A Fault Diagnosis Model Based On Full Vector Spectrum And Feature Engineering

Huanting Cui¹, Hong Chen¹,², Sa Xiao¹

¹ Institute of Vibration Engineering Zhengzhou University, Zhengzhou 450001, China
² Zhengzhou Expert Technology Co., Ltd., Zhengzhou 450053, China

cuihuanting@buaa.edu.cn

Abstract. Aiming at the problem of how to quickly and accurately diagnose different types of rotating machinery faults, this paper starts with the combination of full-vector spectrum theory and feature engineering. A model of rotor imbalance faults diagnosis and identifying is constructed in the process of dealing with the data sets of multiple rotor imbalance fault. The result shows that this fault diagnosis model combined with the full vector spectrum theory and feature engineering can be used to correctly distinguish the faults of the equipment and accurately identify the types of the unbalanced faults.

1. Introduction

Rotating machinery is widely used in industrial production and daily life. Large rotating machinery has many kinds of motors, fans, aeroengines, centrifuges, steam turbines and so on. Large rotating machinery will cause huge economic losses once it fails.[1] It even endangers the safety of people's lives and property. More than 90% of the malfunctions come from unbalanced faults and bearing failures. Different types of rotating machines cause different causes of faults.[2] When diagnosing faults, we should pay attention to the investigation. If we fail to find the cause of the faults accurately, we will not be able to eliminate them in time.

2. Methods

The theoretical basis of full vector spectrum analysis is that the whirling phenomenon of the rotor is composed of harmonic frequencies. The whirling intensity of the rotor at each harmonic frequency is the basic basis for judging and identifying the faults.[1] The vortex trajectory is a series of ellipses, each ellipse is composed of the precession circle and the anti-precession circle with the same frequency and the opposite direction of movement. The long axes and half axes of ellipse are defined as the main vibration vectors and the secondary vibration vectors under this harmonic.

Assume \( x_n \) and \( y_n \) (\( n = 0, 1, 2, \cdots, N - 1 \)) They are signals of a certain frequency doubling harmonic. \( x, y \). The discrete sequence in the axis direction constitutes a complex sequence \( Z_n = x_n + jy_n \). It is obtained by Fourier transform \( Z_k = R_k + jI_k \). By Fourier transform, the following results can be obtained as (1).
The change of rotational energy is related to the rotation change of elliptical trajectory, so it can be represented by the rotation area. The specific expression of full vector power spectrum based on the fusion area is as (2).

\[ E_{An}(f) = \frac{2}{T} |Z_n(f, T)|^2 = \frac{N\pi}{8} R_{Ln} R_{Sn} \]  

On the basis of full vector spectrum transformation, combined with the detailed analysis of polyphonic system, the details of local signals can be amplified effectively, and the local fine structure of the spectrum can be observed. As shown in Figure 1, the process of polyppectral refinement of full vector spectrum is shown.

In machine learning, feature engineering is a central task of data preparation. Feature processing is the core part of Feature Engineering, including data preprocessing, quantitative transformation, data dimensionality reduction and so on. Feature selection constructs suitable feature groups from given features to improve the performance of prediction.

Different features are needed when diagnosing different faults. There are qualitative descriptions in fault diagnosis and non-signal characteristics. When performing Feature Engineering, this qualitative description needs to be converted into quantitative figures as input. Discrete assignment is used in this paper. Modern neural network is a nonlinear statistical data modeling tool, which is used to solve
problems that are difficult to be solved by traditional rule-based programming, such as machine vision, speech recognition, Natural Language Processing, etc. the general neural network is composed of input layer, hidden layer and output layer, and the parameters are trained by backpropagation algorithm. In order to make the data model meet the expected accuracy, the hidden layer and output layer consist of many neurons.\cite{4} In this paper, a standard back propagation neural network is used as the training model.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{structure.png}
\caption{Structure of neuron and BP neural networks}
\end{figure}

3. Experiments and examples

3.1. Diagnostic process and features of imbalanced faults

The filtered relevant features of the rotor imbalance fault\cite{2} and diagnostic process are as follows.

\begin{table}
\centering
\caption{The features of imbalanced faults.}
\begin{tabular}{|l|l|}
\hline
Feature & Feature space \\
\hline
Peak-to-peak & high 1, normal 0 \\
Kurtosis factor & high 1, normal 0 \\
Proportion of 1X & high 1, normal 0 \\
Ratio of horizontal 1X to vertical 1X & high 1, normal 0 \\
Number of rotors & one (0,1) two (1,0) three or more (1,1) \\
Axial vibration amplitude & normal 0, obvious 1 \\
Similarity between signal and sine wave & normal 0, obvious 1 \\
Vibration stability & steady (0,0), increasing(1,0), steady after increasing(0,1) \\
Main vibration direction & radial direction (0,1), axial direction (1,0) \\
Precession direction & precession (0,1), back precession (1,0) \\
Variation of vibration with speed & normal (0,0), linear (0,1), nonlinear (1,0) \\
Variation of vibration with oil temperature & normal 0, obvious 1 \\
Variation of vibration with medium & normal 0, obvious 1 \\
Variation of vibration with flow & normal 0, obvious 1 \\
Variation of vibration with pressure & normal 0, obvious 1 \\
Variation of vibration with load & normal 0, obvious 1 \\
Types of rotating machinery & electric motor, fan, aircraft engine, centrifuge, steam turbine \\
\hline
\end{tabular}
\end{table}
3.2. Experiment and analysis
In this paper, the application of the model is illustrated by an example of imbalance fault in rotating machinery. Data are collected by the WTDS experimental platform of Vibration Engineering Institute of Zhengzhou University.

First, the double channel vibration signal is filtered to get the time-frequency characteristics of the vibration signal. Then, the full vector Hilbert method is used to process the vibration data, and the main vibration vector is obtained.

Figure 4 Diagnostic process of rotor imbalanced fault.

Figure 5 Structure of vibration bench for testing.

Figure 6 Filtered double channal vibration signal.
According to the experts' opinion, when the percentage of 1X frequency in the full vector energy spectrum is higher than 60%, it is normal to record high and below 60%.[6] BP neural network is used to build training model, with 17 inputs, 12 middle layers, and 3 outputs. It represents No Fault (NF), Original Imbalance (OI), Sudden Imbalance (SI), Progressive Imbalance (PI), Other Faults (OF).[6] 400 sets of feature data were trained, and 100 sets of data were used as test sets. The results are as follows.

According to the experimental results, large number of other faults are mistaken for imbalanced faults. After analysis, the model cannot distinguish the misalignment faults from other imbalanced faults belonging to other faults. Through practical investigation and experts' recommendation, adding 2 times frequency characteristics into the model can effectively distinguish unbalanced and misaligned faults. The input of 18 neural network with 2 frequency doubling characteristics is added. According to the experts' suggestion, when the proportion of 2X frequency in the full vector energy spectrum is higher than 30%, it is normal to record it as high as 30%.[7] When training is done on the basis of the original parameters, a new neural network model is obtained. The recognition rate of the model for unbalanced fault is 73.3%. In order to contrast, 10 characteristics of the same time and frequency domain characteristics of the same data are taken as input, and the BP neural network model is established, using the same 500 samples to compose data sets. The results are as follows.

The results show that the diagnostic accuracy of the full vector energy spectrum is obviously higher than that of the pure time frequency domain feature analysis network, because the full vector energy spectrum combines the information of the 2 channels and combines with other fault characteristics besides the vibration signal. To sum up, the diagnostic method combining digital quantization fault
feature with full vector energy spectrum is superior to simple signal analysis in accuracy, and the speed of operation is much higher than that of the original signal input network without feature extraction.

![Figure 9](image)

**Figure 9** Confusion matrix of improved model and comparison model

### 4. Conclusions and Prospects

Starting from the engineering practice of rotor unbalance fault, combining the full vector energy spectrum technology with feature engineering, a new method of rotor imbalance fault recognition is put forward by adding features to make the model more accurate. The full vector energy spectrum information fusion and the effect of vibration signals are fully displayed, and compared with the model without adding full vector energy spectrum. The concrete conclusions are as follows.

1. Put forward a fault diagnosis method of rotating machinery combining digital quantization fault feature and full vector spectrum technology, digitizing fault characteristics after expert analysis and combining with full vector spectrum information, and input neural network model to learn, while retaining the source of multiple sources of artificial fault diagnosis.
2. Compared with the method of fault diagnosis based on vibration signal only, this method selects more reasonable fault features, and the information source is more diverse and the diagnosis reliability is stronger. At the same time, the fault is diagnosed by neural network model, which reduces the expert experience needed in diagnosis, and has obvious advantages in diagnosis efficiency.
3. The advantages and disadvantages of the digital quantitative fault feature proposed in this paper and the fault diagnosis method of rotating machinery combined with full vector spectrum technology are very obvious. That is to say, the fault related features can be added or modified at any time, and the network can be continuously optimized to improve the accuracy of diagnosis.

### References

[1] Han Jie, Shi Lai De. Full vector spectrum technology and engineering application [M]. *Beijing: Machinery Industry Press*, 2008.

[2] Zhao Liangliang, Xia Junchao, Zeng Sheng. Extraction and processing of unbalanced signal of small motor rotor [J]. Chemical Machinery, 2007 (01): 8-11+29.

[3] Li Yitao, Lei Mei Hua, Tang Shaozhang. Research on motor rotor imbalance detection based on neural network [J]. mechanical research and application, 2015,28 (05): 184-187.

[4] Zhou Zhihua. Machine learning [M]. Beijing: Tsinghua University press, 2016.

[5] Zhang Xuexin. Research on full vector prediction model of equipment failure based on big data [d]. *Zhengzhou University*, 2017.

[6] S Cacciola, I Munduate Agud, C L Bottasso, Detection of rotor imbalance, including root cause, severity and location. *Journal of Physics: Conference Series*, 2016.

[7] Jianfei Yao, Liang Liu, Fengyu Yang, Fabrizio Scarpa, Jinji Gao. Identification and optimization of unbalance parameters in rotor-bearing systems. *Journal of Sound and Vibration*, 2018.