Expert Systems with Applications 40 (2013) 5787–5795

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Anomaly detection of cooling fan and fault classification of induction motor using Mahalanobis–Taguchi system

Xiaohang Jin, Tommy W.S. Chow *

Centre for Prognostics and System Health Management, City University of Hong Kong, Hong Kong, China
Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China

ARTICLE INFO

Keywords:
Anomaly detection
Cooling fan
Fault classification
Induction motor
Mahalanobis–Taguchi system
Vibration signal

ABSTRACT

A health index, Mahalanobis distance (MD), is proposed to indicate the health condition of cooling fan and induction motor based on vibration signal. Anomaly detection and fault classification are accomplished by comparing MDs, which are calculated based on the feature data set extracted from the vibration signals under normal and abnormal conditions. Since MD is a non-negative and non-Gaussian distributed variable, Box–Cox transformation is used to convert the MDs into normal distributed variables, such that the properties of normal distribution can be employed to determine the ranges of MDs corresponding to different health conditions. Experimental data of cooling fan and induction motor are used to validate the proposed approach. The results show that the early stage failure of cooling fan caused by bearing generalized-roughness faults can be detected successfully, and the different unbalanced electrical faults of induction motor can be classified with a higher accuracy by Mahalanobis–Taguchi system. Such works could aid in the reliable operation of the machines, the reduction of the unexpected failures, and the improvement of the maintenance plan.

1. Introduction

Rotary-machines (such as cooling fans and induction motors) play an important role in modern society. Cooling fans are widely used for thermal management in electronic products. Schroeder and Gibson (2007) reported that cooling fan was one of the top 10 failing components in electronic products. Induction motors, which convert electrical energy into mechanical energy, are the critical component in industrial equipment. Failure of induction motors can lead to their host system breakdowns, loss of production and income, and even casualties (Chow & Hai, 2004; Li, Chow, Tipsuwan, & Hung, 2000; Motor Reliability Working Group, 1985; Thorsen & Dalva, 1999). Therefore, fault diagnosis in these two rotary-machines is important, and plays a key role for the reliable operation of them and their host system, reducing the unexpected failures, and improving the maintenance.

Research on fault diagnosis of rotary-machines has gained increasing interests world-wide, and is still a hot topic today, as evidenced by Prognostics and System Health Management conferences over recent years. Several types of signals are used for this purpose. For examples, acoustic emission signals are analyzed to estimate the degree of cooling fan bearing degradation (Oh, Azarian, & Pecht, 2011); sound pressure level could differentiate a new and failed cooling fan effectively (Oh, Shibutani, & Pecht, 2012); motor current signals are used to monitor the health condition of induction motor (Niu et al., 2008); and oil-based approaches are used to predict the residual life of plant (Wang, 2009). Although there are many other methods developed for fault diagnosis of rotary machines based on different signals, vibration signal analysis is the most common, effective, and reliable method (Caesarendra, Niu, & Yang, 2010; Chow & Hai, 2004; Gan, Zhao, & Chow, 2009; Randall & Antoni, 2011; Wang, Tse, Guo, & Miao, 2011). Based on vibration signals, the signal processing approaches such as time domain analysis (Heng & Nor, 1998; Martin & Honarvar, 1995), frequency domain analysis (Courrech, 2000; Miao, Cong, & Pecht, 2011; Miao, Azarian, & Pecht, 2011), and time–frequency domain analysis (Tse, Peng, & Yam, 2001; Wang, Zi, & He, 2009), are used for fault diagnosis. In this study, vibration signals from cooling fan and induction motor are also measured and analyzed. Features sets that can reveal the characteristics of time domain and frequency domain of vibration signals are constructed. A health index, Mahalanobis distance (MD), is used and enhanced by Mahalanobis–Taguchi system (MTS) for anomaly detection and fault classification. Then, anomaly detection and fault classification are done by comparing MDs of normal and abnormal conditions.

* Corresponding author at: Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China. Tel.: +852 34424829; fax: +852 34420272.

E-mail addresses: xiaohajin2-c@my.cityu.edu.hk (X. Jin), eetchow@cityu.edu.hk (T.W.S. Chow).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.04.024

MTS combines the MD and Taguchi methods to deal with multivariate problems. It has been successfully applied into different areas, such as damage detection in civil engineering, bankruptcy and financial crisis prediction, and health condition checking (Cheung et al., 2008; Cho, Hong, & Ha, 2010; Lee & Teng, 2009; Taguchi & Jugulum, 2002; Wang, Su, Chen, & Chen, 2011; Yang & Cheng, 2010). Recently, MTS and MD are introduced into fault diagnosis and prognosis of rotary-machines in the following papers: Soylemezoglue, Jagannathan, and Saygin (2011) presented an MTS-based fault diagnosis and prognosis scheme for centrifugal pump failures; Wang, Wang, Tao, and Ma (2012) used MTS to diagnose bearing single-point faults. Jin, Ma, Cheng, and Pecht (2012) presented a health monitoring scheme based on MD with minimum redundancy maximum relevance feature selection. To our best knowledge, it is the first time to use MTS for anomaly detection of cooling fans due to bearings generalized-roughness faults, and fault classification of the unbalanced electrical faults of induction motor.

The main contributions of this paper can be summarized as below: (1) A health index, MD, is proposed to indicate the health condition of cooling fan and induction motor; (2) MTS is employed to make the proposed health index robust; (3) The effectiveness of the proposed health index is validated by two experimental data sets. Results show that early stage of failure of cooling fan caused by bearing generalized-roughness faults can be detected successfully, and different unbalanced electrical faults of induction motor can be classified at a higher accuracy by using MTS.

The rest of this paper is organized as follows. In Section 2, MTS is briefly reviewed, and the way of determination of the ranges of MDs corresponding to different health conditions are introduced. The proposed anomaly detection and fault classification approach is introduced in Section 3. Anomaly detection of cooling fan is reported in Section 4, while fault classification of induction motor is presented in Section 5. Finally, the conclusions are drawn in Section 6.

2. Mahalanobis–Taguchi system

MTS is a pattern recognition technology that is widely used for data classification (Taguchi & Jugulum, 2002). It combines the MD and Taguchi methods together. MD is a generalized distance that is useful for determining the similarities between unknown and known sample data sets. It uses a scalar value to represent a multivariate system. Taguchi methods are statistical methods used to improve engineered quality (Taguchi & Jugulum, 2002) and make the system more robust (Taguchi, Chowdhury, & Wu, 2001).

2.1. Four steps in MTS

Generally, there are four steps in a MTS, as shown in Fig. 1 (Taguchi & Jugulum, 2002).

Step I: Mahalanobis space (MS) construction

Feature data from the healthy products are collected to form the normal data set, and their MDs constitute a reference space that is also known as the MS. Their MDs are around one. In this study, feature data set constructed from the vibration signal from normal rotary machines is used to form the MS.

The normal data set is denoted as P; p_{ij} is the ith observation on jth feature, where $i = 1, 2, \ldots, m$, and $j = 1, 2, \ldots, n$. $\overline{P_j}$ and S_j are the mean and the standard deviation, respectively, of the jth feature (P_j), where $j = 1, 2, \ldots, n$. Each individual feature in each data vector (P_j) is normalized by the mean ($\overline{P_j}$) and the standard deviation (S_j). Thus, the normalized values are as follows:

$$z_{ij} = \frac{p_{ij} - \overline{P_j}}{S_j}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n \tag{1}$$

where

$$\overline{P_j} = \frac{1}{m}\sum_{i=1}^{m} p_{ij} \tag{2}$$

$$S_j = \sqrt{\frac{\sum_{i=1}^{m}(p_{ij} - \overline{P_j})^2}{m-1}} \tag{3}$$

The MDs of the normal dataset are calculated using the following equation:

$$\mathrm{MD}_i = \frac{1}{n} z_i C^{-1} z_i^T \tag{4}$$

where $z_i = [z_{i1}, z_{i2}, \ldots, z_{in}]$, z_i^T is the transpose vector of z_i, and C^{-1} is the inverse of the covariance coefficient matrix C. C is calculated as:

$$C = \frac{1}{m-1}\sum_{i=1}^{m} z_i^T z_i \tag{5}$$

One issue faced by MD is multicollinearity (strong correlations among features). The problem of multicollinearity will lead to an approximate singular covariance coefficient matrix, an inaccurate inverse matrix of the covariance coefficient matrix, consequently, an inaccurate MD (Taguchi & Jugulum, 2002). MD corresponding to the adjoint matrix of the covariance coefficient matrix, which is denoted as MDA, can be used to handle this problem.

$$\mathrm{MDA}_i = \frac{1}{n} z_i C_{adj} z_i^T \tag{6}$$

where C_{adj} is the adjoint matrix of the covariance coefficient matrix C. Since $C^{-1} = C_{adj}/|C|$, the relationship between MD and MDA is shown as below.

$$\mathrm{MD}_i = \frac{1}{|C|}\mathrm{MDA}_i \tag{7}$$

Step II: Validation of MS

Observations of abnormal conditions are selected out first. Their feature data sets are normalized using the mean and standard deviation of the normal data set. Then their MDs are calculated using the normalized feature data and the covariance coefficient matrix of the normal data set. MDs corresponding to the abnormal conditions will be out of the MS, if the MS is appropriately constructed. In other words, these abnormal conditions associated MDs will have higher values.

Step III: Identify the useful features

The useful features are selected out using orthogonal arrays (OAs) and signal-to-noise ratios (S/N ratios). In MTS, OAs are used to identify the important features by minimizing the different combinations of the original set of features. The number of columns in OA depends on the number of features. Two-level factors are used: Level-1 means including the feature, while Level-2 means excluding the feature. S/N ratios, which are calculated using abnormal conditions only, are used to measure the accuracy of the MS for predicting. The formula for calculating the S/N ratio (η_i) corresponding to the ith run of the OA is defined as below:

$$\eta_i = -10\lg\left(\frac{1}{t}\sum_{j=1}^{t}\frac{1}{\mathrm{MD}_j}\right) \tag{8}$$

where t is the number of abnormal conditions, and MD_j is the MD of the jth abnormal condition.