

High-Level Feature Extraction based on Correlogram for State Monitoring of Rotating Machinery with Vibration Signals

Shaohua Yang, Guoliang Lu^{*}, Aiqun Wang, and Peng Yan

Key Laboratory of High Efficiency and Clean Mechanical Manufacturing of MOE, National Demonstration Center for Experimental Mechanical Engineering Education, School of Mechanical Engineering, Shandong University, Jinan, 250061, China

Abstract

Vibration analysis is one of the most popular methods for state monitoring of rotating machines, and feature extraction is of much importance in the design of the monitoring system. In this paper, a new high-level feature extraction method based on correlograms for vibration signal analysis is proposed, and it includes two phases. Firstly, in the learning process, a codebook is created from training data using the k -means algorithm. Next, in the testing process, for a given data stream collected from a monitoring rotating machine, the correlogram in each cycle is obtained by comparing every data point with all codewords in the codebook at first; the entropy is then computed to form final high-level features to measure the state of the machine. A change decision can be made finally based on features extracted from null hypothesis testing. Based on an experimental setup used in our previous work, the proposed method is evaluated with application to the speed change monitoring of a rotating machine. Experimental results demonstrate the excellent performance and the priority of the method compared with ten typical features.

Keywords: rotating machinery; vibration analysis; speed monitoring; feature extraction; correlogram

(Submitted on October 12, 2018; Revised on November 16, 2018; Accepted on December 8, 2018)

© 2019 Totem Publisher, Inc. All rights reserved.

1. Introduction

State monitoring of rotating machines has attracted much attention in recent years, and the detection of structural changes in the machine state has been a long-standing problem in this area [1-3]. This technique can provide an indication for possible failures or unexpected/undesired operation behaviors at an early stage so as to prevent potential failures and improve equipment/product reliability, safety, and quality, as well as productivity in industries [4]. For the purpose of extracting machine state information, the vibration signal is widely employed for its superiority in data collection and its quick interpretation/response as changes occur [5-7].

Most methods to analyze the vibration signal are either frequency-domain or time-domain based. However, although frequency-domain analysis based on Fourier transform has been widely employed in many diagnostic applications, a fatal defect of this method is that it is time-consuming in computation [8], which is not suitable for real-time prognostic applications especially when the system needs to make a response or take appropriate actions as soon as possible as an unexpected/undesired change occurs. Moreover, when the machinery operates in a lower or non-stationary state, the detection based on frequency-domain is inaccurate [9]. As a result, existing approaches for change detection are based on time-domain analysis [10-12].

In general, time-domain analysis based change detection includes two steps: feature extraction and change decision making. In particular, feature extraction plays an important role in the design of detection algorithms and thus directly influences final results [13]. Existing time-domain features such as root mean square (RMS), crest factor, standard deviation (SD), mean square error (MSE), and variance only consider the amplitude of the signal; thus, they are easily affected by noise. As a result, some efforts have been made to extract more reliable and accurate features by taking into account the structural/ordering information in the data from [14-16].

^{*} Corresponding author.

E-mail address: luguoliang@sdu.edu.cn

Inspired by these works, this paper presents a new high-level feature for vibration signal analysis based on correlograms. The correlogram has been widely adopted in many fields, such as quality control [17] and image processing [18], and it has been demonstrated to be effective in the structural analysis/expression of data [19]. To the best of our knowledge scope, this is the first attempt using it to extract features in the analysis of machine vibration signals. As shown in Figure 1, the proposed method includes two steps: learning process and testing process. In the learning process, the codebook is formed from training data where the codewords are generated using a clustering algorithm, typically the k -means algorithm. In the testing process, for a given data stream collected from a rotating machine, the correlogram is computed in each cycle by comparing every data point in the cycle with all codewords firstly; the entropy [20] is then computed from the correlogram to form the high-level feature to measure the state of the machine. We applied the proposed feature method to speed change monitoring based on an experimental setup used in our previous work and compared it with ten typical features in detection performance. Experimental results demonstrated its effectiveness and priority over other features for comparison.

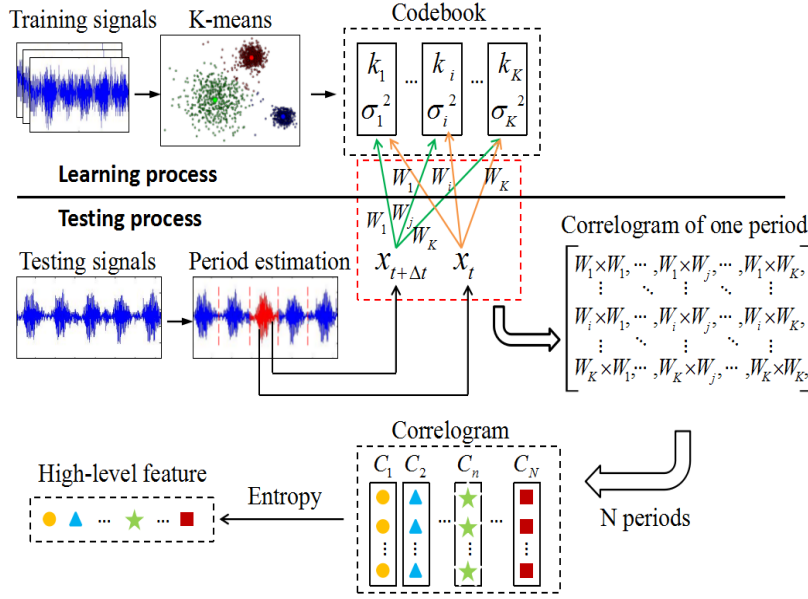


Figure 1. Schematic diagram of the high-level feature extraction based on correlogram

Moreover, in this paper, Section 1 gives a brief introduction about this paper; next, a method of high-level feature extraction is described in Section 2. Section 3 introduces a method of change decision making to estimate the feature extraction method. In Section 4, we provide the experimental results. Finally, some conclusions are given in Section 5.

2. Methodology

2.1. Correlogram Calculation

Given a time series of vibration signals $\chi = \{x_z\}$, $z = 1, 2, \dots, Z$, where Z is the length of the time series, we firstly assume that it has been divided into individual cycles. Then, the data stream can be denoted by $\chi = \{X_1, \dots, X_n, \dots, X_N\}$, where X_n represents the data of the n^{th} cycle and N is the total number of cycles. In particular, for each cycle, the data can be denoted by $X_n = \{x_t\}$, $t = \{1, 2, \dots, T\}$, where T is the period length and x_t is the numerical value of the vibration signal at t . The element of the correlogram in the n^{th} cycle can be calculated by

$$c_n(i, j; \Delta t) = \sum_{t=1}^{T-\Delta t} W_{i,t} \times W_{j,t+\Delta t} \quad (1)$$

Where Δt indicates the time offset and $W_{i,t}$ is the projection of the t^{th} data point in the current n^{th} period to a specific adopted kernel that is computed as a distance between the data and the kernel. In this paper, we adopt the Gaussian kernel function to obtain $W_{i,t}$ because it has good robustness, and $W_{i,t}$ is calculated by

$$W_{i,t} = \exp\left(-\frac{\|x_t - k_i\|^2}{2\sigma_i^2}\right), \quad i = 1, 2, \dots, K \quad (2)$$

Where $\{k_i, \sigma_i^2\}$ is the i^{th} codeword in a pre-designed codebook, which can be obtained through clustering pre-collected training samples using a clustering algorithm (the details of codeword generation will be presented in the following section), and K is the size of codebook. As shown in Figure 2, for each pair of data in the n^{th} cycle with time interval Δt , we can calculate W_i and W_j . Based on this, the correlogram can be constructed in which each element in it is the multiplication of W_i and W_j , which are the projections of data in the n^{th} cycle to the i^{th} and j^{th} codewords, respectively. Finally, for each cycle of vibration data, we can now represent them with a correlogram with a size of $K \times K$, and subsequently the whole time series of vibration signal χ can be represented by a time series of resulting correlograms, i.e., $\chi = \{C_1, \dots, C_n, \dots, C_N\}$, where C_n is the correlogram corresponding to the n^{th} cycle.

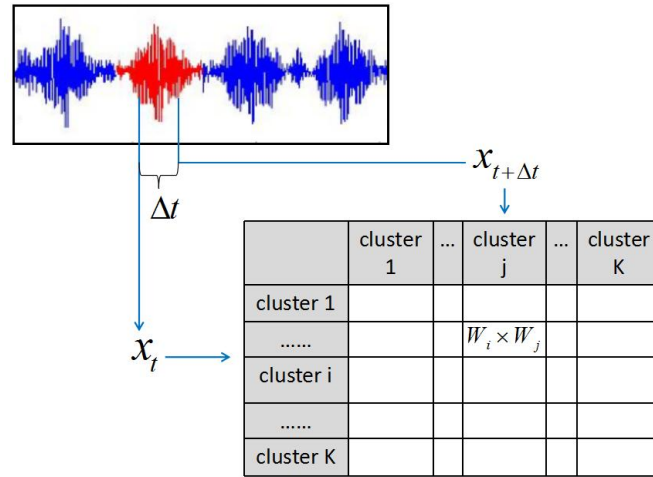


Figure 2. Illustration of construction of a correlogram

2.2. Codebook Generation

As mentioned in the above section, construction of the correlogram depends on a pre-designed codebook that is composed of a series of codewords. For this purpose, a k -mean based clustering algorithm is adopted to generate the codebook, which is performed as follows:

- Collection of sufficient vibration signals from a monitoring rotating machine as the training samples;
- Performing the k -means algorithm on all training samples, and as such, the resulting clustering centers along with their corresponding standard deviations, i.e., $\{k_i, \sigma_i^2\}$, $i = 1, 2, \dots, K$ are considered as the codewords to compose the codebook.

Moreover, it is noted that the cluster number K has great influence on the clustering results [21]. More specifically, when the value of K is small, the robustness of the codebook to noise will be high, but the computation resolution to the feature extraction will meanwhile decrease; similarly, when the value of K is larger, the computation resolution will increase, but the robustness will deteriorate. It is difficult to automatically determine the K value, and thus it is generally determined by a prior estimation or the experience of the user. In this paper, we use a prior estimation to confirm the value of K compatible with our application.

2.3. High-Level Feature Extraction based on Correlogram

Based on the correlograms calculated in Section 2.1, a high-level feature is extracted in this section. The entropy is used to extract the state of the machine from the resulting correlograms. Specifically, for the correlogram of the n^{th} cycle, i.e., C_n , the entropy is calculated by

$$s_n = -q \sum_{i=1}^K \sum_{j=1}^K p_{i,j;n} \ln(p_{i,j;n}) \quad (3)$$

Where q is a constant corresponding to a choice of measurement units and p_{mn} represents the proportion of the m^{th} element of the correlogram of the n^{th} cycle. They are calculated respectively as follows:

$$q = \frac{1}{\ln(N)} \quad (4)$$

$$p_{i,j;n} = \frac{c_n(i, j; \Delta t)}{\sum_{j=1}^K \sum_{i=1}^K c_n(i, j; \Delta t)}, \quad i = 1, 2, \dots, K, \quad j = 1, 2, \dots, K \quad (5)$$

Where $c_n(i, j; \Delta t)$ is the element of the correlogram of the n^{th} cycle.

Coupling with these formulations, a high-level feature is finally extracted. Note that as a cycle-level feature, the high-level feature has two essential merits compared with the data-point level features (such as kurtosis, skewness, variance, etc.): (1) the proposed feature has high computational efficiency since it is now computed in a *cycle-by-cycle* way instead of *point-by-point*; (2) it is believed to have a better robustness to noise since we employ the projections of vibration signals to codebook instead of original data values.

In summary, the proposed feature can be obtained as follows: for a given data stream X , as seen in Figure 3(a), for example, we firstly represent it as a time-series of correlograms using the method of Section 2.1 (refer to Figure 3(b)), and then we use the entropy to extract the final features from these time-sequential correlograms for the purpose of extraction of the state of the monitoring machine state, as seen in Figure 3(c).

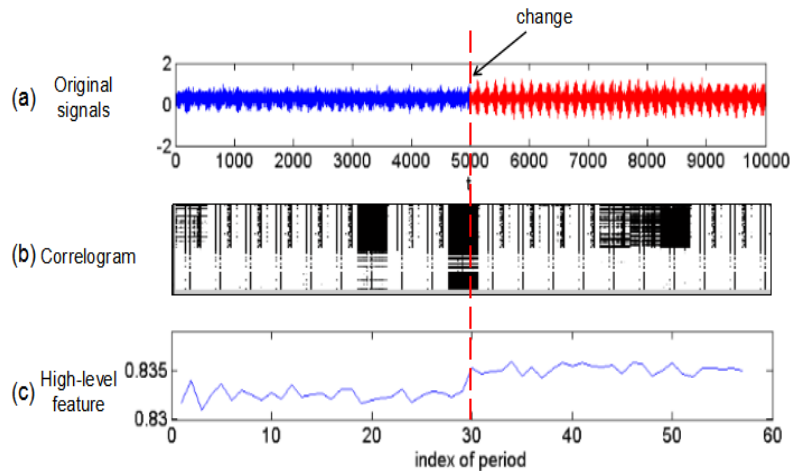


Figure 3. High-level feature extraction based on correlogram by using entropy

3. Change Decision Making

Based on the final resulting features, appropriate change detection tests can be performed to make change decisions. The existing methods can be divided into two groups: retrospective detection and real-time detection. Generally, retrospective detections such as the likelihood ratio test (LRT) [22], Hellinger distance [23], and Kullback-Leibler Divergence (K-L Divergence) [24] accommodate applications that allow certain delays, while real-time detection targets applications require immediate responses.

As for the state monitoring problems studied in this paper, since it is expected that the system obtains the result as soon as possible after a change occurs, we use the real-time detection method in this paper. More specifically, we use the $3\sigma'$ control chart for inspecting whether a change occurs at the current n^{th} cycle. The corresponding hypothesis testing is given

as follows:

$$\begin{aligned} H_0 : |s_n - \bar{s}| < 3\sigma' \\ H_1 : |s_n - \bar{s}| \geq 3\sigma' \end{aligned} \quad (6)$$

Where H_0 means that no change occurs on the n^{th} cycle as long as $|s_n - \bar{s}| < 3\sigma'$, and H_1 denotes that a change occurs when $|s_n - \bar{s}| \geq 3\sigma'$. Here, $\{\bar{s}, \sigma'\}$ is the representation of assumed Gaussian distribution, where \bar{s} is the mean of the distribution and σ' is the standard deviation, which can be calculated as below:

$$\bar{s} = \frac{1}{N} \sum_{n=1}^N s_n \quad (7)$$

$$\sigma' = \sqrt{\frac{1}{N} \sum_{n=1}^N (s_n - \bar{s})^2} \quad (8)$$

4. Experiment

4.1. Implementation

We applied the proposed feature to speed change monitoring based on an experimental setup as shown in Figure 4. Vibration signal was firstly collected by sensor mounted on the gearbox and then sent to PC. During the data collection, we first set the initial system speed as v and then changed it to Δv to simulate a change. With this setting, we can obtain the vibration signal as the testing data. The testing values of v and Δv are given as follows:

On top of this, we have 15 parameter combinations with different speed changes as shown in Table 1, and for each combination, testing data were collected at five times. Finally, we formed our testing database including a total number of 75 testing signals.

Table 1. Simulated change of $v \rightarrow \Delta v$

$\Delta v \setminus v$	250	300	350
50	250→250+50	300→300+50	350→350+50
100	250→250+100	300→300+100	350→350+100
150	250→250+150	300→300+150	350→350+150
200	250→250+200	300→300+200	350→350+200
250	250→250+250	300→300+250	350→350+250

To verify the effectiveness of our proposed extracted feature, three common indicators of precision, recall, and F_1 are employed and are calculated as

$$Precision = \frac{u_2}{u_1} (\%) \quad (9)$$

$$Recall = \frac{u_2}{u} (\%) \quad (10)$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} (\%) \quad (11)$$

Where u represents the number of true changes, u_1 is the number of changes detected, and u_2 is the number of changes that are correctly detected. The accuracy of the detection can be reflected in the precision, and the ability of detecting all change points can be reflected in the recall. F_1 is a composite indicator that considers both precision and recall.

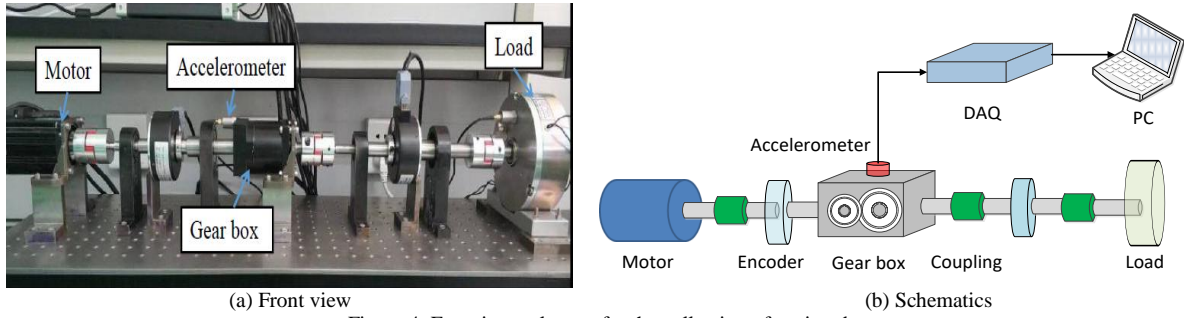
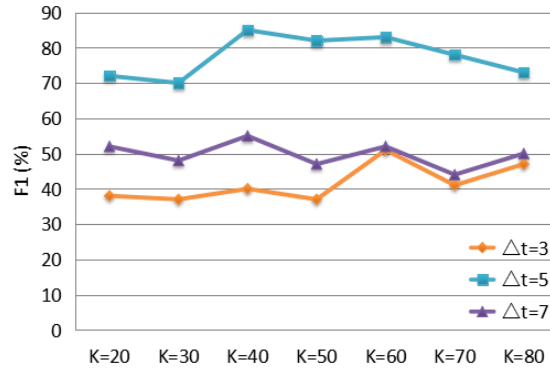


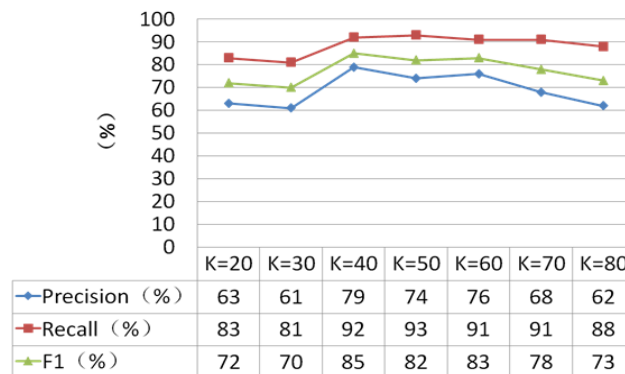
Figure 4. Experimental setup for the collection of testing data

4.2. Result and Analysis

Since the time offset Δt and the cluster number K can greatly influence our proposed feature extraction method, we tested three values of Δt , i.e., $\Delta t = 3, 5, 7$, and values of K from 20 to 80 with an interval of 10 for feature extraction. Notice from Figure 5, when $\Delta t = 5$, F_1 is the highest, which indicates the feature extraction method has achieved the best performance. The main reason is that a smaller value of Δt will lead to less robustness to local noise, while a larger value of K will result in lower computation resolution for the resulting features. In other words, when $\Delta t = 5$, the proposed method can achieve the best balance between robustness and computation resolution for the features computed by our method.

Figure 5. Detection results for different values of Δt

To further investigate the influence of K , we provide the detailed results of different K values for $\Delta t = 5$. From Figure 6, it can be seen that the overall detection performance in terms of precision and recall increases with an increasing value of K before $K = 40$ but decreases subsequently. A clearer observation can be found in the term of F_1 where the proposed method achieves the best performance, i.e., $F_1 = 85\%$ when $K = 40$. The main reason of these observations is the argument stated in Section 2.2, that is, K can affect the final detection results.

Figure 6. Detection results for different values of K with $\Delta t = 5$

Furthermore, to evaluate the priority of our method, the proposed feature is compared with ten typical features used in vibration signal analysis: RMS, crest factor, kurtosis, waveform, skewness, quadratic mean (QM), SD, MSE, variance, and maximum peak (MP), which are shown in Table 2. For these features, we perform the same hypothesis testing in Equation (6) to detect possible changes in each testing data.

Table 2. Computation of selected ten typical features

Features	Formulations
RMS	$X_{RMS} = \sqrt{\frac{\sum_{z=1}^Z x_z^2}{Z}}$
Crest factor	$X_C = \frac{ x_z }{X_{RMS}}$
Kurtosis	$X_K = \frac{1}{Z} \sum_{z=1}^Z \left(\frac{x_z - \bar{x}}{X_{SD}} \right)^4$
Waveform	$X_W = \frac{X_{RMS}}{\frac{1}{Z} \sum_{z=1}^Z x_z }$
Skewness	$X_{SK} = \frac{Z \sum_{z=1}^Z (x_z - \bar{x})^3}{(Z-1)(Z-2)X_{SD}^3}$
Mean	$\bar{x} = \frac{1}{Z} \sum_{z=1}^Z x_z$
SD	$X_{SD} = \sqrt{\frac{1}{Z} \sum_{z=1}^Z (x_z - \bar{x})^2}$
MSE	$MSE = \frac{1}{Z} \sum_{z=1}^Z \varepsilon_z^2, \varepsilon_z = observed_z - predicted_z$
Variance	$X_V = \frac{1}{Z} \sum_{z=1}^Z (x_z - \bar{x})^2$
MP	$X_{MP} = \max(x_z)$

Here, it is worth mentioning that since the proposed method has demonstrated the best performance when $\Delta t = 5$ and $K = 40$, we compared this result with those by the compared features. Figure 7 give two examples of change detection with different speed change by using our method as well as other methods, where from top to bottom are the testing data, the detection results by the proposed feature, as well as the detection results obtained by the compared features. In particular, we used the red cycles for denoting the detected change points for each method. Some conclusions are summarized as below: (1) by using the proposed features, changes can be detected successfully and accurately without any false alarm; (2) for the compared features, including RMS, waveform, skewness, mean, SD, MSE, variance, and MP, changes can be detected successfully, but there are many false alarms as shown in the figure; (3) for other features, such as kurtosis and crest factor, changes cannot be detected successfully but many false alarms have been reported. These observations demonstrate the effectiveness and the priority of our feature over ten compared features.

Moreover, a comprehensive comparison is provided in Figure 8 for all testing data, where for each compared feature, we count all the changes in a period as a change point for giving a fair comparison with our method, considering our method gives the cycle-level detection as mentioned in Section 2.3. On this basis, three indicators of the high-level feature based on our method as well as the compared features are calculated. Here, note that our proposed features achieved the best detection performance in all terms of the three indicators clearly.

In summary, from the above experiments, the proposed feature has shown its potential in real engineering applications.

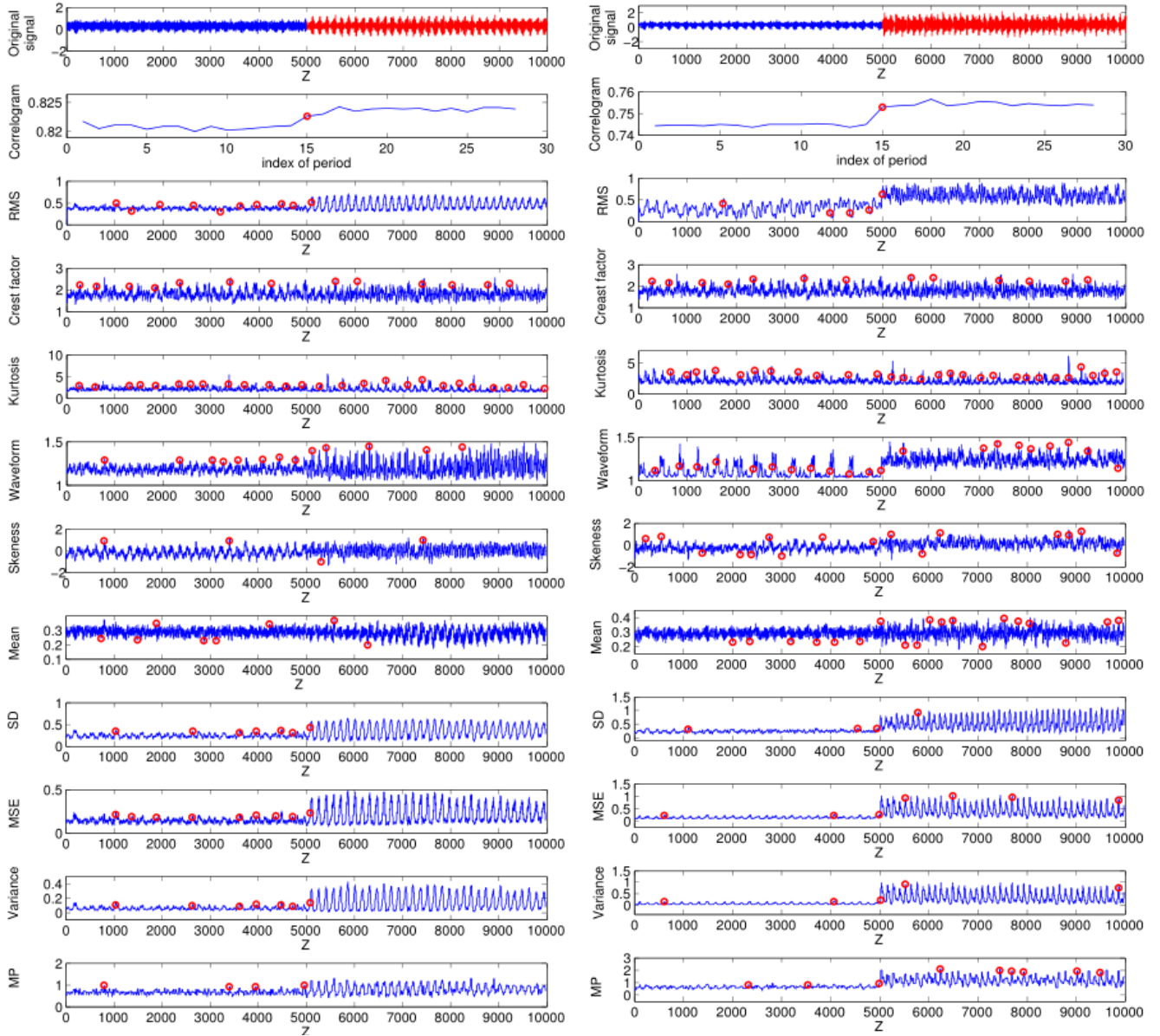


Figure 7. Two examples of the detection results by our proposed feature as well as using ten compared features

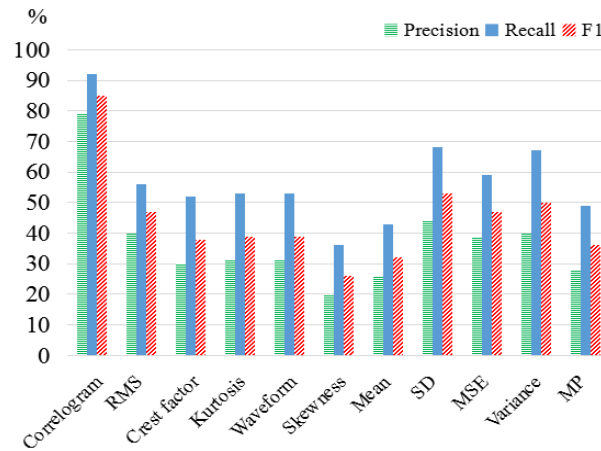


Figure 8. Three indicators for evaluating the methods of feature extraction

5. Conclusions

In this paper, a high-level feature extraction method is proposed based on correlograms for vibration signal analysis. In the method, as a pre-processing, the codebook is firstly created from training data using a k -means algorithm. In the testing process, for a given data stream collected from a monitoring rotating machine, the correlogram of each cycle is calculated by comparing every data point with all codewords in the codebook, and the entropy is calculated to form the final features to represent the state of the machine. Based on the final resulting features, hypothesis testing is performed to detect potential changes in the machine state. As shown by the experiment results, the proposed feature demonstrates excellent detection performance that outperforms ten compared typical feature methods, suggesting its potential in real engineering applications.

In future works, we will focus on further optimizing the computation method to achieve higher computational efficiency.

Acknowledgements

This work is supported by the Young Scholars Program of Shandong University (YSP-SDU).

References

1. D. Goyal and S. B. Pabla, "The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review," *Archives of Computational Methods in Engineering*, Vol. 23, No. 4, pp. 585-594, 2016
2. G. Lu, Y. Zhou, C. Lu, and X. Li, "A Novel Framework of Change-Point Detection for Machine Monitoring," *Mechanical Systems & Signal Processing*, Vol. 83, pp. 533-548, 2017
3. F. Pozo, I. Arruga, L. E. Mujica, M. Ruiz, and E. Podivilova, "Detection of Structural Changes through Principal Component Analysis and Multivariate Statistical Inference," *Structural Health Monitoring*, Vol. 15, No. 2, pp. 127-142, 2016
4. Y. Ding, W. He, B. Chen, Y. Zi, and I. W. Selesnick, "Detection of Faults in Rotating Machinery using Periodic Time-Frequency Sparsity," *Journal of Sound and Vibration*, Vol. 382, pp. 357-378, 2016
5. M. Cerrada, R. V. Sánchez, D. Cabrera, G. Zurita, and C. Li, "Multi-Stage Feature Selection by using Genetic Algorithms for Fault Diagnosis in Gearboxes based on Vibration Signal," *Sensors*, Vol. 15, No. 9, pp. 23903-23926, 2015
6. J. Dybala and R. Zimroz, "Rolling Bearing Diagnosing Method based on Empirical Mode Decomposition of Machine Vibration Signal," *Applied Acoustics*, Vol. 77, No. 3, pp. 195-203, 2014
7. G. Lu, J. Liu, and P. Yan, "Graph-based Structural Change Detection for Rotating Machinery Monitoring," *Mechanical Systems and Signal Processing*, Vol. 99, pp. 73-82, 2018
8. A. Ghods and H. H. Lee, "Probabilistic Frequency-Domain Discrete Wavelet Transform for Better Detection of Bearing Faults in Induction Motors," *Neurocomputing*, Vol. 188, pp. 206-216, 2016
9. Y. Yang, X. J. Dong, Z. K. Peng, W. M. Zhang, and G. Meng, "Vibration Signal Analysis using Parameterized Time-Frequency Method for Features Extraction of Varying-Speed Rotary Machinery," *Journal of Sound and Vibration*, Vol. 335, pp. 350-366, 2015
10. W. Li, S. Zhang, and S. Rakheja, "Feature Denoising and Nearest-Farthest Distance Preserving Projection for Machine Fault Diagnosis," *IEEE Transactions on Industrial Informatics*, Vol. 12, No. 1, pp. 393-404, 2016
11. C. Wang, M. D. Prieto, L. Romeral, Z. Chen, F. Blaabjerg, and X. Liu, "Detection of Partial Demagnetization Fault in PMSMs Operating under Nonstationary Conditions," *IEEE Transactions on Magnetics*, Vol. 52, No. 7, pp. 1-4, 2016
12. X. Zhang, Y. Liang, and J. Zhou, "A Novel Bearing Fault Diagnosis Model Integrated Permutation Entropy, Ensemble Empirical Mode Decomposition and Optimized SVM," *Measurement*, Vol. 69, pp. 164-179, 2015
13. G. He, K. Ding, and H. Lin, "Fault Feature Extraction of Rolling Element Bearings using Sparse Representation," *Journal of Sound and Vibration*, Vol. 366, pp. 514-527, 2016
14. W. He, Y. Zi, B. Chen, F. Wu, and Z. He, "Automatic Fault Feature Extraction of Mechanical Anomaly on Induction Motor Bearing using Ensemble Super-Wavelet Transform," *Mechanical Systems and Signal Processing*, Vol. 54, pp. 457-480, 2015
15. N. Lu, Z. Xiao, and O. P. Malik, "Feature Extraction using Adaptive Multiwavelets and Synthetic Detection Index for Rotor Fault Diagnosis of Rotating Machinery," *Mechanical Systems and Signal Processing*, Vol. 52, pp. 393-415, 2015
16. W. Li, Z. Zhu, F. Jiang, G. Zhou, and G. Chen, "Fault Diagnosis of Rotating Machinery with a Novel Statistical Feature Extraction and Evaluation Method," *Mechanical Systems and Signal Processing*, Vol. 50, pp. 414-426, 2015
17. A. S. Rathore, M. Pathak, R. Jain, and G. P. Jadaun, "Monitoring Quality of Biotherapeutic Products using Multivariate Data Analysis," *Aaps Journal*, Vol. 18, No. 4, pp. 1-8, 2016
18. A. Nigam and R. C. Tripathi, "Trademark Image Retrieval using Weighted Combination of Sift and HSV Correlogram," *International Journal of Computer Applications in Technology*, Vol. 54, No. 1, pp. 61-67, 2016
19. K. Yildiz, "Dimensionality Reduction-based Feature Extraction and Classification on Fleece Fabric Images," *Signal, Image and Video Processing*, Vol. 11, No. 2, pp. 317-323, 2017
20. H. H. Bafroui and A. Ohadi, "Application of Wavelet Energy and Shannon Entropy for Feature Extraction in Gearbox Fault Detection under Varying Speed Conditions," *Neurocomputing*, Vol. 133, pp. 437-445, 2014
21. E. Mooi and M. Sarstedt, "Cluster Analysis," *A Concise Guide to Market Research*, pp. 273-324, Springer, Berlin, Heidelberg, 2014
22. V. Shahsavari, L. Chouinard, and J. Bastien, "Wavelet-based Analysis of Mode Shapes for Statistical Detection and Localization

- of Damage in Beams using Likelihood Ratio Test,” *Engineering Structures*, Vol. 132, pp. 494-507, 2017
23. R. Lourenzutti and R. A. Krohling, “The Hellinger Distance in Multicriteria Decision Making: An Illustration to the TOPSIS and TODIM Methods,” *Expert Systems with Applications*, Vol. 41, No. 9, pp. 4414-4421, 2014
 24. C. D. Fuh and Y. Mei, “Quickest Change Detection and Kullback-Leibler Divergence for Two-State Hidden Markov Models,” *IEEE Transactions on Signal Processing*, Vol. 63, No. 18, pp. 4866-4878, 2015