Mathematical Morphology and Deep Learning-based Approach for Bearing Fault Recognition

Yang Ge* and Xiaomei Jiang

Changshu Institute of Technology, Changsha, 215500, China

Abstract

A fault feature extraction method for rolling element bearings based on mathematical morphology is proposed in this paper. In order to obtain more useful features, this paper attempts to mix mathematical fractal features into time-frequency domain features and wavelet packet energy features. Using the mixed features, support vector machine and deep learning are performed to recognize operation conditions of bearings. It is found that mixed features can improve the conditions recognition accuracy. The comparison results show that deep learning performs better than support the vector machine and is able to predict bearing conditions with a mean accuracy of 99.19%. Therefore, it is concluded that the mixed features and deep learning method are effective for bearing operation conditions recognition.

Keywords: feature extraction; mathematical morphology; deep learning; fault recognition; rolling bearing

© 2017 Totem Publisher, Inc. All rights reserved.

1. Introduction

BEARINGS are widely used as vital components in rotary machines. The occurrence of bearing faults will result in significant breakdown time, elevated repair cost, and even a major safety accident. They are prone to various faults, such as inner ring fault, out ring fault, cage train fault and rolling elements fault [14]. Vibration signals analyses are often used to recognize operation conditions of mechanical products in recent researches. Generally, the process of fault recognition mainly consists of data acquisition, feature extraction, and conditions recognition, among which feature extraction is essential to accurately predict the conditions of bearing [22].

In recent years, various fault feature extraction methods of mechanical vibration signals have been proposed and developed, such as time domain features [25,27], frequency domain features [2], entropy features [1,5], and wavelet packet energy features [15]. Features of vibration signals and various conditions recognition methods are proposed too, such as support vector machine (SVM) [2,13,28], artificial neural network (ANN) [13,17,18], Bayesian classification [21,27], genetic algorithm [6], deep learning [11,24], and k-nearest neighbor (KNN) [9,26]. Each kind of feature may contain multiple parameters, and each parameter has a different sensitivity to the condition of the machine. In general, varieties of feature parameters are adopted to diagnose machine conditions at the same time. Because of the correlation of multiple feature parameters, using too many feature parameters will increase the calculation time of condition classification and reduce classification accuracy. To solve this problem, many researchers employ some data dimension reduction methods, such as principal component analysis (PCA) [7,8], singular value decomposition (SVD) [29], or independent component analysis (ICA) [30].

In recent years, image recognition and speech recognition technologies have quickly developed, with deep learning algorithms significantly contributing to their evolution. Deep learning has been a hot research topic in the field of machine learning because of its better learning performance and capability for unsupervised feature extraction on massive raw data [16]. Thus, it is ideal for processing and classifying mechanical vibration signals.

* Corresponding author.
E-mail address: gy090@163.com
In the studies mentioned above, machine conditions recognize only one kind of signal features, either time domain features, frequency domain features, entropy features, wavelet packet energy features, and so on. In this case, machine conditions recognition accuracy may be not ideal, because signal feature extraction is not comprehensive. In this paper, we attempt to mix the mathematical fractal feature into time-frequency domain features and wavelet packet energy features. We then use SVM and deep learning to recognize the conditions of bearings and make comparisons and analyses between different feature combinations and different recognition methods.

2. Signals features extraction based on fractal dimension

2.1. Fractal dimension of mathematical morphology Structure

Fractal dimensions can quantitatively describe some characteristics of natural morphology, and they are widely used in the image processing field. The fractal box dimension is the most widely used among various fractal dimensions. Many researchers have employed it to process vibration signals, and it has provided an effective method for fault identification of mechanical products. However, fractal box dimension frequently has inaccurate calculations. In this paper, we adopt fractal dimension of mathematical morphology, which can effectively solve the problem of inaccurate calculation in box dimension [23]. Compared with traditional methods, the method uses one-dimensional morphological coverage instead of grid division, making calculation results more stable and accurate, achieving good application effects in mechanical signal processing [16,19].

Mathematical morphology includes two basic operations of erosion and dilation. Assume that \( f(n) \) is a one-dimensional original signal, and \( g(m) \) is a structuring element signal, both of which are discrete time signals, where \( n = 1, 2, \ldots, N \), \( m = 1, 2, \ldots, M \), and \( N \geq M \). The erosion and dilation operations of \( f(n) \) and \( g(m) \) are defined as Equation (1) and Equation (2) respectively.

Erosion:

\[
(f \ominus g)(n) = \min\{f(n + m) - g(m)\}
\]  \hspace{1cm} (1)

Dilation:

\[
(f \oplus g)(n) = \max\{f(n - m) + g(m)\}
\]  \hspace{1cm} (2)

where \( m = 1, 2, \ldots, M \), \( \ominus \) denotes the operator of erosion and \( \oplus \) denotes the operator of dilation.

The structuring element at scale \( \lambda \) is defined as Equation (3)

\[
g^{\oplus \lambda} = g \oplus g \oplus \cdots \oplus g \tag{3}
\]

The morphological cover \( A_g(\lambda) \) at scale \( \lambda \) can be defined as Equation (4)

\[
A_g(\lambda) = \sum_{n=1}^{N} [(f \oplus g^{\oplus \lambda})(n) - (f \ominus g^{\oplus \lambda})(n)] \tag{4}
\]

Morphological cover \( A_g(\lambda) \) and scale \( \lambda \) satisfy the following formula Equation (5)

\[
\ln \left( \frac{A_g(\lambda)}{\lambda^2} \right) = D \ln \left( \frac{1}{\lambda} \right) + c \tag{5}
\]

where \( D \) is the fractal dimension of mathematical morphology, and \( c \) is constant.

Let \( x = \ln \left( \frac{A_g(\lambda)}{\lambda^2} \right) \), \( y = \ln \left( \frac{1}{\lambda} \right) \), Equation (5) becomes as Equation (6)

\[
x = Dy + c \tag{6}
\]
Then, we can use the least square method to calculate the value of $D$. For more information about morphological covering, please refer to the literature [24].

### 2.2. Fractal dimension features extraction of bearing vibration signals

In this paper, the bearing vibration acceleration data is from Case Western Reserve University (CWRU) Bearing Data Center [20]. We choose the Normal Baseline Data and 12k Drive End Bearing Fault Data at 1750r/min motor speed as our research object. The data contains the rolling ball fault, inner race fault, outer race fault and the four normal conditions. Rolling ball fault, inner race fault, outer race fault three conditions including 0.007”, 0.014”, 0.021” three types of fault diameters. Thus, there are 10 kinds of data in all. To facilitate presentation, the 10 data are numbered as shown in Table 1.

<table>
<thead>
<tr>
<th>Data number</th>
<th>Failure parts</th>
<th>Fault Diameter</th>
<th>Data number</th>
<th>Failure parts</th>
<th>Fault Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ball</td>
<td>0.007”</td>
<td>6</td>
<td>Inner race</td>
<td>0.021’’</td>
</tr>
<tr>
<td>2</td>
<td>Ball</td>
<td>0.014”</td>
<td>7</td>
<td>Outer race</td>
<td>0.007’’</td>
</tr>
<tr>
<td>3</td>
<td>Ball</td>
<td>0.021’’</td>
<td>8</td>
<td>Outer race</td>
<td>0.014’’</td>
</tr>
<tr>
<td>4</td>
<td>Inner race</td>
<td>0.007’’</td>
<td>9</td>
<td>Outer race</td>
<td>0.021’’</td>
</tr>
<tr>
<td>5</td>
<td>Inner race</td>
<td>0.014’’</td>
<td>10</td>
<td>Normal data</td>
<td></td>
</tr>
</tbody>
</table>

Original time domain waveforms of bearing vibration acceleration under 10 conditions are shown in Figure 1. We can see that the waveform is obviously different under different conditions.

![Figure 1. Original time domain waveform of bearing vibration acceleration](image)

Setting the largest scale $\lambda_{\text{max}} = 10$ and choosing the flat structure element $[0,0,\cdots,0]$, the 10-data fractal dimension $D$ at different structures length $L$ can be calculated as shown in Figure 2.

![Figure 2. Fractal dimension $D$ at different structures length $L$](image)
3. Deep learning method

Deep learning method is essentially a machine learning method with a multi-layer nonlinear structure network. The common method of deep learning includes deep belief network (DBN), convolutional neural network (CNN), automatic encoder (AE), de-noising automatic encoder (DAE), and so on [3]. In this paper, we use multi-layer automatic encoder as a training algorithm.

3.1. Automatic encoder

AE is a kind of three-layer unsupervised neural network, which is composed of an encoding network and decoding network. The structure of AE is shown in Figure 3, and input data and output targets are the same. High dimensional input data is converted into a low dimensional encoding vector by an encoding network, and the low dimensional encoding vector can refactor input data by the decoder activation function. Therefore, the encoding vector can be deemed a feature of input data.

![AE model structure](image)

Assume unlabeled input samples \(\{x^m\}_{m=1}^M\), each sample \(x^m\) can be converted to vector \(h^m\) by the encoding function \(f_\theta\).

\[
h^m = f_\theta(x^m) = s_f(Wx^m + b)
\]  
(7)

Where \(s_f\) is the encoding activation function, \(\theta = \{W, b\}\) is the encoding parameters set, encoding vector \(h^m\) is the reciprocal transformed to a reconstruction of \(x^m\), which is denoted as \(\hat{x}^m\) by decoding function \(g_{\theta'}\).

\[
\hat{x}^m = g_{\theta'}(h^m) = s_g(W'h^m + d)
\]  
(8)

Where \(s_g\) is the decoding activation function and \(\theta' = \{W', d\}\) is the decoding parameters set. AE completes network training through minimizing reconstruction error \(L(x^m, \hat{x}^m)\) between \(x^m\) and \(\hat{x}^m\). \(L(x^m, \hat{x}^m)\) can be obtained from Formula (9).

\[
L(x^m, \hat{x}^m) = \frac{1}{M} \sum_{m=1}^M \|x^m - \hat{x}^m\|^2
\]  
(9)

As shown in Figure 2, the fractal dimension of every bearing condition data is sensitive to change of the structure element length \(L\). Fractal dimensions of DATA3 and DATA8 are relatively close. As shown in Figure 1, the original waveforms are very similar too. Nevertheless, there are small differences between the fractal dimensions of the two, which can be distinguished. Other fractal dimensions are obviously different, so fractal dimension can be used as a feature of data. We can see that the distinction of the 10-fractal dimension is most obvious when the structure element length \(L\) is in the interval [13, 19]. Fractal dimensions are overlapping or interlaced with other intervals, making it harder to distinguish. Therefore, we chose 16 as the structure element length in the following research.

3.2. Pre-training and fine tuning of deep neural network

The main idea of deep neural network (DNN) is that multiple unsupervised AE layers are stacked to form a DNN hidden layer structure, as shown in Figure 4. First of all, use sample \(x^m\) to train \(AE_1\), and encode \(x^m\) as
\[ h^m_1 = f_{\theta_1}(x^m) \] (10)

where \( \theta_1 \) is the parameters of \( AE_1 \) and \( h^m_1 \) can reconstruct \( x^m \), so it includes the main information of \( x^m \). Then, use \( h^m_1 \) to train \( AE_2 \), and encoding is denoted as \( h^m_2 \). Repeat the above process until \( AE_N \) training is completed. The encoding of the last layer is

\[ h^m_N = f_{\theta_N}(h^m_{N-1}) \] (11)

Pre-training connects multiple \( AE \) reciprocally, forms the DNN hidden layer structure, and realizes bearing condition information extraction layer by layer. After completing pre-training, in order to identify conditions of bearing, an output layer with a classification function is added to DNN. After the classification results are obtained, back propagation (BP) is used to fine-tune DNN parameters layer by layer until classification error reaches the minimum. At last, the output of DNN can be expressed as

\[ y^m = f_{\theta_{N+1}}(h^m_N) \] (12)

where \( \theta_{N+1} \) denotes the parameters of output layer. Assume that the condition classification label of \( x^m \) is \( d^m \), DNN finishes fine-tuning by minimizing classification error \( \phi_{DNN}(\theta) \).

\[ \phi_{DNN}(\theta) = \frac{1}{M} \sum_m L(y^m, d^m) \] (13)

where \( \theta \) denotes the parameters set by DNN, and \( \theta = \{ \theta_1, \theta_2, \ldots, \theta_{N+1} \} \).

3.3. Bearing conditions recognition method based deep learning

The process of bearing conditions data training and testing by the deep learning method are shown in Figure 5. First, the features vectors of training samples and test samples are extracted separately, such as time domain features, frequency domain features, wavelet energy features, and so on. Second, features vectors of training samples without labels are input into the DNN network for training, and then the labels are taken for fine-tuning. The parameters set \( \theta \) of DNN can be obtained. At last, we can use the parameters set \( \theta \) to predict the conditions of the test sample.
4. Experimental verification

4.1. Experimental background

The experimental data comes from the CWRU bearing data center, and data classification is shown in Table 1. A continuous 1000 samples are cut as one group of data, and each condition has 120 groups of data. 60 groups of data are randomly selected from each condition data as the training data, and the remaining 60 groups of data are the test data. So, we have 600 groups of training data and 600 groups of test data in total, and each group of data is individually labeled as a data number in Table 1.

4.2. Feature extraction

In order to facilitate comparative analysis, we adopt three methods to extract data features.

1. Time-frequency domain features (TFDF)

16 time-domain features and 15 frequency-domain features are extracted, such as variance, root-mean-square, kurtosis, peak-peak value, average frequency, frequency center, mean square root frequency and so on. The calculation method of each feature is detailed in the literature [11] and [12]. Through PCA [20], the cumulative contribution rate of the first 13 features is 99.4%, so the first 13 features can be selected as final features of the bearing vibration signal to shrink calculation.

2. Wavelet packet energy features (WPEF)

Bearing vibration acceleration data is decomposed by 5 layers ‘db5’ of wavelet packets [19], and then the decomposed energy matrix is extracted as a final feature. Each group of samples has 32 features.

3. Mathematical morphology fractal dimension feature (MMFDF)

According to the computing method of MMFDF in section 1.2, setting the largest scale as $\lambda_{max} = 10$ and the structural element length as $L = 16$, the MMFDF of each group is calculated separately. Because mathematical morphology fractal dimension has only one feature value, it will be combined with other features for final recognition.

4.3. Feature normalization

Because the units of extracted features are not unified, before training and recognition, all training and test samples need to be normalized to eliminate differences in units. Normalized mapping is as follow:

$$f: x \rightarrow y = \frac{x - x_{min}}{x_{max} - x_{min}}$$  \hspace{1cm} (14)

where $x, y \in R^n$, $x_{min} = \min(x)$, $x_{max} = \max(x)$. After normalization, all features values are normalized to the range of [0, 1].

4.4. Recognition method

Multi-classification SVM is a nonlinear mapping classification method for high-dimensional data, which is widely used in classification of mechanical vibration signals. Many researchers think of it is an effective classification method. Recently, DNN has become a hot topic in many recognition areas. In order to compare which method is better for different features, 8 combination projects are given as follow:

1. TFDF+SVM (T+S)

The 13 time-frequency domain features whose dimensions are reduced by PCA are extracted from each training sample and test sample. After normalization, the features are input into a LIBSVM [4] toolbox to calculate recognition accuracy with respective labels.

2. TFDF+DNN (T+D)

As in (1), after features extraction and normalization, the features are input into Deep Learn Toolbox [10] to calculate recognition accuracy with respective labels.

3. MMFDF+TFDF+SVM (M+T+S)

Mathematical morphology fractal dimension feature and the 13 time-frequency domain features whose dimensions are reduced by PCA are extracted from each training sample and test sample, so each sample has a total of 14 features. After normalization, the features are input into a LIBSVM toolbox to calculate recognition accuracy with respective labels.
(4) MMFDF+TFDF+DNN (M+T+D)
As in (3), after features extraction and normalization, the features are input into a Deep Learn Toolbox to calculate recognition accuracy with respective labels.

(5) WPEF+SVM (W+S)
Each training sample and test sample are decomposed by 5 layers ‘db5’ of wavelet packets, and then the decomposed energy matrix is extracted as features, so each sample has a total of 32 features. After normalization, the features are input into a LIBSVM toolbox to calculate recognition accuracy with respective labels.

(6) WPEF+DNN (W+D)
As in (5), after features extraction and normalization, the features are input into a Deep Learn Toolbox to calculate recognition accuracy with respective labels.

(7) MMFDF+WPEF+SVM (M+W+S)
Mathematical morphology fractal dimension feature and the 32 wavelet packet energy features are extracted from each training sample and test sample, so each sample has a total of 33 features. After normalization, the features are input into a LIBSVM toolbox to calculate recognition accuracy with respective labels.

(8) MMFDF+WPEF+DNN (M+W+D)
As in (7), after features extraction and normalization, the features are input into a Deep Learn Toolbox to calculate recognition accuracy with respective labels.

4.5. Recognition result

According to the 8 projects above, each project is repeated 15 times, and bearing conditions recognition accuracies are shown in Figure 6.

Mean value and standard deviation of recognition accuracy for each combination are shown in Table 2.

<table>
<thead>
<tr>
<th>No.</th>
<th>Projects</th>
<th>Mean recognition accuracies (%)</th>
<th>Standard deviation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T+S</td>
<td>88.09</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>T+D</td>
<td>88.79</td>
<td>0.29</td>
</tr>
<tr>
<td>3</td>
<td>M+T+S</td>
<td>91.13</td>
<td>0.40</td>
</tr>
<tr>
<td>4</td>
<td>M+T+D</td>
<td>96.00</td>
<td>0.64</td>
</tr>
<tr>
<td>5</td>
<td>W+S</td>
<td>91.76</td>
<td>1.25</td>
</tr>
<tr>
<td>6</td>
<td>W+D</td>
<td>97.56</td>
<td>0.53</td>
</tr>
<tr>
<td>7</td>
<td>M+W+S</td>
<td>97.16</td>
<td>0.51</td>
</tr>
<tr>
<td>8</td>
<td>M+W+D</td>
<td>99.19</td>
<td>0.41</td>
</tr>
</tbody>
</table>
From Figure 6 and Table 2, we can see that:

1. Under the same recognition method, recognition accuracy rate of WPEF is higher than TFDF, which shows that wavelet packet decomposition has good applicability for mechanical vibration signal features extraction.

2. Regardless of TFDF or WPEF, under the same recognition method, recognition accuracy rate of features that including MMFDF are higher compared to those without MMFDF, thereby illustrating that MMFDF has a good distinction function in bearing vibration signal recognition.

3. For the same features, recognition accuracy rate of DNN is higher than SVM. The mean recognition accuracy rate can reach 99.19% with mixed features of MMFDF and WPEF.

5. Conclusions

In this paper, mathematical morphology fractal dimension is employed into feature extraction of bearing vibration signal, TFDF, a combination of WPEF and MMFDF, and SVW and DNN. The results show that the mixed features and deep learning method have very good effects on bearing operation conditions recognition. If MMFDF is mixed into features, the recognition accuracy is higher than those without MMFDF. That means MMFDF is sensitive to defect development and the propagation process.

The MMFDF presented in this paper provides a new approach for machinery heath state feature extraction. For future work, we plan to research the prediction method of remaining useful life with MMFDF. In addition, it also would be interesting to work on the best maintenance strategy based on RUL prediction.

Acknowledgments

We would like to thank the Bearing Data Center of Case Western Reserve University for providing the free original bearing data. We also would like to express our gratitude to the reviewers for their valuable comments and constructive suggestions.

References


**Yang Ge** is a Lecturer at the Changshu Institute of Technology. He got his PHD from Shijiazhuang Mechanical Engineering College in 2015. His research is in machine reliability.

**Xiaomei Jiang** is an Associate Professor at Changshu Institute of Technology. She got her PHD from Suzhou University in 2012. Her research is in elevator reliability.