Neurocomputing 166 (2015) 14–25

Contents lists available at ScienceDirect

Neurocomputing

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

Knowledge extraction using data mining for multi-class fault diagnosis of induction motor

Pratyay Konar [a], Jaya Sil [b], Paramita Chattopadhyay [a,*]

[a] Department of Electrical Engineering, Indian Institute of Engineering, Science and Technology, Shibpur, Howrah, India
[b] Department of Computer Science, Indian Institute of Engineering, Science and Technology, Shibpur, Howrah, India

A R T I C L E I N F O

Article history:
Received 5 September 2014
Received in revised form
30 January 2015
Accepted 11 April 2015
Communicated by Hongli Dong
Available online 25 April 2015

Keywords:
Induction motor
Vibration monitoring
Hilbert transform (HT)
Continuous wavelet transform (CWT)
Genetic algorithm (GA)
Rough set theory (RST)

A B S T R A C T

The feature extraction capability of rough set theory (RST) and genetic algorithm (GA) are used to extract knowledge from radial frame vibration signal for fault diagnosis of induction motor. This knowledge can assist in selecting scales for continuous wavelet transform (CWT) and mono-components required for Hilbert transform (HT) to extract fault related information from the vibration signal. Thus, the computational complexity of the signal processing tools is considerably reduced making both CWT and HT hardware friendly and suitable for real-time applications. For machine learning based automatic multi-class fault diagnosis, the performance of the classifiers are also considerably improved with significant reduction in computational burden since the redundant and irrelevant information can be effectively removed. The information obtained using data mining technique is successfully used to detect six types of induction motor faults. The results obtained are also verified in presence of high level of noise which has not been attempted earlier. The main contribution of the paper is to combine the advantages of two powerful signal processing tool like CWT and HT to extract hidden information from vibration signal in conjunction with data mining technique making them computationally efficient and easy to implement.

1. Introduction

Induction motors (IM) are extensively used in industries because of its simple structure, reliability, ruggedness, cost effective design and ease of control [1,2]. The faults in induction motors are due to mechanical and electrical stresses. With the growing demand of high quality production, the machine reliability has gained utmost importance. Consequently, condition monitoring is essential for reducing the maintenance cost and the unexpected failure of the induction motors [3]. Vibration analysis has been used to successfully identify mechanical and electrical faults and offers the most accurate fault diagnosis compared to any other technique [4]. However, automatic diagnosis of different induction motor faults by vibration analysis is still a challenging task.

At present with increasing use of power electronics ac drives in industries, the condition monitoring and fault diagnosis strategies demands intelligent monitoring techniques. The monitoring algorithms should be computationally efficient and hardware friendly to diagnose multiple induction motor faults. Thus, the artificial intelligence based techniques are now replacing the traditional techniques since these are good candidates for the automation of the induction motor diagnostic procedures. But in dealing with real world data even these techniques fails to take decision accurately since they are often vague and redundant [5].

In any AI based technique identifying the most characterizing features (or attributes) of the observed data has become very critical to minimize the classification error as well as computational complexity. Different data mining and feature selection techniques are used to reduce the number of features for fault classification [6,7]. There has been constant application of various feature selection tools to extract useful information for multi class fault diagnosis of induction motor.

Keeping the above viewpoints in mind the objective of this research work is to develop an efficient algorithm, suitable for online monitoring, by combining the advantages of present day signal processing techniques in conjunction with emerging machine learning and data mining techniques. In the proposed diagnostic strategy signal processing tools are used to extract the information hidden in the motor frame vibration signal. This information are reduced into useful knowledge with the help data mining tools and finally the detection of different kinds of motor faults is done with the help of various machine learning techniques.

1.1. Signal processing

Signal processing is the most important part of any condition monitoring and fault diagnosis strategy. To extract the incipient

* Corresponding author. Tel.: +91 9231664811.
E-mail addresses: pratyaymaithon@gmail.com (P. Konar), chattopadhyay@yahoo.com (P. Chattopadhyay).

http://dx.doi.org/10.1016/j.neucom.2015.04.040
0925-2312/

fault related features a good signal processing tool is required. The selection of a signal processing tools depends on the type of signal being analysed. As already mentioned motor vibration is one of the most popular choices for rotating electrical machines condition monitoring and fault diagnosis. Most signal processing tools typically involves frequency selective techniques, which are highly inadequate for analysis of non-stationary vibration signals. Moreover, the vibration signals collected from sensors are often contaminated by some noise which makes the extraction of the fault feature a bit challenging, especially at the early stage. The development of automatic diagnostics methods requires feature extraction. The incipient fault features is often weak and buried in the background signal, so it is difficult to detect them using the traditional signal processing methods [4]. Different induction motor faults are associated with different characteristic frequencies [6]. When a vibration problems occurs due to a fault, finding all the possible causes for the particular identified fault frequency of vibration needs good systematic and analytical approach. Moreover, some faults either electrical or mechanical in nature can cause vibration at the same or similar frequencies like the $1\times$RPM component and for a 2-pole induction motor, $2\times$line frequency and $2\times$RPM are very close especially on light load. Thus, identification of the root cause of vibration becomes elusive [5,6]. Further analysis is needed to distinguish such faults having similar vibration signatures like Bowed Rotor and Unbalanced Rotor. Research being carried out in this aspect is still in infancy. Another aspect in analysis of induction motor vibrations is that there are many harmonic components present in the electromagnetic force caused by electrical motors along with the slot frequency components. Thus, the motor vibration contains all the components arising from the cross products of the fundamental wave with itself and its harmonics and with slot frequency components. These harmonic components also induce the structural vibration and noise making accurate fault diagnosis more challenging.

There has been a substantial amount of research over the past 15 years on the development of condition monitoring of induction motors using various signal processing technique [7,8]. By reviewing the past work it was noted that the wavelet transform has emerged as one of widely used time-frequency analysis method in the area of fault diagnosis of induction motor [9] because of its ability to analyze non-stationary and transient signals. DWT is considerably easier to implement when compared to the CWT. In recent years significant research works is reported in literature on discrete wavelet transform (DWT) [10–13] based approaches for analyzing the vibration signals. DWT only concerns the information on some discrete time domains, whereas, CWT extracts detailed information more efficiently by choosing the proper scale parameter [14]. The CWT decomposition co-efficient thus obtained reflect the original signal in some particular aspects and can be used to extract quantitative information. CWT is more suitable for the feature extraction task in which one expects to obtain some transformation features, which can significantly differentiate different classes. Computation of continuous wavelet coefficients at every possible scale requires a lot of time and space [15]. Moreover, if the scales are not selected properly or selected at random, the resulting scales may emphasize one original aspect of data, but other aspects are inevitably lost. Thus, the selection of CWT scales that offers the optimum results is a challenging task [14,15]. A feature selection technique is necessary to find the most relevant scales that can correlate the wavelet coefficients obtained with the original signal. It is also important to note that apart from determination of scale/level, the mother wavelet also plays a very important role for fault feature extraction [13,16].

The wavelet transform has certain inherent deficiencies like border distortion and energy leakage that generate a lot of small undesired spikes all over the frequency/scales making the results confusing and difficult to interpret [17]. For detecting different types of fault which spreads over a wide frequency range, large size samples are required to be analyzed. Since, computation of wavelet transform is somewhat time-consuming it seems not suitable for large size data analysis. Another limitation of wavelet transform reported, is that it cannot achieve fine resolutions in both time domain and frequency domain simultaneously due to the limitation of Heisenberg–Gabor inequality [18]. Thus, although the wavelet transform has good time resolution in high frequency region, it often cannot separate those impacts, where time interval between them is very small. Thus, for a comparative study Hilbert transform (HT) is used instead of CWT because of its better time–frequency resolution [18]. However, for Hilbert transform all the mono-component signals have to be extracted from the vibration signal. Thus, for effective automatic fault diagnosis, prior information about the relevant fault frequencies is required, to extract the mono-component signals for Hilbert transform (HT).

1.2. Data mining and knowledge extraction

Like signal processing, the machine learning too plays an important role in developing an efficient automatic multi-class fault diagnostic strategy [19]. Without assistance of effective knowledge from the signal processing the good classifiers cannot make a generalized robust fault diagnostic system. A large feature space with vague and redundant information always create problem to the classifiers to take decision accurately and increases the computational complexity. A good data mining technique is thus required to analyse this data set and summarize it into useful information or knowledge. The number of features and computational complexity can thus be considerably reduced along with considerable improvement of the classifier performance. It has been successfully used by Rafiee et al. [20] to find the order of '*daubechies*' wavelet function, decomposition level of wavelet packet and number of hidden layer of the neural network for detection of gearbox fault with the help of genetic algorithm (GA). Data-mining has also gained importance in other fields as well. Recently, rough-set theory (RST) has been successfully used for dimensionality reduction. Hui-Ling Che et al. [21] has used RST to determine the most informative features for classifying the breast cancer. RST is also used to extract the faulty features by Ning Li et al. [22] for detection of various fault conditions of gearbox and valve trains on a gasoline engine. Some of the recent research works [22–27] are furnished in Table 1. However, the feature extraction capability of RST and GA has never been used to find the fault frequencies which can assist in selecting relevant CWT scales or mono-components for HT to extract the important fault related information. Thus, the knowledge obtained using RST/GA can be effectively used to diagnose different induction motor faults and also reduce computational complexity.

2. Proposed methodology

The Schematic representation of the proposed fault diagnosis technique is shown in Fig. 1. Continuous wavelet transform (CWT) and Hilbert transform (HT) are used in this work to extract information from motor frame vibration signal. In order to remove the redundant and irrelevant information from the vast feature pool of CWT and HT and to extract useful fault related knowledge, RST and GA based data mining techniques have been used here. The general framework of the algorithm used to find the relevant features to diagnose multi-class induction motor faults is illustrated in Fig. 2.

2.1. Experimental setup and data acquisition

Spectra quest's machinery fault simulator (MFS) [28] as shown in Fig. 3 is used here as the experimental set-up to extract the vibration