

Trend Feature Extraction in Condition Monitoring by a New Piecewise Linear Representation Method

Changfeng Yan¹, Cheng Yi¹, Lixiao Wu¹ & Jianfang Fang²

¹School of Mechanical & Electrical Engineering, Lanzhou University of Technology, Lanzhou, 730050, China

²Cold-rolled Silicon Steel Factory, Shanxi Taigang Stainless Steel Co. LTD., Taiyuan, 030003, China

Keywords: Trend Feature Extraction, Condition Monitoring, Piecewise Linear Representation, Local Maximum Minimum and Slope.

Abstract. To extract the trend of condition monitoring data set effectively, piecewise linear representation is one of the most potential strategies. A new trend feature extraction method is presented based on the local maximum minimum and slope in this paper. The results of experiments are shown that it can extract the multiscale of trend feature and ensure acceptable fitting error. The process of extracting trend feature focuses on identification of rules and patterns and discovering the relevant and interesting information from condition monitoring system. It can provide a certain reference for condition trend analysis.

Introduction

Massive time series data sets are common place in a variety of online condition monitoring in mechanical engineering fields [1]. This kind of time series is huge, multidimensional, dynamic and complex. The detailed analysis of each and every time point in time series data is often neither practical nor necessary [2]. Approximate representation of the data to extract trend feature is to identify the significant patterns or models and extract the relevant trend information from database of condition monitoring. The accuracy of the trend extraction is the key factor for the condition monitoring system [3]. Because of the large number of signals monitored on a complex process, tuning a trend extraction method for each signal may be a tedious task. Thus, a trend extraction method would be well received by users if it requires little intervention to obtain accurate information from any signal. Parameters should either be easy to tune or ideally be self-tuned by using simple and useful methods. It is a tradeoff between the accuracy of the trend extraction and the tune of the parameter.

Several high-level representations of process trends have been proposed, including frequency domain representation [4], qualitative trend analysis [5], piecewise linear representation (PLR) [1]. In these representations, PLR is one of widely used methods because of its simplicity and practical. The trend can be reduced to the simplest representations by using 3 primitives only increasing, decreasing and steady. It can not only compress the data but also keep the primary trend and ignore trivial details.

Keogh et al. [6] proposed the algorithm of the piecewise aggregate approximation (PAA), which can be used to divide the datasets into the same time width and each subsequence is represented by the average value of the subsequence. Three most common PLR algorithms are as follows Top-Down, Bottom-Up and Sliding Window. Zhan et al. [7] proposed a new piecewise linear representation of time series based on slope extract edge point (SEEP). Charbonnier et al. [8] presented an effective trend extraction procedure based on segmentation algorithm. It is useful for complex system monitoring and decision support. Then, online abrupt change detection is added to the algorithm to detect step variations with precision and to remove artifacts [3]. The local maximum and minimum points of each subsequence are extracted to approximate the feature of local variations of the primary trend [9].

In this paper, trend feature is extracted based on a new piecewise linear representation method of Local Maximum Minimum and Slope (LMMS) from the database of condition monitoring. The

points with large fluctuation can be extracted by judging the variation amplitude of slope. They can reflect the feature of process trend accurately. The method is able to de-noise and not require prefiltering of the data. The extracted episodes emphasize the trend changes occurring on the data, and let the operator focuses on variant of process data.

Trend Feature Extraction Based on LMMS Algorithm

Algorithm Description. Condition monitoring datasets is a sequence set of elements which consists of sample value and sample time. It is represented as $X = \{x(t_1), x(t_2), \dots, x(t_n)\}$, where $x(t_i)$ is the sample value at the sample time t_i . The piecewise linear representation uses the straight line connecting the two endpoints of one segment to represent the data points in the subsequence and generates continuous piecewise trend lines. The key of the method is to extract the trend behavior of significant changes between two segmenting points in the subsequence, and represent approximately by a simple function with acceptable errors. Accuracy of trend feature extraction is measured by using the fitting error function. The fitting error of this algorithm can be calculated by Eq. (1).

$$E = \sqrt{\sum_{i=1}^n (x_i - x_i')^2} \quad (1)$$

where x_i is the original data sets, x_i' is the fitting sequence.

The scale of trend feature extraction of the data is measured by compression ratio which can be calculated by Eq. (2).

$$C = (1 - \frac{m}{n}) \times 100\% \quad (2)$$

where n is the length of original time series and m is the length of linear fitting sequence.

Integrating trend feature point of local maximum and minimum with the variation amplitude of slope, a new trend feature extraction method is proposed. Meeting the requirement of compression ratio and obtaining the least subsequences with the least amount of fitting error, the algorithm of PLR is reasonable. If the number of segments is too large, it is difficult to get the main trend of subsequence because of the noise and the accumulated numerical errors. On the other hand, if the number of segments is too small, the trend feature can become inaccurate and the fitting error can be very large. The most important question is how to choose the segment points which can reflect local variation trend under a suitable observed scale.

Extracting the local maximum and minimum points can reflect the global trend feature of subsequence. The data in the produced segment is always monotonically increased, decreased or steady. The trend feature points represent the change of condition. The large compression ratio can meet the requirement of storage. However there are some data models that adjacent data points constitute a pattern of extreme points with the little variation amplitude and adjacent data points are maximum or minimum points. It will lead to increase the fitting error of the method. It is necessary to judge the variation amplitude of slope, and extracted the feature points of reflecting accurate trend. The slope represents the inclination of straight line. The slope of two segments which are connecting the same point (except the two endpoints of the datasets) can be calculated. It can enhance the accuracy of trend feature extraction and obtain the true trend of mechanical condition.

Trend Feature Extraction Process. The flow chart of the trend feature extraction process can be shown in Fig.1.

Step 1: Load the condition monitoring data X and normalize it.

Step 2: Used the piecewise function and input the number of segments N based on the compression ratio, the subsequences are X_i .

Step 3: The maximum function and the minimum function is used to get the maximum x_{im} and the minimum x_{in} of each subsequence, then save them into the predefined array X'' in accordance with the order of the subscripts.

Step 4: Set the threshold value of variation amplitude of slope D . Add the starting point x_1 and end point x_n into the array of maximum and minimum X'' , one by one from the second point to

calculate the variation amplitude of slope and compare with the threshold value D , then save the point which is meeting the requirement into the predefined array of trend feature points X_s in accordance with the order of the subscripts. Tuning the number of segments N and the threshold value D , it will find the optimal solution of trend feature extraction.

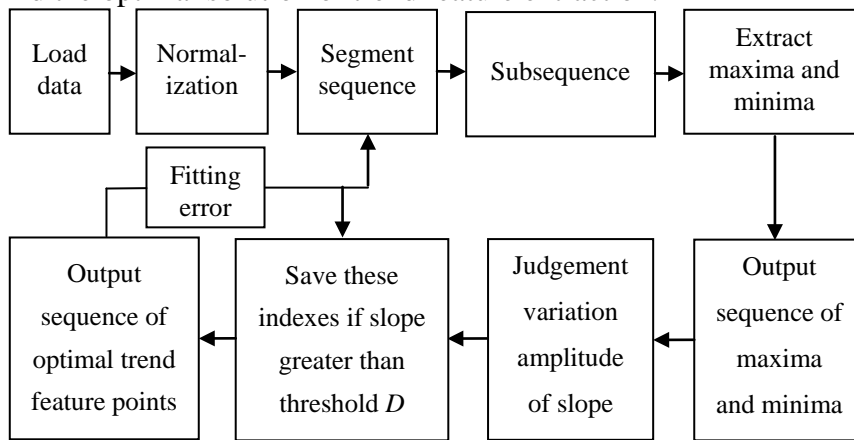


Fig.1: Flow chart of method

Experiment Results and Discussion

Experimental Data. The empirical comparisons are used to validate the performance of the method. Random samples from different areas of condition monitoring data are selected with different length. The bearing vibration dataset is the public roller bearing vibration signals from University of Brunel. The turbine vibration dataset is the shaft's amplitude of vibration on the steam turbine generator set in a power plant. All of them are shown in Fig.2(a)-2(b). The datasets in the experiment are sampled from different fields and the difference of numerical value is very large. In order to compare the results of experiment conveniently, the data values is normalized within range $[0, 1]$.

In order to analyze the characteristic of the method which is based on Local Maximum Minimum and slope to extract trend feature of condition monitoring dataset, the datasets in Fig.2 are used to compare their performance by the representative methods PAA, SEEP and LMMS under the different compression ratios.

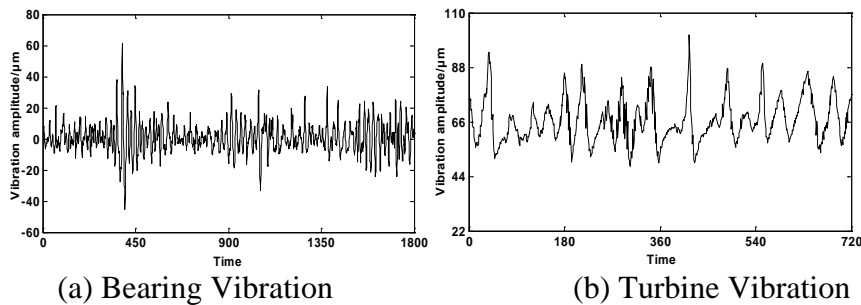
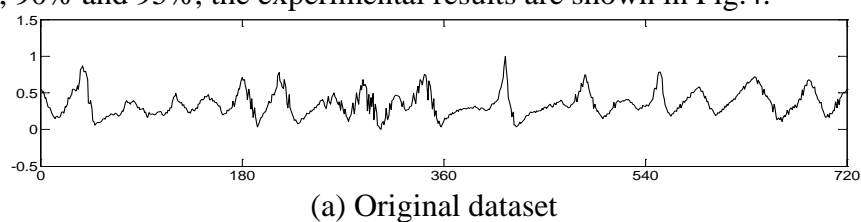


Fig.2: Experiment datasets

Results and Discussion. Under different compression ratios, the fitting error of three algorithms of PLR is compared at each dataset. With different compression ratios of multiscale trend feature extraction of turbine vibration dataset are shown in Fig.3. The value of the compression ratios are 75%, 80%, 85%, 90% and 95%, the experimental results are shown in Fig.4.



(a) Original dataset

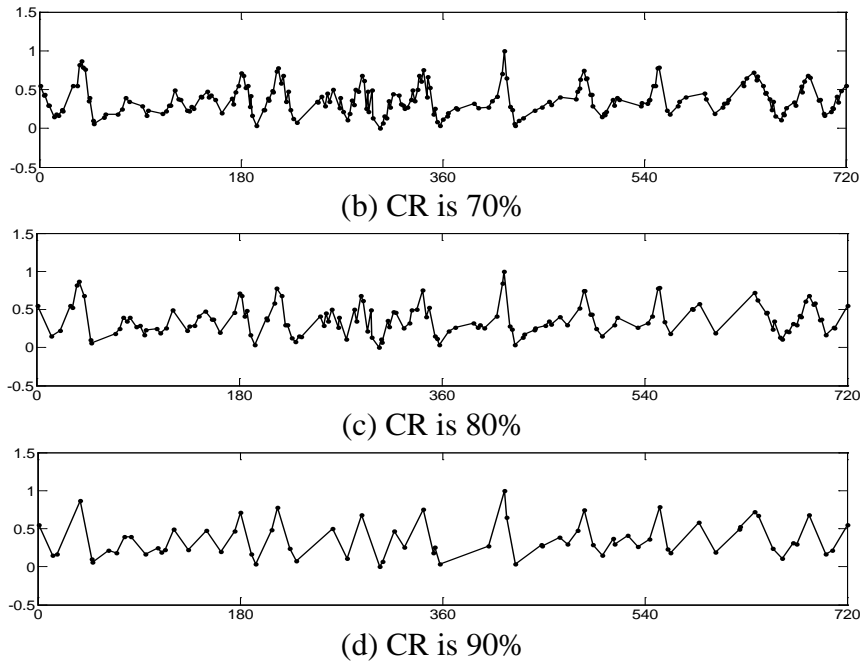


Fig.3: Different compression ratios of multiscale trend feature extraction

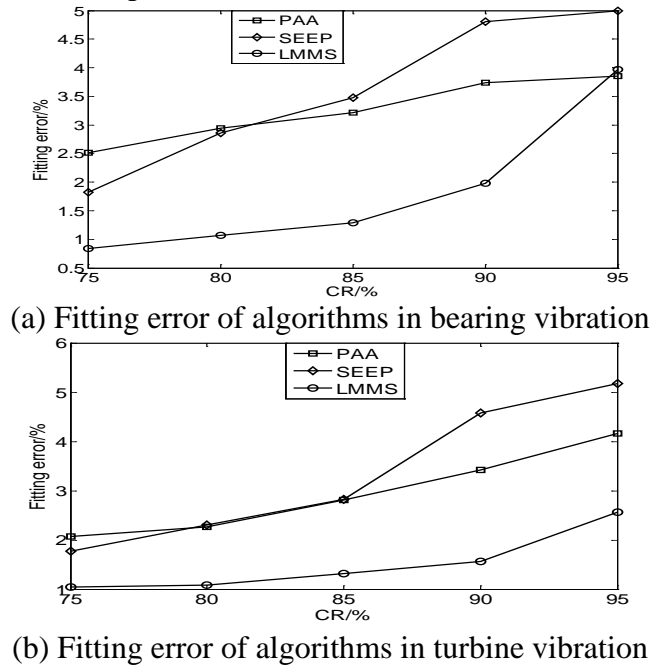


Fig.4: Fitting error of algorithms in different areas of time series data

The results show that the method can extract the trend feature of process datasets accurately and ensure a high fitting quality. The compression ratio will determine the scale of the trend. The trends provide a useful summary of what has happened on the process. An increasing or decreasing trend is detected if the variations observed on the signal are significantly higher than the level of the background noise. It can be helpful for the operator to supervise the process and prognose the trend evolution of mechanical condition. The fitting error increases as the compression ratio increases. It is a tradeoff between the fitting error of algorithm and the size of the compression ratio. The fitting error of the method is influenced by the whole volatility of the data set. According to the characteristics of the data and the requirement of compression ratio, it is a suitable algorithm to improve the accuracy and reliability of trend feature extraction.

The rotor misalignment dataset is the rotor's amplitude of vibration with the fault of misalignment. It is shown in Fig.5(a). The amplitude of rotor misalignment dataset change very quickly as well as the gear broken dataset. The fitting error may be very large using the PLR method to extract trend feature and the trend of amplitude is useless to fault diagnosis. So the trend

of the process can be described by time-frequency analysis. The vibration signal of the rotor misalignment is transformed into frequency signal by Hilbert-Huang Transform. The trend of the amplitudes of each frequency component can be extracted by the PLR method, which is important to fault diagnosis and prognosis. The amplitude changes of $2f$ are shown in Fig.5(b). Compared the performance of algorithm of PLR under the different compression ratios on the dataset, the result of experiment is shown in Fig.6. It is a good way to observe the trend of mechanical condition.

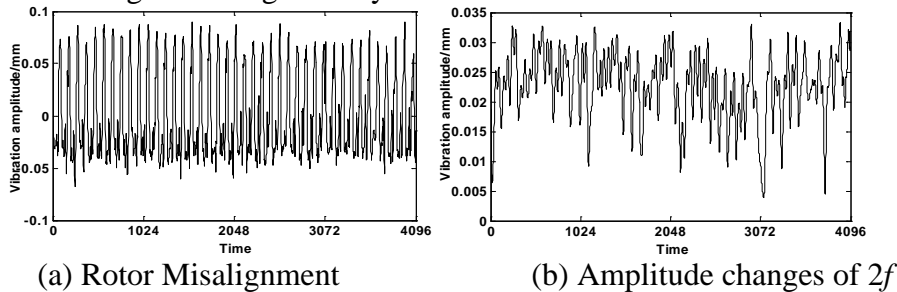


Fig.5: Experiment datasets

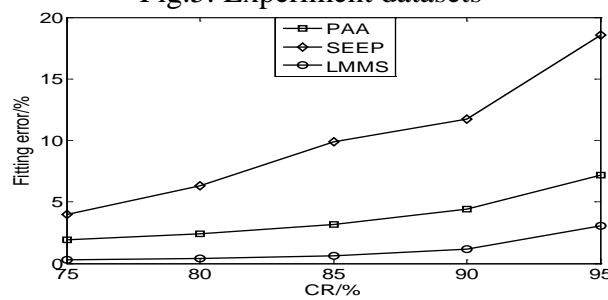


Fig.6: Fitting errors of algorithms

Conclusions

A new trend feature extraction method based on the method of local maximum minimum and slope was presented to filter the signal and to extract the trend from condition monitoring systems in this paper. The results of experiment showed that the method can extract the multiscale trend feature and ensure small fitting error. The fitting error was influenced by the compression ratio of requirement and the whole volatility of the data set. The trend of the high non-stationary vibration signal of each frequency component can be extracted to identify meaningful and interesting rules and patterns in condition monitoring data set. It is helpful to predict behavior of the observed variable in the future on the basis of history data. The efficiency and effectiveness of trend feature analysis might be improved by using the method.

Acknowledgements

This work was supported by Program of Natural Science Foundation of China (No. 51165018).

References

- [1] Keogh E. & Smyth P., A probabilistic approach to fast pattern matching in time series databases. *In Knowledge Discovery in Databases*, pp.24-30, 1997.
- [2] Liu X., Lin Z. & Wang H., Novel online methods for time series segmentation. *IEEE Transactions on Knowledge and Data Engineering*, 20(12), pp.1616-1626, 2008.
- [3] Charbonnier S. & Portet F., A self-tuning adaptive trend extraction method for process monitoring and diagnosis. *Journal of Process Control*, 22, pp.1127-1138, 2012.
- [4] Chan K. P. & Fu A. W. C., Efficient time series matching by wavelets. *In Proc. 15th International Conference on Data Engineering, Sydney, NSW, AUS*, pp.126-133,1999.

- [5] Maurya M. R., Paritosh P. K. et al., A framework for on-line trend extraction and fault diagnosis. *Engineering Applications of Artificial Intelligence*, 23(6), pp.950-960, 2010.
- [6] Keogh E., Chakrabarti K., Pazzani M., et al., Dimensionality reduction for fast similarity search in large time series databases. *Knowledge and information Systems*, 3(3), pp.263-286, 2001.
- [7] Zhan Y., Xu R. & Chen X., Time series piecewise linear representation based on slope extract edge point. *Computer Science*, 33(11), pp.139-142, 2006 .
- [8] Charbonnier S., Garcia-Beltan C., Cadet C., et al., Trends extraction and analysis for complex system monitoring and decision support. *Engineering Applications of Artificial Intelligence*, 18(1), pp. 21-36, 2005.
- [9] Yan C. F., Fang J. F., Wu L. X., et al., An Approach of Time Series Piecewise Linear Representation Based on Local Maximum Minimum and Extremum. *Journal of Information & Computational Science*, 10(9), pp.2747-2756, 2013.