Research Journal of Applied Sciences, Engineering and Technology 5(22): 5314-5319, 2013
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: November 08, 2012 Accepted: January 05, 2013 Published: May 25, 2013

A Novel Fault Feature Extraction Method of Analog Circuit Based on Improved KPCA

[1]He Xing, [1]Wang Hong-li, [1]Lu Jing-hui and [2]Sun Guo-qiang
[1]Xi'an Research Institute of Hi-Tech., Hongqing Town, Xi'an 710025, China
[2]Aviation University of Air Force, Changchun, 130022, China

Abstract: The Kernel Principal Component Analysis (KPCA) extracts the principal components by computing the population variance, which doesn't consider the difference between one class and the others. So, it makes against the fault diagnosis. For solving this problem, the study introduced Fisher classification function into The KPCA and proposed an improved FKPCA with the class information. Then, the algorithm was applied in analog-circuit fault feature extraction and the neural network was applied to diagnose the faults. The results indicate the classification effect of the principal components extracted by the algorithm is more better. It improves the rate of fault diagnosis and reduces the test time.

Keywords: Between-class scatter matrix, feature extraction, fisher criterion, KPCA, within-class scatter matrix

INTRODUCTION

Feature extraction is one of the most important processes in fault diagnosis of analog circuits, so researches concern themselves about how to extract the feature efficiently, which result in lower computational cost and better result. Recently, the Kernel Principal Component Analysis (KPCA) method has been proposed for fault diagnosis, which is a novel nonlinear multivariate statistical analysis method (Scholkopf *et al*., 1998). KPCA was proposed to use for fault identification of process monitoring firstly(). The KPCA method has exhibited superior performance compared to the linear principal component analysis method in processing nonlinear systems (Cho *et al*., 2005; Chin and Suter, 2007; Choi *et al*., 2005). The KPCA is a promising feature extraction method, which can eliminate the correlation of one feature and others and solve the problems such as oversize raw data dimension, unknown disturbances and low SNR(signal to noise ratio), etc (Scholkopf *et al*., 1999). However, there is a classification problem in the KPCA-associated method. In the KPCA method, it analyses all samples as a whole, only considers the population variance instead of category information and class difference, which make the principal component extraction blind and partial (Xiao and He, 2011). Consequently, it could influence the fault diagnosis effect in practice.

To solve the problem, the Fisher linear discriminant function is introduced, then, an improved KPCA, FKPCA (Fisher KPCA), is presented in this study, which reconstructs eigenvector projection space by minimizing between-class scatter and maximizing within-class scatter to make the extracted principal components consist of class information. Subsequently, the proposed FKPCA is used for fault feature extraction of analog circuit, in the end, the extracted fault feature are put into neural network for fault diagnosis. The numerical results show that the suggested approach can improve significantly the performance of fault diagnosis rate of neural network.

The study is organized as follows. Section 2 provides a brief presentation of KPCA. In Section 3, our proposed FKPCA is detailed introduced. Section 4 presents numerical results for fault diagnosis of analog circuit using the FKPCA and neural network. Finally, Section 5 contains a discussion of the results as well as directions for future work.

MATERIALS AND METHODS

Brief review of KPCA: PCA is a powerful technique for extracting intrinsic structure from high dimensional data set (Bishop, 1995). However, PCA is a linear technique and cannot capture nonlinear structure in a data set. Therefore, nonlinear generalizations have been proposed and especially KPCA based on kernel theory was introduced for computing the principal components of the data set mapped nonlinearly into some high-dimensional feature space. KPCA is a nonlinear PCA method. The implementation of KPCA seems to be equivalent to the implementation of the following process: all the samples are first transformed into a new space by using a nonlinear mapping. Then PCA is performed in the new space and extracts the lower dimensional features of samples in the new space. However, KPCA indeed does not need to explicitly

Corresponding Author: He Xing, Xi'an Research Institute of Hi-Tech., Hongqing Town, Xi'an 710025, China

perform the nonlinear mapping. Instead, KPCA implicitly obtains the nonlinear mapping by exploiting the kernel trick. This enables KPCA to have a promising computational cost in comparison with a general nonlinear feature extraction method.

It is also seen that KPCA is an equivalent implementation of PCA in the feature space (i.e. the new space mentioned above). KPCA is briefly presented as follows. Let vectors $x_1, x_2, \ldots, x_N$ be N training sample have been transformed into the feature space by a nonlinear function ϕ. As a result, we can use $\phi(x_1) \ldots \phi(x_N)$ to denote the training samples in the feature space. If the samples in the feature space have zero mean, then the covariance matrix is:

$$\Gamma(\phi) = \frac{1}{N}\sum_{i=1}^{N}\phi(x_i)\phi(x_i)^T$$

We also refer to $\Gamma(\phi)$.as the generation matrix of the feature space. According to the PCA methodology, the most useful eigenvectors of the feature space should be the eigenvectors corresponding to large eigenvalues of $\Gamma(\phi)$.That is, the most useful eigenvectors should be the solutions u_i corresponding to large λ_i of $\Gamma(\phi).ui = \lambda_i u_i$. By exploiting the kernel function $k(x_i, x_j)$ to denote the dot product, i.e., $k(x_i, x_j) = \phi(x_i)^T \phi(x_i)$, the following eigenvalue equation can be derived (Scholkopf *et al*., 1997):

$$K\alpha = \lambda\alpha \tag{1}$$

where K is the so-called Gram matrix that has the entry $(K)_{ij} = k(x_i, x_j)$, α is the eigenvector. The principal component analysis method based on the eigenvalue Eq. (1) is referred to as KPCA.

Basic principle of the proposed FKPCA: First of all**,** definition Parameter. Before introducing the Fisher criterion function, some essential parameters are defined firstly (Wang *et al*., 2006). Setting fault sample as $x \in R^n$, n is the dimension of raw feature set; c is the number of fault class, each class consists of N_j samples, where $j = 1, 2, \ldots, c$ and the sum of samples fulfil the condition: $\sum_{j=1}^{c} N_j = N$.

$P(w_j)$ is the prior probability of training samples of the type j, w_j denotes the j^{th} training sample; $m_j = 1/N_j$ ($m_j \in R^n$) is the mean of training samples of the type j and $m = 1/N\sum_{i=1}^{N} x_i$ ($m \in R^n$) is the mean of population training samples;

Population between-class scatter matrix marks with S_w, which can be calculated by Eq. (2):

$$\begin{aligned} S_w &= \sum_{j=1}^{c} P(w_j) S_{wj} \\ &= \frac{1}{N}\sum_{j=1}^{c}\sum_{i=1}^{N_j}(x_i - m_j)(x_i - m_j)^T \end{aligned} \tag{2}$$

Population within-class scatter matrix marks with S_b, which can be calculated by Eq. (3):

$$\begin{aligned} S_b &= \sum_{j=1}^{c} P(w_j)(m_j - m)(m_j - m)^T \\ &= \sum_{j=1}^{c}\frac{N_j}{N}(m_j - m)(m_j - m)^T \end{aligned} \tag{3}$$

and population scatter matrix marks with S_t, which can be calculated by Eq. (4):

$$S_t = S_w + S_b = \frac{1}{N}\sum_{i=1}^{N}(x_i - m)(x_i - m)^T \tag{4}$$

The between-class scatter matrix represents diffusion condition of every sample points around their mean and the within-class scatter matrix represents the distribution of class distance, which all depend on the character and division of the sample class. But the population scatter matrix is independent of sample division and class character. The KPCA acquires projection space using the population scatter matrix with no class information; consequently, the principle components extracted by KPCA make the class very similar, which maybe bad for diagnose different type of fault. So, it is necessary to study how to introduce class information in the process of KPCA to improve the performance of diagnosis of different fault.

Then, the basic ideas of improvement are introduced. the Fisher linear discriminant function can be acquired from (Bian and Zhang, 2009), which has the form as:

$$J_F(w) = \frac{w^T S_b w}{w^T S_w w} \tag{5}$$

it can see from the above equation that in order to obtain the best classification ability after projection, the bigger is the within-class scatter S_b and the less is the between-class scatter S_w, the better for classification result, in other words, the different types of samples should be dispersed as far as possible, meanwhile, the samples in the same type should be dense as near as possible. Consequently, this problem is reduced to calculate the corresponding value w^* of variable w, which makes the function $J_F(w)$ maximum.

It can obtain the simplified form by using Lagrange multiplier method (Xiao and He, 2011) to solve Eq. (5):

$$S_b w^* = \lambda S_w w^* \tag{6}$$

where w^* is the extremal solution. As S_w is nonsingular, Eq. (7) can be obtain by left multiply by S^{-1}_w on each side of Eq. (6):

$$S_w^{-1} S_b w^* = \lambda w^* \quad (7)$$

From Eq. (7), It could consider obtaining the extremum of $J_F(w)$ as solving the eigenvalue of the general matrix, S^{-1}_b S_b. See not hard, for S_w and S_b containing between-class information and within-class information respectively, if the population covariance matrix $\Gamma(\emptyset)$ is replaced by S^{-1}_b S_b in KPCA, the eigenvalues would also include each class information. Consequently, it would make the projected sample easier to divide.

After original sample space has been projected into eigenvector space, the following problem is how to obtain the between-class scatter matrix S_b and the within-class scatter matrix S_w in the process of KPCA. The detailed steps are given in the next section.

Finally**,** according to the above analysis, the detailed achievement process of FKPCA can be obtained. When raw sample x subjects to non-linear mapping, the between-class scatter matrix S^{Φ}_w and the within-class scatter matrix S^{Φ}_b within eigenvector space are given as:

$$S_w^{\Phi} = \frac{1}{N} \sum_{j=1}^{c} \sum_{i=1}^{N_j} (\Phi(x_i) - m_j^{\Phi})(\Phi(x_i) - m_j^{\Phi})^T \quad (8)$$

$$S_b^{\Phi} = \sum_{j=1}^{c} \frac{N_j}{N} (m_j^{\Phi} - m^{\Phi})(m_j^{\Phi} - m^{\Phi})^T \quad (9)$$

where,

$$m_j^{\Phi} = \frac{1}{N_j} \sum_{i=1}^{N_j} \Phi(x_i), \quad j = 1,2,...,c$$

$$m^{\Phi} = \frac{1}{N} \sum_{i=1}^{N} \Phi(x_i) = \frac{1}{c} \sum_{j=1}^{c} m_j^{\Phi}$$

It can be seen from the above analysis, in the eigenvector space, the expression of Fisher criterion function becomes:

$$J_F^{\Phi}(w) = \frac{w^T S_b^{\Phi} w}{w^T S_w^{\Phi} w} \quad (10)$$

and the solution vector w of all the kernel learning methods, which marked by inner product sums of the image Φ (x_i) of sample vectors in the eigenvector space, i.e:

$$w = \sum_{i=1}^{N} \alpha_i \Phi(x_i) = \Phi(X)\alpha \quad (11)$$

where,

$$\Phi(X) = [\Phi(x_1), ..., \Phi(x_N)], \quad \alpha = [\alpha_1, ..., \alpha_N]^T$$

From Eq. (11), it can see that:

$$w^T m_j^{\Phi} = \frac{1}{N_j} \sum_{i=1}^{N} \sum_{k=1}^{N_j} \alpha_i k(X_i, X_k^{(\omega_j)}) = \alpha^T M_j, \quad j = 1,2,...c \quad (12)$$

Define M_j as a N×1 matrix and:

$$(\boldsymbol{M}_j)_i = \frac{1}{N_j} \sum_{k=1}^{N_j} k(X_i, X_k^{(\omega_j)}), \quad j = 1,2,...c \ ; \ i = 1,2,...,N \quad (13)$$

From Eq. (8) and (9), two equations can be derived (Scholkopf *et al*., 1997):

$$w^T S_b^{\Phi} w = w^T \sum_{j=1}^{c} \frac{N_j}{N} (m_j^{\Phi} - m^{\Phi})(m_j^{\Phi} - m^{\Phi})^T w = \alpha^T T \alpha \quad (14)$$

$$w^T S_w^{\Phi} w = w^T \frac{1}{N} \sum_{j=1}^{c} \sum_{i=1}^{N_j} (\Phi(x_i) - m_j^{\Phi})(\Phi(x_i) - m_j^{\Phi})^T w = \alpha^T H \alpha \quad (15)$$

where,

$$T = M(I_c - \frac{1}{c} L_c) M^T \quad (16)$$

$$H = \sum_{i=1}^{c} K_i (I_{N_i} - \frac{1}{N_i} L_{N_i}) K_i^T \quad (17)$$

in the Eq. (16) and (17), $M = (M_1, M_2, \dots, M_c)$, K_i is a $N \times N_i (i = 1, 2, \dots, c)$ matrix with $(K_i)_{p,q} = k(X_P, X^{(\omega i)}_q)$, where $p = 1, 2, \dots, N$; $q = 1, 2, \dots, N_i$, i.e., K_i is the type i kernel matrix, I_c, I_{Ni} are $c \times c$ and $N_i \times N_i$ identity matrix respectively, where ($i = 1, 2, \dots, c$) and L_c, L_{Ni} are $c \times c$ and $N_i \times N_i$ matrix consisted of 1 respectively, where ($i = 1, 2, \dots, c$).

From Eq. (14) and (15), it can be seen that Eq. (10) is equivalent to the equation:

$$J(\alpha) = \frac{\alpha^T T \alpha}{\alpha^T H \alpha} \quad (18)$$

Then, according to the same as Fisher linear discriminant function, the following equation can be obtained:

$$H^{-1} T \alpha = \lambda \alpha \quad (19)$$

After obtained the Eq. (19), eigenvector space can be calculated and the projections of the mapping data in it are the non-linear principal components.

According to the above analysis, solving steps of FKPCA are as:

Step1: Calculate the between-class scatter matrix and the within-class scatter matrix according to Eq. (8) and (9)

Step2: Calculate eigenvalue and eigenvector according to Eq. (19)

Step3: Sort all the eigenvalue by descending order and then sort the corresponding eigenvector

Step4: Calculate the projection of mapping data in eigenvector space, then, the non-linear principal components with class information would be obtained

Step5: Select the principal components according to accumulative contribution to reduce the dimension of raw data. The accumulative contribution can calculated by Eq. (20):

$$\mathrm{AC} = \sum_{i=1}^{k} \lambda_i \Big/ \sum_{i=1}^{N} \lambda_i \tag{20}$$

k is the number of the front eigenvalues under descending when AC>0.85.

SIMULATION RESULTS AND DISCUSSION

In this study, a simulation circuit of band-pass filter is chosen to verify the validity of proposed FKPCA method. The simulation experiment is carried out in Windows XP with MATALB2007a and the circuit simulation software Pspice9.2. Figure 1 shows circuit diagram of the band-pass filter.

Diagnosis process: for the circuit, let the tolerances of resistances and capacitances are ±5% and ±10% respectively. Here, take R1=10kΩ with ±5% tolerance for example, which have three following cases:

- When $R1 \in [9.5k, 10.5k]$, it indicates that R1is in the normal condition , i.e., fault-free.
- When $R1 > 10.5k$, it indicates that R1 overstep the normal upper limit, that is, soft fault happened; furthermore, limit case, $R1 = \infty$ is considered as stuck-open hard fault.

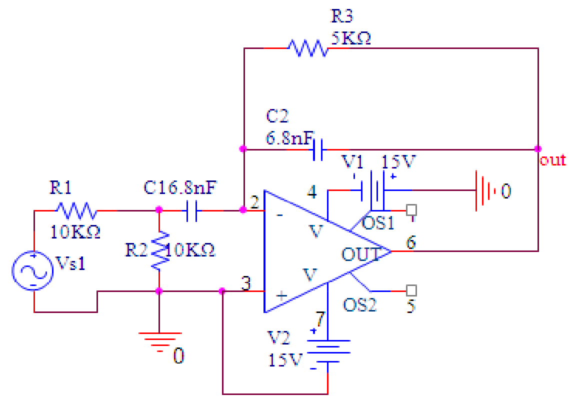

Fig. 1: Band-pass filter

Table 1: Setting of fault pattern

Fault component	Fault mode (out-of-tolerance)	Fault class
	Normal	1
R1	45%	2
R1	-50%	3
R2	50%	4
R2	-40%	5
R3	40%	6
R3	-45%	7
C1	50%	8
C1	-50%	9
C2	45%	10
C2	-50%	11

- When $R1 < 9.5k$, it indicates that R1 overstep the normal lower limit, that is, another soft fault happened; furthermore, limit case, $R1 = 0$ is considered as short circuit hard fault.

After sensitivity analysis using the function in Pspice9.2, the components affected the output mostly would be chosen. The output voltages in seven frequency points (or more points) with obvious difference between the amplitude frequency respond are extracted as the original features of fault modes. Suppose single soft fault is happened each time, so, there are eleven fault modes (include normal condition) in all. Here, the out-of-tolerance range of each component is random. Table 1 shows all the fault modes.

For each fault mode, taking 500 times Monte-Carlo(M-C) analysis by Pspice, then, the original training sample set (20 samples per fault) and original test sample set (10 samples per fault) are obtained.

Table 2: Results comparison of FKPCA and KPCA

	FKPCA			KPCA		
SNE	λ_i	P	AC	λ_i(10-5)	P	AC
1	42.6557	0.7085	0.7085	0.9856	0.8886	0.8886
2	15.7403	0.2614	0.9699	0.1178	0.1062	0.9948
3	1.1536	0.0192	0.9891	0.0050	0.0045	0.9993
4	0.2016	0.0033	0.9924	0.0004	0.00035	0.9996
5	0.1548	0.0026	0.9950	0.0004	0.00035	1.0000
6	0.0879	0.0015	0.9965	0.0000	0.0000	1.0000
7	0.0703	0.0012	0.9977	0.0000	0.0000	1.0000

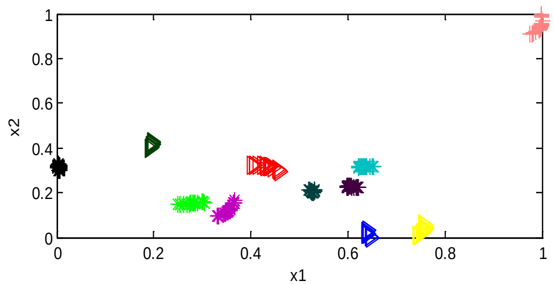

Fig. 2: The principal components projection of FKPCA

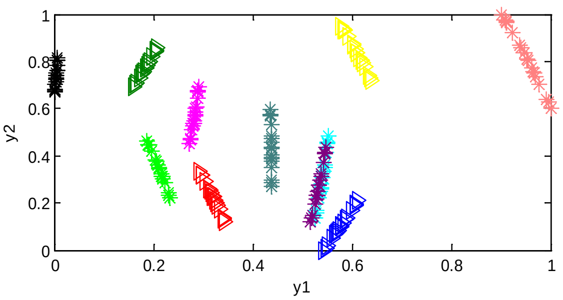

Fig. 3: The principal components projection of KPCA

When the aforementioned process is carried out, then, the fault features can be extracted by proposed KPCA

Result analysis: in the study, Radial Basic Function (RBF) is used as the kernel function and the parameters of RBF are decided by extensive trials. Ultimately, $\sigma_{FKPCA} = 2$ and $\sigma_{KPCA} = 0.25$ are chosen for FKPCA and KPCA respectively. Table 2 shows the simulation results using proposed FKPCA and KPCA with these parameters. Because the numbers of eigenvalues are equal to the numbers of training samples (220-dimension), for purposes of comparison , only the first seven eigenvalues are listed in Table 2, where SNE stands for sequence number of eigenvalues and AC stands for accumulative contribution, $P = \lambda_i / \sum_{i=1}^{n} \lambda_i$.

It can see from Table 2, the accumulative contributions of the first two principle components of FKPCA and KPCA are all overtaken 95%(exceed 85%), so, the first two eigenvectors are enough to construct the new feature subspace, which instead of the old seven ones. In this way, eigenvalue dimensions decrease to 2-dimension, Fig. 2 and 3 show the principle components projection obtained from FKPCA and KPCA respectively.

From Fig. 2 and 3, due to introduction of the class information in FKPCA, the extracted principle components in the same classes are converged and that in the different classes are dispersed, which are helpful for exact classification by classifier, furthermore, helpful for diagnosing various faults. After obtaining the data from the above two methods, they are all sent to BP net with variable rate learning method, RBF net and PNN net respectively for training, the trained net can be used to diagnose the test data. Table 3 shows the results.

Table 3: The results of faults diagnosis

	FKPCA		KPCA	
Diagnosis method	Recogni-tion ratio	Diagnosis time	Recogni-tion ratio	Diagnosis time
BP	100%	11.167s	89.90%	13.875s
RBF	100%	0.109s	96.36%	0.641s
PNN	100%	0.078s	95.45%	0.625s

The simulation results show when use the features extracted from the proposed FKPCA, all networks could diagnose faults perfectly, improve the fault recognizing rate and reduce training and recognizing time. So, through the experiments, it indicates that the proposed FKPCA method could compensate definitely for the deficiency of KPCA method without class information.

CONCLUSION

Aim to the problem of excluding class information in the process of KPCA, a improved KPCA method-FKPCA, is proposed. FKPCA constructs new eigenvector projection space by introducing Fisher linear discriminant function to make the extracted principle components include class information which is helpful for fault diagnosis. At last, the simulation experiment results indicate that the proposed method can make the classification of principle components more better and improve the performance of fault classifier. However, how to introduce the class information quantitatively is the next research work.

REFERENCES

Bian, Z.Q. and X.Z. Zhang, 2009. Pattern Recognition. 2nd Edn., Tsinghua University Press, Beijing.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, New York.

Chin, T.J. and D. Suter, 2007. Incremental kernel principal component analysis. IEEE T. Image Process., 16(6): 1662-1674.

Cho, J.H., J.M. Lee, S.W. Choi, D. Lee and I.B. Lee, 2005. Fault identification for process monitoring using kernel principal component analysis. Chem. Eng. Sci., 60(1): 279-288.

Choi, S.W., C. Lee, J.M. Lee, J.H. Park and I.B. Lee, 2005. Fault detection and identification of nonlinear processes based on kernel PCA. Chemomet Intell. Lab. Syst., 75(1): 55-67.

Scholkopf, B., A. Smola and K.R. Muller, 1997. Kernel principal component analysis. Artificial Neural Networks-ICANN'97, Lect. Notes Comput. Sc., 1327: 583-588.

Scholkopf, B., A. Smola and K.R. Muller, 1998. Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput., 10(5): 1299-1319.

Scholkopf, B., A. Smola and K.R. Muller, 1999. Kernel Principal Component Analysis. In: Advances in Kernel Methods-Support Vector Learning. MIT Press, Cambridge, MA, pp: 327-352.

Wang, X.F., J. Qiu and G.J. Liu, 2006. Research on fault feature extraction based on kernel K-L transformation. Mech. Sci. Technol., 25(3): 288-291.

Xiao, Y.Q. and Y.G. He, 2011. A novel approach for analog fault diagnosis based on neural networks and improved kernel PCA. Neurocomputing, 74: 1102-1115.