Data Fusion and Multi-fault Classification Based On Support Vector Machines

GAO Guohua[1], ZHANG Yongzhong[2], ZHU Yu[1], DUAN Guanghuang[1]

[1] Department of Precision Instrument & Mechanology at the Tsinghua University, Beijing, China, ghgao@mail.tsinghua.edu.cn
[2] Mechitrone Engineering College at the China University of Mining & Technology, Xuzhou, Jangsu, China, zyz@sina.com.cn

Abstract

As a new general machine-learning tool based on structural risk minimization principle, Support Vector Machines (SVM) has the advantageous characteristic of good generalization. For this reason, the application of SVM in fault diagnosis field has becomes one growing reach focus. In this paper, data fusion strategy based on multi-class SVMs is proposed to diagnose the gear fault. The fault features extracted from vibration signals with various analysis methods are transferred into the SVM in the feature fusion level. The signal analysis methods include Power Spectrum, Cepstrum, wavelet, etc. Data fusion improves the reliability of the diagnosis results. The SVM is originally designed for two-class classification. In order to satisfy the need of multi-fault classification, one three-class SVMs is built to combine the outputs of the feature fusion levels and to classify the four fault states of the gear. The actually diagnosis results show that the fault classification performance of the multi-class SVMs is evidently powerful and precise.

Keywords: Support vector machines, data fusion, multi-fault classification, fault diagnosis

1. Introduction

The complexity of engineering systems is permanently growing due to their growing size and the degree of automation, and accordingly increasing is the danger of failing and aggravating their impact for man and the environment. Therefore, increased attention has to be paid to reliability, safety and fault tolerance. [1]

Following the development of the information technique, artificial intelligence, computer technique, fault diagnosis has become an important interdisciplinary subject. The process of fault diagnosis commonly includes fault detection and classification.

Fault feature extraction is the key step in the fault detection process. Combining the relevant fault information in one most efficient manner for improving the veracity and reliability of the diagnosis is important and necessary. Multi-sensor Data fusion refers to intelligent processing of an array of 2 or more sensors that have cooperative, complimentary, and competitive qualities. As long as the sensor array does not contain totally independent sensors, arrays usually contain various levels of these three qualities. Cooperative sensors are those that work together to create a new piece of diagnostic information, while a complimentary array creates a more complete picture of a problem. Finally, a competitive array provides unrelated measurements of the same physical phenomena for improved reliability. [2] In this paper, feature fusion is realized with the theory of Support Vector Machines (SVM). SVM is a new generation learning system based on recent advance in statistical learning theory.

The purpose of the diagnosis process is to realize the fault types classification and identification. In fault classification task, SVM has been found to give better generalization. Comparing with other conventional classifiers, SVM has the advantage that so-called structural misclassification risk is to be minimized in training the SVM, whereas conventional classifiers are usually trained so that the empirical risk is minimized. [3, 4] One multi-fault classifier is built based on SVM theory in this paper. The classifier is composed of three sub-SVMs and has the ability to identify four states of the gear.

This paper is organized as follows. Based on the brief review of the SVM theory in Section 2, the structures of feature fusion level and multi-class SVMs classifier are proposed in Section 3. In Section 4, according to the characteristic that gear faults show different features in various signal analysis methods, one feature fusion level is built based on SVM. In Section 5, the multi-fault classification is realized with one three-class SVMs. Finally, some important conclusions are drawn in Section 6.

2. Review of SVM

The SVM are introduced by Vapnik in the late 1960s on the foundation of statistical learning theory. It can be considered to create a line or a hyperplane between two sets of data for classification. The hyper-plane is defined by a number of *support vectors*, which are a subset of the training data available for both cases, and is used to define the boundary between the two classes. Complex boundaries can be created with the support vectors. [3, 4, 5]

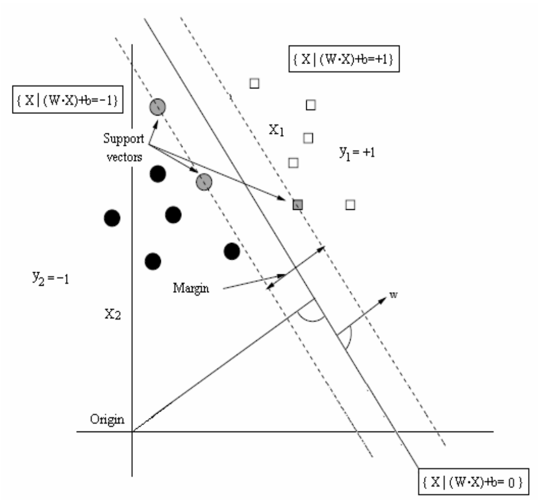

Fig. 1. Classification of data by SVM

In case of two-dimensional situation, the action of the SVM can be explained easily without any loss of generality. In Fig. 1, a series of data points for two different classes of data are shown, class A (white squares) and class B (black circles). The SVM attempts to place a linear boundary between the two different classes, and orientate it in such a way that the *margin* (represented by the dotted lines) is maximized. In other words, the SVM tries to orientate the boundary in such a way as to ensure that the distance between the boundary and the nearest data point in each class in maximal. The boundary is then placed in the middle of this margin between the two points. The nearest data points are used to define the margin, and are knows as *support vectors* (represented by the grey circles and square). Once the support vectors have been selected, the rest of the featureset is not required, as the support vectors contain all the information need to define the classifier.

The boundary can be expressed in terms of

$$(\mathrm{w}\cdot\mathrm{x})+b=0\,,\quad \mathrm{w}\in R^N\,,\qquad b\in R \qquad (1)$$

where the vector w defines the boundary, x is the input vector of dimension N and b is a scalar threshold. At the margins, where the SVs are located, the equations for class A and B, respectively, are

$$(\mathrm{w}\cdot\mathrm{x})+b=+1 \qquad (2)$$

$$(\mathrm{w}\cdot\mathrm{x})+b=-1 \qquad (3)$$

As SVs correspond to the extremities of the data for a given class, the following decision function holds good for all data points belonging to either A or B:

$$f(x)=\mathrm{sign}((\mathrm{w}\cdot\mathrm{x})+b) \qquad (4)$$

We can scale w and b so that this is equivalent to

$$y_i((\mathrm{w}\cdot x_i)+b)\geq 1,\qquad i=1,\ldots,l \qquad (5)$$

where l is the number of training sets.

The optimal separating hyperplane is the one which the margin defined by the distance between the hyperplane and the closest point of S is maximum. Moreover, it has the smallest generalization error.

To maximize the margin the task is therefore:

$$\text{minimize}\quad \Phi(\mathrm{w})=\frac{1}{2}(\mathrm{w}\cdot\mathrm{w}) \qquad (6)$$

Using a Lagrangian, this optimization problem can be converted into maximizing a

$$L=\frac{1}{2}(\mathrm{w}\cdot\mathrm{w})-\sum_{i=1}^{l}\alpha_i(y_i(\mathrm{w}\cdot x_i-b)-1) \qquad (7)$$

where α_i are Lagrangian multipliers.

The derivatives with respect to w and b give:

$$\frac{\partial L}{\partial b}=\sum_{i=1}^{l}y_i\alpha_i=0 \qquad (8)$$

$$\frac{\partial L}{\partial \mathrm{w}}=\mathrm{w}-\sum_{i=1}^{l}\alpha_i y_i x_i=0 \qquad (9)$$

From (7), (8) and (9), get the dual representation of the optimization problem:

$$L=\sum_{i=1}^{l}\alpha_i-\frac{1}{2}\sum_{i=1}^{l}\alpha_i\alpha_j y_i y_j(x_i x_j) \qquad (10)$$

which must be maximized with respect to the α_i subject to the constraint:

$$\sum_{i=1}^{l}\alpha_i y_i=0\ ,\quad \alpha_i\geq 0 \qquad (11)$$

When the maximal margin hyperplane is found, only points which lie closest to it have $\alpha_i>0$ and these points are called Support Vectors (SVs). The resulting decision function is obtained as follows:

$$f(x)=\mathrm{sign}\Big(\sum_{i=1}^{l}y_i\alpha_i^{*}(x\cdot x_i)+b^{*}\Big) \qquad (12)$$

where α_i^{*} is the solution of the constrained maximization problem.

3. The structures of feature fusion level and multi-class classifier based on SVM

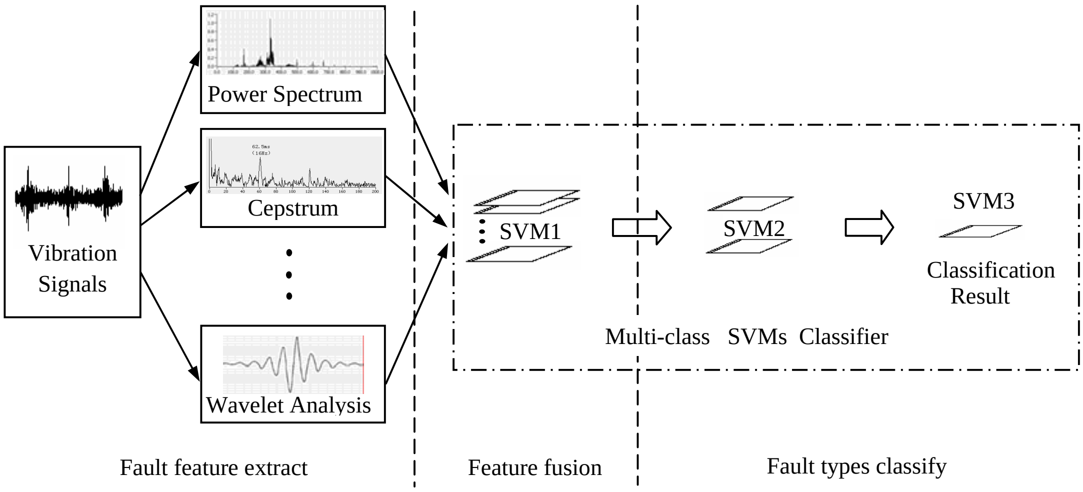

Fig. 2. The structures of feature fusion level and multi-class classifier based on SVM

In order to improve the veracity and reliability of the fault diagnosis results, it's necessary and important to collect and integrate the information from the different sources. The fundamental fusion of data, recorded from a multiple sensor system, to obtain a more precise perception of the physical phenomena under analysis is known as data fusion.

The data fusion can take place at different levels of representation, namely: [3]

1) Raw data fusion at the signal/pixel fusion, where the raw data is robustly and redundantly merged or sensors are validated.

2) Feature fusion at the feature level, where a characteristic is extracted before fusion occurs.

3) Decision fusion at the symbol level, where measured data with or without pre-processing is combined with processed data or a priori knowledge.

In this paper, the data fusion is realized at the feature fusion level shown in Fig. 2.

At first, the vibration signals sampled from the different sensors are processed by several signal analysis methods, such as Power Spectrum, Cepstrum and Wavelet analysis, etc. The feature values obtained by these analysis methods include the amplitudes on the specific frequency (gear rotating frequency, gear mesh frequency, harmonic frequencies, etc.) in Power Spectrum, the specific rotating period in Cepstrum and the Wavelet packet decomposed components energy values, etc. The detail about the fault feature extracting and feature fusion process will be indicated in Section 4.

Then, these fault feature values obtained through above methods are all transferred into the SVM1 and feature fusion is realized.

Multi-fault classification always is the research focus in the fault diagnosis field. There are many techniques for fault classification and identification, such as Dempster-Shafer combination, Artifical Nerual Network and Rough Set Theory. As one growing new technique, the SVM theory has the special advantage than these conventional methods.

As described in Section 2, SVM is originally developed to perform binary classification. In order to fill the need of the classification of data into more than two classes, multi-class SVMs from binary SVM have been proposed in many ways.

When the direct extension of a binary method into a multi-class one is not possible, a general strategy to build multi-class classifiers based on a set of binary classifiers is always possible, the so called *error correcting output coding* (ECOC) strategy, originally proposed by Dietterich and Bakiri in (Dietterich and Bakiri, 1995). [4]

In this paper, there have four gear states to be classified including gear surface pitting, gear fracture tooth, gear fatigue wear and normal state. Base on the ECOC strategy, three-class SVMs have been built shown in Fig. 2.

4. Feature fusion based on SVM

Gear mechanisms are the important element in various mechanical systems. Vibration signal analysis is always the main diagnosis method for gear fault. Conventional methods include crest factor, kurtosis, Power Spectrum, Cepstrum, time-domain averaging and demodulation, which are now well established and have proved to be very effective in fault diagnosis. Their main drawback is that they are based on the assumption of stationary of the analysed vibration signal. For the non-stationary of the gear fault vibration signals, Wavelet Analysis possesses particular advantages because of its characteristic that signals at different localization levels in time is as well as frequency domain.

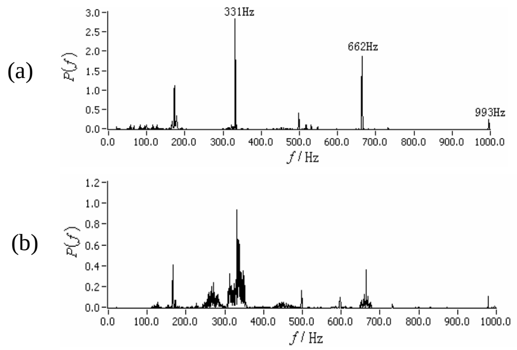

Fig. 3. The comparison of the Power Spectrums between two gear fault types. (a) fatigue wear. (b) fracture tooth.

Fig.3 shows the comparison of the Power Spectrums between two gear fault types, fatigue wear and fracture tooth. It's clear that two fault both stimulate the amplitudes at the frequencies of gear mesh frequency and its harmonics. So it's difficult to classify these two fault types only from Power Spectrum.

Fig. 4 shows the comparison of the Cepstrums between these two fault types. For the fracture tooth fault, there's a pulse at the 307*ms* (3.25Hz, it's the gear rotating frequency.). The reason for this phenomenon is that the facture tooth stimulates the modulated harmonics with gear rotating frequency on the both sides of the gear mesh frequency. Because the fatigue wear is a kind of distributed fault, it will not cause the modulated harmonics. According to this comparison and integrating the Power Spectrums, it's feasible to classify these two fault types.

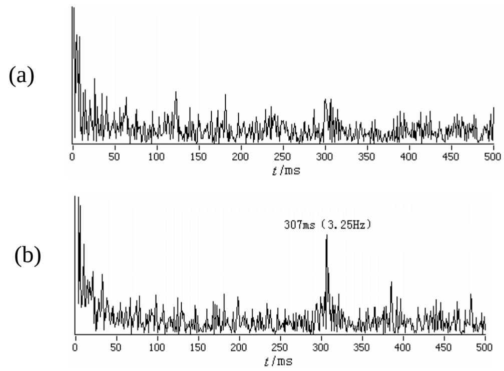

Fig. 4. The comparison of the Cepsrums between two gear fault types. (a) fatigue wear. (b) fracture tooth.

For the gear surface pitting, Wavelet Packet Transform (WPT) is an effective feature extracting method. Fig. 5 shows the comparison of the WPT components between two gear states, normal state and gear surface pitting fault. When pitting appears on the gear surface, the impacts caused by the pits will stimulate the vibration of the gear system in the gear natural frequency (in this paper, the natural frequency of the fault gear is 2210Hz). Compared with the normal gear, the energy of the high frequency part of the pitting gear is enhanced. According this phenomenon, using WPT, the vibration signal can be decomposed into different frequency bands. The energy changes of those frequency bands (W6) including the gear natural frequency will be the evidence for the pitting identification.

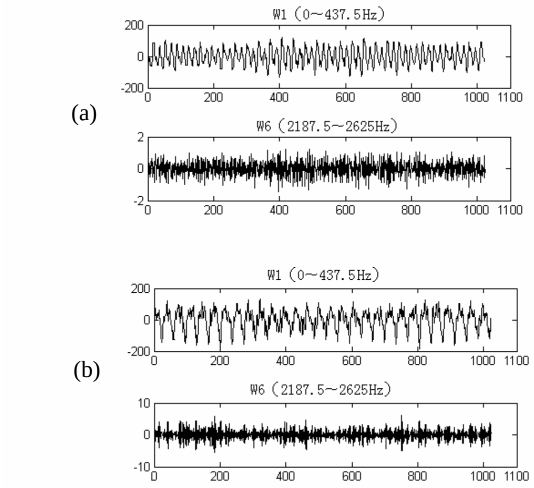

Fig. 5. The comparison of the WPT components in W1 and W6 frequency bands between two gear states. (a) normal state. (b) gear surface pitting.

In this paper, the specific values of the Power Spectrum, Cepstrum and WPT are obtained as the extracted features for the SVM fusion level.

5. Multi-fault classification based on SVM

There are four gear states to be classified in current work, normal state and three fault states: surface pitting, fracture tooth and fatigue wear. According to the characteristics of gear fault vibration signals, the features of these fault types have been extracted with different signal analysis methods narrated in Section 4. Based on the ECOC strategy (Section 3), three-class SVMs is developed to classify these four states and shown in Fig. 6.

The first SVM (SVM1) is trained to separate the gear surface pitting fault from other three gear states. When the feature input is a surface pitting fault state sample, the output of SVM1 is set to -1; otherwise +1.

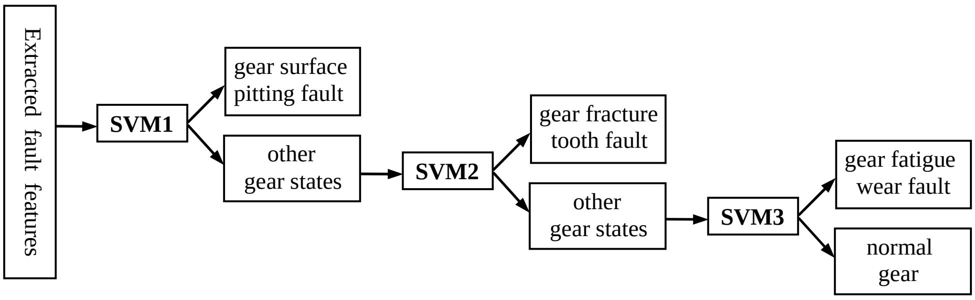

Fig. 6 The scheme of gear multi-fault classification based on multi-layer SVMs

The second SVM (SVM2) is trained to classify the fracture tooth fault and other two gear states. When the feature input is a fracture tooth sample, the output of SVM2 is set to -1; otherwise +1.

The third SVM (SVM3) is trained to classify the gear fatigue wear fault and normal gear state. When the feature input is a fatigue wear sample, the output of SVM3 is set to -1; otherwise +1 indicates the gear normal state.

The SVMs training experiments are conducted on a small data set (20 vibration signals include 5 signals for four different gear states apiece). The classification results are shown in table 1and and show the veracity and reliability of the fault diagnoses.

Table 1. Multi-fault classification

signals SN	actual fault style	SVM 1	SVM 2	SVM 3	classification results
1-5	surface pitting	-1			surface pitting
6-10	fracture tooth	+1	-1		Fracture tooth
11-15	fatigue wear	+1	+1	-1	fatigue wear
16-20	normal state	+1	+1	+1	normal state

6. Conclusion

As a new general machine-learning tool for classification, SVM is powerful for the practical problem with small sampling, nonlinear and high dimension.

Based on SVM theory, a method for gear multi-fault classification is proposed in this paper. The fault features extracted by various signal analysis methods are transferred into SVM and realize the feature fusion. This improves the reliability of the fault diagnosis. One three-class SVMs classifier is built to identify the gear fault types and the performance is excellent.

The results of the experiment prove that the veracity of the SVM classifier is very significant for the multi-fault classification.

7. References

[1] Jozef Korbicz Jan, M. Koscielny, *Fault diagnosis: models, artificial intelligence, applications.* Berlin; New York: 2004.

[2] Brooks, R. R., Iyengar, S. S, *Multi-Sensor Fusion.* Upper Saddle River, New Jersey: 1998

[3] Poyhonen S., Arkkio Antero, Jover P., Hyotyniemi H., "Coupling pairwise support vector machines for fault classification", *Control Engineering Practice*, v 13, n 6, pp. 759-769, 2005.

[4] Samanta B., "Gear fault detection using artificial neural networks and support vector machines with genetic algorithms", *Mechanical Systems and Signal Processing*, v 18, n 3, pp. 625-644, May 2004.

[5] Hannah, P., Starr, A., Ball, A., "Decisions in state monitoring-an examplar for data fusion architecture", *Information Fusion, Proceedings of the Third International Conference*, V1, pp. 10-23, July 2000.

[6] Pakorn Watanachaturaporn, Pramod K. Varshney; Manoj K. Arora, "Multisource Fusion for Land Cover Classification using Support Vector Machines", *2005 7th International Conference on Information Fusion*, pp. 614-621, 2005.

[7] Hu Zhonghui, Cai Yunze, He Xing, "Fusion of multi-class support vector machines for fault diagnosis", *Proceedings of the American Control Conference*, v 3, pp. 1941-1945, 2005.

[8] Yen, Gary G., Lin, Kuo-Chung, "Wavelet packet feature extraction for vibration monitoring", *IEEE Transactions on Industrial Electronics*, v 47, n 3, pp. 650-660, 2000.

[9] Wowk Victor. , *Machinery vibration: measurement and analysis.* New York: 1991