Fault diagnosis method of gearbox based on sampling point optimization and features fusion

Huiling Liu

School of Mechanical, Jinzhong University,
Jinzhong 030619, Shanxi, China

Abstract
To gain the fault information and determine the fault position, monitoring point should be selected properly, otherwise, it usually fails to fulfill the task effectively and accurately in monitoring and diagnosing the gearbox. A method based on sampling point optimization and features fusion is proposed to improve the classification accuracy of the gearbox fault diagnosis . First the vibration signal of gearbox is processed by empirical mode decomposition, energy characteristic parameters based on empirical mode decomposition are extracted as the fault feature vector, and the decision table is established according to the gearbox fault mechanism and Rough Sets Theory. Then a attributes discretization algorithm based on PSO is put forward to discrete decision table, a method based on the discernibility matrix and attribute frequency is presented to calculate the minimal attribute reduction sets without computing core attribute. The most sensitive signal monitoring point is achieved through analyzing the final reduction sets. Finally, fault diagnosis system based on rough sets theory and support vector machine is constructed, the classification result is consistent with that of the sampling point optimization based on attributes reduction obviously. The experiment results show the method is effective and could improve the gearbox fault diagnosis performance, which provides a relatively generic solution for the similar mechanical system.
Key words: SAMPLING POINT OPTIMIZATION, FEATURE FUSION, ROUGH SETS THEORY, SUPPORT VECTOR MACHINE, GEARBOX, FAULT DIAGNOSIS

1. Introduction

Gearbox is one of the most important transmission parts in the mechanical systems, but, for some uncertain reasons, the failure rate is higher. While monitoring and diagnosing the gearbox, selecting monitoring points improperly usually fails to gain the fault information effectively, as well as determine the fault position. More information could be collected, if some sensors are installed on the gearbox, but the signal processing would be longer. Up to now, it is still not suitable for real-time online diagnosis, so, how to select the most sensitive monitoring point becomes a challenging and important issue for fault diagnosis [1].

Now, the related research methods can be classified into three kinds, respectively, the method based on analytical model [2], knowledge [3] and signal processing [4]. The first method includes parameter estimation, equivalent space, etc., although theoretical research has made some achievements to certain problems, the accurate mathematical model is hard to set up, which limits its application. The method based on knowledge contains fault tree, directed graph, and the graph theory, which needs to determine relevant parameters according to the prior knowledge, while prior knowledge is usually difficult to obtain. The last method includes Fourier transform, wavelet, empirical mode decomposition, rough sets theory, neural network, particle swarm optimization technology, etc., which values the importance and sensitivity of every point directly according to the relationship between the vibration signal and fault types, so it is widely used.

Z. Pawlak proposed Rough Sets Theory[5] first, which is a new mathematical tool Dealing with the fuzzy and uncertainty knowledge. The advantage is that it needs neither any prior knowledge nor correcting the inconsistency manifested in data. In recent years, Rough Sets Theory has been applied in fault diagnosis field more and more, especially the condition attribute reduction technology in the fault characteristic parameter set optimization has a great result [6]-[10].

This paper aims to study the condition attribute discretization and reduction technology and apply it to optimize the monitoring points of the gearbox, and further obtain the most sensitive and effective sampling points. Section 2 presents the vibration characteristics of gearbox, energy characteristic parameters based on empirical mode decomposition are extracted. In section 3, an attributes discretization algorithm based on PSO and a method based on the discernibility matrix and attribute frequency are presented. In section 4, fault diagnosis system based on rough sets theory and support vector machine is constructed. Section 5 gives a conclusion to the whole work.

2 Characteristics of gearbox

The vibration signals of gearbox often shows nonlinear and non-stationary characteristics and the fault information is often embedded in the strong background noise, so it is very difficult to get the sensitive feature parameters. If the monitoring points were selected improperly, it could lead to fail to gain the fault information effectively and determine the fault position accurately.

The fault diagnosis experiments are done on JZQ-250 gearbox fault diagnosis test apparatus, seen in Figure 1, which consists of the input shaft, the intermediate shaft, the output shaft, three pairs of bearings and two pairs of straight gears. Three kinds of single faults including tooth fracture, outer ring crackle and inner ring pitting corrosion, two kinds of compound faults including tooth fracture and outer ring crackle, tooth fracture and inner ring pitting corrosion are set in the intermediate shaft, so three monitoring points are selected according to the gearbox fault mechanism and prior experience, respectively the measuring points on two bearings of the intermediate shaft and that near the load of the input shaft in Figure 2.

The parameters are: the rated speed 1200r/min, the sampling frequency 4000Hz. Six states are studied: normal state and five kinds of faults above. The time domain signals from JZQ-250 gearbox surface are collected. Space is limited, so here only lists the vibration signals of the sampling point Ⅲ in Figure 3, including time domain and power spectrum, which both cannot distinguish various operating conditions. For power spectral density is core by Fourier transform that is not suitable for non-linear non-stationary vibration signal.

Empirical mode decomposition is proposed by Norden E.Huang et al in 1998, which is suitable for nonlinear and non-stationary signal. It can decompose the complex original signal into several simple components which are intrinsic mode functions(IMF). In fact, every intrinsic mode function represents one kind of vibration mode, every component is decomposed in order of descending frequency. When the gearbox fails, the energy of each intrinsic mode function would change, so the energy distribution of the various intrinsic mode functions can reflect the different types of faults. In the test, energy characteristic parameters based on EMD are extracted as the fault feature vector. The extract method is described in detail in the references [11]. Figure 4 lists every IMF in time domain of every state. The number of IMF of the normal state is 11, that of five kinds of faults above are 12,10,10,11,11 successively. In fact, decomposition process will produce the false IMF component, which has nothing to do with the original signal, so it is necessary to calculate the correlation coefficient of the IMF component and the original signal to obtain an effective component. In the paper, the critical value is 0.06 , the components whose correlation coefficient is greater than 0.06 are used to extract the gearbox fault feature vector. Finally the first six IMF components are used to calculate energy characteristic parameters taking into account all the conditions.

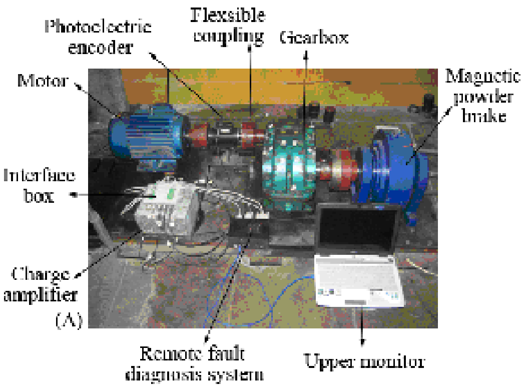


Figure 1. Gearbox fault diagnosis test apparatus

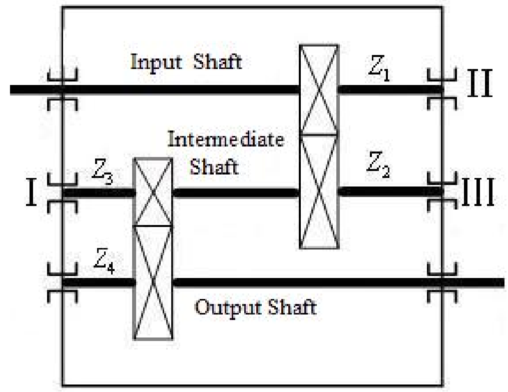
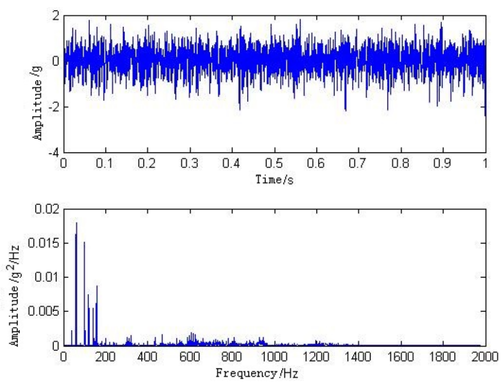
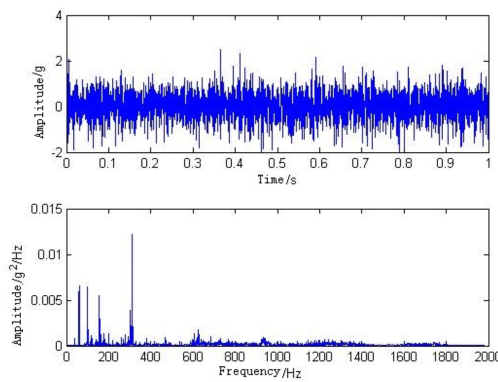


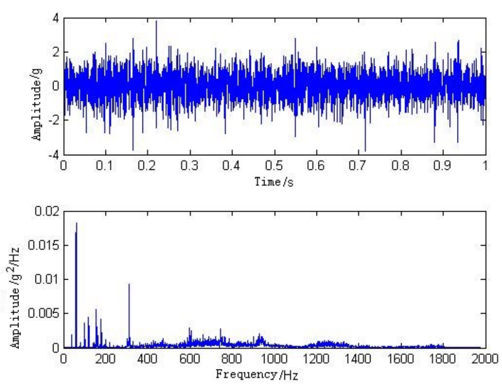
Figure 2. Monitoring points layout



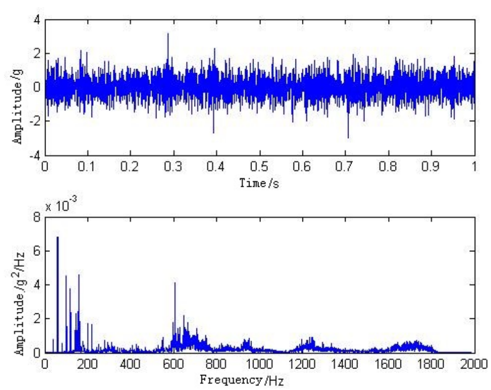
(a) Normal state



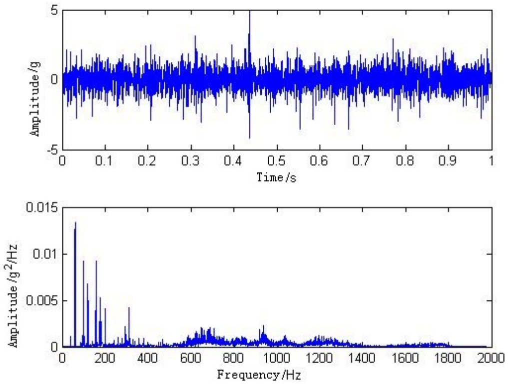
(b) Tooth fracture



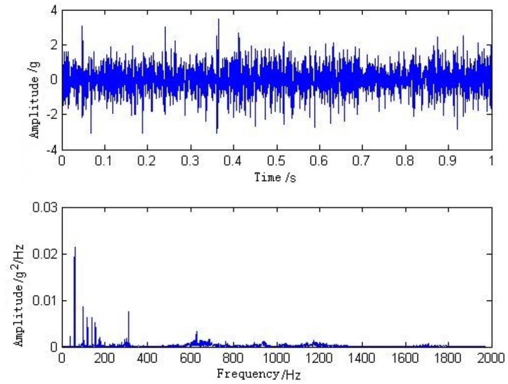
(c) Outer ring crackle



(d) Inner ring pitting corrosion

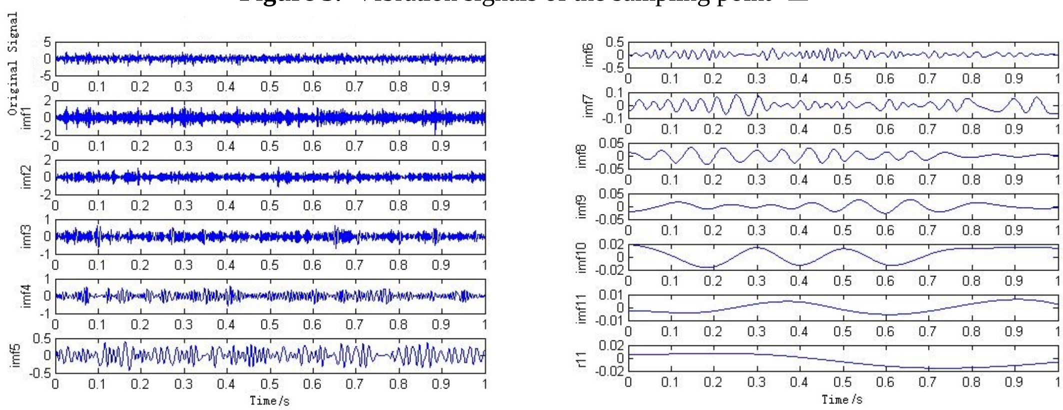


(e) Tooth fracture and outer ring crackle

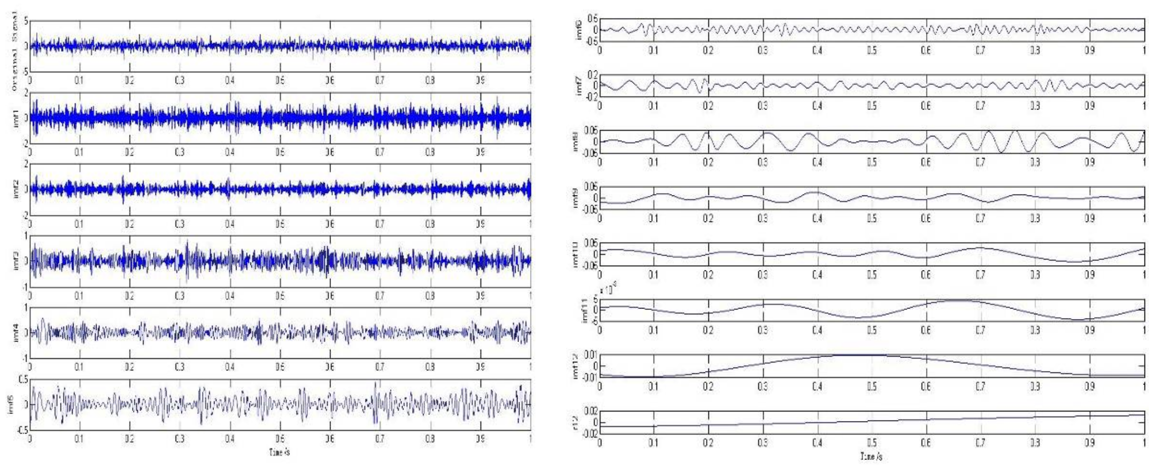


(f) Tooth Fracture and Inner Ring Corrosion

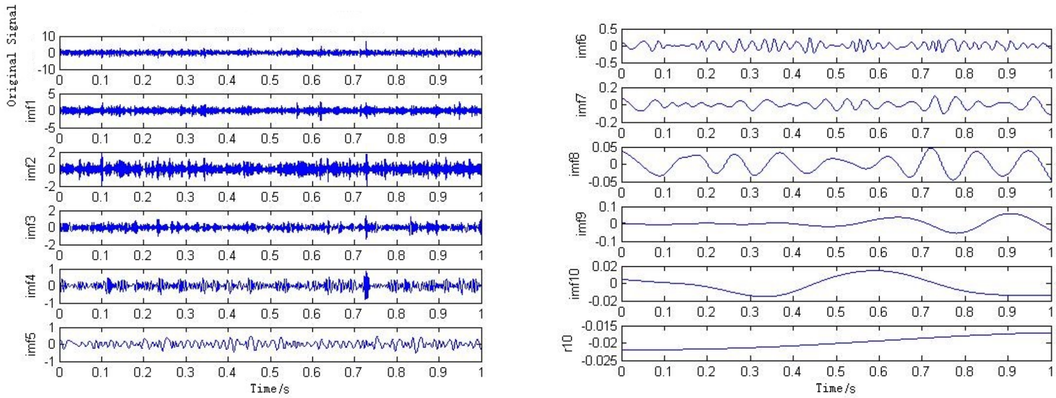
Figure 3. Vibration signals of the sampling point Ⅲ



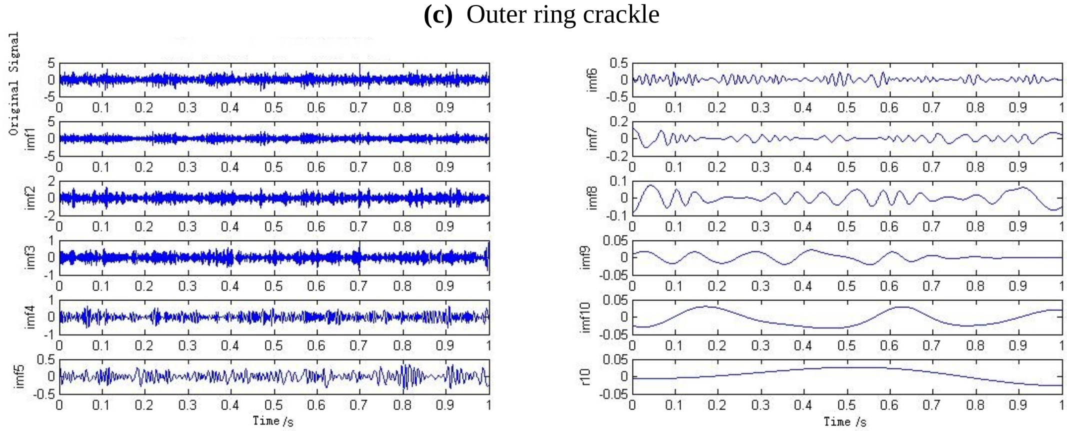
(a) Normal state



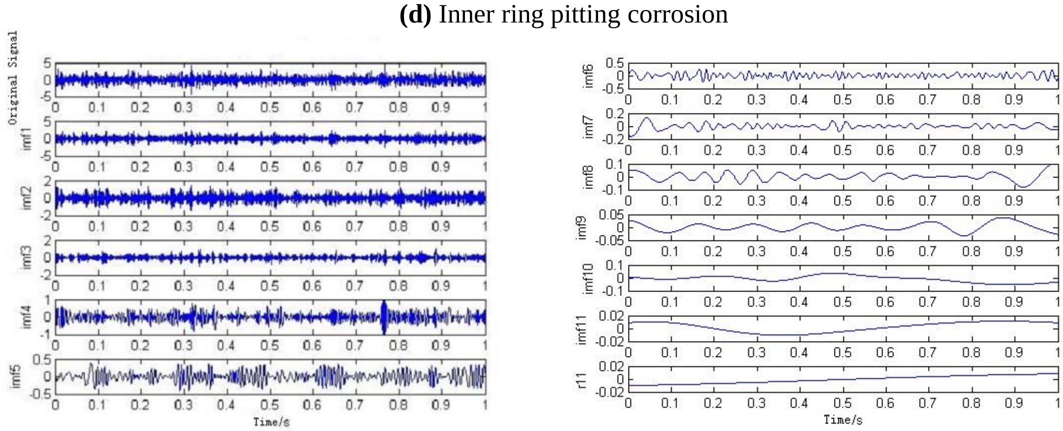
(b) Tooth fracture



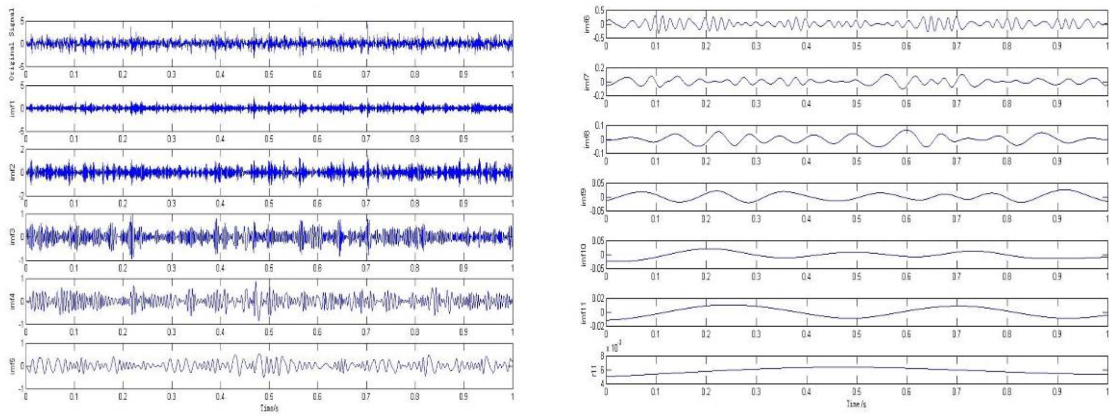
(c) Outer ring crackle



(d) Inner ring pitting corrosion



(e) Tooth fracture and outer ring crackle



(f) Tooth fracture and inner ring pitting corrosion

Figure 4. EMD results of every state of the sampling point III

3 Rough sets theor

3.1 Basic Concepts

Rough Sets Theory is a new mathematical tool dealing with the fuzzy and uncertainty knowledge. The core idea is to delete the redundant attributes under the condition of keeping classification ability. The information system is described in detail in the works of Pawlak and Yasdi [12]. The related basic concepts are discussed here.

If $S=\left(U,\Omega,V_q,f_q\right)$ represents an information system, U is the universe, Ω is a set of condition attributes and decision attributes, V_q is the attribute value domain, f_q is a function $f_q:U\to V_q$..

If there is $B\subseteq\Omega$, $X\subseteq U$, $\underline{B}X=\cup\left\{Y\in U/B\,|\,Y\subseteq X\right\}$ represents B-lower approximation of X, which is a set of those elements belonged to the universe completely according to knowledge. $\overline{B}X=\cup\left\{Y\in U/B\,|\,Y\cap X\neq\phi\right\}$ represents B-upper approximation of X, which is a set of those elements belonged to the universe partly according to knowledge.

If there is a condition attribute set $C\subseteq\Omega$, a decision attribute set $D\subseteq\Omega$, $POS_C(D)$ is C positive region of D. Dependence coefficient $k=\gamma_c\left(D\right)=\left|POS_c\left(D\right)\right|/\left|U\right|=\left|\bigcup_{X\in U/D}\underline{C}X\right|\Big/\left|U\right|$ $(0\le k\le 1)$, which can reflect the classification ability of C, especially when $k=1$, D is determined by C completely . At present, whether dependence coefficient changes is used to judge ending the attributes discretization and reduction algorithm.

In order to facilitate research and analysis, six states are noted as 1,2,3,4,5,6 respectively and each state takes 30 samples, 180 samples totally every sampling point. Taking 180 samples as the universe $U=\left\{x_1,x_2,\cdots,x_{180}\right\}$, six indicators as the condition attributes $C=\left\{a,b,c,d,e,f\right\}$ and six states as the decision attribute D, construct the decision table. The decision tables of three monitoring points correspond to Tables 1-3.

3.2. Continuous Attributes Discretization

Rough Sets Theory can only deal with discrete data. Now there are several discrete algorithms, such as Equidistant and equifrequent classification, NaiveScaler (NS) algorithm and Semi-NaiveScaler algorithm. NS and SNS algorithm have often been used to get the candidate breakpoint set, but which usually result in too many points. Here, a new method of the attributes discretization based on Particle Swarm Optimization[4] is proposed.

Particle Swarm Optimization(PSO) is proposed by Dr. Kennedy and Dr. Eberhart in 1995. There is the d-dimensional search space, the population consists of n particles. Velocity of the i-th particle is

$V=(v_{i1},v_{i2},\cdots,v_{id})$, the particle position vector is $X_i=(x_{i1},x_{i2},\cdots,x_{id})$, the global optimal location vector of the population is $P_g=(p_{g1},p_{g2},\cdots,p_{gd})$, as the optimal position vector currently of the i-th particle is $P_i=(p_{i1},p_{i2},\cdots,p_{id})$. The velocity and position of each particle update according to the following formula:

$$v_{id}(t+1)=v_{id}(t)+c_1r_1(p_{id}-x_{id}(t))+c_2r_2(p_{gd}-x_{id}(t)) \quad (1)$$

$$x_{id}(t+1)=x_{id}(t)+v_{id}(t+1) \quad (2)$$

Where, i ranges of [1,m], d is in the range [1,D], r_1, r_2 is any number uniformly distributed in [0,1], c_1, c_2 are acceleration factor, ranging from non-negative constant, in general $c_1=c_2=2$. In practical engineering applications, $v_{id}(t)$ in Eq.(1) is usually multiplied by a factor ω which represents the inertia weight of the current velocity, generally ranging between 0.1-0.9.

Based on the characteristics of the decision table in the rough set theory, in order to facilitate to apply the PSO algorithm in the continuous attributes discretization, suppose that there are T condition attributes, h breakpoints in every attribute, the initial velocity and position vector are both $h\times T$ matrix, so the velocity and position would update according to the Iterative formula as follows:

$$v(t+1)=\omega v(t)+c_1r_1\left(P_{best}-p(t)\right)+c_2r_2\left(G_{best}-p(t)\right) \quad (3)$$

$$p(t+1)=p(t)+v(t+1) \quad (4)$$

where, v represents the velocity matrix, p represents the position matrix, p_{best} represents the optimal position matrix of one particle, G_{best} represents the optimal position matrix of all the particles. The process of attributes discretization algorithm based on PSO is described as follows:

Step 1: Make every condition attribute normalized first. Set parameters as follows: the number of particles n, the maximum number of iterations $N_{\max}$, and initialize the velocity and position matrix of all the particles in the range [0,1].

Step 2: Discrete the decision table based on the initial position matrix of each particle, and calculate the dependence coefficient of decision table as the adaptive value of particles.

Step 3: Compare the current adaptive value of each particle with its prior optimal value p_{best} and update it. At the same time Compare with the global optimal value G_{best} and update.

Step 4: End the algorithm when the dependence coefficient by G_{best} is equal to that in the step (2) or when the number of iterations is maximum, otherwise continue to update.

Step 5: Change the number of breakpoints h, compare G_{best} in different h, and choose the optima G_{best} to discrete the decision table.

Table 1. Part of decision table of sampling point I

U	a	b	c	d	e	f	D
x_1	0.8884	0.4231	0.1511	0.0918	0.0193	0.0077	1
x_{15}	0.8840	0.4071	0.1900	0.1268	0.0268	0.0047	1
x_{30}	0.8883	0.4136	0.1659	0.1082	0.0212	0.0081	1
x_{31}	0.8648	0.4461	0.2118	0.0847	0.0316	0.0104	2
x_{45}	0.8359	0.4624	0.2757	0.1015	0.0319	0.0132	2
x_{60}	0.8811	0.4274	0.1785	0.0913	0.0269	0.0099	2
x_{61}	0.7891	0.4794	0.3355	0.1670	0.0715	0.0444	3

x_{75}	0.8437	0.4396	0.2784	0.1147	0.0542	0.0356	3
x_{90}	0.8075	0.5166	0.2361	0.1373	0.0777	0.0204	3
x_{91}	0.9127	0.3820	0.1170	0.0842	0.0168	0.0104	4
x_{105}	0.8985	0.4066	0.1428	0.0793	0.0234	0.0089	4
x_{120}	0.9007	0.4007	0.1594	0.0482	0.0178	0.0079	4
x_{121}	0.8663	0.4487	0.2020	0.0791	0.0328	0.0080	5
x_{135}	0.8749	0.4173	0.2300	0.0784	0.0349	0.0126	5
x_{150}	0.8802	0.4305	0.1766	0.0896	0.0255	0.0075	5
x_{151}	0.9303	0.3320	0.1411	0.0651	0.0109	0.0030	6
x_{165}	0.9224	0.3377	0.1673	0.0819	0.0199	0.0042	6
x_{180}	0.9241	0.3247	0.1810	0.0830	0.0298	0.0106	6

Table 2. Part of decision table of sampling point II

U	a	b	c	d	e	f	D
x_1	0.8841	0.3526	0.1647	0.2202	0.1317	0.0313	1
x_{15}	0.8923	0.3176	0.1715	0.2393	0.1250	0.0242	1
x_{30}	0.8688	0.3828	0.1926	0.2167	0.1189	0.0195	1
x_{31}	0.8781	0.3105	0.3131	0.1744	0.0536	0.0363	2
x_{45}	0.8797	0.3357	0.2554	0.1888	0.1072	0.0327	2
x_{60}	0.8674	0.3392	0.3040	0.1828	0.0779	0.0246	2
x_{61}	0.9010	0.3734	0.1897	0.1016	0.0473	0.0136	3
x_{75}	0.9157	0.3743	0.1168	0.0690	0.0527	0.0161	3
x_{90}	0.9252	0.3425	0.1215	0.1006	0.0351	0.0223	3
x_{91}	0.8943	0.3884	0.1741	0.1253	0.0538	0.0201	4
x_{105}	0.8679	0.4632	0.1380	0.1000	0.0511	0.0236	4
x_{120}	0.8768	0.4178	0.1994	0.1203	0.0451	0.0191	4
x_{121}	0.9072	0.3522	0.1875	0.1223	0.0515	0.0120	5
x_{135}	0.9372	0.2973	0.1481	0.0985	0.0377	0.0105	5
x_{150}	0.9062	0.3706	0.1817	0.0873	0.0275	0.0070	5
x_{151}	0.8605	0.4315	0.1887	0.1354	0.1341	0.0387	6
x_{165}	0.8865	0.3430	0.2659	0.1235	0.0955	0.0365	6
x_{180}	0.8336	0.4604	0.2172	0.1732	0.1147	0.0529	6

Table 3. Part of decision table of sampling point III

U	a	b	c	d	e	f	D
x_1	0.7788	0.5180	0.1959	0.2311	0.1765	0.0472	1
x_{15}	0.8522	0.3935	0.1995	0.2394	0.1440	0.0322	1
x_{30}	0.8768	0.3518	0.1885	0.2397	0.1162	0.0305	1
x_{31}	0.9141	0.2997	0.2246	0.1341	0.0732	0.0300	2
x_{45}	0.8913	0.3176	0.2596	0.1709	0.0863	0.0244	2
x_{60}	0.9000	0.3225	0.2337	0.1562	0.0781	0.0283	2
x_{61}	0.8791	0.4365	0.1519	0.0970	0.0597	0.0237	3
x_{75}	0.9121	0.3657	0.1493	0.0945	0.0519	0.0191	3
x_{90}	0.9094	0.3796	0.1280	0.0921	0.0579	0.0246	3

x_{91}	0.9253	0.3369	0.1330	0.0948	0.0571	0.0217	4
x_{105}	0.9345	0.3222	0.1143	0.0875	0.0408	0.0231	4
x_{120}	0.9203	0.3508	0.1245	0.1091	0.0483	0.0170	4
x_{121}	0.9089	0.3651	0.1590	0.1133	0.0492	0.0072	5
x_{135}	0.9164	0.3582	0.1386	0.1031	0.0434	0.0112	5
x_{150}	0.9318	0.3275	0.1299	0.0792	0.0337	0.0140	5
x_{151}	0.8894	0.3706	0.2067	0.1396	0.0924	0.0283	6
x_{165}	0.8788	0.3942	0.1765	0.1573	0.1195	0.0456	6
x_{180}	0.8741	0.4258	0.1690	0.1361	0.0695	0.0522	6

Table 4. Breakpoint sets of the sampling point

Sampling Point	C^a	C^b	C^c	C^d	C^e	C^f
Ⅰ	0.84481	0.37281	0.15983	0.07798	0.01929	0.00698
	0.86518	0.41232	0.17581	0.08485	0.02318	0.00813
	0.88847	0.44153	0.19015	0.09779	0.02903	0.00985
	0.91099	0.46325	0.23325	0.11492	0.03838	0.01365
Ⅱ	0.86906	0.34397	0.17193	0.10032	0.05043	0.01376
	0.88192	0.37296	0.20044	0.12386	0.06635	0.02228
	0.89731	0.39932	0.24453	0.17387	0.10783	0.03176
Ⅲ	0.86287	0.30792	0.12385	0.09145	0.04735	0.01324
	0.87965	0.33492	0.14123	0.09935	0.05380	0.01876
	0.89958	0.35786	0.16475	0.12157	0.06931	0.02653
	0.91528	0.38127	0.19580	0.14013	0.08835	0.03621
	0.92793	0.42065	0.21781	0.18961	0.11696	0.04105

In the above algorithm, the updating principle of p_{best} and G_{best} is that, the larger the corresponding dependence coefficient, the better it is. According to the algorithm, the breakpoints of every sampling point are seen in Table 4.

3.3 Reduction Technology Based On Attributes Frequency

Attribute reduction technology is to delete the redundant attributes under the condition of keeping classification ability. At present, several attribute reduction algorithms have been put forward, the basic idea is computing attribute reduction sets with taking the core attributes as the starting point and attribute importance as heuristic information. Although the method can get attribute reduction sets, it cannot guarantee whether they are the minimal attribute reduction sets. Here, a method based on discernibility matrix and attribute frequency is proposed to calculate the minimal attribute reduction sets without computing core attribute. HU XIAOHUA put forward an improved discernibility matrix [13], defined as follows.

$$m_{ij} = \begin{cases} \{a \in C : f(x_i, a) \neq f(x_j, a)\}, & If \quad f(x_i, D) \neq f(x_j, D) \\ \phi & Others \end{cases}$$

(5)

Every discernibility matrix determines the relative discernibility function uniquely. For it is a Boolean function, the relation between the sets is “and”, between the elements in the set is “or”. The method avoids the complexity of calculating core attributes especially for the decision table with no core attributes, but when matrix is large scale, boolean operations are more complex with large amount of calculation. A reduction algorithm based on attribute frequency is presented, the core idea is as follows. The higher the attribute frequency, the more the samples classified, and if the attribute does not appear in the matrix, it will be abandoned, so the number of appearances in the matrix can be used to estimate the importance of attributes. That is to say the attribute with higher frequency has stronger separating ability than that with lower frequency. The process is as follows: first calculate the discernibility matrix of the decision table, find the attributes with the highest frequency from all the elements. Then delete the elements in the matrix that include the attributes with the highest frequency and find the attributes with the highest frequency from the rest elements, and so on until the matrix is empty. In the above algorithms, if there are attributes with the same frequency, they will be put forward respectively. Finally the combination including the least attributes is achieved as a minimal reduction set.

According to the reduction algorithm based on attributes frequency, finally the minimal reduction sets of Table 1-3 are successively $\{b,c,d,f\},\{a,b,c,d,e\},\{a,d,e,f\}$. In order to detect the sensitivity of the sampling points to faults, the minimal reduction sets of three points as condition attributes, six work states as the decision attribute, construct a big decision table together. In order to facilitate computation, thirteen characteristic parameters are denoted as $a_1,a_2,\cdots,a_{13}$ respectively. Similarly, through attributes discretization based on PSO and reduction based on attributes frequency, two minimal attribute reduction sets of the big decision table are got, $\{a_6,a_{10},a_{11}\},\{a_6,a_{10},a_{12}\}$.

According to the final reduction results, the distribution of parameters of the different measuring point can be got. The proportion of attributes appearing in the minimal reduction sets from the point I is 0/6, the ratio from the point II is 2/6, and that from the point III is 4/6. That is to say the point III is most important, the point I is useless for diagnosing faults. In order to realize the real-time online fault diagnosis and save monitoring cost, we can only monitor the point II and III.

In the test, all the failures are set on the transmission parts at the intermediate shaft. On the basis of rough set reduction results, the most sensitive point is the third point which is the nearest to the fault position, which is consistent with the gearbox vibration mechanism. When the attribute reduction technology in Rough Sets is applied to optimize the measuring points, it does not need modeling and dynamic analysis, but selects the best point directly according to the relationship between the vibration signal and fault types. The test results show the method is simple and feasible, and can be applied to fault diagnosis of other mechanical systems.

4 Simulation and discussion

Support vector machine(SVM) seek a compromise between data approximation accuracy and the complexity of the function with the objective of structural risk minimization, it can obtain better generalization ability in the small sample learning. Attribute reduction technique in rough set theory could extract sensitive fault characteristic parameters, support vector machine has strong pattern recognition ability, so in order to make full use of their advantages in the characteristic parameter extraction and pattern recognition, the intelligent fault diagnosis system is constructed based on RS and SVM. The detail process is described in Figure 5.

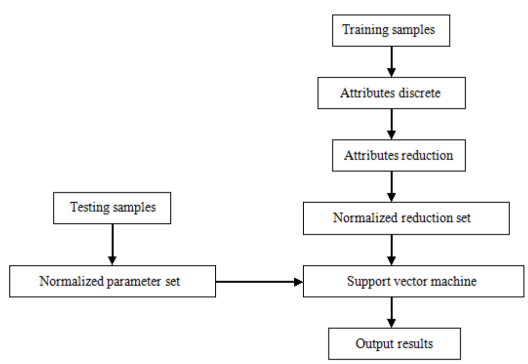


Figure 5. Fault diagnosis system based on RS and SVM

In gearbox fault simulation experiments, every state took 30 samples, totally 180 samples each measuring point, the former 120 samples for training support vector machines, the rest for testing samples, the classification accuracy of each fault monitoring points are 81.63%, 90%, 95%. The classification result is consistent with that of the sampling point optimization based on attributes reduction obviously. In order to improve the classification accuracy, it is necessary to fuse the features of different sampling point. The minimal attribute reduction sets $\{a_6, a_{10}, a_{11}\}, \{a_6, a_{10}, a_{12}\}$ include the components of sampling point II and III, the corresponding classification accuracy reach 100%. Results show that features fusion could improve the gearbox fault diagnosis performance, which provides a relatively generic solution for processing and recognizing nonlinear and non-stationary fault signal.

5 Conclusions

Monitoring the effective sampling point and collecting the fault information would improve the efficiency and accuracy of fault diagnosis. A method based on the attribute reduction technology in Rough Sets is proposed to optimize the sampling points, features fusion of different sampling point is used to improve the classification accuracy with fault diagnosis system based on RS and SVM. First the vibration signal of gearbox is processed by EMD, energy characteristic parameters based on EMD are extracted as the fault feature vector, and the decision table is established according to the gearbox fault mechanism and Rough Sets Theory. Then a attributes discretization algorithm based on PSO is put forward to discrete decision table, a method based on the discernibility matrix and attribute frequency is presented to calculate the minimal attribute reduction sets without computing core attribute. The most sensitive signal monitoring point is achieved through analyzing the final reduction sets. Finally, fault diagnosis system based on RS and SVM is constructed, the classification result is consistent with that of the sampling point optimization based on attributes reduction obviously. The experiment results show that the fault diagnosis method based on sampling point optimization and features fusion is effective and could improve the gearbox fault diagnosis performance, which provides a relatively generic solution for the similar mechanical system. In future, there is the need for further research on more efficient reduction algorithms.

Acknowledgements

This work is supported by Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi(STIP), No. 2013155.

References

1. Haiqing Wang, Song Zhihuan, Wang Hui (2002) Statistical process monitoring using improved PCA with optimized sensor locations. *Journal of Process Control*, 12(6), p.p.735-744.
2. Beck J L, Costas P, Siu-Kui A, Michael W V.(1998) Entropy-based optimal sensor location for structural damage detection. *Proceedings of SPIE conference on Smart Structures and Materials,* San Diego, USA, p.p.161-172.
3. Ali Y, Narasimahan S. (1996) Sensor netwok design for maximizing reliability of bilinear processes. *AIChE Journal*, 42(9), p.p.2563-2575.
4. Pan Hongxia, Huang Jinying, Mao Hongwei (2008) Research on Measuring Point Configuration based on Particle Swarm Optimization Technology in Fault Diagnosis. *Journal of Gun Launch and Control*, 2, p.p.58-62.
5. Z. Pawlak.(1982) Rough sets. *Computer and Information Science,*11(2), p.p.341-356.
6. Shen Lixiang, Francis E.H.T, Qu Liangsheng (2000) Fault diagnosis using rough sets theory. *Computers in Industry*, 43(2), p.p.61-72.
7. Liu Jianmin, Li Xiaolei, Zhang Xiaoming.(2011) Misfire Diagnosis of Diesel Engine based on Rough Set and Neural Networ. *Procedia Engineering*, 16, p.p.224-229.
8. Xu Huijian, Guo Feipeng (2012) Combining Rough Set and Principal Component Analysis for Preprocessing on Commercial Data Stream. *Journal of Convergency Information Technology*, 7(2), p.p.132-140.
9. B.T. Zhao, X.F. Jia, Z.B. Li (2012)The design of Precision Test Turntable Diagnosis System based on Neighborhood Rough Sets. *Journal of*

Theoretical and Applied Information Technology, 43(1), p.p.69-73.

10. Z. Wu (2012) A Method for Diagnosis of Large Aircraft Engine Fault based on Particle Swarm Rough Set Reduction. *Journal of Theoretical and Applied Information Technology*, 42(2), p.p.209-216.
11. Liu Huiling, Pan Hongxia, Liu Xiaojuan (2013) Fault diagnosis of Gearbox Using EMD and Rough Set. *Journal of Convergency Information*
12. R.Yasdi.(1995) Combining Rough Sets Learning and Neural Learning method to deal with uncertain and imprecise information, *Neurocomputing*, 7, p.p.61-84.
13. Hu XiaoHua, Cercone N.(1995) Learning in relational databases:a rough set approach. *Computational Intelligence*, 11(2), p.p.323-337.