Research Article

Wind Turbine Gearbox Fault Diagnosis Method Based on Riemannian Manifold

Shoubin Wang, Xiaogang Sun, and Chengwei Li

School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China

Correspondence should be addressed to Shoubin Wang; wsbin800@126.com

Received 27 February 2014; Accepted 30 March 2014; Published 16 April 2014

Academic Editor: Weichao Sun

Copyright © 2014 Shoubin Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As multivariate time series problems widely exist in social production and life, fault diagnosis method has provided people with a lot of valuable information in the finance, hydrology, meteorology, earthquake, video surveillance, medical science, and other fields. In order to find faults in time sequence quickly and efficiently, this paper presents a multivariate time series processing method based on Riemannian manifold. This method is based on the sliding window and uses the covariance matrix as a descriptor of the time sequence. Riemannian distance is used as the similarity measure and the statistical process control diagram is applied to detect the abnormality of multivariate time series. And the visualization of the covariance matrix distribution is used to detect the abnormality of mechanical equipment, leading to realize the fault diagnosis. With wind turbine gearbox faults as the experiment object, the fault diagnosis method is verified and the results show that the method is reasonable and effective.

1. Introduction

At present, the environmental pollution and the energy shortage become increasingly prominent, while wind power as a representative of the new clean energy has made a leap forward in recent years. Wind power has advantages such as low cost, being clean and pollution-free, short construction period, and small occupation space, which will be more and more obvious in energy issues like curbing carbon dioxide emissions, stabilizing oil price fluctuations, and so on. So wind resources have gotten the more and more attention of governments around the world [1].

With the increase of large wind generator capacity, the system structure is increasingly complex, but at the same time, frequent accidents of wind power, from the common faults such as bearing wear of wind turbine itself, the gear tooth broke, and shaft misalignment to the burning of wind turbine itself, blade drop, and the collapsing of wind machine, not only can cause the blackouts but also can produce serious safety accidents and huge economic losses [2, 3]. In view of the above, the fault diagnosis and the safe operation of wind turbines have gradually become an important research direction in the development of wind power.

Sweden, Finland, and Germany’s wind power failure statistics data show that failures will occur in the wind turbine blade, hydraulic, electrical, and machine driven systems, in which the most common one is the electrical system fault and the major downtime is caused by gear faults. Statistics also point out that the larger the wind turbine size, the higher the frequency of the fault [4, 5].

Gearbox uses the mechanical transmission parts. Whether the operation of the gear box is normal has a great impact on the whole machine or unit. Improper design, manufacture, maintenance, and operation are the main causes of the gear box faults [6]. Therefore to improve the reliability of gear box operation means to improve the level of operation maintenance, to monitor the operation condition of gear box, and to troubleshoot the problems. According to the statistics, in the gear box, gear failure is the main one, about 60% of the total number of all kinds of faults. With the development of related disciplines, new methods of signal analysis and fault diagnosis technology will be applied to gearbox fault diagnosis [7].

Researching the more accurate wind turbine fault diagnosis technology, improving the operation's reliability of wind power turbines, and realizing the conversion from
the preventive maintenance to the state detection make the turbine’s early failure be found in time and prevent downtime accidents. Early failure is made to be found in time and downtime accidents are prevented by researching the more accurate wind turbine fault diagnosis technology, improving the operation's reliability of wind power turbines, and realizing the conversion from the preventive maintenance to the state detection. On the other hand, that can also solve problems such as the lack of equipment regular maintenance, the maintenance of excess, and the improper maintenance, which has profound significance to the development of wind turbines.

Based on the wind turbine operation and the analysis of the structural characteristics, the goal of the researchers becomes how to improve the efficiency of fault diagnosis and visualization. It is necessary to use the new fault diagnosis method. In this paper, by analyzing test data, wind turbine gearbox fault diagnosis methods based on Riemannian manifold are proposed. And the method realized wind turbine gearbox fault diagnosis correctly.

2. Research Status and Method

2.1. Research Status. Fault monitoring and diagnosis technology research originated in the US, which focused on fault mechanism, detection, diagnosis, and prediction in aerospace system at first and then on the important equipment in other industries. Domestic fault diagnosis technology was developed on the basis of the introduction of overseas advanced technology. The research was initially focused on the fault mechanism, diagnosis methods of all kinds of mechanical equipment, and simple monitoring and diagnosis instrument, and then the online monitoring and fault diagnosis devices were gradually developed domestically.
In 1961, the mechanical failure prevention group established by United States Naval Research Laboratory actively engaged in the development of diagnostic technology. In the late 60s and the early 70s, the machine care center headed by R. A. Collacott started the research of fault diagnosis technology in UK. In China, Baoshan Iron & Steel Group Company and Taiyuan Iron & Steel Company became the pilot units to carry out diagnosis technology research in 1983 and vibration, infrared, and iron spectrum laboratories were established firstly, which achieved some results in the fault detection, diagnosis, and prediction. Chinese Society of Vibration engineering established the society of fault diagnosis technology in May 1987. In recent years, some Chinese colleges and universities, such as Xi’an Jiaotong University, Northeastern University, or Shenyang University of Technology, have established the engineering centers of fault diagnosis, contributing positively to the development and popularization of fault diagnosis technology. Fault diagnosis technology absorbs the new achievements of modern science and develops rapidly from theory to practice, which involves lots of fields, such as computer, sensor and testing technology, signal analysis and processing, forecasting, automatic control, system identification, artificial intelligence, machine dynamics and vibration engineering, and other fields, and has the characteristics of strong engineering applicability and close relation to high technology and so forth [8, 9].

Fault diagnosis process generally includes diagnosis objects, sensors, signal processing system, state recognition, diagnosis decision-making, and maintenance. Signal processing technology plays an important role in the correct diagnosis fault, which is the core of diagnosis. So the traditional signal processing methods is introduced briefly in the next section.
2.2. Signal Processing Methods. Signal processing techniques include fault signal detection and processing, where vibration, noise, temperature, pressure, flow, current, and voltage are usually detected. Signal processing theory is based on mathematical statistics and stochastic process. The basis of the correct mechanical fault diagnosis includes the appropriate signal processing method, the accurate processing, and the obvious expression.

Vibration monitoring of rotating machinery and other large machinery is commonly used in practice. It needs the installation of the sensors to measure the vibration of the main parts of wind turbine, such as displacement, velocity, and acceleration, and then to process all kinds of measured data, in order to judge the fault type and location. This machine fault diagnosis method is relatively accurate and can locate the fault more directly, so it is one of the commonly used methods in the field of fault diagnosis. The theoretical basis of vibration monitoring is as follows: during the process of manufacture, installation, and operation, the failure like error, crack, and spalling of the gear, shaft, bearings, and other mechanical parts will accumulate gradually and become the exciting source of vibration expressed in vibration signal with the period of the part’s rotating cycle. The main signal analysis methods are time domain analysis, and frequency-domain analysis.

Traditional signal processing methods include Fourier transform, wavelet analysis, and the empirical mode decomposition (EMD). Fourier transform establishes the bridge between the signals in time domain and frequency domain and provides time domain and frequency-domain analysis. The time domain analysis method is to use time domain signal for wavelet analysis. The time domain analysis method is to use time domain signal to analyze and give the results. In a project, directly measured signal is usually the time domain signal, and the time domain analysis is more effective in the signal containing obvious harmonic, periodic, or instantaneous pulse component. As a result, it is the most simple and most direct monitoring method. Because the generation and
development of failure often cause the dynamic changes of the signal frequency, it often needs to obtain the frequency-domain information of the signal for better understanding of the dynamic characteristics of diagnosed equipment.

Fourier transform can decompose a complex time domain signal into the finite or infinite frequency harmonic component for the profounder level information. Spectrum analysis has become the main content of the mechanical equipment failure vibration diagnosis. Wavelet transform converts the signal in the time-scale plane to observe it in different scales and decomposes it in different frequency bands, leading to the obvious outline and details of the signal. So it has the multiresolution ability. Empirical mode decomposition is to stabilize a signal (or its reciprocal) and to decompose the signal in different scale of fluctuation or trend leading to producing a series of data having the different scale characteristics.

3. Multivariate Time Series Processing Method Based on Riemannian Manifold

Multivariate time series processing method based on Riemannian manifold is based on sliding window and uses the covariance matrix as a descriptor of the time sequence, in which Riemannian distance is used as the similarity measure and the statistical process control (statistical process control, SPC) diagram as evaluation to detect the abnormality of multivariate time series. And the visualization of the covariance matrix distribution is used to detect the abnormality of mechanical equipment, so as to realize the fault diagnosis.

Riemannian manifold $M$ is a local European topological space and a differential manifold having a continuous Riemannian metric. In the manifold $M$ each point has a small neighborhood which is a diffeomorphism with a small neighborhood of the European space. Geometry in the space...
should be based on the distance between the two infinite near points, \((x_1, x_2, \ldots, x_n)\) and \((x_1 + d_1, \ldots, x_n + d_n)\), measured in the view of the positive definite quadratic form identified by the square of differential arc length that is the positive definite symmetric matrices composed of function, which is the Riemannian metric.

3.1. The Standardized Time Series. Because the unstandardized sequence comparison does not make any sense for data mining, before any data processing the first thing to do should be the data standardization, which makes it form a time series with a mean of 0 and a standard deviation of 1. The standardized sequence considers both the dependence on the original observation data in time series and the interference of random fluctuations. What is important is that it can
reduce the computational complexity of algorithm maintaining the information of the original sequence effectively at the same time:

\[ X = \frac{(X_i - \text{mean}(X))}{\text{std}(X)}, \]

where \( X \) is a time serie, \( \text{mean}(X) \) is mean value, and \( \text{std}(X) \) is standard deviation.

3.2. Sliding Window Partition and Covariance Matrix Calculation. Consider an \( N \)-channel time sequence. If the length of the sliding window is \( T \), that means the original sequence is divided into a number of the subsequence with the length of \( T \), each subsequence of the multivariate time series forms a matrix \( X_i = [x_{i+T}, \ldots, x_{i+T-1}] \), and \( X_i \in \mathbb{R}^{N \times T} \). The space covariance matrix of subsequence is estimated by the sample covariance matrix:

\[ P_i = \frac{1}{T-1} X_i X_i^T, \]

where \( P_i \) is the sample covariance matrix and \( T \) is the length of the sliding window.

Because covariance matrix is the symmetric and positive semidefinite structure belonging to positive definite symmetric matrix of Riemannian manifold, a series of operations provided by the Riemannian geometry can be used to deal with the covariance matrix. For spatial information of the time series contained in space covariance matrix, all of the algorithms based on distance can be represented using Riemannian distance.

3.3. The Reference Covariance Matrix and the Riemannian Distance Calculation. After the estimation of the sliding window sample covariance matrix with (1), the reference matrix \( \overline{P} \) is estimated by self-adaption. The formula is as follows:

\[ \overline{P}_{t+1} = \left( P_t \right)^{1/2} \left[ \left( P_t \right)^{-1/2} P \left( P_t \right)^{-1/2} \right]^{1/\alpha} \left( P_t \right)^{1/2}, \]

where \( \overline{P}_t \) is the reference matrix of the previous iterations, \( P \) is the current covariance matrix, and \( \alpha \) is a coefficient of self-adaptive speed. The Riemannian distance between each covariance matrix \( P_t \) and reference covariance matrix \( \overline{P} \) is defined as

\[ d_R(p, \overline{P}) = \sqrt{\sum_{n=1}^{N} \log^2 (\lambda_n)}, \]

where \( d_R \) is the Riemannian distance and \( \lambda_n \) is the eigenvalue of \( P^{-1/2} \overline{P} P^{-1/2} \).
3.4. The Distance Threshold Calculation according to the 3σ Principle of Statistical Process Control. Statistical process control diagram is composed of three parts:

UCL: up control line, the values of $\mu + 3\sigma$;
CL: center line, the values of $\mu$;
LCL: low control line, the values of $\mu - 3\sigma$,

where two important parameters are the following.

$\mu$ is position parameter and average, the center of the distribution, and expectations. It reflects the overall comprehensive ability.

$\sigma$ is scale parameter, degree, and standard deviation of the distribution. It reflects the degree to which the real value deviates from the expected value. The higher the value, the more scattered the data.

Control chart principle belongs to the small probability event principle: if any point is out of the bound, the abnormal judgment is obtained. If the small probability events do not occur actually and the above abnormal phenomenon is found, the control chart is the graphical method of hypothesis test.

According to the 3σ principle, the boundary value $th$ ($th = \mu + 3\sigma$) can be estimated by using the mean $\mu$ and standard deviation $\sigma$ of the Riemannian distance of the covariance matrix and the reference matrix. But according to the experience, when whether the abnormal phenomenon occurs is detecting, the boundary value $th$ usually is defined as

$$th = \mu + 2.5\sigma, \quad (5)$$

where $th$ is boundary value, $\mu$ is mean value, and $\sigma$ is standard deviation.

If $d_R > th$, it proves that time series is abnormal.

**Figure 19:** Get the time delay of autocorrelation method.
4. The Experimental Results and Analysis

The equipments used in the test are two low-frequency vibration acceleration sensors and two normal acceleration sensors.

Consider the experiment mainly aimed at the gear box; four sensors were located to detect equipment failure more fully, in which the low-frequency vibration acceleration sensors were placed on the input shaft and the annular gear and the ordinary vibration acceleration sensors were placed on the intermediate shaft and high speed shaft. And the sampling frequency of 500 Hz, long time sampling, and the total sampling time more than 10 min were used for the two channels of the input shaft and the annular gear of gear box; the sampling frequency of 2000 Hz and the total sampling time close to 10 min were used for the two channels of the intermediate shaft and the high speed shaft.

In order to verify the fault diagnosis methods, this paper chooses the gearbox data in normal and fault state for analysis.

Because of the large amount of data, they are divided into 7 groups and the analysis and fault diagnosis are mainly based on number 2 group and number 3 group given below. The fault diagnosis steps are as follows:

1. the parts working condition analysis according to the time domain signal diagram,
2. the parts working condition determination based on the spectrum diagram,
3. the parts working condition judgment based on the process control chart of Riemann manifolds,
4. the system working condition determination considering the time domain analysis, the frequency spectrum analysis, and the process control chart of Riemann manifolds.

4.1. Analysis of Data 2

4.1.1. Time Domain Analysis of Data 2. The time domain waveform graphs of the four time series belonging to Data
2 are shown as in Figures 1, 2, 3, and 4. It can be seen that the vibration signal is flat in the time domain diagram and the impact vibration signal does not occur. It leads to the preliminary judgment that no failure occurs.

4.1.2. Spectrum Analysis of Data 2. The responding spectrograms are shown as in Figures 5, 6, 7, and 8. According to them, the trouble-free characteristic information can be determined and the working condition of system is normal.

4.1.3. Process Control Charts Analysis of Data 2 Based on Riemann Manifolds. At first, the phase space is reconstructed for each measured data. For comparison, let the embedding dimension be 3, and the delay time tau is determined by the autocorrelation method.

The process control diagrams based on Riemannian manifolds are shown in Figures 9 and 10. The diagrams show that the Riemannian distance \( d_g \) is less than a boundary value \( th \), leading to the trouble-free judgment of system. According to the time domain diagrams, spectrum diagrams, and Riemann control charts, there is not any failure or malfunction detected in Data 2.

4.2. Analysis of Data 3

4.2.1. Time Domain Analysis of Data 3. The time domain waveforms graphs of the four time series belonging to Data 3 are shown as in Figures 11, 12, 13, and 14. The time domain diagrams show that vibration signal produces an impact phenomenon, but the amplitude is small and the phenomenon is not obvious. The system working condition is hardly determined.

4.2.2. Spectrum Analysis of Data 3. The responding spectrograms are shown as in Figures 15, 16, 17, and 18. It can be seen that different simple harmonic waves appear on the measured points on the input shaft and the high speed shaft from the spectrum diagram. It is particularly serious on the measured point of the input shaft and the fault can be determined.

4.2.3. Process Control Charts Analysis of Data 3 Based on Riemann Manifolds. Control chart based on Riemannian manifold is shown in Figures 19 and 20. In Riemann control charts, the abnormal points appear on the intermediate shaft and the high speed shaft.

Considering the analysis of the time domain, frequency spectrum, and the Riemann control charts. Data 3 is judged in the fault state. The method can diagnose faults effectively and accurately, and the judgment is consistent with the experimental results.

5. Conclusion

Based on the sliding window, this paper uses the covariance matrix as a descriptor of the time sequence to realize the spatial filtering of the multidimensional time series and Riemannian distance as the similarity measure to realize the anomaly detection of multivariate time series by the assistance of the statistical process control charts in projects and the visualization of the distribution of covariance matrix. By the analysis of the failure data of wind turbine gearbox vibration testing, it is proved that the method based on Riemannian manifold can diagnose faults effectively and quickly.

Riemannian manifold method has a certain advantage in the treatment of multivariate time series, but the researches on fault diagnosis of mechanical system is less. The proposed method is the new method for fault diagnosis and expands the research scope. What is more, it also can be applied to the fault diagnosis of other rotating machinery.

The future research should combine this method with the advanced control algorithm (such as neural network, genetic algorithm, wavelet analysis, and Fourier transform), to eliminate, analyze, and process the data, leading to the more intuitive and effective results. In addition, the wind turbines are the complex mechanical transmission system. So the research on the feature information extraction and mechanism of the failure should be strengthened.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References


