2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

The Transformer Fault Diagnosis Combing KPCA with PNN

Chenxi Dai
the School of Electrical Engineering, Southwest Jiaotong University
Chengdu , China
sakuro106@126.com

Zhigang Liu
the School of Electrical Engineering, Southwest Jiaotong University
Chengdu , China
Liuzg_cd@126.com

Yan Cui
the School of Electrical Engineering, Southwest Jiaotong University
Chengdu , China
cy_kuaile123@126.com

***Abstract*—The probabilistic neural network (PNN) can detect the complex relationships and be used to develop its basis for the interpretation of dissolved gas-in-oil data that can identify the fault types. An efficient algorithm known as the kernel principle component analysis (KPCA) is applied to increase features in order to get higher detection accuracy. KPCA reflects the nonlinear or high order features that permit to represent and classify the varying states. More features can be obtained by the nonlinear transformation of KPCA, which can realize the biggest between-class margin of the classifiers. In this paper, we apply the method of combining KPCA with PNN in transformer fault diagnosis. The method has more superior performance than traditional PNN alone method. The property of the nonlinear extension of original data of KPCA can obtain the higher diagnosis accuracy, which can achieve better classification and diagnosis.**

Keywords—Transformer Fault Diagnosis, Dissolved Gas Analysis, Probabilistic Neural Network, Kernel Principle Component Analysis

I. INTRODUCTION

The power transformer is an essential apparatus in power systems and its failure may interrupt power supplies and diminish profits [1]. Minimizing the risk of power outage needs to detect the incipient faults inside power transformers immediately. Efficient detecting methods are necessary as the importance of the incipient fault diagnosis of power transformers. Dissolved gas analysis (DGA) is to obtain the amount of dissolved gases in the oil by sampling and testing the insulation oil of transformers periodically which can indicate the deterioration of the insulating materials inside. These gases are hydrogen (H2), methane (CH4), ethylene (C2H4), ethane (C2H6), acetylene (C2H2), carbon monoxide (CO), and carbon dioxide (CO2) [2-3]. Many approaches are applied to the fault diagnosis of power transformers. Development in artificial intelligence (AI) technique has rapidly improved the transformer fault diagnosis in recent years [4]. Experts system derives the decision rules from the previous experience while the fuzzy-set represents the decision rules by using vague quantities [5-6]. With artificial neural network (ANN), the complex problems can be solved by using the highly nonlinear mapping nature of neural networks.

The calculation of these methods is simple, and they tend to work well in diagnosing severe faults. Their defects show that they could be insensitive to more subtle faults. They are usually used as the general guideline. In an actual diagnostic process, other information such as the variability of dissolved-gas data and the effect of loading and environmental factors on these data is usually also taken into consideration. The single original approach cannot achieve the expected effect, so that some improvement and combination are necessary. Kernel principle component analysis (KPCA) is one of the ways which can effectively increase characteristics to form new classifiers. Combining KPCA with ANN, we can achieve the better effects on transformer fault diagnosis. In this paper, we focus on the application of kernel principle component analysis combined with probabilistic neural network in transformer fault diagnosis.

This paper is organized by 7 sections. In Section II, we give a brief of dissolved gas analysis. In Section III, we introduce the transformation process of KPCA. And in section IV, we review the principle and structure of probabilistic neural network. Then in section V, the process of KPCA combined with PNN is shown. Example verification of the proposed method is in Section VI. Conclusive remarks are in Section VII.

II. DISSOLVED GAS ANALYSIS

The DGA is one of the most common techniques used for incipient fault diagnosis because the transformer cannot be de-energized. The DGA requires routine oil sampling and modern technologies for on-line gas monitoring. The key step in using gas analysis for fault detecting is correctly diagnosing the fault that generated the gases. Abnormal electrical or thermal stresses cause insulation oil breaking down and releasing small quantities of gases. The composition of these gases depends on the fault type. The detection of certain level of gases generated in oil-filled transformer in service is often the first available indication of the transformer failure. Possible mechanisms of gas generation include arcing, corona discharge, low energy sparking, overheating of insulation due to severe overloading, and failure of forced cooling systems.

978-1-4799-1484-5/14/$31.00 ©2014 IEEE

Faults in oil-filled transformers can be identified according to the gases generated, the gases that are typical or predominant at various temperatures. Each fault type produces gases that are generally combustible. An increase in total combustible gases (TCG) that correlates with an increase in gas generating rates may indicate the existence of any one or a combination of thermal, electrical or corona faults.

The three-ratio method is the most common representation in DGA, TABLE I is IEC/IEEE codes for the interpretation of DGA method. In this chart the fault classification of the three-ratio method is showed. The three characteristic values of gases dissolved in transformer oil are taken as input feature vector of the diagnostic model system. Then the proposed diagnosis model has 3 input nodes which represent 3 characteristic values of CH_4/H_2, C_2H_2/C_2H_4 and C_2H_4/C_2H_6. In this paper the fault type is 6 based on our real data.

TABLE I.
IEC/IEEE CODE FOR THE INTERPRETATION OF DGA METHOD

Fault code	Fault type	Ratios of characteristic gases		
		C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
0	no fault	<0.1	0.1-1.0	<0.1
1	<150°C thermal fault	<0.1	0.1-1.0	1.0-3.0
2	150°C -300°C thermal fault	<0.1	>1.0	<0.1
3	300°C -700°C thermal fault	<0.1	>1.0	1.0-3.0
4	2700°C thermal fault	<0.1	>1.0	>3.0
5	low energy partial discharges	<0.1	<0.1	<1.0
6	high energy partial discharges	0.1-3.0	<0.1	<1.0
7	low energy discharges	<0.1	0.1-1.0	>1.0
8	high energy discharges	0.1-3.0	0.1-1.0	>3.0

III. KERNEL PRINCIPAL COMPONENT ANALYSIS

A. Principal Component Analysis

Principal component analysis (PCA) is a multivariable statistical method that can be used for damage detection of structures or fault diagnosis in mechanical systems [7]. It is known as an efficient method to compress large sets of the random variables and to extract interesting features from a dynamical system. However, this method is based on the assumption of linearity [8]. To some extent, many systems show a certain degree of nonlinearity, therefore, the detection necessitates methods that are able to study nonlinear systems. Kernel principal component analysis (KPCA) is what we need to solve the proposed problem mentioned above.

B. Kenel Principal Component Analysis

Kernel principal component analysis (KPCA) is a nonlinear extension of PCA built to authorize features with nonlinear dependence between variables [9]. The method is ‘flexible’ in the sense that different kernel functions may be used to better fit the testing data.

The key idea of KPCA is first to define a nonlinear map $x_k \mapsto \phi(x_k)$ with $x_k \in R^n$, (k=1, $\cdots$, M) which represents a high dimensional feature space F, and then to apply PCA to the data in space F.

With the assumption of centered data, i.e. a, the covariance matrix in the space F is given by:

$$C = \frac{1}{M}\sum_{i=1}^{M}\phi(x_i)\phi(x_i)^T . \quad (1)$$

Principal components may be next extracted by solving the eigenvalue equation:

$$\lambda V = CV . \quad (2)$$

By defining the kernel matrix K of dimension $M \times M$ such that:

$$M\lambda\alpha = K\alpha \quad (3)$$

where α identifies the eigenvector V after normalization.

The eigenvectors identified in the feature space F can be considered as kernel principal components (KPCs) which characterize the dynamical system. Note that, since the number of eigenvectors (i.e. nonlinear PCs) is the same as the number of samples, which is higher than the number of (linear) PCs given by PCA. The KPCA method is termed ‘‘nonlinear’’ since the feature mapping in space F is achieved by a nonlinear function. Because of this property, the extracted KPCs are able to reflect nonlinear or high order features which permits representation and classification varied states. KPCA has the aptitude to use more nonlinear PCs to collect structural features than noise.

Regarding the kernel functions, they can be chosen for instance as follows:

- polynomial kernel,

  $K(x_i, x_j) = (x_i \bullet x_j + 1)^d$, where d is a positive integer; (4)

- radial basis function(RBF),

$$K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2) \quad (5)$$

where $2\sigma^2 = \omega$ is the width of the Gaussian kernel.

It is worth noting that in general, the above kernel functions give similar results if appropriate parameters are chosen. The radial basis function may present advantages owing to its flexibility in choosing the associated parameter. For instance, the width of the Gaussian kernel can be very small (<1) or quite large. In contrast the polynomial kernel requires a positive integer for the exponent.

IV. ARTIFICIAL NEURAL NETWORK ALGORITHM

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

This work was supported by National Nature Science Foundation of China(U1134205, 51377136).

A. Artificial Neural Network

ANN knowledge is discreetly distributed all over the network, based on the sample learning, and not stored in a knowledge base. Obviously, ANN has a great capacity for obtaining knowledge. Very complex systems can be characterized with very little explicit knowledge using ANNs. The relationship between gas composition and incipient-fault condition is learned by the ANN from actual experience (through training samples) [10]. Obvious and not so obvious (hidden) relationships are detected by the ANN and used to develop its basis for interpretation of dissolved gas-in-oil data. Through training process, ANN can reveal complex mechanism that may be unknown to experts. Theoretically, a neural network could represent any observable phenomenon.

An ANN design includes selection of input, output, network topology (structure, or arrangement of nodes), and weighted connections of the nodes. Input feature (information) selection constitutes an essential first step. The feature space needs to be chosen very carefully to ensure that the input features will correctly reflect the characteristics of the problem. Another major task of the ANN design is to choose network topology. This can be done experimentally through a repeated process to optimize the number of hidden layers. Figure 1 illustrates an overall ANN design process with step by step adjustments to achieve desired structure and feature space. The corresponding connection weights are also determined in the process. Once the process is done, all weights will be retained and the ANN is now "trained" and ready for use. New samples will be fed into the trained ANN and predicted values can be readily obtained as shown in Figure 1.

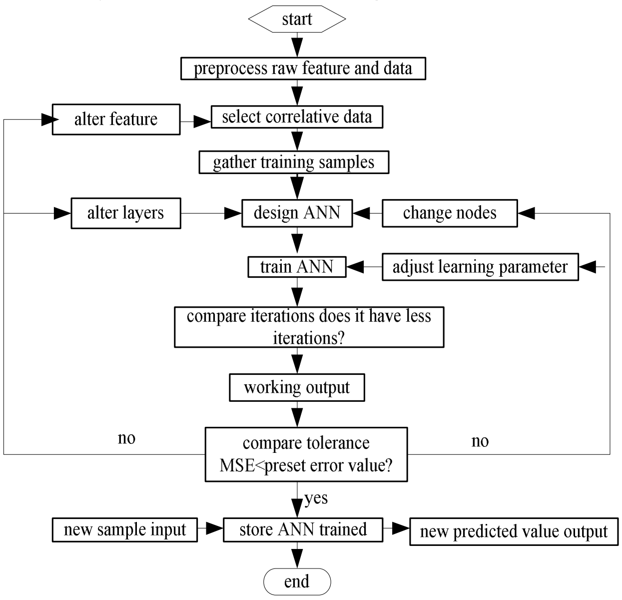


Fig.1. ANN Design and Training Flow Chart

B. Probabilistic Neural Network

The PNN is a type of radial basis network originally developed for radar classification. It was first proposed by Specht in 1989 [11-12]. Based on the statistical theory it equals to the well-known Bayesian strategy in classification. The essence of PNN is a parallel algorithm developed from the Bayesian minimum risk criteria. Compared with traditional BP neural network, PNN does not need to set the number of hidden layer neurons, takes less time in training and does not need to train again when more training data are added or reduced. Moreover, it can always obtain the optimal solution in Bayesian criterion when the training data are sufficient, regardless of the complexity of the classification problem [13]. The PNN network is simply a parallel 4-layer structure: input layer, pattern layer, summation layer and output layer.

The input layer receives and normalizes input vector, its input neurons equal with dimension of the sample vector. Each unit in pattern layer calculates the input feature vector and training focused on the match between the models. The pattern layer neurons are equal to the sum of each category training samples. The transfer function of the pattern layer is Gaussian function, and then the output of every pattern layer can be expressed as:

$$f(X,W_i) = \exp[-\frac{(X-W_i)^T(X-W_i)}{2\delta^2}] \quad (6)$$

Where W_i is the connection weight of the input layer and model layer, and δ is the smoothing parameter.

Summation layer computes the summation of each pattern and multiplies the loss factor. Each neuron of the output layer respectively corresponds to a data type which is failure mode. The output layer calculate the probability of each mode appearing, the largest output of probability density function of the neuron which correspond to the kind of recognition model of sample category is 1, all other neurons output is 0.

V. METHOD BASED ON KPCA AND PNN

Dissolved gas analysis (DGA) can only extract five characteristics that can't satisfy forming enough classifiers to diagnose transformer faults accurately. In addition, the property of classifier and the difference between classifiers is the dominant factor for classifier to recognize fault which also needs more character. Suppose there are X training sample set divided into L fault classes, each fault class X_i has M_i samples, the total number of samples is $M=\sum_{i=1}^{L} M_i$. Using KPCA to analyze X, we can get increased features by the nonlinear transformation. Considering the different property of different fault classes, use KPCA to analyze different fault classes respectively, then we can get transformation matrix T_i, $i=1,2,\dots,L$. The number of analyzed Kernel features is equal to the number of samples in theory. If we select two features from each class to construct a classifier randomly, different combinations of $\prod_{i=1}^{L} C_{M_i}^2$ features can be obtained.

In this diagnosis model, we use PNN to train and analyze formed classifier group, then we can get the final result of diagnosis in the form of probability. Calculation steps are as follows, shown in Fig.2:

Step 1: obtain the dissolved gases in the oil and adopt DGA to the classification of fault diagnosis.

Step 2: Normalization processing for the dissolved gases data.

Step 3: adopt KPCA to achieve nonlinear transform to get more kernel features which is in high order.

Step 4: For each state of operation of the transformer, we built 5 classifiers. Each classifier is trained randomly using 2 features of those generated after KPCA.

Step 5: adopt PNN to train and analyze the data which is the outputs of the classifier group above.

Step 6: obtain the final result of diagnosis.

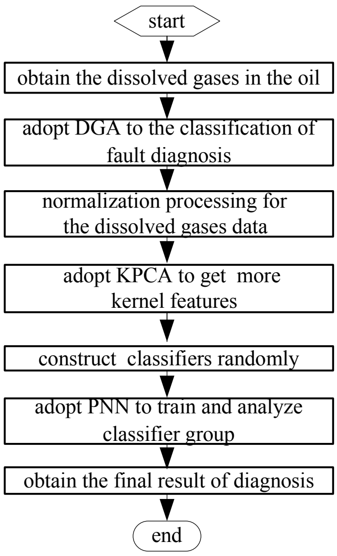


Fig.2. Method of KPCA and PNN Flow Chart

VI. EXAMPLE VERIFICATION

A. Modeling

Failure types can be divided into 6 categories including normal state, hyperthermia, high energy, hypothermia, low energy and mid-temperature, represented by F1, F2,⋯, F6. The samples used for training are shown in TABLE II.

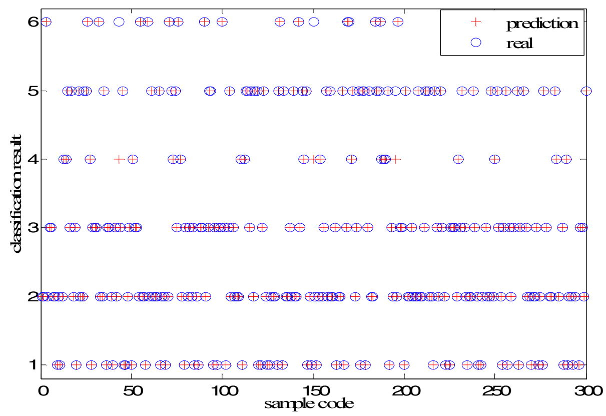
Select $h(x,y)=(x\bullet y+1)^d$ as a kernel function in kernel principal component analysis, d is the default function parameters. Taking all aspects into account, the value of d is 3. The training sample set after the high dimensional transformation is also in TABLE II. Comparing the dimension of characteristics before and after the transformation we can see the effect of dimension extension is obvious. Select two kernel features randomly from each fault class subset to construct a PNN classifier, adding up to 5 classifiers. The amount of kernel features we can choose is larger than classifiers that can ensure every classifier constructed by different kernel features.

TABLE II. Training Sample Set

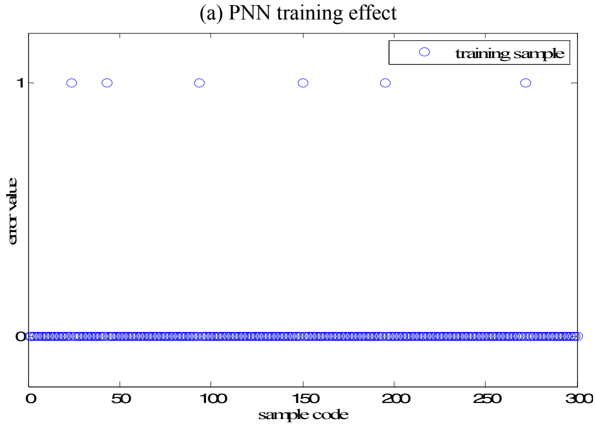
Failure type	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**
Number of the samples	58	111	71	21	61	29
The dimension of original characteristics	5	5	5	5	5	5
The dimension of l characteristics after KPCA	9	9	9	7	7	9

B. Simulation

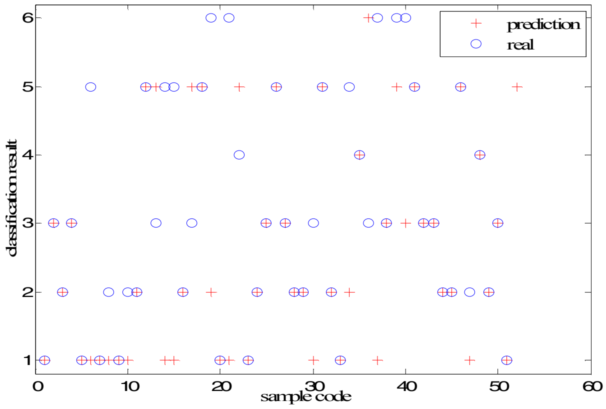
The number of total samples is 351. We choose 300 samples as the training sample, and the rest is as the test sample. The blue circle in the figure represents the original real samples and the red plus sign represents the prediction samples after training. Training, test and error resultsobtained by using PNN alone are shown in Figure 3. Combining KPCA with PNN, better results can be obtained correspondingly in Figure.4. In Figure 3(a), the meaning of the horizontal axis is training sample code and the meaning of the vertical axis is classification result of fault. In Figure 3(b), the meaning of the horizontal axis is training sample code and the meaning of the vertical axis is the prediction error. The meaning of zero of the vertical axis represents the accurate diagnosis and one represents the false diagnosis in Figure 3(b, d). Figure 3(a, b) show the training property while Figure 3(c, d) shows the testing property. Figure 4 is as the same.



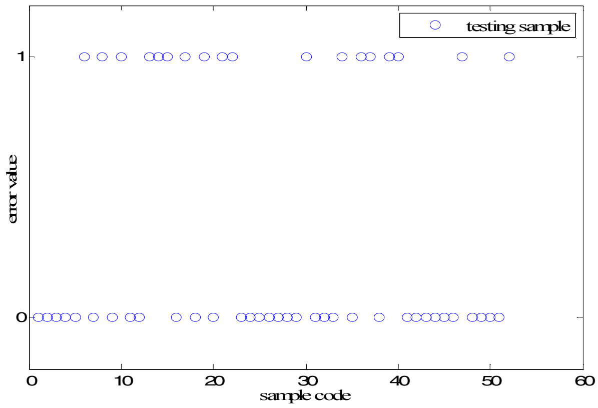
(a) PNN training effect



(b) PNN training error



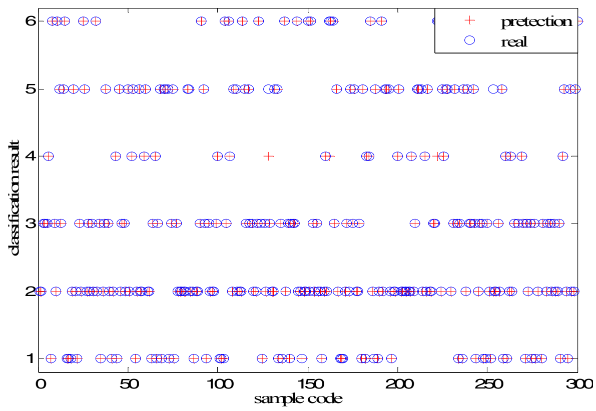
(c) PNN prediction effect



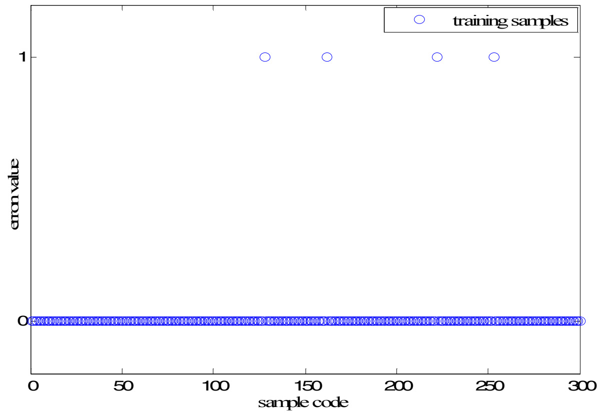
(d) PNN prediction error

Fig.3. PNN training and test in transformer fault diagnosis

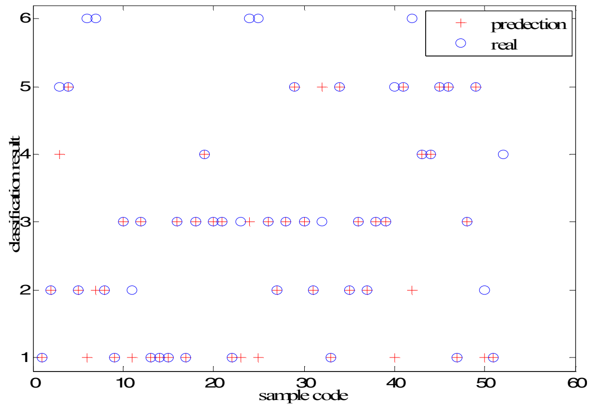
Figure 3 and Figure 4 show that the two methods can achieve transformer fault diagnosis. Their training effect and the training accuracy are well. However, the prediction accuracy of KPCA combined with PNN is higher than PNN alone method shown in TABLE III. The higher diagnosis accuracy is obtained since KPCA has exacted more features that can form more classifiers. Based on the above comparison the result can be obtained that the PNN combined KPCA method is more effective in transformer fault diagnosis.



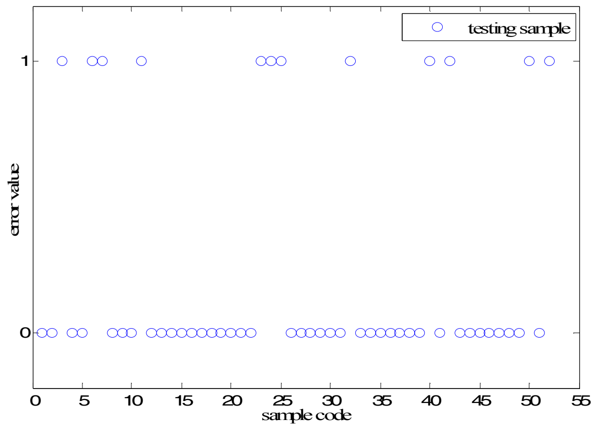
(a)PNN training effect after KPCA



(b) PNN training error after KPCA



(c) PNN prediction effect after KPCA



(d) PNN prediction error after KPCA

Fig.4. PNN training and test after KPCA in transformer fault diagnosis

TABLE III. DIAGNOSIS PREDICTION ACCURACY

Method	**KPCA combined with PNN**	**PNN**
Accuracy	0.7692	0.6538

C. Verification

The gas relay of a transformer (OD-12600/56) actuated in service and the circuit breaker switched off and caused a supply interruption. The DGA data after this fault is shown in TABLE IV.

TABLE IV. COMPONENTS OF DISSOLVED GAS IN A TRANSFORMER

Gas type	**H_2**	**CH_4**	**C_2H_6**	**C_2H_4**	**C_2H_2**
Value	833	3167	390	5793	1697

What can we obtain by the chart above is that the fault type of this transformer is F2 with the method KPCA combining with PNN. It means that the temperature of the transformer is too high. The result is the same with the inspection result of hanging core. KPCA combing with PNN really work well in transformer fault diagnosis by the example of verification.

VII. Conclusion

KPCA is a nonlinear extension of original data that maps data into high-dimensional space. Any dimension can be obtained by a nonlinear function. KPCA can help to get more characteristics to form enough classifiers to increase the fault diagnosis accuracy of PNN, and is valuable for the practical application under current requirement. Through the simulation results above, the conclusion is easy to obtain. The prediction accuracy of KPCA combined with PNN is higher than using PNN alone. This method is really suitable for transformer fault diagnosis. However, studies need to be done for further development and improvement, in order to get better results.

REFERENCES

[1] Huo-Ching Sun, Yann-Chang Huang, Chao-Ming Huang. "A Review of Dissolved Gas Analysis in Power Transformers," Energy Procedia 14 (2012) 1220-1225.

[2] C. E. Lin, J. M. Ling, C. L. Huang. "An Expert System for Transformer Fault Diagnosis Using Dissolved Gas Analysis," IEEE Transactions on Power Delivery, vol. 8, no. 1, January 1993.

[3] Diego Roberto Morais, Jacqueline Gisèle Rolim. "A Hybrid Tool for Detection of Incipient Faults in Transformers Based on the Dissolved Gas Analysis of Insulating Oil," IEEE Transactions on Power Delivery, vol. 21, no. 2, April 2006.

[4] P. Purkait, S. Chakravorti. "Time and Frequency Domain Analyses Based Expert System for Impulse Fault Diagnosis in Transformers," IEEE Transactions on Dielectrics and Electrical Insulation vol. 9, no. 3, June 2002.

[5] Q. Su, C. Mi, L. L. Lai, et al. "A Fuzzy Dissolved Gas Analysis Method for the Diagnosis of Multiple Incipient Faults in a Transformer," IEEE Transactions on Power Systems, vol. 15, no. 2, MAY 2000.

[6] W. Xu, D.Wang, Z.Zhou. "Fault Diagnosis of Power Transformers: Application of Fuzzy Set Theory, Expert Systems and Artificial Neural Networks," IEE Proceedings Science, Measurement and Technology, vol. 144, no. 1, January 1997.

[7] Richard P. Good, Daniel Kost, and Gregory A. Cherry. "Introducing a Unified PCA Algorithm for Model Size Reduction," IEEE Transactions on Semiconductor Manufacturing, vol. 23, no. 2, May 2010.

[8] Mingtao Ding, Zheng Tian, Haixia Xu. "Adaptive Kernel Principal Component Analysis," Signal Processing 90 (2010) 1542–1553.

[9] Paul Honeine. "Online Kernel Principal Component Analysis: a Reduced-Order Model," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 9, September 2012.

[10] Fábio R. Barbosa, Otacílio M. Almeida, Arthur P.S.Braga, et al. "Application of an Artificial Neural Network in the Use of Physicochemical Properties as a Low Cost Proxy of Power Transformers DGA Data," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 19, no. 1, February 2012.

[11] Song L, Xiu-ying L, Wen-xu W. Fault diagnosis of transformer based on probabilistic neural network[C]//Intelligent Computation Technology and Automation (ICICTA), 2011 International Conference on. IEEE, 2011, 1: 128-131.

[12] N. Perera, A. D. Rajapakse. "Recognition of Fault Transients Using a Probabilistic Neural-Network Classifier," IEEE Transactions on Power Delivery, vol. 26, no. 1, January 2011.

[13] D. Evagorou, A. Kyprianou, P.L. Lewin, et al. "Feature Extraction of Partial Discharge Signals Using the Wavelet Packet Transform and Classification with a Probabilistic Neural Network," IET Science, Measurement and Technology, vol. 4, no. 3, pp. 177–192, 2010.