J Intell Manuf (2013) 24:1213–1227
DOI 10.1007/s10845-012-0657-2

Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network

Zhenyou Zhang · Yi Wang · Kesheng Wang

Received: 14 December 2011 / Accepted: 26 April 2012 / Published online: 12 May 2012

Abstract This paper proposes a method for classification of fault and prediction of degradation of components and machines in manufacturing system. The analysis is focused on the vibration signals collected from the sensors mounted on the machines for critical components monitoring. The pre-processed signals were decomposed into several signals containing one approximation and some details using Wavelet Packet Decomposition and, then these signals are transformed to frequency domain using Fast Fourier Transform. The features extracted from frequency domain could be used to train Artificial Neural Network (ANN). Trained ANN could predict the degradation (Remaining Useful Life) and identify the fault of the components and machines. A case study is used to illustrate the proposed method and the result indicates its higher efficiency and effectiveness comparing to traditional methods.

Keywords Diagnosis · Prognosis · Wavelet packet decomposition · Fourier transform and artificial neural network

Z. Zhang · K. Wang (✉)
Department of Production and Quality Engineering, Norwegian University of Science and Technology, S. P. Andersensveien 5, Valgrinda, 7491 Trondheim, Norway
e-mail: Kesheng.wang@ntnu.no

Z. Zhang
e-mail: Zhenyou.zhang@ntnu.no

Y. Wang
School of Materials, The University of Manchester, Sackville St Blg, Manchester M13 9PL, UK
e-mail: yi.wang-2@manchester.ac.uk

Introduction

During a system failure, only a small fraction of the downtime is spent to maintain or repair the components that cause the fault. Up to 80 % of it is spent to locate the source of the fault (Kegg 1984). In case of complex installation such as automotive manufacturing plant, 1 min downtime may cause as high as $20,000 cost (Spiewak et al. 2000). Early fault diagnosis is crucial for avoiding major malfunction and massive loss in economy and productivity. In diagnosing rotating machinery, sound emissions or vibration signals are used to monitor the performance of the machine and could be used to judge whether the machine is failure or degrading. Many useful techniques for signal analysis have been applied. These techniques can be classified into three types: time domain (Wang et al. 2010; Chen et al. 2008), frequency domain such as Fast Fourier Transform (Corinthios 1971; Liu et al. 2010; Rai and Mohanty 2007) and time-frequency domain such as the Short Time Fourier Transform (Portnoff 1980), Hilbert-Huang Transform (Yu et al. 2007), Wigner-ville distribution (Andria et al. 1994; Staszewski et al. 1997; Wang et al. 2008) and Wavelet Transform (Chen et al. 2005; Lin and Qu 2000; Prabhakar et al. 2002; Serhat and Emine 2003; Tse et al. 2004; Wu and Chen 2006; Zheng et al. 2002; Wu and Liu 2009). Autoregressive model method can also be used to extract features of a machine or component for fault diagnosis and prognosis (Li et al. 2009). Wavelet transform is the best of these tools because the short time Fourier transform only provides a constant time-frequency resolution, and Wigner-ville distribution produced interface terms on the time-frequency domain in a critical condition (Wu and Chen 2006). It has particular advantages for characterizing signals at different localization levels in time as well as signal processing, image processing, pattern recognition, seismology and machine fault diagnosis.

After processing vibration signals and extracting the features, the more important thing is identifying the fault and predicting the remaining useful life. There are many methods could be used in this area. Support vector machine (SVM) learning is a popular machine learning application due to its high accuracy and good generalization capabilities (Saravanan et al. 2008). Li and Wu (2005) proposed a hidden Markov model (HMM)-based fault diagnosis in speed-up and speed-down process for rotary machinery. In the implementation of the system, one PC was used for data sampling and another PC was used for data storage and analysis. Wu and Chow (2004) presented a self-organizing map (SOM) based radial-basis-function (RBF) neural network method for induction machine fault detection. The system was implemented by utilizing a PC and additional data acquisition equipment. Many methods based on ANN have been developed for online surveillance with knowledge discovery, novelty detection and learning abilities (Kasabov 2001; Marzi 2004; Markou and Singh 2003). ANN, Fuzzy Logic System (FLS), Genetic Algorithms (GA) and Hybrid Computational Intelligence (HCI) systems were applied in fault diagnosis and a case of centrifugal pump was utilized to show how the methods work (Wang 2002). Decision tree method was used to identify fault in of mean shifts in bivariate processes in real time (He et al. 2011). Probability based Bayesian network methods was used to identify vehicle fault which can be used to diagnose single-fault and multi-fault (Huang et al. 2008). Lee et al. (2006) developed an intelligent prognostics and e-maintenance system named "Watchdog Agent" with the method of Statistical matching, and performance signature and Support Vector Machine (SVM) based diagnostic tool.

There exist some literatures integrating these techniques for fault diagnosis and prognosis. Momoh and Button integrated FFT and ANN to analyze and identify the fault of aerospace DC arcing (Momoh and Button 2003). Fourier transform and wavelet transform were integrated to detect and identify the fault of induction motor using stator current information (Lee 2011). Wavelet analysis techniques and ANN were integrated for fault diagnosis in induce motors (Lee 2011), automotive generator (Wu and Kuo 2009) and gear box (Saravanan and Ramachandran 2010) and the results were pretty good. However, there are no literatures integrating three techniques of WPD, ANN as well as FFT in condition monitoring or fault diagnosis and prognosis. This paper proposes a method applying these three techniques for fault classification and prediction. This method extracts features using wavelet transform and Fourier transform from pre-processed vibration signals and, then these features could be used to train SBP neural network which could classify and predict fault, and further to predict the remaining useful life. These results can be used to support the maintenance decision making and optimizing the scheduling. This paper is organized as the following: "Data acquisition experiment" presents the experiment setup of data acquisition; wavelet packet transform and Fourier transform are introduced briefly in "Features extraction" and "Fault diagnosis and prognosis integrating WPD, FFT and ANN" respectively; a case study and its results with discussions are presented in "Case study"; the conclusions and future research are presented in last section.

Data acquisition experiment

To research how to diagnose the fault type and prognose the condition of the monitored equipment, a simple experimental setup is established in Knowledge Discovery Laboratory (KDL) in NTNU.

Experimental setup

Figure 1 shows the hardware of the experimental setup which includes a blower, three vibration sensors, power supply for sensors, connector, DAQ card and a computer. In this setup, the blower is selected as our monitoring object and a kind of vibration sensor (Kistler: Type 8702B100) is chosen to collect the signals from the blower. Three sensors are setup on the blower in three directions which can collect the vibration signals in different directions (Fig. 2). The signals are collected from the sensors and processed using some processing method like filter, de-noising and compression. Then the features are extracted in different domain which can be used to train and query ANN. After training, the system can judge the real states of monitored components using real time signals.

Experimental procedure

In the present study, four different degradations of unbalance are simulated using three different parts (Fig. 3) which are mounted in the axis end of the blower. The unbalance degradation (condition) contains 0, 0.3, 0.6 and 1 which represent the performance states from perfect (condition 0) to absolutely failure (condition 1). In the first case, power on the blower, collect and store signals from the sensors without amounting any simulation part. Next, power off the blower and mount first part in the axis end and power on the blower, then collect and store the signals from sensors. Repeat this process until collect all the degrading signals simulated by simulation parts. Figure 4 shows the signals from perfect state to absolutely failure.

Features extraction

Features extraction is very significant in fault diagnosis and prognosis process. In this paper, WPD and FFT are used to

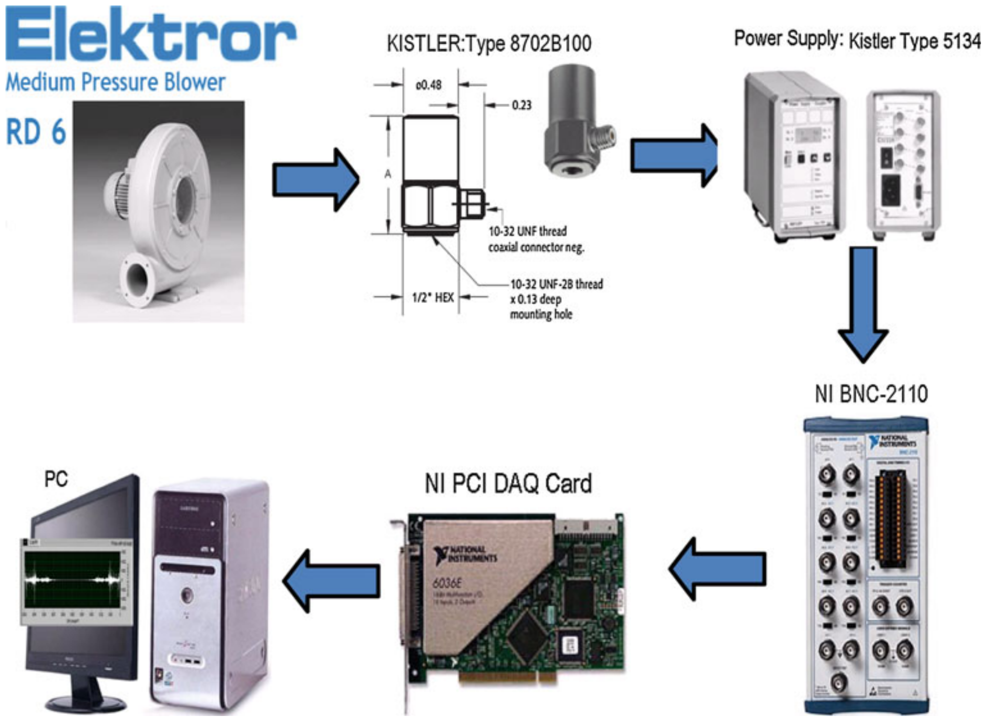


Fig. 1 Hardware of experimental setup

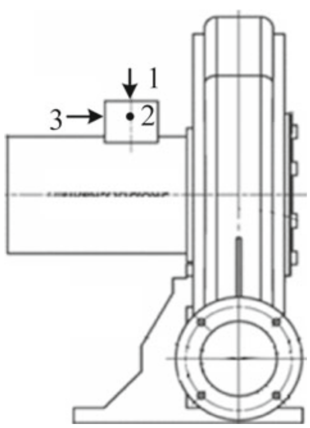


Fig. 2 Sensors setup on blower

extract features from preprocessed signals. In this part, the principle of WPD, FFT and how to apply these techniques in extracting features will be described briefly.

Wavelet transform is a time-frequency decomposition of a signal into a set of “wavelet” basic function. Wavelet analysis has proved its great capabilities in decomposing, denoising, and signal analysis which made the analysis of non-stationary signals achievable as well as detecting transient feature components as other methods were inept to perform since wavelet can concurrently impart time and frequency structures. Wavelet Transform (WT) gives good time and poor frequency resolution at high frequencies, and good frequency and poor time resolution at low frequencies. Analysis with wavelets involves with breaking up a signal into shifted and

Fig. 3 Parts for simulation degradation

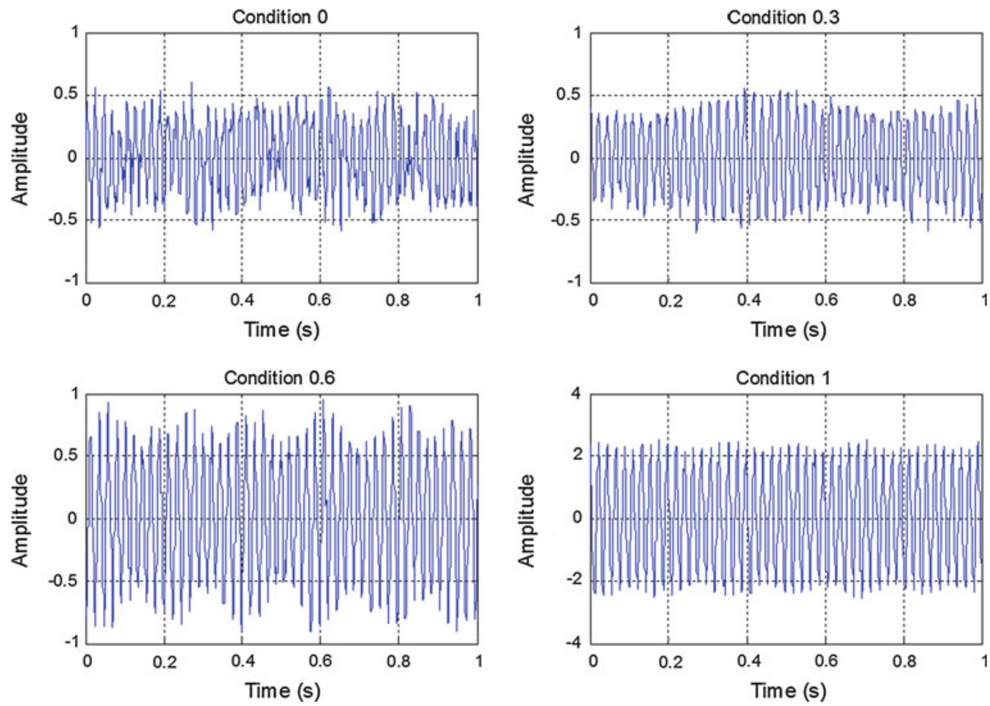


Fig. 4 Raw signals with different degradations

scaled versions of the original (or mother) wavelet, i.e., one high frequency term from each level and one low frequency residual from the last level of decomposition. There are three categories of this transformation: Continuous Wavelet Transform (CWT), Discrete Wavelet Transform (DWT) and WPD.

Continuous wavelet transform

A CWT is used to divide a continuous-time function into wavelets. Unlike Fourier transform, the continuous wavelet transform possesses the ability to construct a time-frequency representation of a signal that offers very good time and frequency localization. The continuous wavelet transform of a time function $x(t)$ is given by following equation:

$$CT(a,b) = \int_{-\infty}^{\infty} x(t)\psi^{*}_{(a,b)}(t)dt \quad (1)$$

where $\psi^{*}_{(a,b)}(t)$ is a continuous function in both the time domain and the frequency domain called the mother wavelet and * represents operation of complex conjugate. $\psi^{*}_{(a,b)}(t)$ can be expressed as:

$$\psi^{*}_{(a,b)}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad \text{where} \quad a, b \in \mathrm{R},\ a \neq 0 \quad (2)$$

The main purpose of the mother wavelet is to provide a source function to generate the daughter wavelets which are simply the translated and scaled versions of the mother wavelet. As seen in Eq.(2), the transform signal $CT(a,b)$ is defined on $a-b$ plane, which a and b are used to adjust the frequency and the time location of the wavelet in Eq. (2). A small a produces a high-frequency wavelet when high frequency resolution is needed and the reverse is also true. The WT's superior time-localization properties stem from the finite support of the analysis wavelet: as b increases, the analysis wavelet transverses the length of the input signal, and a increases or decreases in response to changes in the signal's local time and frequency content. Finite support implies that the effect of each term in the wavelet representation is purely localized. This sets the WT apart from the Fourier Transform, where the effects of adding higher frequency sine waves are spread throughout the frequency axis.

Discrete wavelet transform

In numerical analysis and functional analysis, DWT is a wavelet transform for which the wavelet $\psi_{(a,b)}$ is discretely sampled. As with CWT, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency and location information (location in time). Usually, the DWT can be derived from discretization of CWT. The most common discretization is dyadic method:

$$DT(a,b) = \int_{-\infty}^{\infty} x(t)\psi^{*}_{(j,k)}(t)dt \quad (3)$$

$$\psi^{*}_{(j,k)}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^j k}{2^j}\right) \quad (4)$$

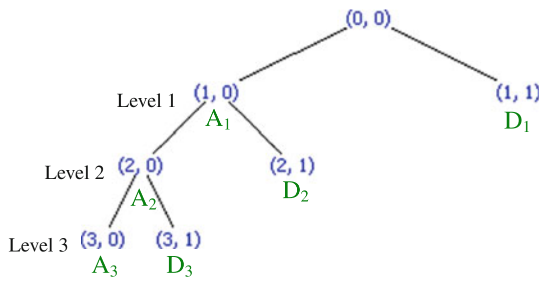


Fig. 5 3-Layer structure of wavelet packet decomposition

where a and b are replaced by 2^j and $2^j k$ respectively (Daubechies 1988; Mallat 1989). An efficient way to implement this scheme using filters was developed by Mallat (1989). The original signal $x(t)$ passes through two complementary filters and emerges as low frequency called approximations and high frequency called details. The decomposition process can be iterated, with successive approximations being decomposed in turn, such that a signal can be broken down into many lower-resolution components.

Wavelet packet decomposition

The structure of wavelet packet transform (WPT) is similar to DWT. Both have the framework of multi-resolution analysis. The main difference in the two techniques is the WPT can simultaneously break up detail (D_i) and approximation (A_i) versions while DWT only breaks up as an approximation version. Therefore, the WPT have the same frequency bandwidths in each resolution and DWT does not have this property. The mode of decomposition does not increase or lose the information within the original signals. Therefore, the signal with great quantity of middle and high frequency signals can offer superior time-frequency analysis. The WPT suits signal processing, especially non-stationary signals because the same frequency bandwidths can provide good resolution regardless of high and low frequencies.

As discussed above, WPT can decompose the signal into two parts: low-frequency A_1 and high frequency D_1. In the process of decomposition, the lost information belonging to the low frequency part was captured by the high frequency part. In the next level of decomposition, this method will also decompose A_1 into two parts: low-frequency A_2 and high frequency D_2. The lost information belonging to low frequency A_2 was capture by the high-frequency D_2, and thus, a deeper level decomposition can be done. WPD is more effective, it can decompose not only the low-frequency part, but also high-frequency. The 3-layer structure of signal based on WPD is shown in Fig. 5 in which only approximation version is decomposed.

For the case in this paper, the sample frequency is 512 Hz, and thus D_1, D_2, D_3 and A_3 represent the frequency 256–512, 128–256, 64–128 and 0–64 Hz respectively in Fig. 5. In this experiment, only these four parts are analyzed to judge the degradation of the performance. The decomposed signals by WPD from the different degrading signals are shown in Figs. 6, 7, 8 and 9.

Fast Fourier transform

Fourier transform is a mathematical operation that decomposes a signal into its constituent frequencies. The original signal depends on time, and therefore is called the time domain representation of the signal, whereas the Fourier transform depends on frequency and is called the frequency domain representation of the signal. The term Fourier trans-

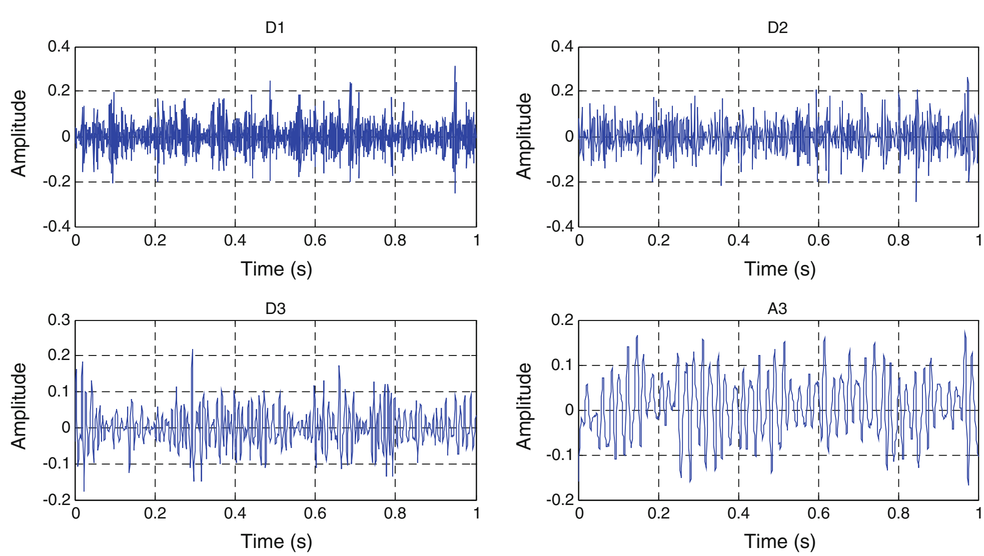


Fig. 6 Decomposed signal of condition 0

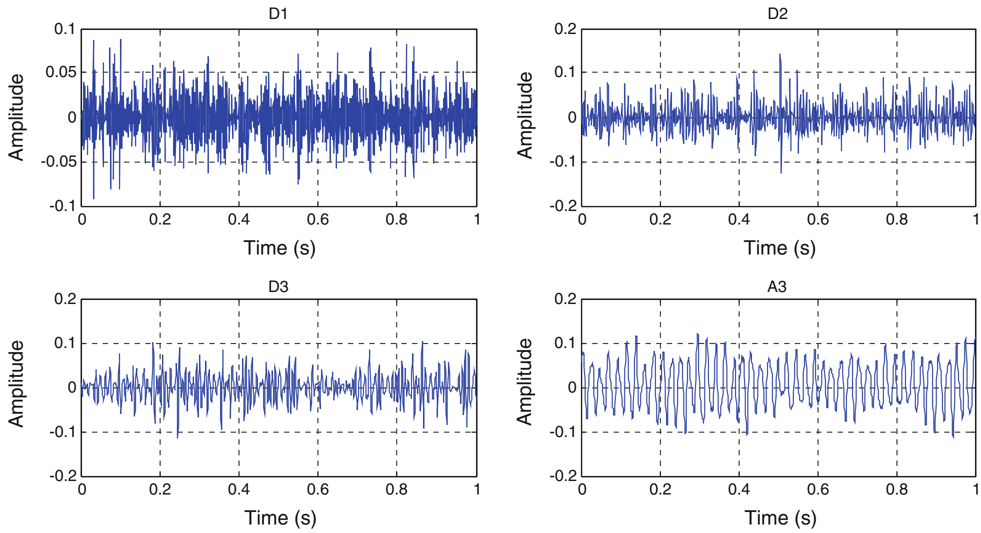


Fig. 7 Decomposed signal of condition 0.3

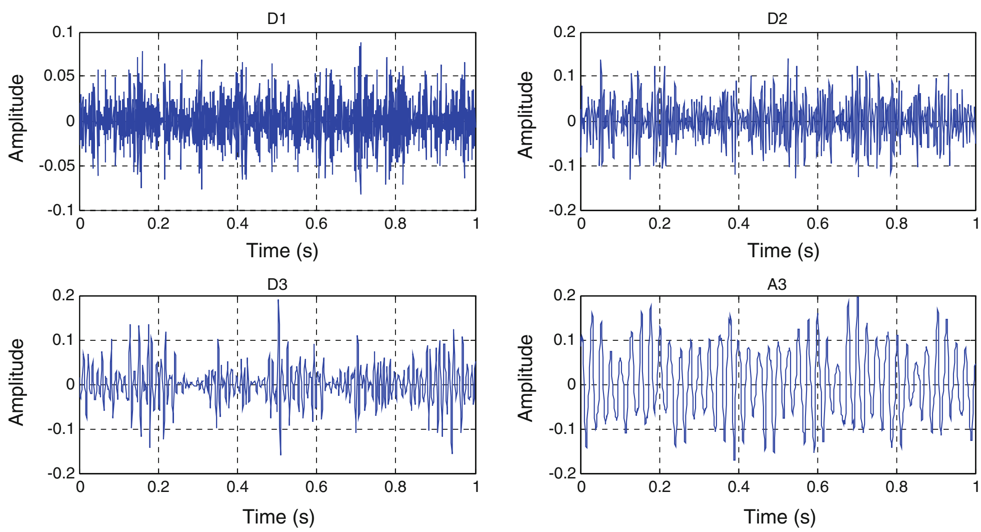


Fig. 8 Decomposed signal of condition 0.6

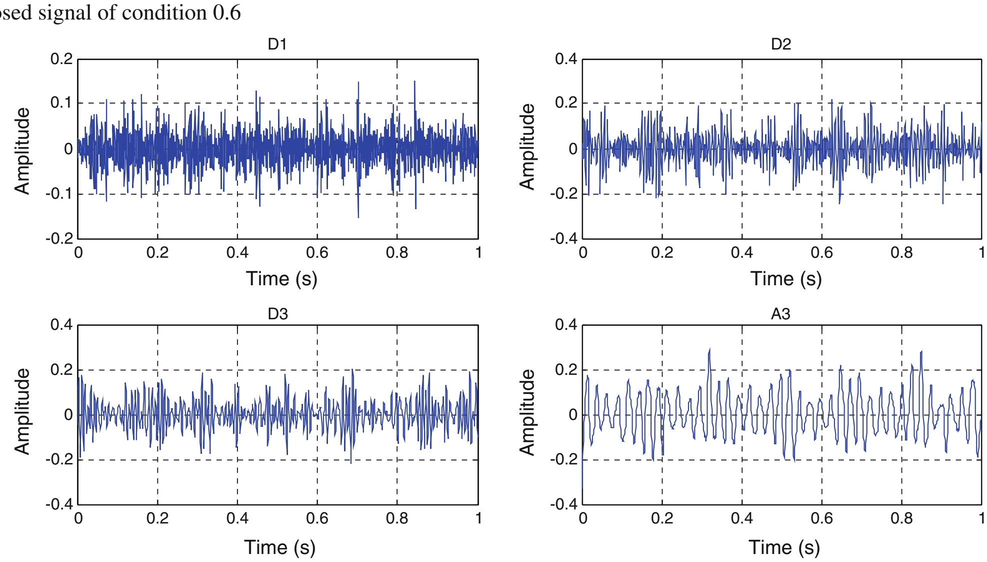


Fig. 9 Decomposed signal of condition 1

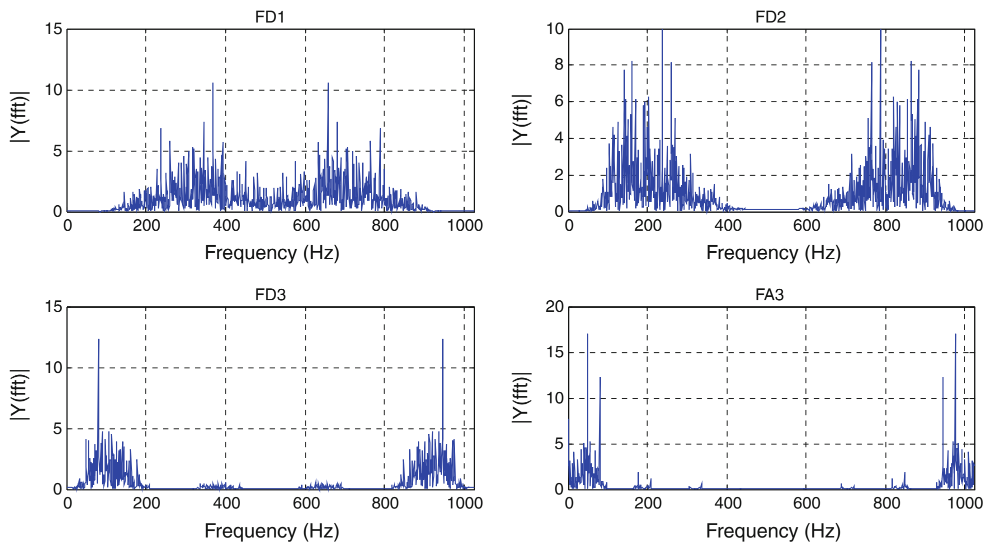


Fig. 10 FFT for each version signal of condition 0

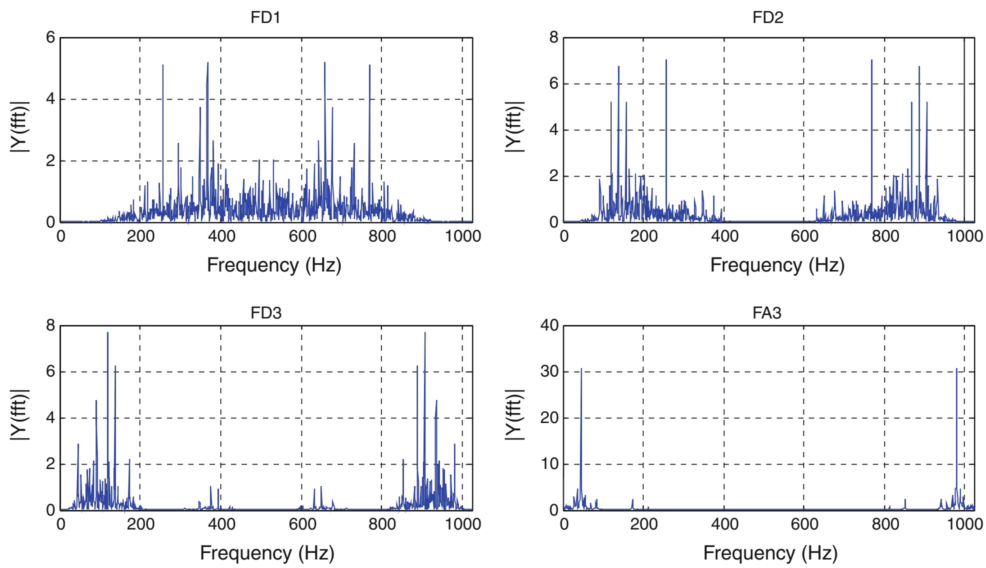


Fig. 11 FFT for each version signal of condition 0.3

form refers both to the frequency domain representation of the signal and the process that transforms the signal to its frequency domain representation. A FFT is an efficient algorithm to compute the discrete Fourier transform (DFT) and it's inverse. Commonly, the FFT of a signal can be calculated by the following equation (Sohn et al. 2010):

$$\mathrm{FFT}(k) = \sum_{n=1}^{N} x(j)\omega_N^{(n-1)(k-1)} \tag{5}$$

$$\omega_N = e^{(-2\pi i)/N} \tag{6}$$

where N is the number of samples for one signal and ω_N is an Nth root of unity. In "Wavelet packet decomposition", the original signal was decomposed as on approximation and details. Then, the decomposed signals are transformed with FFT which are shown in Figs. 10, 11, 12 and 13 which present different conditions from condition 0 to condition 1. From the result of FFT, some kinds of features can be chosen. In this paper, the peaks for each part are selected as features to judge the condition of monitored equipment.

Fault diagnosis and prognosis integrating WPD, FFT and ANN

The pattern classification theory has become a key factor in fault diagnosis and prognosis. Some classification methods for equipment performance monitoring use the relationship

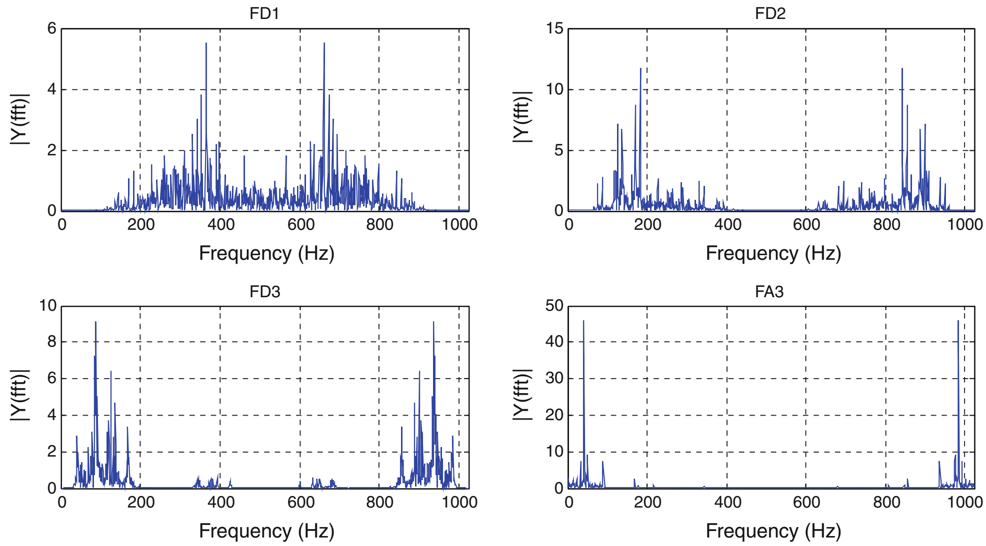


Fig. 12 FFT for each version signal of condition 0.6

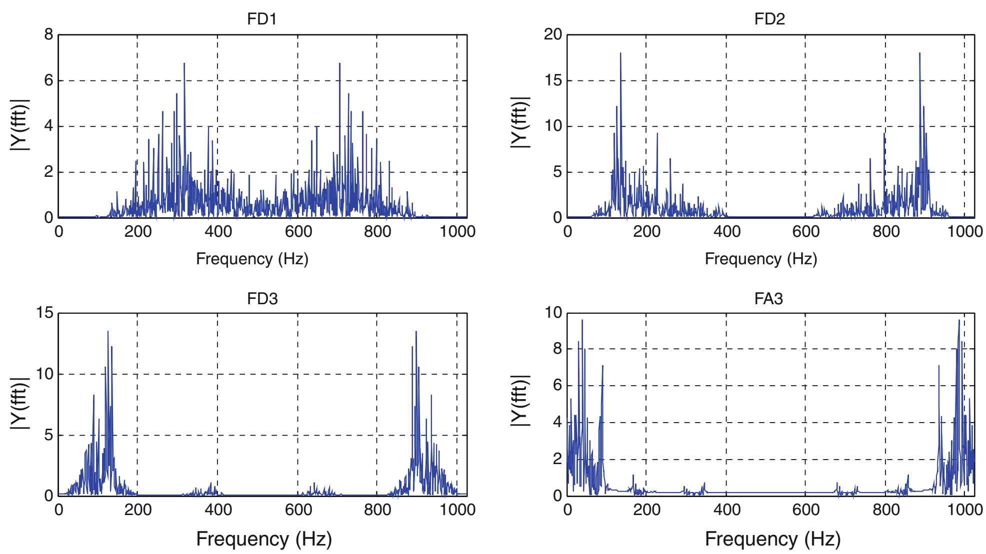


Fig. 13 FFT for each version signal of condition 1

between the type of fault and a set of patterns which is extract from the collected signals without establishing explicit models. Currently, ANN is one of the most popular methods in this domain. ANN is a model that emulates a biological neural network (Wang 2005). The origin of ANN can be traced back to a seminar paper by McCulloch and Pitts (1943) that demonstrated a collection of connected processors, loosely modeled on the organization of brain, could theoretically perform any logical or arithmetic operation. Then, the development of ANN techniques is very fast which is extensive to many categories containing Back-propagation (BP), Self-organization Mapping (SOM) and Radial Basis Function (RBF), etc. The application of artificial neural network models lies in the fact that they can be used to infer a function from observations. This is particularly useful in applications where the complexity of the data or task makes the design of such a function by hand impractical. This attribution is very nontrivial in diagnostic problems. BP neural network is a main type of ANN used to solve fault diagnosis and prognosis problems.

ANN can deal with complex non-linear problem without sophisticated and specialized knowledge of the real systems. It is an effective classification techniques and low operational response times needed after training. The relationship between the condition of component and the features is not linear but non-linear. BP neural network does not need to know the exact form of analytical function on which the model should be built. This means neither the functional type nor the number and position of the parameters in the

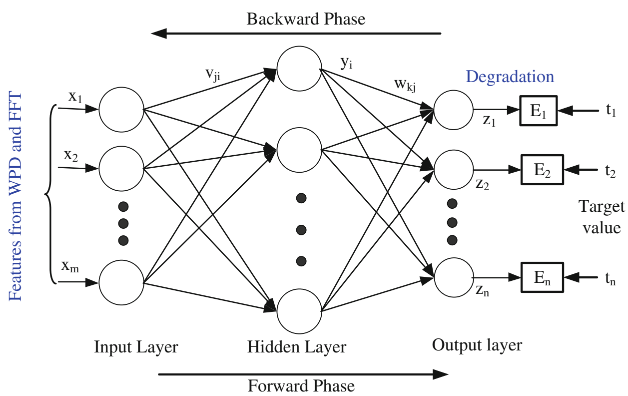


Fig. 14 A BP neural network with single hidden layer

model-function need to know. It can deal with multi-input, multi-output, quantitative or qualitative, complex system with very good abilities of data fusion, self-adaptation and parallel processing. Therefore, it is very suitable to select as a method of fault diagnosis and prognosis. There are many papers dealing with the use of ANN and most of their contributions are ANN training efficiency and strategies for ANN itself. ANN proposed in this paper is work with other two methods together as a completed process of diagnosis and prognosis. Wavelet analysis is a better method than FFT and STFT for signal process and feature extraction. There are many papers selecting entropy of coefficients or energy of sub-signals of WT as features to train ANN. In more complicated system, these features do not contain enough information of the machine. The frequency properties for every sub-signal could be very useful to judge the component condition, and thus the FFT is selected to get these kinds of information. By integrating the WT, FFT and ANN, the condition can be diagnosed and predicted using trained ANN. Therefore, this paper proposes a method applying these three techniques for fault classification and prediction. In this section, BP neural network is introduced briefly followed by the procedure of fault diagnosis and prognosis integrating WPD, FFT and ANN.

BP neural network

BP neural network which is the most widely used neural network model currently was proposed by Rumelhart et al. in 1986. It is a multilayer feed-forward network usually containing the input layer, hidden layer, and output layer (Fig. 14), which trained by an error back propagation algorithm. The biggest advantage of ANNs trained by back propagation is that there isn't need to know the exact form of analytical function on which model should be built. So it's not necessary have neither the function type not even the number and position of the parameters in the model function. Moreover, BP network can learn and store a lot of input-output model mapping without mathematical equations which describing this mapping. The learning method of BP is the steepest descent method which is adjusting the weights and thresholds of the network to minimize the sum of squared errors. The general procedure of BP network training can be summarized as follows (Wang 2005):

(1) Initialize the weights to small random vales $(-1, 1)$;
(2) Select a training vector pair (input and the corresponding desired output) from the training set and present the input vector to the inputs layer of the ANN;
(3) Calculate the actual outputs (forward phase);
(4) Adjust the weights w_{ji} to reduce the difference between actual output and target (backward phase);
(5) Return to step 2 and repeat for each pattern p until the total error has reached an acceptable level;
(6) Stop.

Figure 14 shows a BP network structure with a single hidden layer. $\vec{\mathbf{x}}$ and $\vec{\mathbf{t}}$ are input and target of training data respectively. $\mathbf{v_{ji}}$ and $\mathbf{w_{kj}}$ are weights between input and hidden layer, and between output and hidden layer respectively. $\mathbf{y_i}$ and $\mathbf{z_k}$ are outputs of hidden and actual output of output layer. The objective of ANN training is to obtain all the suitable weights to meet the input and the target of training data. After the training of BP network, for each set of test data or query data, there is a set of output calculated by the final updated weights. For a specific application in fault diagnosis and prognosis, after training by features extracted from processed historic data, the BP network can classify the fault and predict the states of the monitored components or machine units. In this case, the input features are peak values of FFT series from decomposed signals using WPD while the output means the degradation for each monitored equipment or component.

The procedure of fault diagnosis and prognosis

The peak values of FFT for decomposed signals with WPD of vibration signals are used to estimate the condition of components and machines. BP neural network made up of one input layer, one output layer and one hidden layer. And it has been proved that such three layers' BP neural network model can approach any continuous functions at any precision. The structure of the BP network is shown in Fig. 14. The values of output are from 0 to 1 which represent from perfect condition to complete failure of specific kinds of fault as mentioned in "Data acquisition experiment".

Because of convenience of handling the signal collection, a signal processing and interface thing, Labview is selected as program software in this project. However, the capability of mathematical calculation of Labview is not as good as Matlab. Therefore, both kinds of software are combined to apply in this study. The procedure of diagnosis and

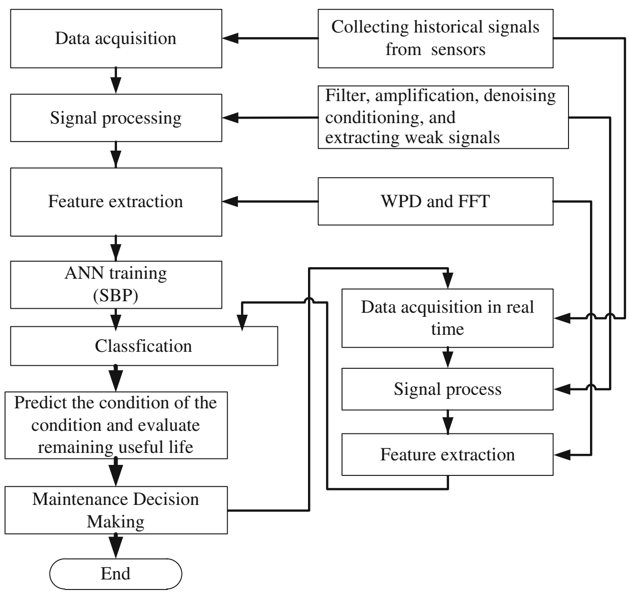


Fig. 15 The procedure of diagnosis and prognosis

prognosis is shown in Fig. 15. The historic data is collected and processed which are first two steps. Then, the processed signals can be decomposed by WPD. Each part of decomposed signals can be transformed using FFT and the peak value for each of them is selected as feature to train BP network. After training, the signals in real time are collected and used to query the BP network, and then the condition of the monitored components can be obtained. Finally, the remaining useful life is evaluated for decision making of maintenance according to the condition.

Case study

A framework called Intelligent Blower Fault Diagnosis and Prognosis is developed in KDL to show how apply proposed method in real system and to validate the correctness of proposed techniques. This framework is a part of SFI-Norman project called Condition based Maintenance in order to achieve near zero-breakdown manufacturing and further to reach zero-defect manufacturing. In this section, some experiments are done to certify the correctness, robustness and precision of proposed methods and comparison of the different result among using different inputs.

The hardware of the framework was shown in Fig. 1 while the monitored component was shown in Fig. 2. Figure 16 shows the interface of software in which the condition of the blower can be seen in real time. The raw signals collected from sensors in real time are displayed on the top right of the interface. One can set the number of training data. After training, the conditions of the monitored components which are presented in the condition textbox can be calculated by ANN. The results are displayed graphically as seen on the left side of this figure. The area inside the blue circle represent safe condition; the area between blue circle and yellow circle represent warning condition while the area between yellow circle and red circle represent failure condition. The points located in radial line represent the conditions of the specific components. From this figure, the conditions of monitored components can be presented clearly.

Experiment and results

In this case study, four conditions for the monitored component are defined which include 0, 0.3, 0.6 and 1 which represent from perfect performance to completely failure discretely. For each condition, 200 training signals were collected and processed. The training signals are preprocessed firstly and then decomposed by WPD. For each part of decomposed signal, calculating the peak value in its frequency domain transformed using FFT which called PFD1, PFD2, PFD3, and PFA3. In this case, there are three

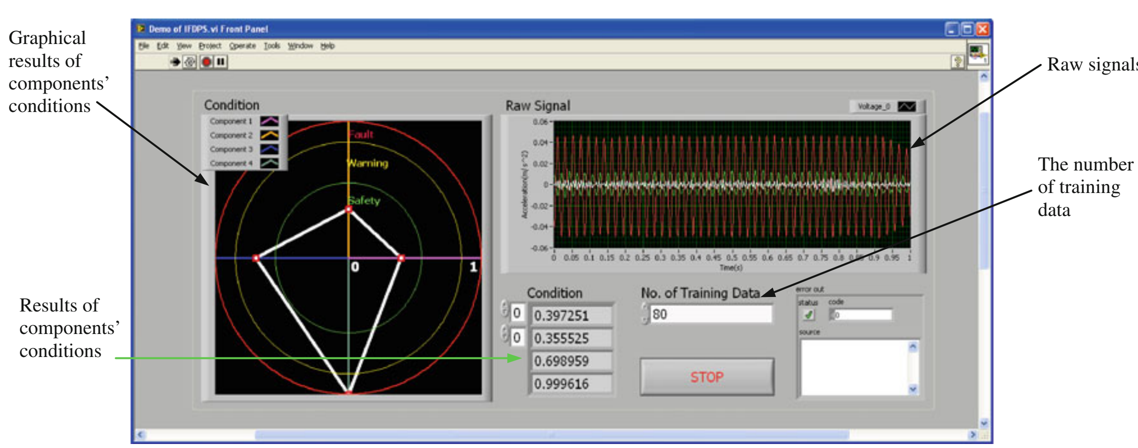


Fig. 16 The interface of the system

Table 1 Part of training data

Sensor 1				Sensor 2				Sensor 3				C
PFD1	PFD2	PFD3	PFA3	PFD1	PFD2	PFD3	PFA3	PFD1	PFD2	PFD3	PFA3	
4.20	3.18	3.768	49.05	4.07	3.756	3.26	95.17	4.325	4.323	2.816	101.08	0
4.46	2.965	2.788	20.38	4.54	3.404	3.346	102.0	3.891	4.108	3.248	107.36	0
4.58	4.039	3.874	317.6	6.15	3.603	3.704	4,170	4.663	3.55	5.447	1,094.9	0.3
3.42	3.802	3.227	314.3	3.46	3.765	3.659	4,220	4.261	3.42	4.659	1,132.5	0.3
4.87	4.238	5.951	482.2	5.19	4.184	4.617	6,975	3.523	3.845	2.723	1,889.8	0.6
4.49	3.745	4.178	395.6	4.03	4.412	4.289	6,828	4.781	3.705	3.022	1,861.4	0.6
6.41	3.007	18.46	1,933	4.94	3.053	5.048	2,035	5.095	2.919	5.601	6,189.9	1
4.54	4.304	18.43	1,936	4.72	4.23	4.73	2,103	4	4.506	6.062	6,391.8	1
…	…	…	…	…	…	…	…	…	…	…	…	…

Table 2 Test data and the results

Sensor 1				Sensor 2				Sensor 3				Results		Deviation
PFD1	PFD2	PFD3	PFA3	PFD1	PFD2	PFD3	PFA3	PFD1	PFD2	PFD3	PFA3	NC	TC	
3.71	3.382	2.941	37.608	4.636	4.582	3.018	99.027	4.095	3.719	4.749	107.685	0	0.04	0.038
4.755	3.079	3.049	30.693	6.705	4.092	3.187	81.135	3.951	3.525	3.583	105.698	0	0.05	0.046
4.29	4.083	2.418	20.416	4.258	4.069	3.583	85.343	6.563	3.09	4.254	100.095	0	0.04	0.041
4.803	3.506	3.056	39.464	4.561	3.792	2.952	76.756	4.398	4.545	4.052	111.304	0	0.03	0.033
4.114	3.856	2.557	32.6	4.77	4.005	3.258	112.054	4.752	3.517	3.48	116.38	0	0.04	0.038
4.133	3.684	3.162	275.02	4.114	3.736	3.227	4, 135.45	4.701	3.452	4.835	1, 199.69	0.3	0.28	0.017
4.433	3.475	3.589	280.2	3.944	3.532	3.121	4, 174.94	4.215	3.701	3.25	1, 223.66	0.3	0.28	0.025
5.301	3.352	3.539	280.78	4.89	4.044	4.023	4, 280.03	5.667	4.422	4.008	1, 259.53	0.3	0.29	0.014
5.346	6.322	4.353	257.69	4.934	3.491	3.361	4, 175.62	5.982	5.922	3.044	1, 212.37	0.3	0.28	0.02
3.852	3.516	3.548	303.43	4.874	3.825	3.852	4, 193.24	3.817	3.952	3.428	1, 233.34	0.3	0.29	0.011
4.699	3.421	3.31	311.82	4.327	4.911	5.273	7, 101.38	4.158	3.558	3.183	2, 098.35	0.6	0.61	0.005
4.087	4.644	3.008	278.09	3.865	3.482	5.644	7, 211.46	5.392	4.981	3.51	2, 147.95	0.6	0.54	0.059
3.978	3.719	3.463	286.09	4.321	3.635	5.177	7, 122.77	4.196	3.682	3.883	2, 094.52	0.6	0.6	0.002
4.347	2.976	3.434	279.05	5.284	5.405	4.546	7, 157.75	3.978	4.449	3.333	2, 126.02	0.6	0.57	0.033
4.44	3.505	3.345	262.11	5.521	3.628	4.63	7, 080.4	3.991	3.587	4.037	2, 082.84	0.6	0.57	0.031
3.603	4.235	8.397	910.92	5.633	5.258	9.274	21, 669.1	3.839	3.875	5.985	6, 824.31	1	1	0.002
5.451	3.87	6.187	885.07	4.922	6.128	12.67	21, 416.9	5.687	3.643	8.407	6, 661.59	1	1	0
5.957	3.575	8.829	918.7	5.764	5.873	8.867	21, 244.5	4.995	4.161	4.444	6, 594.82	1	1	0.002
4.818	3.049	8.352	882.03	6.918	5.931	10.12	20, 684.6	5.035	3.511	8.835	6, 461	1	1	0.002
3.745	3.083	8.32	885.89	6.902	5.976	10.9	20, 605	5.183	3.216	8.642	6, 455.84	1	1	0.002

sensors and thus there are 12 parameters are input to input nodes of ANN and one output value which represents condition of the monitored component (Called C). A part of training data is shown in Table 1. After training, test data or query data obtaining from real system can be used to test or query ANN. In this case, 20 sets of test data (Table 2) are used to test ANN.

There is no mathematical method to select the best structure of the SBP network, but the three layers SBP structure was validated its powerful function to build a complex model. The SBP structure in this experiment is set to three layer $12\times20\times1$ networks. 12 means the number of input parameters (features in this experiment), 20 means the number of the hidden layer nodes and 1 means only one output in this ANN structure (condition). Its maximum training epoch is set to 5,000. For each condition, 80 training sets are used to train ANN and 20 sets of features are chosen to test it. Table 2 shows the results of the test data. As mentioned before, there are 20 sets of test data in which there are 5 sets of them for each condition. The nominal condition is called NC while the output condition of test is called TC in this table. From this table, the results are 100 % correct in the above parameter

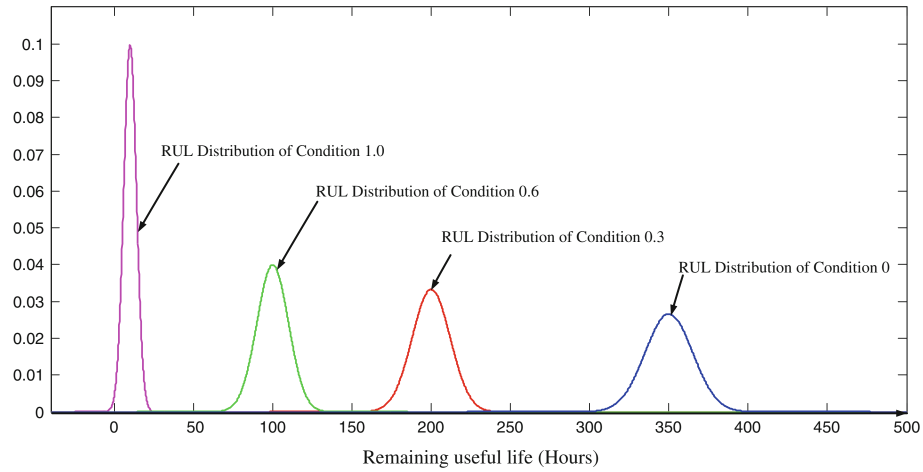


Fig. 17 Remaining useful life distribution for each condition for simulated component

sets. However, the output is not exactly the same as the nominal condition and there are deviations between them. The precision of the output is discussed next part.

After the condition of the component is predicted, the remaining useful life can be evaluated according to the condition. Most current RUL estimation methods are based on the event data or condition monitoring data which want to find the relationship between RUL and time the component used or RUL and feature values (Si et al. 2011; Lee et al. 2006; Tian 2006). This paper tries to find the relationship between the RUL and the condition of a component that is evaluating RUL by the condition and RUL distribution for each condition (Fig. 17). The distributions of RUL are obtained by the statistical methods. For example, if the condition of a component is 0, the remaining useful is 350 h with a certain standard deviation. When the condition is 1.0, the RUL is much closed to 0 which means the component has to be maintained or repaired.

It is easy to understand that the later method (proposed this paper) need far less history datasets to train neural network than the former because the life cycle could be a very long time and need huge data to express its process while there only some of conditions and only a few data can express its process. The later method's result is clearer and sometimes the customers can obtain satisfactory result even if only know the condition because it has a direct relation with RUL. From Fig. 17, the RUL distribution become narrow that means the RUL evaluation is more accuracy when the condition closed to failure. Therefore the confidence value of RUL increase with the condition deterioration.

Discussions

In this section, three issues will be discussed. The first one is how many training sets should be used in order to achieve enough accurate condition of the machine from SBP network. The second one is attempting to discuss the relationship between the accuracy and the number of hidden layer nodes. The last issue is convergent time of the BP network training.

To discuss the first issue, the numbers of training sets for each condition are changed from 1 to 200. The number of hidden layer nodes is set to 20 and the number of training epoch is set to 5,000. For each testing data, compare the output of ANN and the nominal value which is called "error from nominal value" which is average value of testing data for each condition. The values of these errors are shown in Fig. 18. We can see from this, the result is believable whatever the condition of the component is when the number of training data is larger than 20. For condition 0 and condition 1, the result is still believable even if the number of training data is smaller than 20. It is clear that the result will be believable if there are only two conditions (0 and 1 or good and fault) even if the number of training data is very small. But if there are more conditions, the number of training data should be increased. Therefore, the number of conditions should be considered in designing of how many training sets are used to trained SBP neural networks.

To discuss the second issue, the number of hidden layer nodes is changed from 5 to 135. The number of training data is set to 80 and the number of maximum training epoch is set to 5,000. For each training process, several test sets for every condition are used to test the trained SBP networks. The results are shown in Fig. 19. From the figure, with the increasing of the number of hidden layer nodes, the fluctuations of the output for each condition are small. So the changing of the number of hidden layer nodes does not affect the accuracy of the output. What's more, there is no mathematical method to prove what the number of it is best. Therefore, the number of hidden layer nodes does not need to be considered much.

To discuss the last issue, the number of hidden layer nodes is set as 20 and the training epoch is set as 2,000. Figure 20 shows the BP network training time with the number of training data increasing from 10 to 200. From the figure, training time is not apparently increasing with the increasing of

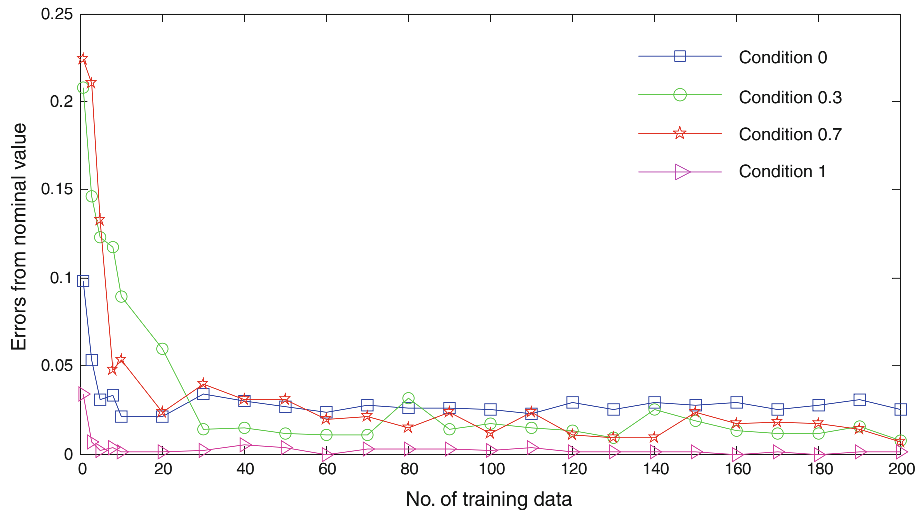


Fig. 18 Errors for each condition

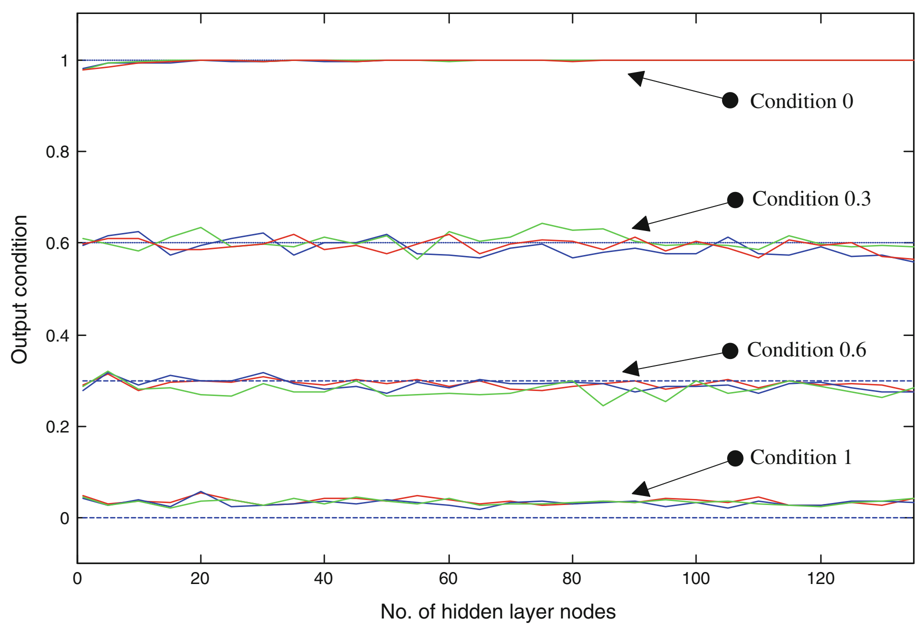


Fig. 19 Output with different number of hidden layer nodes

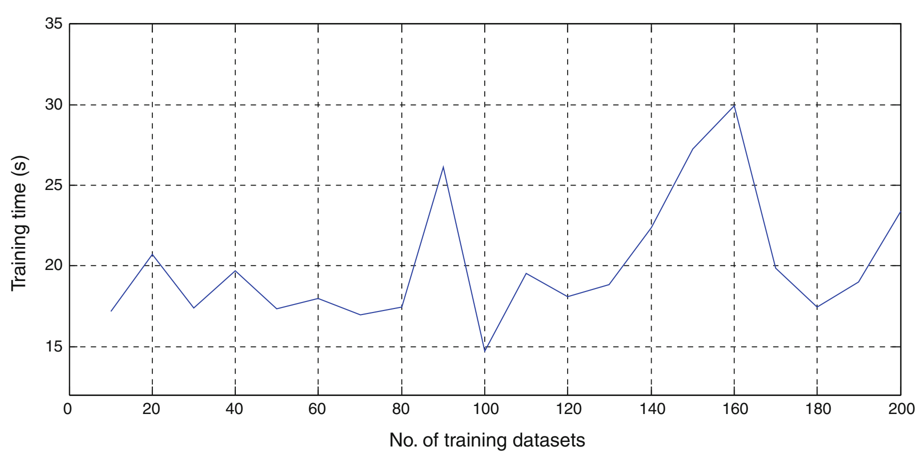
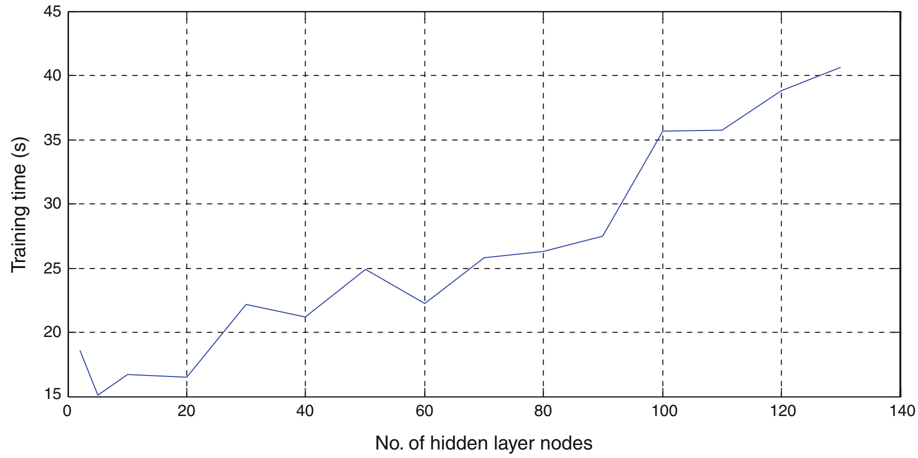


Fig. 20 ANN training time with the increasing of training data

Fig. 21 ANN training time with the increasing of hidden layer nodes



training data sets. Therefore, when the BP network is need, we should use as many as possible data sets to complete the training. Figure 21 shows the training time changes with the increasing of hidden layer nodes. The number of training data sets is set as 200 and the training epoch is set as 2,000. From Fig. 21, the training time increase gradually with the increasing of hidden layer nodes. Therefore, when a BP network need to trained, the number of hidden layer nodes should be considered. However, from the experience of previous work, the numbers of the hidden layer neurons depends both on the input layer number and the output layer neuron number but the numbers could not be too many (Meng and Meng 2010).

Conclusions and future research

In this paper, a new method applying WPD, FFT and BP neural network for fault diagnosis and prognosis was proposed. To verify the correctness and effectiveness of this method, a framework called Intelligent Blower Fault Diagnosis and Prognosis System was established as a case study. From the case, the proposed method for diagnosis is very effective and efficient. This method can also predict the degradation and further to predict remaining useful life of monitored component as well. Finally, this method has many advantages because of using artificial neural network such as the ability to easily deal with complex problems without sophisticated and specialized knowledge, the ability to carry out classifications, the ability to deal with non-linear systems and low operational response times after the learning phase.

In this paper, the minimum bandwidth 0–64 Hz is chosen in WPD because of the fundamental frequency of the vibration signal is the 47.5 Hz. In a real system, the minimum bandwidth of WPD (which means how many levels should be decomposed) should be selected according to the real fundamental frequency. The peak values of FFT for decomposed signals are selected as features to judge the degradation of the monitored machine. In a real application, some other parameters may be chosen as features to judge the condition according to what kinds of faults need to prognose and/or prognose. In the case study of this paper, there is only one kind of fault was simulated. In the future, multi-fault diagnosis and prognosis should be a research topic. The degradation information could be very useful for maintenance decision making and thus, how to apply this degradation information in maintenance decision making should be a research issue as well in the future.

Acknowledgments This paper is a result of Norwegian Manufacturing Future project in Center of Research and Innovation of Norwegian Manufacturing Future (SFI Norman), which is financially supported by Norwegian Research Council.

References

Andria, G., Savino, M., & Trotta, A. (1994). Application of Wigner distribution to measurements on transient signals. *IEEE Transactions on Instrumentation and Measurement, 43*, 187–193.

Chen, C., Sun, C., Zhang, Y. & Wang, N. (2005). Fault diagnosis for large-scale wind turbine rolling bearing using stress wave and wavelet analysis. In *ICEMS 2005 proceedings of the eighth international conference on electrical machines and systems* (Vol. 3, pp. 2239–2244).

Chen, G., Liu, Y., Zhou, W., & Song, J. (2008). Research on intelligent fault diagnosis based on time series analysis algorithm. *The Journal of China Universities of Posts and Telecommunications, 15*(1), 68–74.

Corinthios, M. J. (1971). A fast fourier transformation for high-speed signal processing. *IEEE Transactions on Computers, 20*, 843–846.

Daubechies, I. (1988). Ortho-normal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics, 41*, 909–996.

He, S., He, Z.,& Wang, G. (2011). Online monitoring and fault identification of mean shifts in bivariate processes using decision tree learning techniques. *Journal of Intelligent Manufacturing*, 1–10. doi:10.1007/s10845-011-0533-5.

Huang, Y., Mcmurran, R., Dhadyalla, G., & Jones, R. P. (2008). Probability based vehicle fault diagnosis: Bayesian network method. *Journal of Intelligent Manufacturing, 19*(3), 301–311.

Kasabov, N. (2001). Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning. *IEEE*

Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 31, 902–918.
Kegg, R. L. (1984). On-line machine and process diagnostics. *Annals of the CIRP, 32*(2), 469–573.
Lee, I. (2011). Fault diagnosis of induction motors using discrete wavelet transformation and artificial neural network. In C. Stephanidis (Ed.), *HCI international 2011—Posters' extended abstract*, Vol. 573, (pp.510–514).
Lee, J., Ni, J., Dragan, D., Qiu, H., & Liao, H. (2006). Intelligent prognostics tools and e-maintenance. *Computers in Industry, 57*, 476–489.
Li, Z. N., & Wu, Z. T. (2005). Hidden Markov model-based fault diagnostics method in speed-up and speed-down process for rotating machinery. *Mechanical Systems and Signal Processing, 19*(2), 329–339.
Li, R., Sopon, P., & He, D. (2009). Fault features extraction for bearing prognostics. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-009-0353-z.
Lin, J., & Qu, L. (2000). Feature extraction based on Morlet wavelet and its application for mechanical diagnosis. *Journal of Sound and Vibration, 234*(1), 135–148.
Liu, Y., Guo, L., Wang, Q., An, G., Guo, M., & Lian, H. (2010). Application to induction motor faults diagnosis of the amplitude recovery method combined with FFT. *Mechanical Systems and Signal Processing, 24*, 2961–2971.
Mallat, S. G. (1989). A theory for multi-resolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*, 674–693.
Markou, M., & Singh, S. (2003). Novelty detection: A review–part 2: Neural network based approaches. *Signal Processing, 83*, 2499–2521.
Marzi, H. (2004). Real-time fault detection and isolation in industrial machines using learning vector quantization. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 218*, 949–959.
McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics, 5*, 115–133.
Meng, X., & Meng, X. (2010). Nonlinear system simulation based on the BP neural network. In *2010 3rd international conference on intelligent networks and intelligent systems (ICINIS)*, pp. 334–337.
Momoh, J. A.,& Button, R. (2003). Design and analysis of aerospace DC arcing fault using fast fourier transformation and artificial neural network. In *Power engineering society general meeting*, pp. 788–793.
Portnoff, M. (1980). Time-frequency representation of digital signals and systems based on short-time Fourier analysis. *IEEE Transactions on Acoustics, Speech and Signal Processing ASSP, 28*, 55–69.
Prabhakar, S., Mohanty, A. R., & Sekhar, A. S. (2002). Application of discrete wavelet transforms for detection of ball bearing race faults. *Tribology International, 35*, 793–800.
Rai, V. K., & Mohanty, A. R. (2007). Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform. *Mechanical Systems and Signal Processing, 21*, 2607–2615.
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representational by error propagation. In D. E. Rumenhart & J. L. McCelland (Eds.), *Parallel distributed processing: Exploration in the microstructure of cognition* (pp. 318–362). Cambridge, MA: MIT Press.
Saravanan, N., & Ramachandran, K. I. (2010). Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for fault extraction and fault classification using artificial neural network (ANN). *Expert Systems with Applications, 37*, 4168–4181.
Saravanan, N. K., Sidddabattuni, V. N. S., & Ramachandran, K. I. (2008). A comparative study on classification of features by SVM and PSVM extracted using Morlet wavelet for fault diagnosis of spur bevel gearbox. *Expert System with Applications, 35*, 1351–1366.
Serhat, S., & Emine, A. (2003). Feature extraction related to bearing damage in electric motors by wavelet analysis. *Journal of the Franklin Institute, 340*, 125–134.
Si, X., Wang, W., Hu, C., & Zhou, D. (2011). Remaining useful life estimation—A review on the statistical data driven approaches. *European Journal of Operational Research, 213*(1), 1–14.
Sohn, R. H., Son, J. S., Hwang, H. J., IM, C. H., & Kim, Y. H. (2010). SSVEP-based functional electrical stimulation system For motor control of patients with spinal cord injury. In *Proceedings of 6th word congress of biomechanics* (Vol. 31. pp. 655–658).
Spiewak, S. A., Duggirala, R., & Barnett, K. (2000). Predictive monitoring and control of the cold extrusion process. *CIRP Annals—Manufacturing Technology, 49*(1), 383–386.
Staszewski, W. J., Worden, K., & Tomlinson, G. R. (1997). Time-frequency analysis gearbox fault detection using the Wigner Ville distribution and pattern recognition. *Mechanical Systems and Signal Processing, 11*(5), 673–692.
Tian, Z. (2006). An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-009-0356-9.
Tse, P. W., Yang, W. X., & Tam, H. Y. (2004). Machine fault diagnosis through an effective exact wavelet analysis. *Journal of Sound and Vibration, 277*, 1005–1024.
Wang, K. (2002). *Intelligent condition monitoring and diagnosis systems*. Amsterdam: IOS Press.
Wang, K. (2005). *Applied computational intelligence in intelligent manufacturing systems*. Advanced Knowledge International Pty Ltd, Australia.
Wang, C., Zhang, Y., & Zhong, Z. (2008). Fault diagnosis for diesel valve trains based on time–frequency images. *Mechanical Systems and Signal Processing, 22*, 1981–1993.
Wang, C., Kang, Y., Shen, P., Chang, Y., & Chung, Y. (2010). Applications of fault diagnosis in rotating machinery by using time series analysis with neural network. *Expert Systems with Applications, 37*, 1696–1702.
Wu, J. D., & Chen, J. C. (2006). Continuous wavelets transform technique for fault signal diagnosis of internal combustion engines. *NDT & E International, 39*, 304–311.
Wu, S. T., & Chow, T. W. S. (2004). Induction machine fault detection using SOM based RBF neural networks. *IEEE Transactions on Industrial Electronics, 51*(1), 183–194.
Wu, J., & Kuo, J. (2009). An automotive generator fault diagnosis system using discrete wavelet transform and artificial neural network. *Expert Systems with Applications, 36*, 9776–9783.
Wu, J., & Liu, C. (2009). An expert system for fault diagnosis in internal combustion engines using wavelet packet transform and neural network. *Expert Systems with Applications, 36*, 4278–4286.
Yu, D., Yang, Y., & Cheng, J. (2007). Applications of time–frequency entropy method based on Hilbert-Huang transform to gear fault diagnosis. *Measurement, 40*, 823–830.
Zheng, H., Li, Z., & Chen, X. (2002). Gear fault diagnosis based on continuous wavelet transforms. *Mechanical Systems and Signal Processing, 16*(2–3), 447–457.