

An Effective Method to Improve Electronic Equipment Condition Monitoring Based on KPCA-EDA and MMSH-SVDD

Yang Sen, Meng Chen, Lv Meng

Department of Missile Engineering, Ordnance Engineering College
Shijiazhuang, Hebei, China, 050003

Abstract

In order to improve the validity of electronic equipment condition monitoring, overcome the shortage of normal KPCA (Kernel Principal Component Analysis) and SVDD (Support Vector Data Description) monitoring model, a method on electronic equipment condition monitoring based on KPCA-EDA (KPCA- Estimation of Distribution Algorithm) and MMSH-SVDD (Maximal Margin Separating Hypersphere SVDD Mode) is put forward. Firstly, the feature of original monitoring data is extracted by KPCA-EDA algorithm, and a group of features with enough state identifying information are obtained; then the MMSH-SVDD model is trained by the normal state and a little bit of fault state features, and the unknown state feature is applied to the trained model; Finally, a filter circuit is taken as an example in simulations, the result shows that this method is the effective method improve the performance of electronic equipment condition monitoring.

Keywords: PCA-EDA, MMSH-SVDD, electronic equipment, condition monitoring, method

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

There are many kinds of electronic equipment in electronic field, and their characteristics are different, most of the existing literatures researches on the electronic equipment fault diagnosis show little concentrate on circuit condition monitoring. But actually in the process of circuit operation, we need to monitor the circuit state online and pay more attention to the health degradation of the circuit, so as to judge the actual health state of the circuit [1, 2].

Starting from the point of electronic equipment condition monitoring, a monitoring method based on KPCA-EDA and MMSH-SVDD is put forward in the paper. First, it adopts EDA (Estimation of Distribution Algorithm) to choose the KPCA principal component feature information, so as to retain more identifying information under the premise of adequate state feature information; and then on the basis of traditional SVDD, the MMSH-SVDD monitoring model is proposed according to the sample data after the feature extraction. The model can makes a trade-off between the minimum volume description and the maximum class interval, increase the generalization ability of the model; the filter circuit is taken as the example in simulations, the analysis results show that the method is effective.

2. Feature Extraction Based on KPCA-EDA

The feature extraction algorithm based on KPCA is a nonlinear promotion to PCA (Principal Component Analysis) [3], with high ability of nonlinear feature extraction, and it has been widely used in pattern recognition field. But it still has shortcomings in the process of feature extraction, during the analysis process; KPCA takes all of the samples as a whole and seek direction with maximum divergence. However, in many cases the direction with maximum divergence is not consistent with the identification of the most favorable direction [4]. Therefore, it is necessary to select the principal component obtained after KPCA, with the purpose of improving the sample recognition rate [5, 6].

In this paper, the feature extraction process of the monitoring data, it firstly uses KPCA algorithm to extract the feature of original samples, the larger neigenvalues ($n \leq N$, N is the

total number of samples) and corresponding eigenvectors are obtained through solving the nuclear matrix characteristic equation; Then according to some criteria, it adopts EDA to select the eigenvectors, the principal component combination with optimal identifiable information is obtained; Finally, the projecting monitoring data to this principal component and the final feature data are got. The algorithm schematic is as shown in Figure 1.

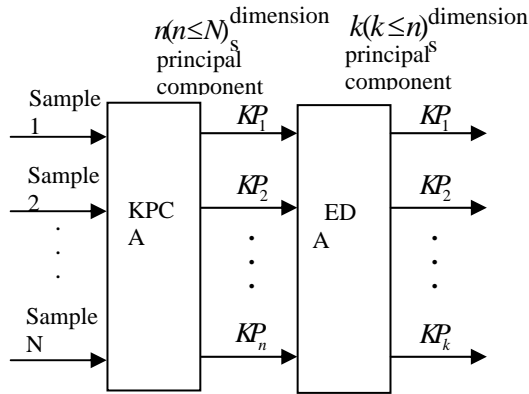


Figure 1. Feature Extraction Process Based on KPCA-EDA

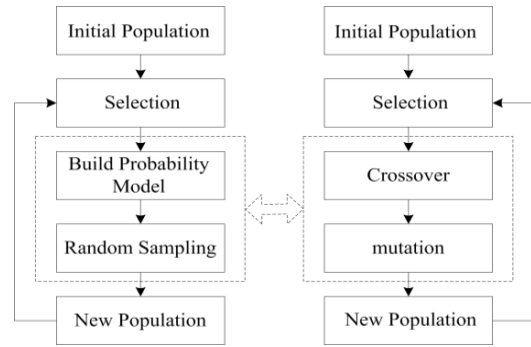


Figure 2. Comparison of EDA and GA

2.1. Principle of EDA Algorithm

EDA algorithm is proposed as a new type of intelligent optimization method by *Mühlenbein* in 1996. There are no crossover and mutation operations in the distributed estimation algorithm and the traditional genetic algorithm [7-10].

The basic idea of EDA algorithm is that, in the first step, a portion of individuals are preferred from the current population and generates statistical information according to these preferred individuals, its probability distribution is estimated; Then the next generation of individuals taking is produced according to the advantage of this probability distribution. EDA algorithm and genetic algorithm are compared as shown in Figure 2.

2.2. Feature Extraction Algorithm Based on KPCA-EDA

The main steps of the feature extraction algorithm based on KPCA-EDA are as follows:

Step 1 is to produce initial population $D_l (l = 0)$. Use KPCA algorithm to analyze the principal component of monitoring data in feature space, and N eigenvalues and corresponding eigenvectors α^r are obtained, where $r = 1, 2, \dots, N$ (N is the total number of monitoring data samples). Then the eigenvalues are sorted in descending order, and the eigenvectors are chosen which are corresponding to the larger $n (n \leq N)$ eigenvalues as the candidate principal component. If we set $KP = (a^1, \dots, a^n)$ and $\Lambda_{KP} = \text{diag}(\lambda^1, \dots, \lambda^n)$, and takes $\alpha^r (r = 1, 2, \dots, n)$ as the random variable, and the initial population D_0 of M individuals will be generated. The binary encoding way is obtained in encoding the n eigenvectors α^r , every code bit represents that whether the corresponding eigenvector has been selected or not, 1 denotes it is selected, 0 denotes it is unselected. The length of coding is n , and the probability of the variable is 1 or 0 is the same during the initial population process.

Step 2 is the Calculation of the fitness, and choosing the excellent individual D_l^{Se} . The design of fitness function is the key of KPCA-EDA algorithm, the excellent individual is chosen in the statistical analysis in generating new samples. In KPCA-EDA algorithm, we need to consider both the identification ability and adequate information content, so the fitness function is designed as follows:

$$fitness_i = J(KP_i) \times F(KP_i) \tag{1}$$

Where $J(KP_i)$ represent the identification ability of the chosen eigenvectors, $J(KP_i)$ represent the information contain ability.

Through equations we know that after feature extraction through KPCA, every sample x_i^j is changed according to $\tilde{x}_i^j = (w^r)^T \varphi(x_i^j) = (\alpha^r)^T \delta_i^j = (KP)^T \delta_i^j$, the centre of class j samples m_j will be satisfied with $\tilde{m}_j = (w^r)^T m_j^F = (\alpha^r)^T \psi_j = (KP)^T \psi_j$, and the centre of the whole samples m is turned into $\tilde{m} = (w^r)^T m^F = (\alpha^r)^T \psi = (KP)^T \psi$.

As the process of condition monitoring is based on classification ideology in essence, the criterion based on distance in classification definition is introduced. In the paper, we take $J(KP_i)$ as the within-class and the class distance criterion, the former is represented by within-class scatter matrix S_w , and the latter is represented by between-class scatter matrix S_b . The expression of the criterion based on within-class and between-class distance is as follows [12]:

$$J(KP_i) = tr(S_b) / tr(S_w) \quad (2)$$

Among them,

$$S_b = \frac{1}{N} \sum_{j=1}^c N_j (\tilde{m}_j - \tilde{m})(\tilde{m}_j - \tilde{m})^T \quad (3)$$

$$S_w = \frac{1}{N} \sum_{j=1}^c \sum_{i=1}^{N_j} (\tilde{x}_i^j - \tilde{m}_j)(\tilde{x}_i^j - \tilde{m}_j)^T \quad (4)$$

When the value of $J(KP_1)$ and $J(KP_2)$ are almost the same, the individual corresponding to the larger eigenvalue should be chosen, which includes the larger amounts of information. So the function $F(KP_i)$ is designed as follows:

$$F(KP_i) = e^{a(\eta-b)}, a \geq 0, 0 \leq b \leq 1 \quad (5)$$

Where $\eta = (\sum_{k=1}^m \lambda_{rk}) / (\sum_{j=1}^n \lambda_j)$, m indicates the number of chosen eigenvectors, n is the length of code size, which is the number of candidate eigenvectors; b represents the threshold of η ; a represents the adjustment coefficient, which is used to adjust the weight of $F(KP_i)$ in the whole fitness function.

After design process of the fitness function, the number $Se (Se \leq M)$ of individuals are selected, usually the fittest ones.

Step 3 is to build the probability model $p_i(x)$ according to the excellent individual. Since the eigenvectors between each others are statistically uncorrelated, the UMDA algorithm with independent variable is chosen as the EDA probability model. Every marginal distribution is decided by the frequency of 1 in every bit l. The probability distribution is as follows [13]:

$$p_i(x) = p(x | D_i^{Se}) = \prod_{i=1}^n p_i(x_i) = \prod_{i=1}^n \frac{\sum_{j=1}^{Se} \delta_j(X_i = x_i | D_i^{Se})}{Se} \quad (6)$$

Where,

$$\sum_{j=1}^{Se} \delta_j(X_i = x_i | D_i^{Se}) = \begin{cases} 1, & X_i = x_i \\ 0, & other \end{cases} \quad (7)$$

Step 4 is the random sampling. After M times sampling from probability vector $p_l(x)$, the new generation population D_{l+1} is obtained. Then the fitness value of each new individual can be calculated.

Step 5 is the termination condition judgment. If a termination condition is satisfied, it will go to next step; Else $l = l + 1$, and turn back to step 2.

Step 6 is to output results. The population D_{l+1} is the solution.

The algorithm flow is shown in Figure 3.

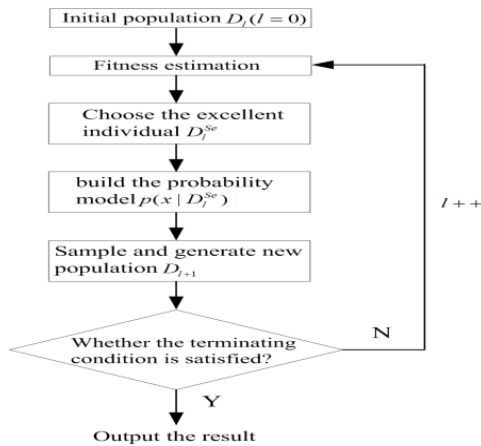


Figure 3. The Flow of Algorithm

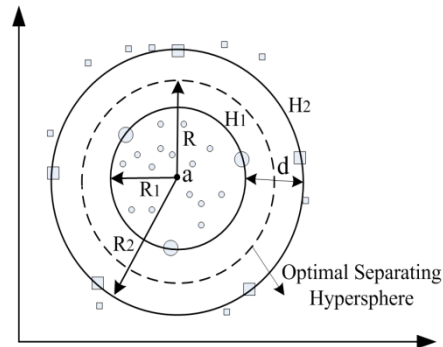


Figure 4. MMSH-SVDD Model

3. Condition Monitoring Based on MMSH-SVDD

The between-class margin and within-class cohesion are two important factors impacting on the performance of classifiers [14]. Normal SVDD model only take the largest of within-class into consideration, and does not take the between-class margin into account [15, 16]. In this section, we consider both factors above and put forward a new SVDD model based on maximal margin separating hypersphere, which is shown in Figure 4.

Two concentric supraspheres of different radius are built, the target samples are in the small one and non-target samples are outside the big one. Make sure that the small suprasphere is minimized, and the distance between target and non-target samples is largest, while the generalization ability of the model is improved.

The realization process of the model is as follows, first make the assumption that two classes training samples are $(x_1, y_1), \dots, (x_l, y_l), x \in R^n, y \in \{-1, 1\}$, where the number of target samples are m, non-target samples are n, the total number is $l(l = m + n)$. Through nonlinear mapping ϕ , the training samples are projected to the high dimension feature space, in which two concentric supraspheres S_1 and S_2 of different radius are built, the centre of spheres is a, and the radius is R_1 and R_2 ($R_1 \leq R_2$), H_1 and H_2 are the boundary surface of S_1 and S_2 , d is the distance between H_1 and H_2 . The optimal goal is to minimize S_1 and maximize d .

As shown in Figure 4, the between-class margin $d = R_2 - R_1$, maximized d is equivalent to maximized R_2 and minimized R_1 , or maximized R_2^2 and minimized R_1^2 . Then the optimization problem can be described as:

$$\min \mu R_1^2 - R_2^2 + C_1 \sum_{i=1}^m \xi_i + C_2 \sum_{j=1}^n \xi_j \tag{8}$$

$$\| \phi(x_i) - a \|^2 \leq R_1^2 + \xi_i \quad 1 \leq i \leq m \tag{9}$$

$$\|\phi(x_j) - a\|^2 \geq R_2^2 - \xi_j \quad 1 \leq j \leq n \quad (10)$$

Where μ is the trade-off parameter of radius, which is used to compromise R_1^2 and R_2^2 ; ξ_i, ξ_j are the slack variable; C_1, C_2 are the penalty coefficient; R_1 is the radius of the small suprasphere and R_2 is the big one.

In order to solve the optimization problem in (8)–(10), we construct the Lagrangian:

$$\begin{aligned} L = & \mu R_1^2 - R_2^2 + C_1 \sum_{i=1}^m \xi_i + C_2 \sum_{j=1}^n \xi_j - \sum_{i=1}^m a_i [R_1^2 + \xi_i - (\phi(x_i) \cdot \phi(x_i) \\ & - 2a\phi(x_i) + a^2)] + \sum_{j=1}^n a_j [R_2^2 - \xi_j - (\phi(x_j) \cdot \phi(x_j) - 2a\phi(x_j) + a^2)] \\ & - \sum_{i=1}^m \beta_i \xi_i - \sum_{j=1}^n \gamma_j \xi_j \end{aligned} \quad (11)$$

We set the partial derivatives to zero, and use kernel function $k(x_i, x_j)$ to replace the inner product $\phi(x_i) \cdot \phi(x_j)$, the quadratic dual problem of the optimization problem is as follows:

$$\begin{aligned} \min \sum_{i=1}^m a_i k(x_i, x_i) - \sum_{j=1}^n a_j k(x_j, x_j) - \frac{1}{\mu-1} [\sum_{i=1}^m \sum_{k=1}^m a_i a_k k(x_i, x_k) \\ + \sum_{j=1}^n \sum_{p=1}^n a_j a_p k(x_j, x_p) - 2 \sum_{i=1}^m \sum_{j=1}^n a_i a_j k(x_i, x_j)] \end{aligned} \quad (12)$$

$$\text{st. } \sum_{i=1}^m a_i = \mu \sum_{j=1}^n a_j = 1, 0 \leq a_i \leq C_1, 0 \leq a_j \leq C_2 \quad (13)$$

Through solving Equation (12) and (13), we can get a_i and a_j . According to related condition, the large fraction of the weights a_i and a_j turns to 0, only a few non-zero weights are called the support vectors (SVs).

When $0 \leq a_i \leq C_1$, the SVs are on H_1 , which is the boundary surface of S_1 , by using any SV $\phi(x_k)$ we can get radius R_1 :

$$R_1^2 = \|\phi(x_k) - a\|^2 \quad (14)$$

When $0 \leq a_j \leq C_2$, the SVs are on H_2 , which is the boundary surface of S_2 , by using any SV $\phi(x'_k)$ we can get radius R_2 :

$$R_2^2 = \|\phi(x'_k) - a\|^2 \quad (15)$$

According to (14) and (15), we can get the final radius R as shown in (16).

$$R = \frac{R_1 + R_2}{2} \quad (16)$$

If we assume the test sample is z , then if $\|\phi(z) - a\| \leq R$, z is the target sample; and if $\|\phi(z) - a\| > R$, z will be the non-target sample.

4. Simulation

The biquadratic filter circuit as an example in the simulation, the circuit is as shown in Figure 5. The amplitude of frequency points in amplitude-frequency response curves is taken as the candidate features, sampling interval is in [100Hz, 10KHz]. The total characteristic number is 21, and the amplitude-frequency response curve of normal output is shown in Figure 6.

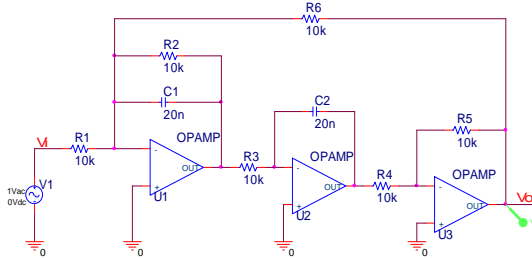


Figure 5. Biquadratic Filter Circuit

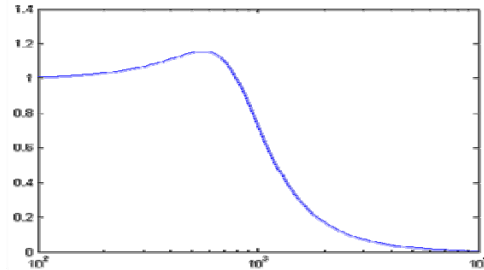


Figure 6. Amplitude-frequency Response Curve of Normal Output

Through the sensitivity analysis, we find that the changing of R_1 and C_2 can influence the output voltage of the circuit. If we suppose that the single soft fault is happened in the circuit, including no fault, R_1 , C_2 positive and negative 50%, the total is five state types. The type of each single soft fault is shown in Table 1.

Table 1. Type of Single Soft Fault

Num	Fault type
F1	NF
F2	$R_1 = 15k\Omega$
F3	$R_1 = 5k\Omega$
F4	$C_2 = 30nF$
F5	$C_2 = 10nF$

In the Simulation, the normal state of this circuit is simulated by Monte-Carlo. The number of the total original samples is 780, and the dimension of them is 21. Choosing 2/3 samples of each state to form the training sample sets, the rest of them is taken as the testing sample sets.

4.1. Feature Extraction Based on KPCA-EDA

KPCA-EDA algorithm is applied in extracting the feature of the sample sets. The step of feature extraction is as follows:

Step 1: Choose RBF function as the kernel function $k(x_i, x_j)$ of the samples [17].

Through calculating kernel matrix K, then we can get \bar{K} , and the kernel parameter $\gamma = 0.5$.

Table 2. Candidate Eigenvalues

Num	eigenvalues	Num	eigenvalues
1	10.692	8	0.7089
2	5.4723	9	0.4775
3	3.1001	10	0.2601
4	2.9716	11	0.1438
5	1.2631	12	0.0902
6	1.1433	13	0.0323
7	0.9898		

Step 2: Calculate the eigenvalues and corresponding eigenvectors of \bar{K} . We take the eigenvectors which accumulating contribution rate is 98% as the candidate principal eigenvectors (the total number of them is 13). The corresponding candidate eigenvalues are as shown in Table 2.

Step 3: Taking the candidate eigenvectors as the random variables, and the initial population is D_0 . Through binary encoding way to encode it, and the length of coding is 13, and the number of initial population is 150.

Step 4: Choose the excellent eigenvectors by EDA. The fitness function of EDA is $F(KP_i) = e^{10(\eta-0.98)}$, where $a=10, b=0.98$. The relationship between the iterations and average fitness of each generation is as shown in Figure 7.

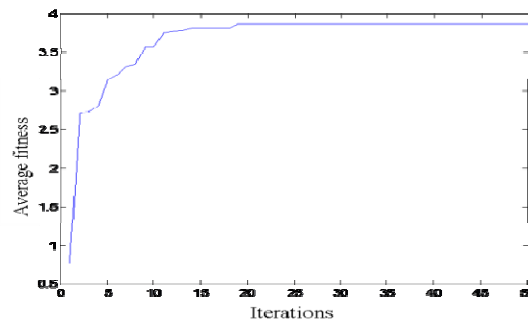


Figure 7. Relationships between Iteration and Average Fitness

From Figure 7 we can see that the globally optimal solution is obtained after 19 times iteration simulation, and the corresponding fitness value is 3.8735. The best individual is 1110011001010, which indicates that the eigenvectors of number 1, 2, 3, 6, 7, 10, 12 are selected; the total selected eigenvectors are 7. Then calculating the accumulating contribution rate of this eigenvector combination, the result is 93%, which is general conformed to the requirements.

Step 5: Project the training sample and testing sample to the chosen eigenvectors, and then the final feature training sets and feature test sets by Equation (10) are got.

4.2. Condition Monitoring Based on MMSH-SVDD

The feature training sets are used in training MMSH-SVDD model and normal SVDD model, the KPCA and KPCA-EDA are also obtained, and then testing of the condition monitoring ability of the models is made through using feature testing sets.

The parameters are chosen by 10 fold cross validation method, and the model performance is evaluated by accuracy test. The test results are shown in Table 3.

Model	Feature extraction	Accuracy
SVDD	KPCA	92.31%
	KPCA-EDA	94.23%
MMSH-SVDD	KPCA	94.62%
	KPCA-EDA	97.69%

From Table 3, we can see that by using MMSH-SVDD model and KPCA-EDA feature extraction algorithm, the accuracy of testing sample sets is the highest which is about 97.69%.

5. Conclusion

In this paper, it analyzes the electronic equipment condition monitoring method based on KPCA-EDA and MMSH-SVDD. Aiming at the shortage of KPCA lack of identifying information, EDA is used to choose the principle component further, so as to reduce the characteristic dimension, and retain more identifying information; in order to solve fitting problem of normal SVDD, a few non-target samples are imported to the model training process, and the condition monitoring model based on MMSH-SVDD is built. Finally, a filter circuit is taken as an example to validate the improved algorithms; the result shows that the performance of this method is more effective than normal SVDD and KPCA.

References

- [1] Sun Zhixin, Li Zili. *The world electronic information equipment*. Beijing: Defense Technology Publishing. 2001: 15-28.
- [2] Ma Sasa, Chen Guoshun, FANG Xingqiao. Research on Prognostic and Health Management System of Complex Equipment. *Computer Measurement & Control*. 2010; 18(1): 1-4.
- [3] Jolliffe IT. *Principal component analysis*. Springer Series in Statistics, Springer, second edition.2002.
- [4] Cao LJ, Chua KS, Chong WK. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. *Neurocomputing*. 2003; 55: 3211-336.
- [5] Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis*. New York: Prentice Hall. 2001: 55-74.
- [6] Xia Guoen. Customer churn prediction on kernel principal component analysis feature abstraction. *Computer Applications*. 2008; 28(1): 149-151.
- [7] Wang Xinfeng, Qiu Jing, LIU Guanjun. Research on Fault Feature Extraction Based on Kernel K-L Transformation. *Mechanical Science and Technology*. 2006; 25(3): 288-291.
- [8] Na Chen, Shaopu Yang, Cunzhi Pan. Application of Fault Detection based on Hybrid Intelligent Methods. *Indonesian Journal of Electrical Engineering*. 2012; 10(8).
- [9] Yanjun Ren, Guanghua Wen, Xiuyun Li. An SVM based Algorithm for Road Disease Detection using Accelerometer. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(9): 5169-5175.
- [10] Qu Zhou, Weigen Chen, Xiaoping Su, Shudi Peng. Quantitative Recognizing Dissolved Hydrocarbons with Genetic Algorithm-Support Vector Regression. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(9): 5509-5516.