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers and editors.

References

[1] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, T. Zhang, Deep model based domain adaptation for fault diagnosis, IEEE Transactions on Industrial Electronics 64(3)(2017) 2296-2305.

[2] T. Bruin, K. Verbert, R. Babuska, Railway track circuit fault diagnosis using recurrent neural networks, IEEE Transactions on Neural and Learning Systems 28(3)(2017) 523-533.

[3] D. Cao, J. Fu, Y. Li, Fault diagnosis of actuator of flight control system based on analytic model, in: Proc. 2016 IEEE Chinese Guidance, Navigation and Control Conference, 2016.

[4] Q. Wang, Y. Guo, H. Yu, Applications of the bionic pattern recognition in the mechanical faults diagnose, in: Proc. 13th International Conference on Ubiquitous Robots and Ambient Intelligence, 2016.

[5] R. Guo, P. Yang, X. Pan, Design of fault diagnosis expert system for civil aircraft based on FTA-BDD analysis, in: Proc. 27th Chinese Control and Decision Conference, 2015.

[6] X. Jiao, B. Jing, Y. Huang, W. Liang, G. Xu, A fault diagnosis approach for airborne fuel pump based on EMD and probabilistic neural networks, in: Proc. 2016 Prognostics and System Health Management Conference, 2016.

[7] J. Li, Q. Zhang, K. Wang, J. Wang, Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine, IEEE Transactions on Dielectrics and Electrical Insulation 23(2)(2016) 1198-1206.

[8] Q.H. Zhang, G.C. Zhao, L. Shu, A.S. Qin, Y. Zhang, Research of dimensionless index for fault diagnosis positioning based on EMD, Journal of Computers 27(1)(2016) 62-73.

[9] J. Tian, C. Morillo, M.H. Azarian, M. Pecht, Motor bearing fault detection using spectral kurtosis based feature extraction coupled with K-nearest neighbor distance analysis, IEEE Transactions on Industrial Electronics 63(3)(2016) 1793-1803.

[10] M. Van, H.J. Kang, Bearing-fault diagnosis using non-local means algorithm and empirical mode decomposition-based feature extraction and two-stage feature selection, IET Science, Measurement & Technology 9(6)(2015) 671-680.

[11] X.H. Cui, L.Y. Ma, Z.L. Wei, T.H. Wang, Application of wavelet threshold denoising and EMD in instantaneous feature extraction, Machinery and Electronics 32(10)(2014) 12-16.

[12] M.Y. Asr, M.M. Ettefagh, R. Hassannejad, S.N. Razavi, Diagnosis of combined faults in rotary machinery by non-naive Bayesian approach, Mechanical Systems and Signal Processing 85(2017) 56-70.

[13] Q. Zhang, L. Shu, A. Qin, Y. Zhang, Research of dimensionless index for fault diagnosis positioning based on EMD, Journal of Computers 27(1)(2016) 62-73.

[14] A. Aggarwal, H. Malik, R. Sharma, Feature extraction using EMD and classification through probabilistic neural network for fault diagnosis of transmission line, in: Proc. 1st IEEE International Conference on Power Electronics. Intelligent Control and Energy Systems, 2016.

[15] C. Zhao, F. Gao, Fault subspace selection approach combined with analysis of relative changes for reconstruction modeling and multi fault diagnosis, IEEE Transactions on Control Systems Technology 24(3)(2016) 928-939.

[16] Z. Gao, C. Cecati, S.X. Ding, A survey of fault diagnosis and fault-tolerant techniques—Part II: fault diagnosis with knowledge-based and hybrid/active approaches, IEEE Transactions on Industrial Electronics 62(6)(2015) 3768-3774.

[17] S. Jain, V. Bajaj, A. Kumar, Efficient algorithm for classification of electrocardiogram beats based on artificial bee colony-based least-squares support vector machines classifier, Electronics Letters 52(14)(2016) 1198-1200.

[18] B.M. Sherin, M.H. Supriya, Selection and parameter optimization of SVM kernel function for underwater target classification, in: Proc. IEEE Underwater Technology, 2015.

[19] J. Zhang, J. Cao, Fault diagnosis of multivariable dynamic system based on nonlinear spectrum and support vector machine, in: Proc. 34th Chinese Control Conference, 2015.

[20] R. Sandeep, C.R. Kailash, ECG signal analysis using DCT-based DOST and PSO optimized SVM, IEEE Transactions on Instrumentation and Measurement 66(3)(2017) 470-478.

[21] P. Li, D. Xu, Z. Zhou, W. Lee, B. Zhao, Stochastic optimal operation of micro grid based on chaotic binary particle swarm optimization, IEEE Transactions on Smart Grid 7(1)(2016) 66-73.

[22] Y.G. Kim, M.J. Lee, Scheduling multi-channel and multi-timeslot in time constrained wireless sensor networks via simulated annealing and particle swarm optimization, IEEE Communications Magazine 52(1)(2014) 122-129.

[23] P. Yang, R. Guo, X. Pan, J.W. Liu, The fault feature extraction of state signal in control systems based on joint noise reduction method and empirical mode decomposition, Journal of Computers 27(4)(2016) 1-14.

[24] Z.C. Zheng, W.Y. Xu, F. Xu, Z.B. Liu, Forecasting of slope displacement based on PSO-LSSVM with mixed kernel, Rock and Soil Mechanics 33(5)(2012) 1421-1427.