2012 International Conference on Future Electrical Power and Energy Systems

Vibration Fault Diagnosis Method of Centrifugal Pump Based on EMD Complexity Feature and Least Square Support Vector Machine

Zhou Yunlong\textsuperscript{a}, Zhao Peng\textsuperscript{b}

\textsuperscript{a}School of Energy and Power Engineering, Northeast Dianli University, Jilin 132012, China
\textsuperscript{b}School of Energy and Power Engineering, North China Electric Power University, Beijing 102206, China

Abstract

Aiming at the non-stationary and non-linearity characteristics of the vibration signals of centrifugal pump, a new method based on complexity feature of Empirical Mode Decomposition (EMD) and Least Square Support Vector Machine (LS-SVM) is put forward. First of all, the Empirical Mode Decomposition (EMD) method was used to decompose the vibration signals into a finite number of stationary Intrinsic Mode Functions (IMF), and then complexity features of each IMF is extracted as the fault characteristics vectors and served as input parameters of LS-SVM classifier to diagnosis fault. Application results showed that the proposed method is very effective, which can better extract the nonlinear features of the fault and more exactly diagnosis fault.

1. Introduction

Centrifugal pump plays an important role in industries of Electric power, petroleum and chemical, metallurgy, mechanical, and military, as a result, it is necessary to develop fault diagnosis technique of centrifugal pump. It has been discovered that fault types, extent, location, and cause are closely associated with vibration signals which come about during rotating of centrifugal pump, especially with amplitude of vibration, frequency component contained in the vibration signals, presently, it is an effective method of fault monitoring and diagnosis, and widely application \cite{1-3}. Formerly, fault signature extraction of centrifugal pump adopted Fourier transform, it is unavoidable the weakness in time domain analysis. In recent decade, wavelet analysis has been applied in signal processing to extract the fault signature, such as wavelet packet entropy \cite{4}, Autoregressive spectrum \cite{5}, etc. used to extract different fault characteristics. However, vibration signals are non-stationary, non-linear randomness, many problems existing by far. Empirical mode decomposition (EMD) is believe to be a breakthrough in signal processing area in recent
years, it is a self-adaptive signal processing method and suitable for non-stationary signal analysis. It will be more accurate and effective to extract the signature when adopting EMD method in vibration signals processing of centrifugal pump.

This paper adopts EMD method to decompose the vibration signal to calculate the complexity of intrinsic mode function (IMF) on each ranking, and conducts the complexity as input feature vector of least-squares SVM. Experimental indicates that it’s a new method for fault diagnosis and performance effective in estimating centrifugal pump fault types.

2. Empirical Mode Decomposition Method

EMD is a time-frequency analysis method, which proposed by Dr. Huang [6] of NASA in 1998, supposing any signals are comprised by different intrinsic mode function (IMF). IMF is defined by the following conditions: over the entire data set, the number of extreme point must be equal to the number of zero-crossings or differ at most by one; the mean value of the envelope defined by the local maximum value and the envelope defined by the local minimum value is zero.

Specific algorithm refers to reference [6], initial signals \( x(t) \) after EMD processing can be express as:

\[
x(t) = \sum_{i=1}^{n} c_i + r
\]

Where, \( r \) is residual error function, stands for average trend of signals; IMF components \( c_1(t), c_2(t), \ldots, c_n(t) \) contain different elements separately from low frequency to high frequency of signals.

3. Complexity Extraction of Fault Signals Based on EMD

3.1. Complexity

Kolmogorov [8] proposed the symbol sequence complexity for the first time in 1965, but he did not build corresponding algorithm. The algorithm of complexity was provided by Lempel and Ziv [9] in information science research, it made complexity application come true.

The algorithm is as following:

Suppose sequence \( \{x_i\} = x_1, x_2, \ldots, x_N \), reconstruct the sequence, command:

\[
s_i = \begin{cases} 
1, & x_i \geq \bar{x} \\
0, & x_i < \bar{x} 
\end{cases}
\]

Where, \( \bar{x} = (x_1 + x_2 + \cdots + x_N) / N \).

According to express (2), time sequence \( \{x_i\} \) can be mapped to sequence \( \{s_i\} = s_1, s_2, \ldots, s_N \), where \( s_i = 0 \) or 1. The complexity of generating the sequence ‘0’, ‘1’ is marked \( C \). First, command \( Q = s_{r+1} \) and obtain symbol sequence \( SQ \) from characters \( S = S_1, S_2, \ldots, S_r, r < N \) in the sequence ‘0’, ‘1’,and command \( SQ\pi \) is the character strings which obtained from symbol sequence \( SQ \) by subtracting the last character of the sequence, if \( Q \) can be copied form substring in \( SQ\pi \), then add the character to the end, named ‘copy’, if not, called ‘addition’, add \( \tilde{\cdot} \) at the end of \( S = (s_1, \ldots, s_r, s_{r+1}, s_{r+2}) \) to avoid joining together, repeat the steps above. The number of ‘\·’ reflex the complexity \( C(N) \) of symbol sequence.

According to Lempel and Ziv, when \( N \to \infty \), the complexity \( b(N) \) will approach to a value \( b(N) \).
Relative complexity:

\[
C_r(N) = \frac{C(N)}{b(N)}
\]  

Relative complexity \(C_r(N)\) reflects the approaching degree between a time sequence and a random sequence, and the relative complexity \(C_r(N)\) of a completely random sequence is approaching to 1, but it will be approaching to 0 if it is a periodic sequence, the relative complexity between two conditions is defined as \(C_r(N) \in (0, 1)\). Complexity reflects the structural characteristics of the symbol sequence, and can be the characteristic parameter of system state \([10]\).

3.2. Fault Signature Extracting

Calculate complexity of the Intrinsic Mode Functions (IMF) which is obtained from Empirical Mode Decomposition; choose relative complexity \(C_{ri}\) \((i = 1, \cdots, 8)\) of the first 8 IMF components as the characteristic vector \(T\) of different states:

\[
T = [C_{r1} \ C_{r2} \cdots \ C_{r8}]
\]  

4. The Principle of Support Vector Machine Classifier and Parameters Selecting

4.1. The Principle of Support Vector Machine

Least Square Support Vector Machine (LS-SVM) \([11]\) is a new learning method based on Support Vector Machine (SVM), adopting quadratic loss function \([12-13]\) to convert the quadratic programming problem of SVM into liner equations for solution, it reduce the difficulty of calculation and guarantee precision simultaneously, has been widely applied in fault diagnosis, flow pattern identification and power quality classification \([14-16]\).

To LS-SVM, it’s optimizing-target:

\[
\min_{w, b, \xi} J_{LS}(w, \xi) = \frac{1}{2} \|w\|^2 + \frac{\gamma}{2} \sum_{i=1}^{n} \xi_i^2
\]

s.t. \(y_i[w^T \varphi(x_i) + b] = 1 - \xi_i, \ i = 1, 2, \cdots, n\)

Where, \(\gamma\) is error punishing coefficient, used for controlling \(J_{LS}(w, \xi)\), introducing Lagrange function for solution:

Where, \(\alpha_i\) is Lagrange multiplier, positive or negative, derivation of \(w, b, \xi\) and \(\alpha_i\) separately of expression above at extreme point, command them equal to zero. The conditions above make up a liner system, its:
\[
\begin{bmatrix}
I & 0 & 0 & -Z^T \\
0 & 0 & 0 & -Y^T \\
0 & 0 & \mathcal{F} & -I \\
Z & Y & 0 & 0
\end{bmatrix}
\begin{bmatrix}
w \\
b \\
\xi \\
a
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
0 \\
I
\end{bmatrix}
\] 

(8)

Where:
\[Z = [\varphi(x_1)^T, \ldots, \varphi(x_n)^T, y_1, \ldots, y_n]^T, \quad Y = [y_1, \ldots, y_n]^T, \quad \tilde{I} = [1, \ldots, 1]^T, \quad \xi = [\xi_1, \ldots, \xi_n]^T, \quad a = [\alpha_1, \ldots, \alpha_n]^T, \quad \mathcal{I} \in R^{n \times n} \text{ is unit matrix.}
\]

After eliminating \( w \) and \( \xi \), expression (8) is simplified:
\[
\begin{bmatrix}
0 & Y^T \\
Y & ZZ^T + \gamma^{-1} \mathcal{I}
\end{bmatrix}
\begin{bmatrix}
b \\
a
\end{bmatrix} = \begin{bmatrix}
0 \\
I
\end{bmatrix}
\] 

(9)

Define \( \Omega = ZZ^T = [q_{ij}]_{n \times n} \); applying Mercer Conditions to matrix \( \Omega \), then the elements of the matrix expressed as following form:
\[q_{i,j} = y_i y_j \varphi(x_i)^T \varphi(x_j) = y_i y_j K(x_i, x_j)
\]

(10)

Where, \( K(x_i, x_j) \) is kernel function. Generally, kernel function contains polynomial function, RBF function and sigmoid kernel function, etc... Radial Basis Function (RBF) kernel function is adopted in this paper, expression:
\[K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right);
\]

To solve the expression (9) with least square method; in the algorithm of LS-SVM, the quadratic programming problem of SVM is converted into liner equations, finally, optimal classification functions:
\[y(x) = \text{sign}\left[\sum_{i=1}^{n} a_i y_i \varphi(x_i) + b\right]
\]

(11)

Although LS-SVM diagnostic model is completeness in theory, it still being choice problem of model parameters in practice. Error punishing coefficient \( \gamma \) and kernel function parameter \( \sigma^2 \) are important to classification precision.

4.2. Polyomous

On classification, EMD considers two type issues: \( y_i = +1 \) stands for one type sample; \( y_i = -1 \) stands for the other type. On multi-categories classification, convert this problem to two categories by four kinds coding scheme: minimum output coding, error correction output coding, one-to-many coding, one-to-one coding, and minimum output coding is adopted in this paper.

Aiming at fault diagnosis, four kind states including normal, incorrect alignment, unbalance and looseness are marked as \( Y = [1 \ 2 \ 3 \ 4]^T \), the codes are as following after minimum output coding:
5. Analysis of Diagnosis Example

Centrifugal pump is 2BA-6A, maximum rotate speed 3000r/min, flow rate 20m³/h, pressure head 25.2m, efficiency 65.6%, height of suction 7.2m, adopting open system; electromotor type is JZS2-51-1, principal voltage 380V, rotate speed 470~2900r/min, frequency 50Hz. To measure radial displacement by non-contact eddy displacement transducer, installed on support of the centrifugal pump in vertical and horizontal direction; the vertical surface of centrifugal pump coupler is measurement surface, to measure axial displacement by install the instrument horizontally. The measurement signal is including normal, mass unbalance, rotor incorrect alignment, and vibration displacement signal of foundation looseness. The rotate speed increase from 500r/min to 2900r/min during the experiment, sampling while increasing every 20r/min by INV306fF data acquisition unit, sampling frequency is 800Hz, sampling numbers 4096. To simulate mass unbalance by fixing a bolt to disk which is install to coupler. To simulate rotor incorrect alignment by disassembling the coupler, make it deviate from centre and connect together. Loosen the bolt a little of electromotor foundation to simulate foundation looseness. The test system is shown in figure 1.

![Figure 1](image1)

1-electromotor, 2-valve, 3-displacement transducer, 4-exit pressure, 5-entrance pressure, 6-centrifugal pump, 7-turbine flow-meter, 8-water tank.

![Figure 2](image2)

Figure 2 The flow chart of fault diagnosis based on EMD and LS-SVM

![Figure 3](image3)

Figure 3 Time domain waveforms of four states

![Figure 4](image4)

Figure 4 The spectral diagram of vibration signals four states

\[ Y_c = \begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \] (16)
The flow chart of centrifugal pump fault diagnosis based on EMD and LS-SVM is shown in figure 2. Fault diagnosis details:

- To obtain the data of centrifugal pump in different states, 15 groups’ data of each state are selected randomly from the whole test data to be training samples, and the others are test data.
- Time domain oscillograph and frequency spectrogram of four states including normal, incorrect alignment, unbalance and looseness are shown in figure 3 and figure 4. It is difficult to distinguish four kind states from frequency spectrogram, and needs feature extraction in further. The complexity of 8 IMFs of four type vibration signals account for 98.77% of total complexity \[ \sum_{i=1}^{n_{\text{max}}} C_{ri} / \sum_{i=1}^{\infty} C_{ri}, \quad n_{\text{max}} \text{ is maximum decomposition number of EMD}. \] It indicates that the first 8 IMFs containing majority information of the signal, the complexity of first 8 IMFs is enough. Two group feature vectors of EMD complexity extracted from four kinds States is shown in table 1. The complexity entropy \( E \) obtained from expression \[ E = -\sum_{i=1}^{8} C_{ri} \log C_{ri} \] based on EMD is shown in figure 5. There are 4 types states of complexity entropy, but it is not very clear. For the purpose of distinguishing the different states reliably and with accuracy, to distinguish the difference by inputting the complexity feature vector to LS-SVM for diagnosis.
- Table 2 list out the fault diagnosis result of EMD energy and EMD singular value feature, by which for the purpose of comparing the effect of diagnosis by EMD complexity with other EMD features.

### TABLE I. Feature vectors

<table>
<thead>
<tr>
<th>States</th>
<th>( C_1 )</th>
<th>( C_2 )</th>
<th>( C_3 )</th>
<th>( C_4 )</th>
<th>( C_5 )</th>
<th>( C_6 )</th>
<th>( C_7 )</th>
<th>( C_8 )</th>
<th>( E )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.8594</td>
<td>0.5156</td>
<td>0.2041</td>
<td>0.1772</td>
<td>0.0752</td>
<td>0.0483</td>
<td>0.0376</td>
<td>0.0376</td>
<td>1.669</td>
</tr>
<tr>
<td>Normal</td>
<td>0.8271</td>
<td>0.5854</td>
<td>0.1772</td>
<td>0.1611</td>
<td>0.0537</td>
<td>0.0376</td>
<td>0.0215</td>
<td>0.0215</td>
<td>1.662</td>
</tr>
<tr>
<td>Incorrect alignment</td>
<td>0.6250</td>
<td>0.6250</td>
<td>0.4492</td>
<td>0.2148</td>
<td>0.0879</td>
<td>0.0586</td>
<td>0.0391</td>
<td>0.0391</td>
<td>2.035</td>
</tr>
<tr>
<td>Incorrect alignment</td>
<td>0.6543</td>
<td>0.5957</td>
<td>0.4004</td>
<td>0.2246</td>
<td>0.1465</td>
<td>0.0781</td>
<td>0.0488</td>
<td>0.0488</td>
<td>1.985</td>
</tr>
<tr>
<td>unbalance</td>
<td>0.7819</td>
<td>0.7477</td>
<td>0.4204</td>
<td>0.2538</td>
<td>0.1299</td>
<td>0.0684</td>
<td>0.0393</td>
<td>0.0256</td>
<td>1.792</td>
</tr>
<tr>
<td>unbalance</td>
<td>0.7708</td>
<td>0.7502</td>
<td>0.4322</td>
<td>0.2632</td>
<td>0.1615</td>
<td>0.0872</td>
<td>0.0504</td>
<td>0.0282</td>
<td>1.888</td>
</tr>
<tr>
<td>looseness</td>
<td>0.6875</td>
<td>0.5693</td>
<td>0.4136</td>
<td>0.2417</td>
<td>0.1772</td>
<td>0.1128</td>
<td>0.0806</td>
<td>0.0537</td>
<td>2.098</td>
</tr>
<tr>
<td>looseness</td>
<td>0.7681</td>
<td>0.5817</td>
<td>0.3062</td>
<td>0.2041</td>
<td>0.1558</td>
<td>0.0913</td>
<td>0.0483</td>
<td>0.0322</td>
<td>2.121</td>
</tr>
</tbody>
</table>

### 6. Conclusion

This paper presents a fault diagnosis method of centrifugal pump based on EMD complexity feature. The complexity of vibration signal varies along with the state of centrifugal pump, firstly, to obtain numbers of stationary Intrinsic Mode Functions (IMF) by EMD method and calculate its complexity, in which the variation of complexity of different frequency sections is captured; then, the complexity are served as vectors of LS-SVM for classification, comparing to EMD energy and singular value, it performance higher accuracy.

![Figure 5 Complexity entropy of four state](image-url)
TABLE II. Diagnosis results

<table>
<thead>
<tr>
<th>feature</th>
<th>Classification accuracy / %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>normal</td>
</tr>
<tr>
<td>Energy</td>
<td>93.3</td>
</tr>
<tr>
<td>Singular</td>
<td>93.3</td>
</tr>
<tr>
<td>Complexity</td>
<td>100</td>
</tr>
</tbody>
</table>

Reference