ELSEVIER

Computers in Industry 43 (2000) 61–72

COMPUTERS IN INDUSTRY

www.elsevier.nl/locate/compind

Fault diagnosis using Rough Sets Theory

Lixiang Shen [a], Francis E.H. Tay [a,*], Liangsheng Qu [b], Yudi Shen [b]

[a] Department of Mechanical and Production Engineering, National University of Singapore, 10 Kent Ridge Crescent, 119260, Singapore
[b] Institute of Diagnostics and Cybernetics, School of Mechanical Engineering, Xian Jiaotong University, Shaanxi, 710049, People's Republic of China

Received 25 February 1999; accepted 29 January 2000

Abstract

The fault diagnosis on diesel engine is a difficult problem due to the complex structure of the engine and the presence of multi-excite sources. Usually, one method of fault diagnosis can only inspect one corresponding fault category. In this paper, a new method, Rough Sets Theory, is used to diagnose the valve fault for a multi-cylinder diesel engine. Through the analysis of the final reducts generated using Rough Sets Theory, it is shown that this new method is effective for valve fault diagnosis. Rough Sets analysis requires discretizing the fault condition attributes. However, in practice, some of the limits of these attributes are unknown. A new discretization method has been created and the method is found to be suitable for discretizing the attributes without a priori knowledge.

Keywords: Fault diagnosis; Rough Sets Theory; Discretization

1. Introduction

Fault diagnosis on machinery has been well researched [1]. There are many effective methods that are used to diagnose accurately and quickly a certain category of faults [2]. For instance, large-scale centrifugal compressor can be diagnosed by holospectrum technique [3]. However, up to now, it is difficult to diagnose more than one category of faults. This is especially so in diagnosing the dynamic characteristics of reciprocating machinery, such as reciprocating compressor and diesel engine. This is due to the complex structure of the reciprocating machinery. Although many methods can be used to determine specific fault category, such as broken valve and cracked crankshaft [4], the results obtained from such fault specific method are not easy to interpret. There is a need to have a method that can diagnose more than one category of faults in a generic manner. In this paper, a method based on Rough Sets Theory is proposed and implemented.

Z. Pawlak (Poland) first proposed Rough Sets Theory in 1982. This theory has been developed and used in many domains, such as medical diagnosis [5], stock market forecast [6], fault diagnosis in engineering domain [7], decision making for bank manager [8] and some other uses [21].

The advantage of Rough Sets Theory is that it needs neither additional information about the data nor is it necessary to correct the inconsistencies manifested in data. Instead, rules generated are categorized into certain rules or possible rules.

In this paper, the Rough Sets Theory [9] is used to analyze the decision table composed of attributes

* Corresponding author. Tel.: +65-874-6818; fax: +65-779-1459.

E-mail address: mpetayeh@nus.edu.sg (F.E.H. Tay).

0166-3615/00/$ - see front matter
PII: S0166-3615(00)00050-6

extracted from the vibration signals, which are collected from a 4135 diesel engine. The remainder of this paper is organized as follows. The characteristics of vibration signal of diesel engine will be analyzed in Section 2, then Rough Sets Theory is introduced in Section 3. The attributes field is established in Section 4, which is used to compose the decision table. A new discretization method, more suitable to discretize continuous attributes without a priori knowledge, is proposed in Section 5. The diagnosis results using Rough Sets Theory are discussed in Section 6 and at the end of this paper, conclusions based on the aforementioned analysis are given.

2. The Characteristics of vibration signals for a 4153 diesel engine

There are many surveillance methods [2,10] proposed to monitor the condition for a diesel engine, based on the processing of the vibration signals. Due to the complex structure and multi-excite sources that exist in diesel engine, the vibration signals collated from the engine surface have the following characteristics:

- Presence of a number of self-exciting vibration and forced vibration in the diesel engine that is running. Therefore, the width of spectrum in frequency domain is very large.
- The vibration signals in the time domain are more complex compared to a large-scale rotational machinery, which is a pure sine curve.

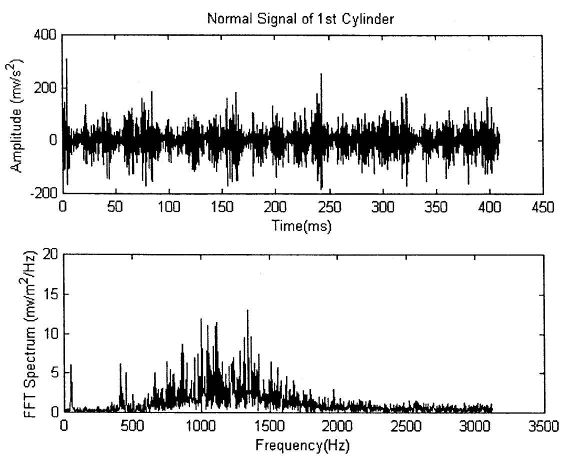

Fig. 1. Normal state (sample point 1).

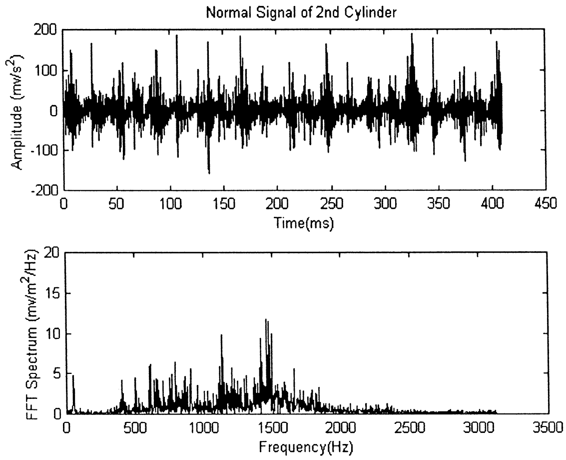

Fig. 2. Normal state (sample point 2).

- In a diesel engine, such as 4135 engine, the stroke cycles are fixed. Therefore, the time series appear periodical. However, in every period, there exist many other periodical vibrations within the stroke cycle.

In Figs. 1–12 some vibration signals, collected from a 4135 engine surface, are presented. The parameters of 4135 diesel engine are:

Rated Engine Power: 80 hp,
Rated Engine Speed: 1500 rpm,
Four states are studied in this paper. They are,

- Normal state
- Intake valve clearance is too small

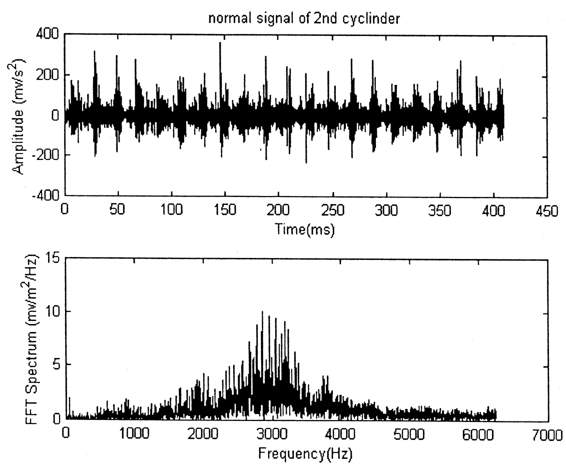

Fig. 3. Normal state (sample point 3).

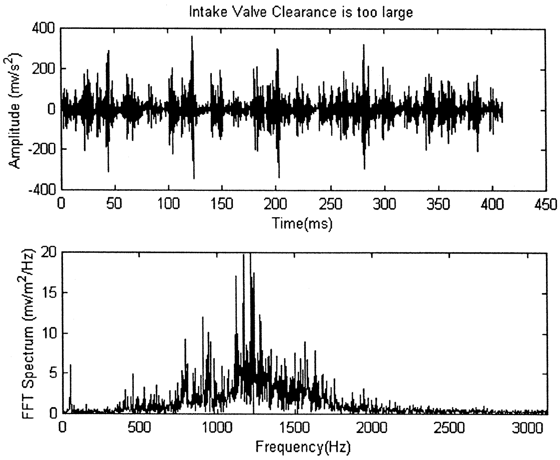

Fig. 4. Intake valve clearance is too large (sample point 1).

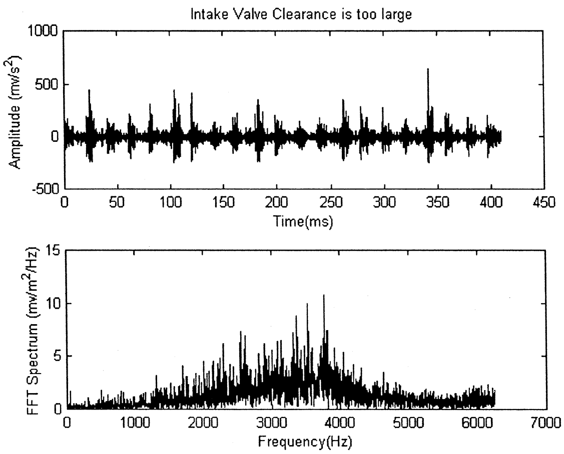

Fig. 5. Intake valve clearance is too large (sample point 2).

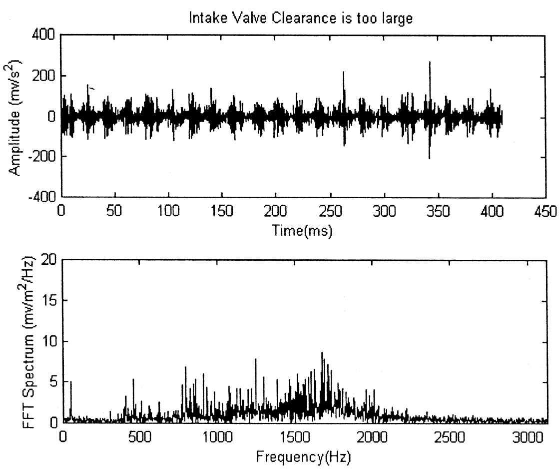

Fig. 6. Intake valve clearance is too large (sample point 3).

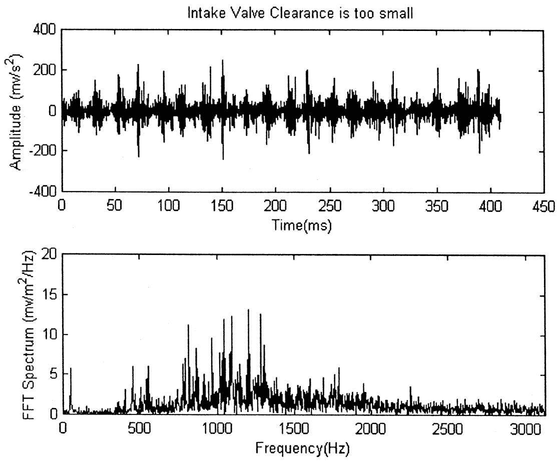

Fig. 7. Intake valve clearance is too small (sample point 1).

- Intake valve clearance is too large
- Exhaust valve clearance is too large.

Among these four states, three fault types were simulated in the intake valve and exhaust valve on the second cylinder head. In this process, three points are selected to collect vibration signals. They are the first cylinder head, the second cylinder head and another one that is at the centre of the piston stroke, on the surface of the *cylinder block*. In each figure, the top portion represents the time series and the

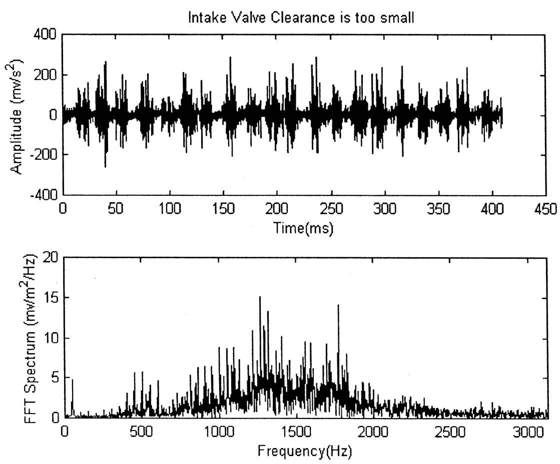

Fig. 8. Intake valve clearance is too small (sample point 2).

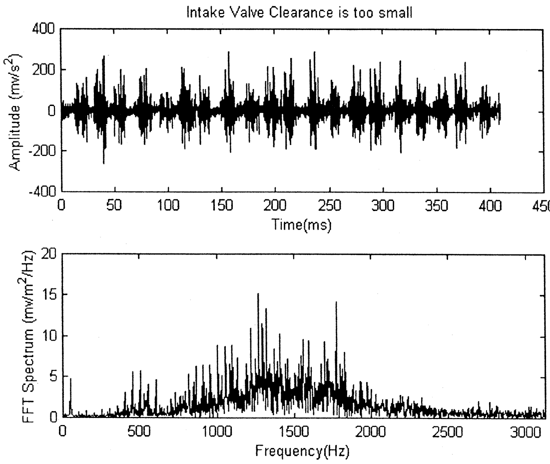

Fig. 9. Intake valve clearance is too small (sample point 3).

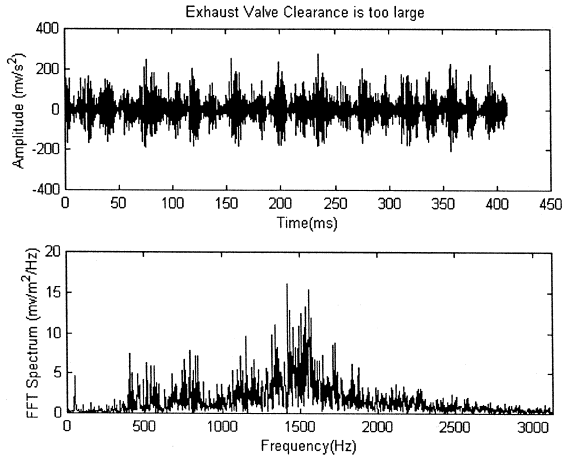

Fig. 11. Exhaust valve clearance is too large (sample point 2).

bottom portion represents the corresponding FFT (Fast Fourier Transformation) spectrum.

According to the diesel engine operating principle, when the clearance of valve increases, the energy distributed in frequency axis will move up to a higher frequency [10]. Thus, theoretically, from the state of small intake valve clearance to normal and to large intake clearance, the center of frequency distribution will move up the axis.

From the FFT spectrum in Figs. 1–12, it is difficult to differentiate the energy change corresponding to the fault type change. The energy spectrum is distributed in a wide range from 50 to 2000 Hz for the vibration signals collected from the cylinder head. The energy spectrum covers the entire frequency axis for the vibration signals collected from the point in the middle of piston stroke.

In the time domain, every waveform peak represents the *uncontrolled burning* or *flame propagation phase* [11]. From the time series waveform, no conspicuous difference exists among the different fault types. Therefore, due to the complex vibration signals, the different valve faults cannot be diagnosed using conventional FFT spectrum.

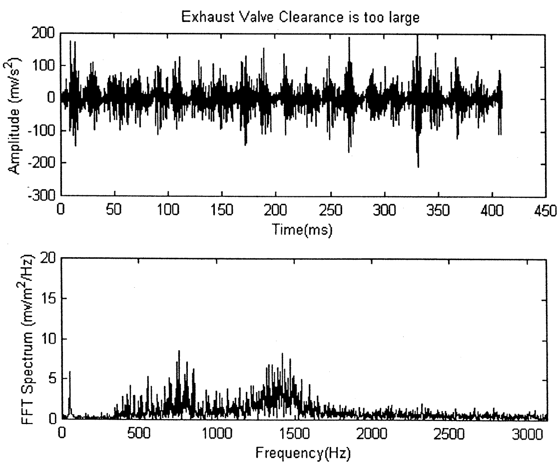

Fig. 10. Exhaust valve clearance is too large (sample point 1).

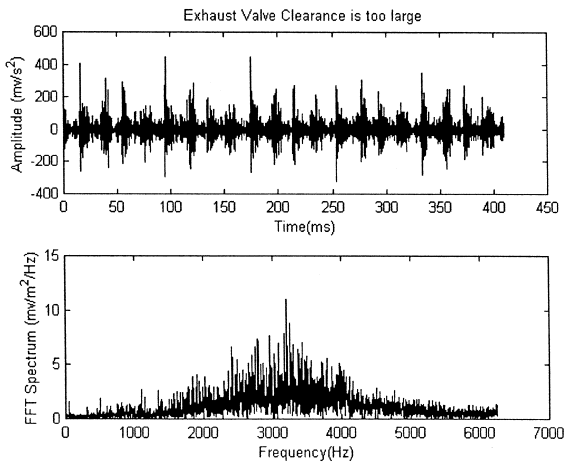

Fig. 12. Exhaust valve clearance is too large (sample point 3).

3. Rough Sets Theory

Rough Sets Theory [9] is a new mathematical tool to handle vagueness and uncertainty inherent in making decisions. Rough Sets Theory finds applications primarily in some branches of artificial intelligence and cognitive sciences, such as machine learning, knowledge discovery from databases, expert systems, inductive reasoning, automatic classification, pattern recognition and learning algorithm.

Since Z. Pawlak proposed the Rough Sets Theory in 1982, this theory has been well studied by many researchers and have made great progress.

Due to its advantage, which include the elimination of the need for additional information about data and the ability to extract rules directly from data itself, this theory has been used in more and more domains. In Section 3.1 the following section, the fundamental knowledge about Rough Sets Theory is introduced.

3.1. Information system and decision table

Information systems are used to represent knowledge. The notion of an information system presented in this paper is described in the works of Pawlak [12] and Yasdi [13]. A summary is included here for completeness.

An information system $S=(U, \Omega, V_q, f_q)$ consists of: U — a nonempty, finite set called the *universe*; Ω — a nonempty, finite set of *attributes*; $\Omega = C \cup D$, in which C is a finite set of *condition attributes* and D is a finite set of *decision attributes*; For each $q \in \Omega$, V_q is called the domain of q; f_q — an information function f_q: $U \rightarrow V_q$.

There are various possible interpretations of *objects* in practical applications, for example, cases, states, processes, patients and observations. *Attributes* can be interpreted as features, variables and characteristic conditions. In this paper, a special case of information systems called *decision table* or *attribute-value table*, is applied into the following analysis. In a decision table, the columns are labeled by attributes, and rows are by objects (states, processes, events, etc.). Example of a decision table (attribute-value table) is given in Table 1.

In Table 1, $U=\{1, 2, 3, 4, 5\}$, $\Omega=\{a, b, c, d, e\}$.

Table 1
A decision table

U^{Φ}	a	b	c	d	e
1	1	0	2	1	0
2	0	0	1	2	1
3	2	0	2	1	0
4	0	0	2	2	2
5	1	2	2	1	0

3.2. Lower and upper approximation

In an *Information System*, to every subset of attributes $B \subseteq \Omega$, a binary relation, denoted by $\mathrm{IND}_{\Omega}(B)$ or $\mathrm{IND}(B)$, called the B-indiscernibility relation, is associated and defined as follows:

$$\mathrm{IND}(B) = \{(x,y) \in U^2 : \text{for every } a \in B, \quad a(x) = a(y) \tag{1}$$

where $\mathrm{IND}(B)$ is an equivalence relation and

$$\mathrm{IND}(B) = \bigcap_{a \in B} \mathrm{IND}(a)$$

Objects x, y satisfying relation $\mathrm{IND}(B)$ are indiscernible by attributes from B.

Consider the subset $B=\{a, b, c\}$ in Table 1, then

$$U/\mathrm{IND}(\{a\}) = \{\{1,5\},\{2,4\},\{3\}\};$$

$$U/\mathrm{IND}(B) = \{\{1\},\{2\},\{3\},\{4\},\{5\}\}.$$

Sets that unions of some classes of the indiscernibly relation $\mathrm{IND}(B)$ are called definable by B.

If $S=(U, \Omega, V_q, f_q)$ is a decision table, $B \subseteq \Omega$ and $X \subseteq U$ then *B-lower* and *B-upper approximation* of X is defined respectively as follows:

$$\underline{B}X = \cup\{Y \in U/\mathrm{IND}(B) : Y \subseteq X\}; \tag{2}$$

$$\overline{B}X = \cup\{Y \in U/\mathrm{IND}(B) : Y \cap X \neq \varnothing\}, \tag{3}$$

where $U/\mathrm{IND}(B)$ denotes the family of all equivalence classes of B(classification of U).

The set $\underline{B}X$ is the set of all elements of U, which can be certainly classified as elements of X, with respect to the values of attributes from B; and the set $\overline{B}X$ is those elements of U which can be possibly defined as elements of X with respect of the value of the attributes from B. Finally, $BN_B(X)$ is the set of elements which can be classified neither in X nor in $\overline{X}$ on the basis of the values of attributes from B.

The set $BN_B(X) = \overline{B}X - \underline{B}X$ is called the B-boundary of X.

From Table 1, the subset of attribute $B = \{a, b\}$ and the subset of objects $X = \{2,3\}$, the following approximations can be derived:

$$\underline{B}X = \{3\};$$

$$\overline{B}X = \{2,3,4\};$$

$$BN_B(X) = \{2,4\};$$

3.3. Quality of approximation

Inexactness of a set (category) is due to the existence of a boundary region. The following *quality of the lower approximation* of X by B on Ω were introduced in Pawlak [12]:

$$\underline{\gamma}_B(X) = \frac{|\underline{B}X|}{|U|} \quad \text{and} \quad \overline{\gamma}_B(X) = \frac{|\overline{B}X|}{|U|}. \tag{4}$$

Thus, the quality of lower approximation of X by B in Ω is the ratio of the number of all certainly classified objects by attributes from B as being in X to the number of all objects in the system. $\overline{\gamma}_B(X)$ is intended to capture the degree of completeness of our knowledge about the set X. It is a kind of relative frequency. The *quality of upper approximation* of X by B in Ω is the ratio of the number of all possible classified objects by attributes from B as being in X to the number of all objects in the system. It is also a kind of relative frequency.

Two measures to describe inexactness of approximation classifications have been defined; the first measure is the *accuracy of approximation* of Ω by B. It expresses the possible correct decisions when classifying objects employing the attribute B.

$$\alpha_B(\Omega) = \frac{\Sigma \text{card}(\underline{B}X_i)}{\Sigma \text{card}(\overline{B}X_i)} \tag{5}$$

The second measure is called the *quality of approximation* of Ω by B. It expresses the percentage of objects, which can be correctly classified into class Ω employing the attribute B:

$$\gamma_B(\Omega) = \frac{\Sigma \text{card}(\underline{B}X_i)}{\text{card}(U)} \tag{6}$$

Table 2
The discernibility matrix for decision Table 1

U	1	2	3	4	5
1					
2	*a, c, d, e*				
3	*a*	*a, c, d, e*			
4	*a, d, e*	*c, e*	*a, d, e*		
5	*b*	*a, b, c, d, e*	*a, b*	*a, b, d, e*	

3.4. The discernibilty matrices and discernibility function

The elements of a discernibility matrix [14] of B is defined as follows:

$$(C_{ij}) = \{a \in B : a(x_i) \neq a(x_j)\}$$

$$\text{for } i,j = 1,2\ldots n \tag{7}$$

and a discernibilty function $f(B)$ is a Boolean function of n attributes defined as:

$$f(B) = \prod \sum_{(x,y)\in U^2} \delta(x,y)$$

$$= \cap\{\cup(c_{ij}) : 1 \leq j \leq i \leq n \cdot n; c_{ij} \neq 0\} \tag{8}$$

where $\prod$ denotes products of partitions.

For the decision table as shown in Table 1, the discernibility matrix is listed in Table 2.

The discernibilty function for this set is:

$$\begin{aligned} f(a,b,c,d,e) &= (a \vee c \vee d \vee e) a (a \vee c \vee d \vee e) \\ &\times (a \vee d \vee e)(c \vee e) \\ &\times (a \vee d \vee e) b (a \vee b \vee c \vee d \vee e) \\ &\times (a \vee b)(a \vee b \vee d \vee e) \\ &= ab(c \vee e) \end{aligned}$$

In this paper, the discernibility matrix and the discernibility function are used to extract the minimal reducts.

4. Specification of attributes field

Before using Rough Sets Theory, the attributes field must be specified to compose the decision table. In this paper, six attributes are extracted from the vibration signal for each sampling point. These six attributes can be divided into two categories, frequency domain and time domain. They are introduced as follows:

4.1. Frequency domain attributes

(a) IF — Waveform Complexity in frequency domain [15]

$$\mathrm{IF} = -\sum_{i=1}^{N/2} X(i)\log X(i), \quad (9)$$

where $X(i)$ — the FFT spectrum.

From the Eq. (9), it can be seen that IF is a frequency domain entropy, reflecting the complexity of FFT spectrum.

(b) CG — the center frequency of spectrum [16]

$$\mathrm{CG}(X) = \sum_{K=1}^{N/2} \frac{K}{N/2}\mu(X(K)), \quad (10)$$

where $\mu(X(K)) = X(K)/\sum_{j=1}^{N/2} X(j)$, $X(K)$ — the FFT spectrum, $k = 1,2,\dots,N$.

4.2. Time domain attributes

(a) IT — Waveform Complexity in time domain [15]

$$\mathrm{IT} = -\sum_{i=1}^{m} \lambda_i \log\lambda_i, \quad (11)$$

where λ_i — the singular value of a time series according to its period; m — the number of periods in a time series. The physical significance of IT is to reflect the complexity of time series. It is time domain entropy.

(b) σ — Nonperiod complexity [16]

$$\sigma = \frac{m}{m-1}\sum_{i=2}^{m}\lambda^2 i \bigg/ \sum_{i=1}^{m}\lambda^2 i, \quad (12)$$

λ_I — the singular value of a time series according to its period.

(c) D_x — the variance of time series

$$D_x = \frac{1}{n}\sum_{i=1}^{n}\left[x(t_i) - \bar{x}\right]^2, \quad (13)$$

n— length of a time series; $\bar{x}$— mean value of the whole series; $x(t_i)$ — time series.

(d) α_4 — Kurtosis of time series.

$$\alpha_4 = \frac{1}{n}\sum_{i=1}^{n}\left[x(t_i)\right]^4. \quad (14)$$

The above six attributes present the information contained in vibration signals both from the frequency domain and time domain. For IT and σ, they reflect the time series periodical characteristic because the single fault type shows the periodicity in time domain and the energy will increase in a certain frequency in the spectrum, which is reflected by attributes, IF and CG. Variance D_x and Kurtosis α_4 are the measures of the data distribution.

5. Discretization of attributes

Having established the attributes field, the crucial problem now is how to discretize them because the Rough Sets Theory cannot be used to deal with the continuous attributes, which is the disadvantage of this method.

Usually, the value of attribute changes monotonically with deterioration of the state. Boundary values to divide its domain into intervals corresponding to different state are selected by experience.

However, when experience to discretize the domain is not present, other discretization methods have to be used.

There are many discretization methods, such as *equal-width-intervals*, *equal-frequency-intervals* [17] and *Minimal Entropy Method* [18]. Among these methods, there are always some disadvantages that make them impractical. In [7], four methods defining attribute limit values were compared. They were called L-, W-, P- and C-method. For these four methods, practical experience is needed to calculate the limit value, so these four methods cannot be used in our case.

In this paper, a new method is proposed to discretize the attributes extracted from vibration signals. This method extracts the salient features from ChiMerge algorithm [19] and ε-indiscernibility relation algorithm [20]. The algorithm is presented below:

Step 1: all the objects are sorted in ascending order of the value of the attributes.

Table 3
The quality of classification for different μ

μ		Class 0	Class 1
$\mu = 0.1$	lower approximation	25	3
	upper approximation	44	22
	accuracy of approximation	0.5682	0.1364
	accuracy of classification	0.4242	
	quality of approximation	0.5091	
$\mu = 0.2$	lower approximation	25	4
	upper approximation	43	22
	accuracy of approximation	0.5814	0.1818
	accuracy of classification	0.4462	
	quality of approximation	0.5273	
$\mu = 0.3$	lower approximation	32	14
	upper approximation	40	22
	accuracy of approximation	0.8000	0.6364
	accuracy of classification	0.7419	
	quality of approximation	0.8364	
$\mu = 0.4$	lower approximation	34	21
	upper approximation	34	21
$\mu = 0.9$	accuracy of approximation	1	1
	accuracy of classification	1	
	quality of approximation	1	

Step 2: calculate the χ^2 value between each consecutive pair. At the beginning, every object constitutes a separate cluster. The calculation of χ^2 value is defined in the following equation:

$$X^2 = \sum_{i=1}^{2} \sum_{j=1}^{k} \frac{\left(A_{ij} - E_{ij}\right)^2}{E_{ij}} \tag{15}$$

where k— number of classes(decision attributes); A_{ij} — number of objects in the ith interval, jth class; R_i — number of objects in the interval in the ith interval, $\sum_{j=1}^{k} A_{ij}$; C_j — number of objects in the jth class, $\sum_{i=1}^{2} A_{ij}$; N— total number of objects; E_{ij} — expected frequency of $A_{ij} = R_i C_j / N$, if either R_i or C_j is 0, E_{ij} is set to 0.1.

Step 3: The pair of clusters for which the computed χ^2 is minimal is found, if the *purity condition* described below is fulfilled, the selected clusters are merged. If for a selected pair of consecutive clusters, the purity condition is not fulfilled, the next pair according to the χ^2 is to be searched for. The process is repeated until there is no pair of clusters fulfilling the purity condition or, if the number of clusters decreases to 2.

The purity condition is defined as follows. During the entire process, the distribution of class frequencies for each cluster (C_j/N) is stored and updated separately. When merging two clusters, their distributions are summed. Based on that distribution, a majority class, that is, the class with maximal share in the cluster, may be easily computed. If the share of majority class in the resulting cluster does not drop below a fixed threshold μ, the clusters considered as candidates for merging is deemed to fulfill the purity condition.

In this paper, the *quality of approximation* $\gamma_B(\Omega)$ is regarded as the standard to evaluate the final discretization result. If the quality of classification remains stable, the corresponding $\mu$is adopted.

Following the above algorithm, a case in Ref. [7] is tested and compared with the L-method mentioned in Ref. [7], the result is presented in Tables 3 and 4.

From the Table 3, when value of $\mu$is larger than 0.4, the quality of approximation reaches 1. For

Table 4
The quality of classification of every attribute for the different μvalue

μ	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}
0.1	0.35	0.29	0.35	0	0.40	0.35	0	0.38	0	0.38	0.36	0.36
0.2	0.35	0.29	0.35	0	0.40	0.35	0	0.38	0	0.38	0.36	0
0.3	0.35	0	0.35	0	0.40	0.35	0	0.38	0.20	0.55	0.36	0
0.4	0.35	0.29	0.35	0	0.65	0.20	0.15	0.60	0.20	0.55	0.36	0.31
0.5	0.45	0.29	0.44	0	0.65	0.20	0.15	0.69	0.47	0.55	0.16	0.31
0.6	0.53	0.36	0.56	0.04	0.65	0.20	0.15	0.69	0.47	0.55	0.16	0.31
0.7	0.53	0.36	0.56	0.04	0.65	0.20	0.15	0.69	0.47	0.56	0.16	0.31
0.8	0.53	0.15	0.56	0.04	0.65	0.20	0.15	0.69	0.47	0.56	0.16	0.31
0.9	0.67	0.56	0.56	0.04	0.65	0.24	0.33	0.69	0.49	0.64	0.20	0.31
L	0	0.04	0.35	0.24	0.02	0.13	0.35	0.13	0.16	0.09	0.09	0.09

Table 5
The decision table composed of indices of vibration signals collected from the surface of a 4135 diesel engine

No	The first sampling point						The second sampling point						The third sampling point						D
	IF	IT	σ	CG	D_x	α_4	IF	IT	σ	CG	D_x	α_4	IF	IT	σ	CG	D_x	α_4	
1	968.63	9.463795	0.000236	0.403311	2330.824	4.392614	920.8692	6.527313	9.52×10^{-5}	0.40796	1253.28	6.014534	1779.608	12.39776	0.000424	0.471584	2829.108	6.281026	1
2	966.0803	9.12031	0.000216	0.405479	2292.812	4.367723	854.1059	6.564167	9.76×10^{-5}	0.399058	1269.885	6.144371	1757.672	12.48108	0.00047	0.464052	2782.227	5.9828	1
3	928.178	9.212937	0.000203	0.408101	2275.398	3.857294	750.0185	7.272025	0.000129	0.386061	1747.084	4.22025	1631.94	11.93877	0.000417	0.481095	3184.857	5.841912	1
4	934.3243	9.652909	0.000266	0.402286	2273.328	3.832025	815.8439	8.049914	0.000175	0.3794	1698.614	4.294675	1689.364	12.26842	0.000451	0.472905	3098.786	5.509754	1
5	906.8166	8.18973	0.000163	0.403422	2393.748	3.969839	929.2076	6.078476	8.79×10^{-5}	0.364983	991.9959	3.886004	1657.002	11.52554	0.000387	0.479898	2819.799	6.387094	1
6	913.2175	8.270162	0.000168	0.40688	2417.229	3.97771	911.7933	5.750949	7.11×10^{-5}	0.367497	998.8395	3.935522	1632.826	12.2223	0.000442	0.468387	2811.04	6.353366	1
7	860.4138	10.25628	0.000291	0.402683	2364.624	3.934964	1206.267	15.42848	0.000812	0.499007	6187.435	6.20678	1842.024	12.03191	0.000434	0.481547	2962.745	6.938825	1
8	854.1437	10.19812	0.000289	0.400282	2298.57	3.966693	1179.122	16.36945	0.000956	0.48497	6201.076	6.220813	1907.579	11.12551	0.000346	0.494687	2997.293	7.18038	1
9	938.1746	10.29018	0.0003	0.410277	2404.253	4.769968	1003.692	5.862651	7.80×10^{-5}	0.420415	740.6773	4.53411	1718.902	11.46342	0.000372	0.497391	2788.227	4.824433	1
10	933.4472	10.92431	0.000347	0.406396	2462.082	4.70767	965.6075	5.571255	7.09×10^{-5}	0.421925	704.0458	4.51455	1700.047	12.2018	0.000442	0.493479	2827.68	5.027644	1
11	748.9689	11.53852	0.00042	0.40151	4115.304	7.406265	1083.173	7.379985	0.00013	0.456184	1332.931	4.590098	1856.117	13.308	0.000564	0.507847	3771.855	9.348355	2
12	759.4642	11.8019	0.000409	0.397719	3878.832	6.721322	1063.689	8.410843	0.000194	0.435923	1386.426	5.777319	1840.721	13.14226	0.00054	0.511589	3821.941	11.40681	2
13	828.2558	10.44427	0.000276	0.39677	2994.575	6.282765	1028.062	8.915699	0.000189	0.457331	1918.919	4.391126	2097.51	13.391	0.000507	0.509141	3497.027	7.669647	2
14	834.6539	10.05399	0.000272	0.397975	2905.803	5.931248	1036.01	8.372194	0.000176	0.461422	1897.649	4.367016	2072.981	13.49212	0.000573	0.50746	3508.434	7.633504	2
15	841.6564	11.49167	0.000393	0.407036	3763.61	7.153016	978.5342	9.329181	0.000239	0.452388	2186.351	9.243974	1943.545	13.8059	0.000604	0.505638	3363.842	8.870944	2
16	856.2049	10.90233	0.000335	0.413575	3721.703	7.245844	980.6038	8.691654	0.000196	0.451483	2234.845	10.66554	1989.373	13.8788	0.000649	0.498117	3307.159	9.017534	2
17	837.6024	12.09892	0.00037	0.38792	2842.585	4.466144	1053.039	16.13031	0.000835	0.465067	6109.673	7.04936	1924.621	14.95749	0.000689	0.519337	4255.616	13.00923	2
18	860.2143	11.44314	0.00039	0.396719	2843.152	4.233111	1095.057	14.70214	0.000702	0.482986	5923.511	6.747132	1972.851	14.48197	0.000701	0.523222	4326.552	12.61506	2
19	960.8651	10.30765	0.000296	0.425713	2322.519	5.68833	986.1711	12.45634	0.000469	0.456184	3991.37	5.155113	2052.661	11.15374	0.000352	0.507847	2259.82	5.365647	3
20	1006.686	10.06968	0.000272	0.438123	2399.944	5.655421	982.3972	11.67106	0.000395	0.435923	3784.813	5.057999	2097.9	11.92495	0.000414	0.511589	2361.131	5.99369	3
21	929.9577	10.6088	0.000322	0.416517	2511.911	6.263797	1010.356	17.34732	0.001111	0.457331	6879.374	4.825951	1983.917	11.47818	0.000379	0.509141	2521.838	5.814606	3
22	950.1628	9.867081	0.000258	0.42326	2532.227	6.000203	1020.085	17.95784	0.001219	0.461422	6777.754	4.746955	2013.993	11.72726	0.000398	0.50746	2546.619	6.290961	3
23	981.588	10.45415	0.00027	0.404788	2434.985	6.452266	1038.908	16.07058	0.000908	0.452388	6593.186	5.43869	2129.193	11.72058	0.000408	0.505638	2541.82	5.930342	3
24	952.3014	11.44342	0.000411	0.425976	2411.399	6.479662	1027.337	16.84694	0.001031	0.451483	6621.677	5.140761	2016.157	11.48792	0.00039	0.498117	2486.469	5.903653	3
25	1042.218	9.264848	0.00022	0.42183	2213.716	6.605749	994.8993	15.01201	0.000744	0.482986	5712.068	5.17196	1951.248	11.63425	0.000396	0.523222	2565.575	6.246081	3
26	978.3475	9.350612	0.000227	0.418596	2090.416	6.983087	1057.175	15.5163	0.000818	0.471835	6341.818	4.144594	2145.83	12.04028	0.000439	0.529551	2382.222	5.827804	3
27	976.9817	9.742837	0.000251	0.407637	2101.55	6.954501	1063.936	15.70303	0.000841	0.457609	6216.18	4.116981	2141.317	12.86789	0.000505	0.534045	2452.92	6.125845	3
28	1070.333	8.090354	0.000163	0.409697	1461.063	5.024917	1010.476	11.91118	0.000424	0.437947	3998.688	3.981336	1891.114	12.35004	0.000454	0.500906	3538.96	8.215332	4
29	1073.929	8.027985	0.000156	0.409937	1444.384	5.531318	1016.647	11.61503	0.000388	0.446539	4078.717	4.054873	1876.782	13.03733	0.000539	0.489654	3529.036	8.094243	4
30	978.7982	9.108615	0.000197	0.418583	1744.463	4.577565	998.7954	16.98574	0.000967	0.431412	7240.758	4.770887	2002.549	13.5434	0.000534	0.486128	3531.318	7.881434	4
31	905.979	7.856989	0.000151	0.422491	1794.124	4.729704	1015.769	16.92916	0.001014	0.432994	7213.852	5.011682	1999.3	13.12159	0.000533	0.495409	3591.101	7.407416	4
32	1030.805	9.255664	0.000225	0.419269	1846.876	5.051797	1082.497	16.11993	0.00084	0.418296	6545.535	5.099336	2043.259	14.48351	0.000687	0.484417	3774.407	7.736985	4
33	1039.858	10.42931	0.000309	0.409654	1798.844	5.115474	1094.823	15.05881	0.000763	0.424336	6527.427	5.136675	1958.081	11.9159	0.000421	0.500597	3757.967	7.628443	4
34	969.4824	7.667061	0.000139	0.380958	1485.435	5.246665	1024.016	18.28751	0.001295	0.424527	8084.484	4.826401	1997.751	12.8451	0.000499	0.500984	3557.557	7.417416	4
35	969.2537	7.65842	0.00014	0.381866	1444.417	5.524306	1051.173	16.56692	0.001003	0.428467	7641.543	5.047001	2030.748	12.80923	0.000514	0.498772	3634.432	7.901189	4
36	862.2621	7.294524	0.001235	0.357076	1812.592	4.667809	972.4308	15.9116	0.000802	0.40641	7219.99	5.356308	2025.661	13.37121	0.00052	0.498935	3330.902	6.197613	4
37	867.5084	8.224223	0.000173	0.353314	1826.878	4.685793	1028.222	15.23855	0.000806	0.412785	6332.31	5.779836	1869.5	13.09545	0.000525	0.49861	3215.49	7.485392	4

every attribute in Table 4, when the value of μ is larger than 0.4, the quality of approximation remain stable. So in this example, 0.4 is assigned to μ.

In the entire calculation, no a priori knowledge is needed. Only the purity condition is used to end its discretization. This advantage overcomes the shortcoming of ChiMerge algorithm, which needs user-defined parameters — to select the minimal and maximal intervals. However, this new algorithm only considers the relationship between a single continuous valued attribute and its decision attribute.

6. Implementation and discussion

In this section, the decision table composed of 18 condition attributes (six attributes for each sampling point) and four decision attributes. The decision attributes correspond to the following:

- normal,
- intake valve clearance is too small,
- intake valve clearance is too large,
- exhaust valve clearance is too large.

The decision table constructed is shown in Table 5.

Having discretized the condition attributes, discernibility matrix and discernibility function are used to get the final reducts. Here the quick reducing method [8] is used to choose the minimal reducts. The procedure of the quick reducing method is described as follows:

- The core or attributes with highest quality of classification is calculated.
- The above core or attributes are augmented by one of the remaining attributes. The pair with the highest quality of classification is chosen.
- The process is repeated until the quality of classification is equal to one or equals to the quality of classification for all the condition attributes.

According to the quick reducing method, the following results are presented in Table 6.

In Table 6, according to the minimal reducts, the distribution of attributes belonging to different sampling points can be calculated. From 0.2 to 0.5 of μ value, the attribute numbers appearing in the final reducts of sampling point 1 (from a_1 to a_6) — the first cylinder head, is 11/28; the attribute numbers of sampling point 2 (from a_7 to a_{12}) — the second cylinder head is 14/28, and the attributes appearing in the final reducts of sampling point 3 (from a_{13} to a_{18}) — the middle point in piston stroke of the second cylinder is 3/28. That is to say that the second sampling point, which corresponds to the second cylinder, is more important than the other points in the final decision.

Table 6
Final reducts according to different μ value

μ	Quality of classification	Final reducts
$\mu = 0.2$	1	$\{a_4, a_9, a_{10}, a_{13}\}$; $\{a_5, a_7\}$; $\{a_5, a_{10}\}$; $\{a_5, a_{13}\}$
$\mu = 0.3$	1	$\{a_5, a_6\}$; $\{a_5, a_{10}\}$; $\{a_{10}, a_{12}\}$
$\mu = 0.4$	1	$\{a_4, a_7\}$; $\{a_6, a_{12}\}$; $\{a_5, a_{11}\}$; $\{a_{10}, a_{12}\}$
$\mu = 0.5$	1	$\{a_6, a_{11}\}$; $\{a_{10}, a_{17}\}$

Where a_1, a_7, a_{13} — IF; a_2, a_8, a_{14} — IT; a_3, a_9, a_{15} — σ; a_4, a_{10}, a_{16} — CG; a_5, a_{11}, a_{17} — D_x; a_6, a_{12}, a_{18} — α_4.

This agrees well in practice. The second cylinder was observed to be the most sensitive to the changing of vibration. Thus, the results obtained from the Rough Sets Theory agrees well with the observation.

From the final reducts, the most important attribute can be chosen according to their appearance times in the final reducts. Attributes a_5 and a_{10} get the highest value, which is equal to 6/28. Attribute a_5 is the variance value of the first sampling point and attribute a_{10} is the spectrum center value of the second sampling point.

7. Conclusions

In this paper, valve fault diagnosis using Rough Sets Theory is presented. Summarizing the formulation; first, the decision table is established. In this process, the attributes field has to be specified according to collected signals. Next, using a discretization method, either with expert experience or not, transform the continuous valued attributes to discrete ones. Finally, the Rough Sets Theory is used to get the final reducts and to extract the rules. These rules are used to distinguish the fault type or to inspect the dynamic characteristic of the machinery.

Through the implementation and results, the following observations and conclusions are made:

- The Rough Sets Theory used to diagnose the valve fault of a 4135 diesel engine has been found to be effective.
- The lack of a priori knowledge in practice for fault diagnosis is very prevalent and therefore the new discretization method proposed in this paper will be very useful.

References

[1] L. Qu, J. Meng, Fault-diagnosis techniques and current advanced sciences, in: The Proceedings of 95' Conference on Equipment Diagnostic Techniques, P.R. China, 1995, pp. 9–34, (in Chinese).

[2] L. Qu, Z. He, Machinery Fault Diagnostics, Shanghai Scientific Press, P.R. China, 1981, (in Chinese).

[3] L. Qu, Y. Shen, Orbit complexity: a new criterion for evaluating the dynamic quality of rotor systems, Journal of Mechanical Engineering Sciences 207 (1993) 325–334.

[4] D. Huang, Vibration analysis and fault diagnosis for diesel engine, Journal of Vibration Engineering 8 (2) (1995) 144–149, (in Chinese).

[5] S. Tsumoto, H. Tanaka, Automated discovery of medical expert system rules from clinical databases based on Rough Sets, in: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96), AAAI Press, Menlo Park, 1996, pp. 63–69.

[6] W. Ziarko, R. Golan, D. Edwards, An application of Datalogic/R knowledge discovery tool to identify strong predictive rules in stock market data, in: Proceedings of AAAI Workshop on Knowledge Discovery in Databases, (Washington D.C.), 1993, pp. 93–101.

[7] R. Nowicki, R. Slowinski, J. Stefanowski, Evaluation of vibroacoustic diagnostic symptoms by means of the Rough Sets Theory, Computers in Industry 20 (1992) 141–152.

[8] R. Slowinski, C. Zopounidis, Application of the rough set approach to evaluation of bankruptcy risk, Intelligent Systems in Accounting, Finance and Management 4 (1995) 27–41.

[9] Z. Pawlak, J. Grzymala-Busse, R. Slowinski, W. Ziarko, Rough Sets, Communications on the ACM 38 (11) (1995) 89–95.

[10] D. Tan, The Vibration Control of Internal Combustion Engine, Southwest Jiaotong University Press, P.R. China, 1995, (in Chinese).

[11] J.F. Dagel, R.N. Brady, Diesel Engine and Fuel System Repair, 4th edn., Prentice Hall Press, 1998.

[12] Z. Pawlak, Rough Sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishing, 1991.

[13] R. Yasdi, Combining Rough Sets Learning and Neural Learning method to deal with uncertain and imprecise information, Neurocomputing 7 (1995) 61–84.

[14] A. Skowron, A. Rauszer, The discernibility matrices and functions in information systems, in: R. Slowinski (Ed.), Intelligent Decision Support, Handbook of Applications and Advances of Rough Sets Theory, Kluwer Academic Publishing, 1992, pp. 331–362.

[15] J. Meng, Some Advanced Techniques in Fault Feature Extraction for Large Rotating Machinery, PhD Thesis, Xi'an Jiaotong University 1996 (in Chinese).

[16] H. Liu, On the Integration and Extraction of diagnostic Information, PhD Thesis, Xi'an Jioatong University 1997 (in Chinese).

[17] N. Shan, H.J. Hamilton, W. Ziarko, N. Cercone, Discretization of continuous valued attributes in classification systems, in: Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery (RFSD'96) (Tokyo, Japan),1996, pp. 74–81.

[18] U.M. Fayyad, K.B. Irani, On the handling of continuous-valued attributes in decision tree generation, Machine Learning 8 (1992) 87–102.

[19] R Kerber, ChiMerge: discretization of numeric attributes, in: Proceedings of AAAI-92, AAAI-Press, 1992, pp. 123–128.

[20] K. Krawiec, R. Slowinski, D. Vanderpooten, Construction of rough classifiers based on application of a similarity relation, in: Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery (RFSD'96) (Tokyo, Japan), 1996, pp. 23–30.

[21] R. Slowinski, Intelligent Decision Support, Handbook of Applications and Advances of Rough Sets Theory, Kluwer Academic Publishing, 1992.

Lixiang Shen is currently a PhD candidate of Mechanical and Production Engineering in National University of Singapore. Born on March 27, 1972 in China, he received the B.Eng. degree in 1993 from the Zhengzhou Institute of Technology and MSc degree in 1998 from Xi'an Jiaotong University, respectively. His current research interests are in Data Mining, Decision Support System, Artificial Intelligence and their applications to financial engineering.

Francis E.H. Tay is currently an Assistant Professor of Mechanical and Production Engineering in National University of Singapore. He graduated with B.Eng. (First Class) and M.Eng. from the National University of Singapore in 1986 and 1991, respectively. Subsequently, he obtained his PhD degree from MIT in 1996. After his B.Eng., he worked with the Singapore Aerospace Manufacturing as an Engineer between 1986 and 1989. His research interests are in Design Automation and Design Optimisation, Total Design Methods, and Data Mining Using Soft Computing.

Liangsheng Qu is currently a professor of School of Mechanical Engineering and the Head of Research Institute of diagnostics and Cybernetics in Xi'an Jiaotong University, P.R. China. Born on January 28, 1931, he received the B.Eng. from the Jiaotong University, Shanghai, in 1952 and MSc degree from the Harbin University of Technology in 1955, respectively. During the many-years research acitivities, he invented Holospectra technology in 1982. His project — Online Condition Monitoring and Diagnostic System for Large Centrifugal Compressors (1993–1996) is awarded the 1998 Prize of National Science and Technology, which is the highest level in science and technology development awards in P.R. China. Now he is a senior member of SME and CSME, member of ISO TC108. His current research interests cover Machinery Surveillance and Diagnostics, Vibration, Noise and Acoustic Emission, Non-linear Diagnostics, Vibration Signal Processing, Precision Machining techniques, Dynamic Field Balancing, Time Series Applications, Modeling and Feature Extraction, Intelligent Manufacturing Process (Joint Research with University of Michigan-Ann Arbor; Joint Research with Nanyang Technological University of Singapore). He is author and co-author of more than 140 papers and books. And he is also editors of many research journals.

Yudi Shen is currently an Associate Professor of School of Mechanical Engineering, Xian Jiaotong University, P.R. China. Born in 1948 in China, she received the B.Eng. degree in 1976 from Huazhong University of Science and Technology. Her current research interests are in the Reciprocating Machinery Surveillance and Diagnostics, Sampling Techniques, Signal Processing and Gearbox fault diagnostics. She is author and co-author of more than 10 papers and books.