Intelligent Fault Diagnosis of Engine Based on PCA-SOM

Yangyang Zhang¹, Yunxian Jia†, Chiming Guo¹ and Weiyi Wu¹

¹Army Engineering University, Shijiazhuang, Hebei, 050003, China
†Corresponding author’s e-mail: 2108007979@qq.com

Abstract. Self-Organizing Feature Map (SOM) is a kind of self-organizing and self-learning network without teacher. It is mainly used for pattern recognition and region classification of input vectors. A fault diagnosis method of engine fuel supply system based on SOM neural network is proposed. The sensor is used to monitor the fuel pressure waveform of a certain engine fuel supply system, time domain analysis and feature extraction are carried out on the waveform, and the feature dimension reduction is realized by PCA to form the input vector of SOM neural network. The SOM neural network is used to establish the diagnosis model and then recognize fault patterns for test samples. The results of pattern recognition show that SOM neural network can identify and classify faults accurately, and it has certain engineering application value.

1. Introduction
With the development of science and technology, the structure of mechanical equipment is becoming more and more complex, and the probability of failure occurring in the working process is relatively increased. The mechanical composition of vehicle engines is more complex and can easily cause some unnoticeable faults in a long time. If these faults can be found or dealt with in time, it will not only ensure personal safety, but also avoid unnecessary losses[1]. The emergence of neural network provides a new way to solve the problem of fault diagnosis. Among them, Self-Organizing Feature Map (SOM) has a high degree of self-organization and self-learning ability, and has become an effective way of fault diagnosis[2-6].

Fuel supply system is an important part of engine. In order to realize intelligent fault diagnosis, a fault diagnosis method of engine fuel supply system based on SOM neural network is proposed. The fuel pressure waveform of the fuel supply system is monitored by sensors, the state characteristics are extracted, and the feature dimension is reduced by Principal Component Analysis (PCA), which is used as the input vector of SOM neural network. The experimental and simulation results show that the proposed method can effectively identify the fault of engine fuel supply system.

2. Theoretical method

2.1. Principal Component Analysis
The core of Principal Component Analysis (PCA) is to make the original index into a new set of independent comprehensive indexes which contain most of the original information through linear combination, so as to achieve the goal of dimensionality reduction. On the premise of ensuring enough information of the original data, the original multi-dimensional index variable data can be dimensionally reduced, the number of indicators can be compressed, the data can be simplified, and the time consumed in calculation can be reduced[7]. The model of this method is as follows:

...
Assuming that the value of the first $j$ index of the first $i$ sample in $m$ samples and $n$ evaluation indexes is $x_{ij}$, the original evaluation index data matrix is constructed:

$$
X = \begin{bmatrix}
  x_{11} & x_{12} & \cdots & x_{1n} \\
  x_{21} & x_{22} & \cdots & x_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{m1} & x_{m2} & \cdots & x_{mn}
\end{bmatrix}
$$  \hspace{2cm} (1)

Let the original variable index is $x_1, x_2, \ldots, x_n$. After dimension reduction, the comprehensive index, i.e. the principal component is $y_1, y_2, \ldots, y_n$.

$$
\begin{align*}
  y_1 &= l_{11}x_1 + l_{12}x_2 + \cdots + l_{1n}x_n \\
  y_2 &= l_{21}x_1 + l_{22}x_2 + \cdots + l_{2n}x_n \\
  \vdots \\
  y_n &= l_{n1}x_1 + l_{n2}x_2 + \cdots + l_{nn}x_n
\end{align*}
$$  \hspace{2cm} (2)

The above formula has the following characteristics: (1) The sum of squares of coefficients of each equation is 1, that is: $l_{1j}^2 + l_{2j}^2 + \cdots + l_{nj}^2 = 1 \ (j = 1, 2, \ldots, n)$.

(2) Any two principal components are independent and independent of each other.

(3) $y_1$ is the largest variance of all linear combinations of $x_1, x_2, \ldots, x_n$.

The steps to solve the principal components from the correlation matrix are as follows:

Step1: Standardizing raw data.

$$
x_{ij}^* = \frac{(x_{ij} - \bar{x}_j)}{s_j} \ (i = 1, 2, \ldots, m ; j = 1, 2, \ldots, n)
$$  \hspace{2cm} (3)

Step2: Computing Pearson Coefficient Matrix between Indicators, $R = (r_{kl})_{mn}$ $(k, l = 1, 2, \ldots, n)$

$$
r_{kl} = \frac{\sum_{i=1}^{m} (x_{ik} - \bar{x}_k)(x_{lj} - \bar{x}_l)}{\sqrt{\sum_{i=1}^{m} (x_{ik} - \bar{x}_k)^2 \sum_{i=1}^{m} (x_{lj} - \bar{x}_l)^2}}
$$  \hspace{2cm} (4)

Step3: Computing eigenvalues and eigenvectors of correlation matrix A. $\lambda_1, \lambda_2, \ldots, \lambda_n$. The unit eigenvector corresponding to the eigenvalue is denoted as $p_1, p_2, \ldots, p_n$

Step4: Determine the number of principal components. For calculating the cumulative contribution rate of principal components, the former principal components corresponding to eigenvalue greater than 1 and cumulative contribution rate of 85%-95% are generally taken. The variance contribution rate of the first principal component is $v_s = \lambda_s / \sum_{s=1}^{n} \lambda_s, s = 1, 2, \ldots, n$.

Step5: Calculate the corresponding score of extracted principal components. The principal component coefficient matrix is: $U = (p_1, p_2, \ldots, p_n)$. If the first n principal components are extracted from the original index, there are: $y_s = X^* p_s = (x_1^*, x_2^*, \ldots, x_n^*) p_s \ (s = 1, 2, \ldots, k)$
2.2. **Self-Organizing Feature Map**

Self organizing feature map network, also known as Kohonen network, was proposed by Teuvo Kohonen in 1981. The network is an unsupervised, self-organizing, self-learning network composed of fully connected neuron arrays[8]. The SOM network structure is shown in Figure 1. The main goal of SOM is to perform a non-linear transformation adaptively, or to map data into two-dimensional space. The steps are as follows:

![Structure Diagram of SOM Neural Network](image)

Figure 1. Structure Diagram of SOM Neural Network

1. **Network initialization.** The connection weights of N input neurons to output neurons are given smaller weights, generally between (0,1).
2. The input vector \( X = (x_1, x_2, \ldots, x_m)^T \) is input to the input layer.
3. In the mapping layer, the Euclidean distance between the weights of each neuron and the input vector is calculated.
   \[
   d_j = \|X - W_j\| = \sqrt{\sum_{i=1}^{n} (x_i(t) - w_{ij}(t))^2}, \quad d_k = \min_j(d_j) \tag{5}
   \]
   By calculating, a neuron with the smallest distance is obtained, which is called the winning neuron, and its adjacent neuron set is given.
4. **Modifying the weights of output neurons and their adjacent neurons**
   \[
   \Delta w_{ij} = w_{ij}(t+1) - w_{ij}(t) = \eta(t)(x_i(t) - w_{ij}(t)) \tag{6}
   \]
5. **Calculate the output value:** \( \alpha_k = f(\min_j \|X - W_j\|) \)
6. If the requirement is met, the algorithm ends; otherwise, step (2) is returned to enter the next round of learning.

3. **Engine Fault Diagnosis**

The common faults of the fuel supply system of an engine are simulated on the test bench with a fuel injection pump. The main faults are 100% fuel supply (normal state \( F_1 \)), 75% fuel supply \( (F_2) \), 25% fuel supply \( (F_3) \), idle fuel \( (F_4) \), pin valve stuck (small amount of oil \( F_5 \) ), pin valve stuck (calibrated amount of oil \( F_6 \) ), pin valve leakage \( (F_7) \), and oil outlet valve failure \( (F_8) \).
The fuel pressure waveform of the engine is monitored by sensors, as shown in Figure 2. Among the main parameters of the waveform, the maximum pressure \( P_1 \), the secondary maximum pressure \( P_2 \), the waveform amplitude \( P_3 \), the rising edge width \( P_4 \), the waveform width \( P_5 \), the maximum afterwave width \( P_6 \), the waveform area \( P_7 \), and the injection pressure \( P_8 \) are selected as the characteristic parameters reflecting the operation of the engine.

![Figure 2. Fuel pressure waveform](image)

After sampling and obtaining the data of sampling points, the dimensionality reduction is realized by principal component analysis, and then normalization method is used to standardize the data, which can improve the accuracy of fault diagnosis and finally obtain the fault samples. The SOM network with output layer is established, and the input fault samples are trained. When the training step is 200, each sample is divided into one class, which indicates that the training has been completed, and the fault diagnosis model based on SOM network has been established. Then the fault diagnosis of the test samples is carried out. The results are shown in Figure 3 below. The diagnostic results show that the above model can effectively identify the fault of the engine fuel supply system.

![Figure 3. Classification of neurons](image)
4. Conclusions
The intelligent fault diagnosis method of engine fuel supply system based on the combination of principal component analysis and SOM neural network is studied. The proposed method is verified by engine fault experiment. The results show that: (1) principal component analysis can effectively extract the main features of fuel pressure waveform, reduce the dimension of sample data, and help to improve the accuracy of fault diagnosis; (2) SOM neural network. The network can accurately identify the faults of the engine fuel supply system, which has certain engineering application value.

Acknowledgments
This work was supported by National Natural Science Foundation of China (71871220),

References
[1] Zhao, B.Y., Zhao, Y.L., Li, C.L. (2018) Research on fault diagnosis of automobile diesel engine based on rough set. Journal of Liaoning University of Technology, 38 (03):187-191.
[2] Gao, Y. (2010) Fault diagnosis of wind power electronic devices based on SOM neural network. Journal of power systems and automation, 22 (03):142-145.
[3] Zhang, Q.Y., Kou, X.Z. (2009) Research on Intelligent Diagnosis of Internal Combustion Engine Faults Based on SOM Neural Network. Small Internal Combustion Engine and Motorcycle, 38 (04): 49-51.
[4] Wang, S.T., Ba, Y.L., Guo, Z.B. (2015) Engine fault diagnosis based on SOM neural network. Automation and Instrumentation, (08): 52-54+57.
[5] Xu, Y.W., Liu, M.M., Wei, et al. (2019) Intelligent diagnosis of Thin-walled bearing fault of robot based on information fusion. Optical Precision Engineering. 27 (07):1577-1592.
[6] Cao, H.L., Huang, L.T., Li, Z.W., et al. (2014) Health assessment of aero engine lubricating oil system based on SOM neural network. Journal of Civil Aviation University of China, 32 (06):19-22.
[7] Guo, L.J., Zhao, M.J., Kang, J.X. (2016) Fault monitoring and diagnosis of petrochemical plant based on PCA and RBF neural network. Journal of Yanshan University. 40 (05): 456-461.
[8] Ye, J.F., Wang, H.M. (2018) AdaBoost detection combined with SOM’s automatic face recognition method. Journal of Harbin Engineering University. 39 (01):129-134.