http://dx.doi.org/10.5755/j01.eee.18.8.2635 ELEKTRONIKA IR ELEKTROTECHNIKA, ISSN 1392-1215, VOL. 18, NO. 8, 2012

Fault Signal Classification using Adaptive Boosting Algorithm

Pei Yao, Zhenbao Liu, Zhongsheng Wang, Shuhui Bu
School of Aeronautics, Northwestern Polytechnical University,
Xi'an 710072, Shaanxi, P.R.China, phone: 86 29 88492344
liuzhenbao@nwpu.edu.cn

Abstract—In recent years, researchers seldom investigate how to boost the classification performance of any learning algorithm for fault signal detection. We propose a fault signal classification method based on adaptive boosting (adaboost) in this paper. Adaboost is able to select an optimal linear combination of classifiers to form an ensemble whose joint decision rule has relatively high accuracy on the training set. First, we extract statistical features from sample signals. And then we make use of a decision tree to identify optimal features, which are used to classify the sample set by adaboost algorithm. To verify its accuracy, we set up the roller bearing experiment. Practical results show that the method can precisely identify fault signals, and be comparable to SVM based traditional method.

***Index Terms*—Classification algorithms, decision trees, fault diagnosis, feature extraction.**

I. INTRODUCTION

Roller bearing is one of the most widely used in rotating machinery. Condition monitoring of such an element is greatly advantageous for economical and security value, and can be regarded as a pattern recognition problem [1]. A basic framework of a roller bearing is composed of rolling elements, inner and outer races, and a cage. The main failure mode of rolling bearings is localized defects, which is dislodging of a sizable piece of contact surface during operation as a result of fatigue cracking under cyclic contact stressing in the bearing metal [2]–[11].

According to the traditional signal processing methods, we might calculate several symptom values under each condition, such as the wavelet energy coefficients [10], the Fourier coefficients [11], the statistical measurements [8], and the frequencies obtained from envelop analysis. However, these symptoms cannot be automatically recognized by computers. Researchers began to be interested in investigating a machine learning method for monitoring the bearing conditions. Jayanta [7] and Samantha [8] use Artificial Neural Network (ANN) to recognize fault types. Some researchers use SVM [5], [6], [9] and Petri Nets [4] as the classifier because of the high accuracy and good generalization capability.

Although several fault diagnosis systems have been proposed, the development of systems is still very limited. Moreover, identification accuracy of existent methods needs to improve.

In this letter, we propose a fault signal classification method based on adaptive boosting (Adaboost). Boosting is a general methodology for improving the performance of any given learning algorithms, and adaboost is a representative one of boosting algorithms, which is firstly introduced by Freund and Schapire [3], and which is a machine learning meta-algorithm for performing supervised learning.

II. ROLLER BEARING FAULT DIAGNOSIS

The block diagram of proposed fault signal diagnosis system is shown in Fig. 1. There are three main phases: feature extraction, feature selection (dimensionality reduction) and classification. In our experiment, we set up a monitoring system of roller bearing where a dataset from a vibration signal was used. The dataset included about 4 conditions including normal condition, inner race defect, ball defect, and outer race defect. And then, we divided signal length on each condition into 25 pieces (samples). The set of database consists of 100 samples in total. Each feature vector was built by computing the 12 statistical features for each sample. We used a decision tree to select important features for classification, and the selection results composed the input of the classifier. Finally, we use an adaboost algorithm for classification.

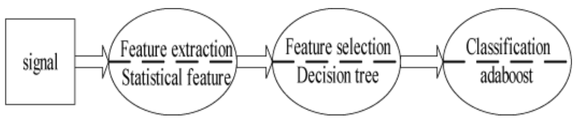

Fig. 1. The block diagram of the proposed method.

III. FEATURE EXTRACTION

The vibration signal of the machine is a typical nonstationary random signal. The statistical features can be considered as the descriptors, which can reflect the change tendency of non-stationary random signal. Several statistical parameters are adopted such as mean, standard error, median, standard deviation, variance, kurtosis, skewness, range, minimum, maximum, and sum [8]. In our research, we added the steepest gradient, which denotes the maximal change rate in neighborhood. The gradient (or gradient vector field) of a scalar function $f(x_1, x_2, \cdots x_n)$ is defined as the vector field

Manuscript received March 19, 2012; accepted May 10, 2012.
This work was partially supported by the National Natural Science Foundation of China (No. 60472116, 61003137 and 61202185) and NWPU fund (No. JC201202, JC201220).

whose components are the partial derivatives of f. That is

$$\nabla f = (\partial f / \partial x_1, \partial f / \partial x_2, \cdots, \partial f / \partial x_n). \quad (1)$$

Table I shows the whole mathematical expressions of the statistical parameters.

TABLE 1. MATHEMATICAL EXPRESSION OF STATISTICAL PARAMETERS.

Statistical parameter	Mathematical expression
Mean	$\bar{x} = \frac{1}{n}\sum_{i=1}^{N} x_i$
Median	$x_{med} = \begin{cases} x_{(N+1)/2} & N \in odd \\ \frac{1}{2}(x_{N/2} + x_{(N/2+1)}) & N \in even \end{cases}$
Standard error	$\sqrt{\frac{1}{n-2}[(y-\bar{y})^2 - \frac{[\sum(x-\bar{x})(y-\bar{y})]^2}{\sum(x-\bar{x})^2}]}$
Standard deviation	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$,where $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$
Kurtosis	$kur = \frac{\frac{1}{n}\sum_{i=1}^{N}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{N}(x_i - \bar{x})^2)^2} - 3$
Skewness	$skew = \frac{\frac{1}{n}\sum_{i=1}^{N}(x_i - \bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{N}(x_i - \bar{x})^2)^{3/2}}$
Range	$range = x_{\max} - x_{\min}$

IV. FEATURE SELECTION

The goal of feature selection is to find such a lower dimensional subset that holds most of the information (suitable for classification or regression). In order to improve the efficiency of adaboost classification, we selected the optimal features from the candidate features in advance by an implementation of J48 decision tree.

The construction criterion of decision tree for feature selection is as follows:

The extracted features are the input of the feature selection algorithm, and the output is a generated decision tree;

Each node represents a subset of classes, which will be partitioned successively in the child nodes;

Each leaf node is associated with a class label;

The branch of the tree denotes a threshold, which originates from the attribute;

In the decision tree, the optimal features are selected by an importance criterion of each node. The importance criterion is based on the theories of entropy reduction and information gain. More details of entropy reduction and information gain are introduced below.

For a sample set S, the information gain of an attribute A is defined as follows:

$$G(S, A) = Entropy(S) - \sum_{v \in value(A)} \frac{|S_v|}{S} Entropy(S_v) \quad (2)$$

The first term denotes the entropy of original sample set S; entropy reflects the homogeneity of the set and is defined as follows:

$$Entropy(S) = \sum_{i=1}^{c} -P_i \log_2 P_i, \quad (3)$$

where c is the number of classes, P_i is the distribution on the proportion of S belonging to class label i.

The next term is the expected value of the entropy using attribute A. $Value(A)$ is the set of all possible values for attribute A, and S_v is the subset of S for which attribute A has value v. $||$ is defined as the number of sample set.

V. CLASSIFICATION BOOSTING

The adaboost algorithm achieves strong classifier by combining many weak learners h_t

$$H(x) = \sum_t \alpha_t h_t(x), \quad (4)$$

where α_t denotes the weight coefficient, and in our study the weak learner is defined as a decision stump. The detailed implementation of adaboost is as follows: First, the weights are initialized $\omega_i = 1/N, i = 1, \cdots N$, and algorithm enters into main loop. Second, we fit a discriminant function $f_m(x) \in \{-1, +1\}$ using ω_i, and then compute learning error $err_m = E_w[I_{y \neq f_m(x)}]$, the weight $\alpha_m = c \cdot \log((1 - err_m)/err_m)$, where c is a constant $0 < c < 1$. ω_i is successively updated $\omega_i = \omega_i \cdot \exp(\alpha_m \cdot I_{(y(i) \neq f_m(x))})$, and the main loop ends after updating. Finally, the strong classifier is realized by a linear combination of weak learners as $H(X) = sign[\sum_{m=1}^{M} \alpha_m f_m(x)]$. The complete algorithm of adaboost is given below:

1) The training set $S_C = \{(x_i, y_i), i = 1, \cdots, N\}$, Where $x_i \in X$ a set of N feature vectors. $y_i \in Y = \{+1, -1\}$, the desired class labels.

2) Initial the sample weights $\omega_0 = 1/N$.

3) For $t = 1, \cdots, T$, Do

a) Train one weak hypothesis $g_m(x_i)$, $m = 1, \cdots, M$ for each feature vector, the training set was sampled according to probability distribution $\omega_{t,i}$, $i = 1, \cdots, N$.

b) Choose the hypothesis g_t with the lowest classification error ε_t.

c) Update the sample weights:

$$\omega_{t+1,i} = \frac{1}{Z_t} \omega_{t,i} e^{-\alpha_t g_t(x_t) y_t}, \quad (5)$$

$$g_t(x_i) = \{+1, -1\}, \tag{6}$$

where $g_t(x_i)$ is correctly or incorrectly classified by the weak hypothesis g_t

$$\alpha_t = 0.6\log(\frac{1-\varepsilon_t}{\varepsilon_t}) \tag{7}$$

and Z_t is the normalizing constant so that ω_{t+1} is the distribution.

4) End

$$f_c(X) = \sum_{t=1}^{T} \alpha_t g_t(X). \tag{8}$$

VI. EXPERIMENTAL RESULTS AND DISCUSSION

In our experiment, we use the 1200ATN self-aligning ball bearing as experimental subject, equipped in test rig of ZHX-3A rotor. The parameters of roller bearing are shown in Table II.

TABLE II. PARAMETERS OF THE ROLLER BEARING.

ORR	IRR	PD	BD	BN	CA
R_0/mm	R_i/mm	D/mm	d/mm	N_b	α /°
15.00	5.00	19.89	4.75	9	11

Note: ORR(Outer Race Radius); IRR(Inner Race Radius); PD(Pitch Diameter); BD(Ball Diameter); BN(Ball Number); Contact Angel(CA)

The vibration signals from YDI-12 sensor are acquired by NI PXI4472 DAQ (data acquisition), which is equipped in PXI-1042Q industrial computer. The sampling frequency is 10 kHz. The sampling length is 8192 for the speed of 900RPM in all conditions. Fig. 2 shows the vibration signals from roller bearing in 4 conditions including normal condition, inner race defect, ball defect, and outer race defect. In this study, we utilize 12 statistical features for each sample, which are reduced to 5 features by J48 decision tree, as shown in Fig. 3. Successively, we implemented adaboost for classification. The iteration number of adaboost algorithm is empirically set 20. The test details are given by comparing classification of one versus one:

1. Normal condition vs inner race defect;
2. Normal condition vs ball defect;
3. Normal condition vs outer race defect;
4. Inner race defect vs ball defect;
5. Inner race defect vs outer race defect;
6. Ball defect vs outer race defect.

We summarized the experimental results as follows.

We integrated twelve statistical features, including mean, standard error, median, standard deviation, variance, kurtosis, skewness, range, minimum, maximum, sum and an introduced feature- steepest gradient.

Decision tree reduces the dimensionality of statistical features, and selects optimal features for the next classification. In the experiment, we found that the decision tree selected five features, including standard variance, mean, min, median, and the steepest gradient due to their significant contributions. Fig. 3 gives the visualized result.

Referring to Table III and Table IV, different speeds (600RPM, 900RPM) for the same test identities are experimented. The reason of using two different speeds is that magnitude of vibration changes is proportional to shaft speed and hence the statistical features changes with speed.

We engaged the adaboost algorithm for classification, and the classification accuracy under different speeds is given in Table III and Table IV. And meanwhile, SVM based fault signal classification[5] is taken for comparison in order to prove the superiority of our algorithm. Here the kernel function of SVM is radial basis function.

The result indicates that the accuracy of the proposed method is superior to that of SVM based fault signal classification.

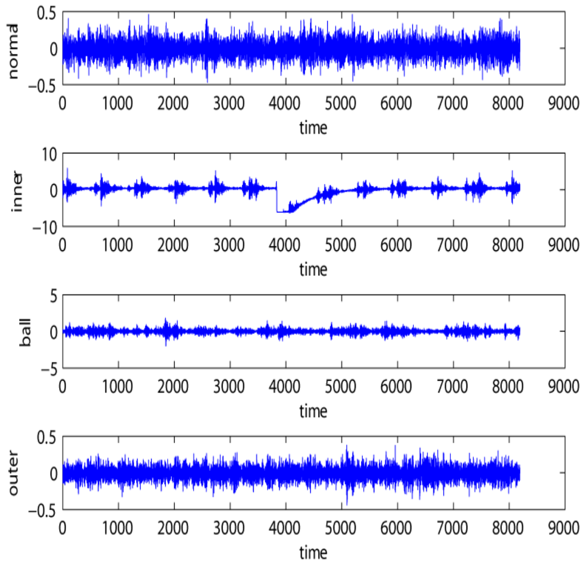

Fig. 2. The time-domain signals at the speed of 900 PRM taken from roller bearings in 4 conditions, including the normal condition, the inner race defect, the outer race defect, and the ball defect.

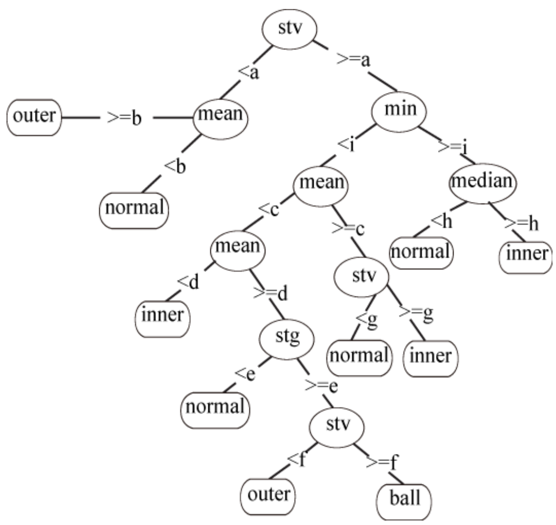

Fig. 3. Five optimal features are selected by using J48 decision tree at the speed of 900RPM. Stv is the abbreviation of standard variance, Stg is steepest gradient respectively. Here, a=0.0103496, b=-0.0074633, c=-0.0024877, d=-0.17447, e=0.26575, f=0.022598, g=0.0284284, h=-0.005, i=-0.4345.

TABLE III. CLASSIFICATION ACCURACY (600RPM).

method	normal vs inner	normal vs ball	normal vs outer
adaboost	100/92	100/96	100/96
SVM	100/88	100/96	92/83

method	inner vs ball	inner vs outer	ball vs outer
adaboost	100/96	100/100	100/96
SVM	96/96	96/88	100/92

Note: training accuracy (%) / test accuracy (%)

TABLE IV. CLASSIFICATION ACCURACY (900RPM).

method	normal vs inner	normal vs ball	normal vs outer
adaboost	100/92	100/96	100/92
SVM	100/88	100/96	92/79

method	inner vs ball	inner vs outer	ball vs outer
adaboost	100/96	100/92	100/96
SVM	96/96	96/83	100/92

Note: training accuracy (%) / test accuracy (%)

VII. CONCLUSIONS

Fault diagnosis of roller bearing is a challenging issue in the field of condition monitoring of rotary machines. The theory and experiment in this paper have demonstrated the ability to make an efficient fault diagnosis for machine system.

REFERENCES

[1] X. Shi, Z. Xu, “A Study on the automatic recognition of vibration signal for ball bearing faults-the FFTAR feature extraction and classification methods”, in *Proc. of the IEEE ATSA-World Scientific*, 1988, pp. 318–321.

[2] C. Li, S.Wu, “Online detection of localized defects in bearing by pattern recognition analysis”, *ASME Journal of Engineering for Industry*, vol. 111, no. 4, pp. 331–336, 1989. [Online]. Available: http://dx.doi.org/10.1115/1.3188768

[3] Y. Freund, A. Schapire, “Decision-theoretic generalization of on-line learning and an application to boosting”, *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997. [Online]. Available: http://dx.doi.org/10.1006/jcss.1997.1504

[4] N. Pamuk, Y. Uyaroglu, “Modeling of Fault Diagnosis in Power Systems using Petri Nets”, *Elektronika ir Elektrotechnika* (*Electronics and Electrical Engineering)*, no. 2, pp. 63–66, 2012.

[5] V. Sugumaran, V. Muralidharan, K. Ramachandran, “Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing”, *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 930–942, 2007. [Online]. Available: http://dx.doi.org/10.1016/j.ymssp.2006.05.004

[6] S. Avidan, “Support vector tracking”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 1064–1072, 2004. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2004.53

[7] B. Jayanta, K. Rajat, K. Sankar, “Unsupervised feature selection using a neuro-fuzzy approach”, *Pattern Recognition Letters*, vol. 28, no. 11, pp. 997–1006, 1998.

[8] B. Samantha, “Artificial neural network based fault diagnostics of rolling element bearings using time domain features”, *Mechanical Systems and Signal Processing*, vol. 17, no. 2, pp. 317–328, 2003. http://dx.doi.org/10.1006/mssp.2001.1462

[9] D. Mattera, F. Palmieri, S. Haykin, “An explicit algorithm for training support vector machines”, *Signal Processing Letters*, vol. 6, no. 9, pp. 243–245, 1999. [Online]. Available: http://dx.doi.org/10.1109/97.782071

[10] B. Paya, I. Esat, M. Badi, “Artificial neural network based fault diagnosis of rotating machinery using wavelet transforms as a preprocessor”, *Mechanical Systems and Signal Processing*, vol. 11, no. 5, pp. 751–765, 1997. [Online]. Available: http://dx.doi.org/10.1006/mssp.1997.0090

[11] B. Li, M. Chow, Y. Tipsuwan, J. Hung, “Neural-network-based motor rolling bearing fault diagnosis”, *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1060–1064, 2000. [Online]. Available: http://dx.doi.org/10.1109/41.873214