Neurocomputing 138 (2014) 271–282

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Life grade recognition method based on supervised uncorrelated orthogonal locality preserving projection and K-nearest neighbor classifier

Feng Li [a,*], Jiaxu Wang [b], Baoping Tang [c], Daqing Tian [a]

[a] School of Manufacturing Science and Engineering, Sichuan University, Chengdu 610065, China
[b] School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China
[c] The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400030, China

ARTICLE INFO

Article history:
Received 1 July 2013
Received in revised form
3 November 2013
Accepted 20 January 2014
Communicated by J. Zhang
Available online 17 February 2014

Keywords:
Time–frequency domain feature set
Supervised Uncorrelated Orthogonal Locality Preserving Projection (SUOLPP)
Feature compression
K-nearest neighbor classifier (KNNC)
Life grade recognition
Rotating machine

ABSTRACT

A novel life grade recognition method based on Supervised Uncorrelated Orthogonal Locality Preserving Projection (SUOLPP) and K-nearest neighbor classifier (KNNC) is proposed in this paper. A time–frequency domain feature set is first constructed to completely extract the feature of different life grades, then SUOLPP is proposed to automatically compress the high-dimensional time–frequency domain feature sets of training and test samples into the low-dimensional eigenvectors with better discrimination, and finally the low-dimensional eigenvectors of training and test samples are input into KNNC to conduct life grade recognition. SUOLPP algorithm considers both local information and label information in designing the similarity matrix, and requires the output basis vectors to be statistically uncorrelated and orthogonal in order to improve the life grade feature extraction power of OLPP. KNNC ranks the test samples' neighbors among the training samples and uses the class labels of similarity neighbors to classify the unknown input test samples, so that it has such advantages as less calculation amount, finer timeliness and higher pattern recognition accuracy compared with support vector machine (SVM) and Fuzzy C-Means Clustering (FCM). The life grade recognition example on deep groove ball bearings demonstrated the effectivity of the proposed life grade recognition method.

1. Introduction

The high maintenance cost of mechanical equipments caused by ineffective or unresponsive failure monitoring for expensive rotating components (such as the bearing, gearbox and blades) has seriously affected the normal utilization of mechanical equipments [1–4]. Consequently, an effective life grade recognition in the complete life spans of rotating components is essential for mechanical equipments maintenance decisions which aim to reduce operating costs. In particular, a timely recognition of the life grade when a fault is to appear or has arisen should be guaranteed since it plays a more important role in avoiding safety accidents [5]. In the condition of complex physical laws, few training samples and heavy uncertain factors disturbance, life grade recognition methods based on information technology are applicable to solving life grade recognition problems for rotating machines, and more and more attentions have been paid to these methods recently [5]. However, these methods are somewhat immature at present for the following two reasons: firstly, they usually adopt the single or single domain signal analysis approach for life grade feature extraction so that it is very difficult to comprehensively dig nonlinear, weak and strongly coupled life grade feature of rotating machines [6]; secondly, life grade recognition methods based on information technology need manual analysis to accomplish the selection and optimization of life grade feature, which means both life grade feature extraction quality and life grade recognition accuracy are mainly determined by professional knowledge and field experience of engineers [7]. Therefore, it is quite hard to realize the high-precision and high-efficiency of life grade recognition using these methods. In view of this, a novel life grade recognition method based on Supervised Uncorrelated Orthogonal Locality Preserving Projection (SUOLPP) and K-nearest neighbor classifier (KNNC) is proposed in this paper. Notably, the feature compression with SUOLPP is the key technique that improves life grade recognition precision, and KNNC is crucial to the enhancement of computation efficiency.

In the existing life grade recognition methods based on information technology, the life grade signals, i.e. vibration signals, are first analyzed by a classical signal processing method such as

* Corresponding author. Tel.: +86 18382385401.
E-mail addresses: lifeng19820501@163.com (F. Li), wjx@scu.edu.cn (J. Wang).

http://dx.doi.org/10.1016/j.neucom.2014.01.037
0925-2312

short-time Fourier transform (STFT), Wigner–Ville distribution (WVD) [8], Wavelet decomposition [9], Empirical Mode Decomposition (EMD) [10], Autoregression (AR) model [11,12] and so on. However, these methods may not synchronously recognize different life grades due to the difficulty of comprehensively extracting different life grade features of rotating machines [6]. Thus, for completely digging the time–frequency domain feature of different life grades, 11 time domain feature parameters and 13 frequency domain feature parameters are proposed and combined into the 24-dimensional time–frequency domain feature set in this paper. 11 time domain feature parameters can characterize the size and distribution of vibrational amplitude and energy of life grade signals in the time domain; meanwhile, 13 frequency domain feature parameters can reflect the size and distribution of vibrational energy, the decentralization or centralization degree of frequency spectrum, and the position changes of main frequency band in the frequency domain.

In order to further improve life grade recognition precision and realize the high-efficiency of the life grade recognition process, it is necessary to further extract main eigenvectors with low dimensions, high sensitivity and good clustering from high-dimensional time–frequency domain feature sets that inevitably contain redundancy and disturbance components by an appropriate dimensionality reduction method. Classical dimensionality reduction approaches involve Multi-Dimensional Scaling (MDS) [13], Principal Component Analysis (PCA) and Kernel Independent Component Analysis (KICA) [14]. However, these approaches are only applicable to the datasets with linear structure and Gauss distribution, but not suitable for life grade data with nonlinear structure and non-Gaussian distribution [15]. Meanwhile, manifold learning is a new and effective theory of nonlinear dimensionality reduction, the purpose of which is to project the original high-dimensional data into a lower dimensional feature space by preserving the local neighborhood structure [16–19]. At present, typical manifold learning methods mainly include Linear Discriminate Analysis (LDA) [20], Locality Preserving Projection (LPP) [21], Orthogonal Neighborhood Preserving Embedding (ONPE) [22], Orthogonal Locality Preserving Projection (OLPP) [23], etc. However, these methods take either local structure or discriminant information separately into consideration rather than combine them together. Besides, the basis vectors obtained from LDA, LPP, ONPE and OLPP are all statistically correlated so that the features extracted by them usually contain some redundant information which may distort the distribution of features [15]. In this paper, based on LPP, we propose a novel manifold learning algorithm called Supervised Uncorrelated Orthogonal Locality Preserving Projection (SUOLPP) for dimensionality reduction of high-dimensional nonlinear time–frequency domain feature sets. SUOLPP considers both local information and class labels in designing the similarity weight matrix and requires the output basis vectors to be statistically uncorrelated and orthogonal, ensuring it can more effectively extract life grade features while compressing the time–frequency domain feature sets.

K-Nearest neighbor classifier (KNNC) is a novel supervised classifier which ranks the test samples' neighbors among the training samples and uses the class labels of similarity neighbors to classify input test samples [24]. Compared with other machine learning algorithms (e.g. support vector machine (SVM) and Fuzzy C-Means Clustering (FCM)), the great strength of KNNC lies in less calculation amount, finer timeliness and higher pattern recognition accuracy. In this paper, KNNC is introduced to establish the mapping relations between life grades and class labels.

Overall, we propose a novel life grade recognition method based on SUOLPP and KNNC. Through our method, the higher precision and efficiency of life grade recognition for rotating machines can be achieved. The rest of this paper is organized as follows. Firstly, the time–frequency domain feature set is constructed in Section 2. Then, SUOLPP algorithm and KNNC are introduced in Sections 3 and 4, respectively. In Section 5, life grade recognition experiment for deep groove ball bearings is carried out to evaluate the proposed method. Finally, some conclusions are drawn in Section 6.

2. Time–frequency domain feature set

Traditional signal processing approaches (e.g. STFT, WVD, Wavelet decomposition, EMD and AR model) are all unfit for life grade feature mining due to the following three reasons. First of all, they are mainly graphical spectrum analysis methods, but life grade recognition of rotating machines is essentially a pattern recognition problem which requires vector condition feature (i.e. eigenvector). Eigenvector is the main form of input data of pattern recognition methods, and choosing the eigenvectors with clear distributed geometry structure will help to improve pattern recognition precision. Secondly, these methods have their own intrinsic shortcomings and deficiencies. For example, STFT is not suitable for analyzing the common multi-scale signals because of its fixed time–frequency window size. WVD usually generates cross-term when analyzing multi-component signals. For wavelet transform, what criterion can be used to select wavelet basis function is still a difficulty in both theory field and application field [9]. EMD involves questions of modes mixture (i.e. failure of Intrinsic Mode Function (IMF) criteria) and end effects [25]. Also AR model coefficients are just applicable to the analysis of mutation signals with certainty, periodicity and energy aggregation [11]; whereas life grade signals of rotating machines, particularly normal state signals before fault appearance are all uncertain, nonlinear and nonstationary ones composed chiefly of background noise. Finally, these signal processing methods are all single or single-domain ones, their working alone presents severe difficulties in comprehensively extracting different life grade features of rotating machines in complex conditions.

Not difficult to discover, as the life grade of rotating machines varies, not only the time domain amplitude and its probability distribution of vibration signals but also the frequency components as well as their energy and spectrum peak position of vibration signals will respond to the change accordingly. Therefore, the amount and distribution of time domain waveform energy and its frequency spectrum can definitely represent the difference of life grades. Based on this, we construct 11 time domain feature parameters and 13 frequency domain feature parameters as shown in Table 1, and further combine them into the time–frequency domain feature set with 24 elements for the sake of digging the life grade features of rotating machines fully and faithfully.

In Table 1, $x(n)$ denotes time domain signal series, where $n=1,2,\ldots,N$, and N is sampling number; $s(k)$ represents the frequency spectrum of $x(n)$, where $k=1,2,\ldots,K$, and K is the spectrum line number; and f_k is the frequency of the kth spectrum line. Time domain feature parameters such as mean value p_1, square amplitude p_2, root-mean-square value p_3, and peak value p_4 reflect the amplitude and energy amount of time domain signal, and standard deviation p_5, skewness index p_6, kurtosis index p_7, peak indicator p_8, margin index p_9, waveform parameter p_{10}, and pulse index p_{11} reflect the time series distribution of time domain signal. Frequency domain feature parameter p_{12} characterizes the amount of frequency spectrum energy, p_{13}–p_{15} and p_{17}–p_{21} characterize the decentralization or centralization degree of frequency spectrum, and p_{16} and p_{22}–p_{24} characterize the position change of main frequency band.

In contrast with traditional signal processing approaches, the proposed 24-dimensional time–frequency domain feature set has three advantages as explained below: (1) its vector form is more popular with pattern recognition problem than traditional