Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 735485, 7 pages
http://dx.doi.org/10.1155/2014/735485

Research Article

Fault Diagnosis of Car Engine by Using a Novel GA-Based Extension Recognition Method

Meng-Hui Wang[1] and Pi-Chu Wu[2]

[1] *Department of Electrical Engineering, National Chin-Yi University of Technology, 411 No. 35, Lane 215, Sec.1, Chung-Shan Road, Taichung County, Taiwan*
[2] *Department of Marketing and Logistics, China University of Technology, 116 No. 56, Sec. 3, Xinglong Road, Wenshan District, Taipei City, Taiwan*

Correspondence should be addressed to Meng-Hui Wang; wangmh@ncut.edu.tw

Received 16 December 2013; Accepted 15 February 2014; Published 17 March 2014

Academic Editor: Her-Terng Yau

Due to the passenger's security, the recognized hidden faults in car engines are the most important work for a maintenance engineer, so they can regulate the engines to be safe and improve the reliability of automobile systems. In this paper, we will present a novel fault recognition method based on the genetic algorithm (GA) and the extension theory and also apply this method to the fault recognition of a practical car engine. The proposed recognition method has been tested on the Nissan Cefiro 2.0 engine and has also been compared to other traditional classification methods. Experimental results are of great effect regarding the hidden fault recognition of car engines, and the proposed method can also be applied to other industrial apparatus.

1. Introduction

Cars are an important tool in human life, and as such the traffic accidents have become a part of human life as well. There are many kinds of traffic accidents with various causes. Sometimes the drivers cause the accidents, and sometimes they are caused by engine faults. An engine fault not only damages the engine itself but can also break the car system. Usually, the component module of the engine generates a natural loss and improper maintenance that will cause a gradual increase of engine oil consumption and lead to an increase in exhaust thickness. The cylinder vibration and the temperature of engine exhaust will become abnormal. This type of the hidden defect is gradually formed, so it is difficult to recognize in the normal inspection. Therefore, it is necessary to know how to recognize the signs of engine faults early and to immediately repair or remove them. The evaluation of safety management has become a crucial issue for enterprises.

In the past, various pattern clustering techniques including expert systems (ES) [1], fuzzy clustering [2], and neural networks (NN) [3] have been extensively used in pattern recognition. Combinations of personal computers (PC), expert systems, and fuzzy systems show the possibilities of automating recognition. However, it is hard to use these rule-based methods to acquire pictorial knowledge, and it is hard to maintain the database of decision rules. The main advantage of an MNN over other classifiers is that it can directly acquire experience from the training data and overcome some of the shortcomings of the expert system. However, the training data must be sufficient and compatible to ensure proper training; its convergence of learning is influenced by the network topology and values of learning parameters. A further limitation of the MNN approach is the inability to produce linguistic output, because it is difficult to understand the content of network memory.

To overcome the limitations of the ES and MNN mentioned above, a new recognition method based on the GA and extension theory is presented for fault diagnosis of car engine in this paper. The concept of extension theory was first proposed by Cai to solve contradictions and incompatibility problems in 1983 [4]. Extension theory consists of two parts, matter-element model and extended set theory. Extension theory is now used in the research field of artificial intelligence (AI) and its relevant sciences [5, 6]. With the combination of extension theory and management science, cybernetics, information science, and computer science, extension-engineering methods have been applied to some engineering fields. The drawback of the extension method is that it needs to adjust the weight and the matter-element model by using the experienced rules for enhancing the accuracy. Therefore, this paper will propose using the GA to adjust the matter-element model of the extension method and to achieve the optimal solution to the diagnostic problem. The proposed diagnostic method has been tested on a practical car engine and has also been compared to other traditional classified methods. The results of the experiment show that the GA-based extension recognition method has high accuracy and is much more suitable as a practical solution to the diagnosis problem [7].

2. Summary of Extension Theory

Extension theory was first introduced in 1983 by a Chinese scholar, Cai W. There are two main points in extension theory that are matter-element model and extension set [8, 9]. The hard core of extension theory is two theoretical pillars that include matter-element theory and extension set theory. The former studies matter-elements and their transformations; it can be easy to represent the nature of a matter. The latter is the quantitative tool of extension theory; it can represent the correlation degree of two matter-elements in the designed correlation functions. The combination of these two pillars with other science generates the respective knowledge, which is the soft part of extension theory. Extension theory makes it possible to develop formalized descriptions for activities of creative thinking, such as knowledge innovation, new product designing, and strategy generation.

2.1. Matter-Element Theory. In extension theory, a matter-element uses an ordered triad as the basic element for describing things as follows:

$$R = (N, c, v), \tag{1}$$

where N represents the matter and c the characteristics; v is N's measure of the characteristics c, where v can be a value or an interval. If we assume that $R = (N, C, V)$ is a multidimensional matter-element, $C = [c_1, c_2, \ldots, c_n]$ is a characteristic vector, and $v = [v_1, v_2, \ldots, v_n]$ is a value vector of C, then a multidimensional matter-element is defined as follows:

$$R = (N, C, V) = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix} = \begin{bmatrix} N, & c_1, & v_1 \\ & c_2, & v_2 \\ & \vdots & \vdots \\ & c_n, & v_n \end{bmatrix}, \tag{2}$$

where $R_i = (N, c_i, v_i)$ $(i = 1, 2, \ldots, n)$ is defined as the sub-matter-element of R. For example,

$$R = \begin{bmatrix} \text{Wang}, & \text{Height}, & 178\,\text{cm} \\ & \text{Weight}, & 75\,\text{kg} \end{bmatrix}. \tag{3}$$

This can be used to state that Wang's height is 178 cm, and his weight is 75 kg. A matter has many characteristics, and one characteristic or characteristic-element can be possessed by many matters. Using the matter-element model, we can describe the quality and quantity of a matter, which is a new concept in mathematical territory.

2.2. Summary of Extension Set Theory. Set theory is a mathematical scheme that describes the classification and pattern recognition about an objective. A cantor set describes the definiteness of matters; a fuzzy set describes the fuzziness of matters. The extension set extends the fuzzy set from $[0, 1]$ to $(-\infty, \infty)$ [10]. As a result, it allows us to define a set that includes any data in the domain. The extension set is composed of two definitions.

Definition 1. Let U be a space of objects and x a generic element of U; then an extension set $\widetilde{E}$ in U is defined as a set of ordered pairs as follows:

$$\widetilde{E} = \{(x, y) \mid x \in U, y = K(x) \in (-\infty, \infty)\}, \tag{4}$$

where $y = K(x)$ is called the correlation function for the extension set $\widetilde{E}$. $K(x)$ maps each element of U to a membership grade between $-\infty$ and ∞. An extension set in U can be denoted by

$$\widetilde{E} = E^+ \cup Z_0 \cup E^-, \tag{5}$$

where

$$E^+ = \{(x, y) \mid x \in U, y = K(x) > 0\} \tag{6}$$

$$Z_0 = \{(x, y) \mid x \in U, y = K(x) = 0\} \tag{7}$$

$$E^- = \{(x, y) \mid x \in U, y = K(x) < 0\}. \tag{8}$$

In (6) and (7), E^+, E^-, and Z_0 are called the positive field, negative field, and zero boundary in $\widetilde{E}$, respectively.

Definition 2. If $X_0 = \langle a, b \rangle$ and $X = \langle f, g \rangle$ are two intervals in the real number field, and $X_0 \cap X$, where X_0 and X are the classical (concerned) and neighborhood domains, respectively, the correlation function in extension theory can be defined as follows:

$$K(x) = \begin{cases} -\rho(x, X_0) & x \in X_0 \\ \dfrac{\rho(x, X_0)}{\rho(x, X) - \rho(x, X_0)} & x \notin X_0, \end{cases} \tag{9}$$

where

$$\rho(x, X_0) = \left| x - \frac{a+b}{2} \right| - \frac{b-a}{2},$$
$$\rho(x, X) = \left| x - \frac{f+g}{2} \right| - \frac{g-f}{2}. \tag{10}$$

The correlation function can be used to calculate the membership grade between x and X_0. The extended membership function is shown in Figure 1. When $K(x) < 0$, it indicates the degree to which x belongs to X_0. When $K(x) < 0$, it describes the degree to which x does not belong to X_0. When $-1 < K(x) < 0$, it is called the extension domain, which means that element x still has a chance to become part of the set if conditions change.

2.3. The Basic Theory of Genetic Algorithm. The best-known evolutionary algorithm (EA) is the genetic algorithm (GA), which transposed the notion of evolution in nature to computers and imitates natural evolution and selection [10, 11]. Basically, a GA finds a solution to a problem by maintaining a population of possible solutions according to the "survival of the fittest" principle. The genetic algorithm constitutes a class of search algorithms especially suited to solving complex optimization problems [11]. In addition to parameter optimization, genetic algorithms are also suggested for solving problems in creative design, such as combining components in a novel creative way. In general, the major advantage of using a GA is that the optimal solution is obtained globally [12]. A genetic algorithm generally includes the following five parts.

(1) Gene Coding. It combines all genes into a chromosome sequence of 0 and 1.

(2) Fitness Function. It describes the capability of a certain individual gene to reproduce and is usually equal to the proportion of the individual's genes in all genes of the next generation.

(3) Selection Mechanism. It is the intentional manipulation by chromosome of the fitness of individuals in a population to produce a desired evolutionary response.

(4) Crossover. It is a process in which chromosomes exchange genes through the breakage and reunion of two chromosomes.

(5) Mutation. It is a change in a gene resulting in new or rearranged hereditary determinants. Mutations are rare, random events in which the base sequence of the gene is changed.

3. The GA-Based Extension Recognition Method

In this paper, the proposed recognized method involves a combination of the genetic algorithm (GA) and extension theory. Extension theory provides a means for distance measurement in the classification process. A genetic algorithm has the ability to search for an optimal solution within a wide space. The proposed GA-based extension recognition method is a kind of supervised learning that finds the best classical domain with better accuracy. This section will present a mathematical description of the proposed recognized method. We need to define several variables before using the algorithm.

3.1. The Training Stage. The chromosomes propagate the next generation of chromosomes to combine the matter-element models in the proposed method. Setting patterns = $\{p_1, p_2, \ldots, p_n\}$ with ith as follows: $p_{ij} = \{c_1, c_2, \ldots, c_k\}$. In the patterns, i is the total number of genes and j is the type of pattern. Using the proposed method can be simply described as follows.

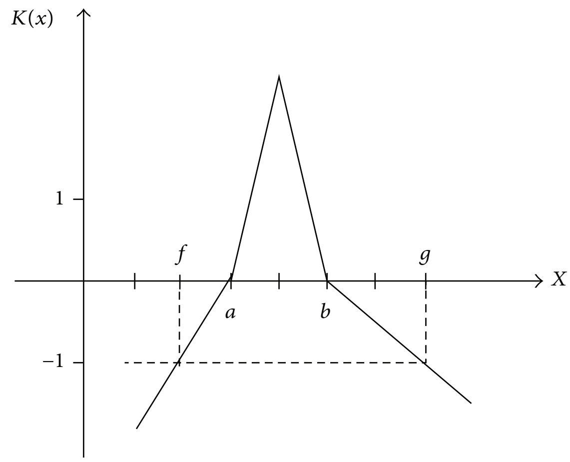

Figure 1: The extended membership function.

Step 1. Set the epoch, the crossover rate C_r, the mutation rate m_u, the tolerance of error rate E_r, and the chromosome rate R_a.

Step 2. Find the genes with lower limit and upper limit value:

$$\begin{aligned} v_a^j &= \min\left(c_{kn}^j\right) \\ v_b^j &= \min\left(c_{kn}^j\right) \\ v^j &= \left\langle v_a^j, v_b^j \right\rangle, \end{aligned} \quad (11)$$

where k is the number of characteristics, v_a is the upper limit, and v_b is the lower limit.

Step 3. Produce new genes with lower limit and upper limit values with the chromosome rate. The chromosome rate is produced by a random generator. Consider

$$\begin{aligned} v_a^j - R_a \le G_L^j \le v_a^j + R_a \\ v_b^j - R_a \le G_L^j \le v_b^j + R_a. \end{aligned} \quad (12)$$

Step 4. The genes make up the chromosome. Consider

$$\text{chrom} = \left\{G_L^{11}, G_L^{11}, G_L^{12}, G_L^{12}, \ldots, G_L^{jk}\right\}. \quad (13)$$

The number of genes in a chromosome is calculated by function $2 * k * j$.

Step 5. Build the matter-element model from the genes. Consider

$$R_j = \begin{bmatrix} N, & c_1, & \left\langle G_L^1, G_U^1 \right\rangle \\ & c_2, & \left\langle G_L^2, G_U^2 \right\rangle \\ & \vdots & \vdots \\ & c_n, & \left\langle G_L^k, G_U^k \right\rangle \end{bmatrix} \quad j = 1, 2, \ldots, m. \quad (14)$$

Step 6. Input the training of data that is the value of gene. Consider

$$x^j = \{c_1, c_2, \ldots, c_k\}. \tag{15}$$

Step 7. Calculate the correlation function. Consider

$$z^k = \left(G_L^k + G_U^k\right)$$

$$K_{nk} = \sum_{i=1}^{n} \left[\frac{\left|x_{nk}^j - z_{jk}\right| - \left(G_U^{jk} - G_L^{jk}\right)/2}{\left|\left(G_U^{jk} - G_L^{jk}\right)/2\right|} + 1 \right]. \tag{16}$$

Step 8. Normalize the value of the correlation function for the matter-element model to be between 1 and −1.

Step 9. Input the next training of data to repeat Steps 6 through 8.

Step 10. Input the next matter-element model and repeat Steps 5 through 9.

Step 11. Calculate the fitness function. Consider:

$$\text{Fitness} = \frac{N_r}{N_a}, \tag{17}$$

where N_r is the right amount and N_a is the total amount.

Step 12. The selection of parental chromosomes is put into the mating pool, and the genes implement a cross over mechanism.

Step 13. Let the next generation of chromosomes replace the chromosomes and implement the mutation mechanism.

Step 14. Calculate the correct rate. Consider

$$E_r = (1 - \text{Fitness}) \times 100\%. \tag{18}$$

Step 15. Continue until training is finished. If the training process is not finished, go to Step 3.

3.2. The Recognizing Stage

Step 1. Build the matter-element model by using an optimization solution. Consider

$$R_j = \begin{bmatrix} N, & c_1, & \left\langle G_L^1, G_U^1 \right\rangle \\ & c_2, & \left\langle G_L^2, G_U^2 \right\rangle \\ & \vdots & \vdots \\ & c_n, & \left\langle G_L^k, G_U^k \right\rangle \end{bmatrix} \quad j = 1, 2, \ldots, m. \tag{19}$$

Step 2. Input the data that is recognized. Consider

$$x^j = \{c_1, c_2, \ldots, c_k\}. \tag{20}$$

Step 3. Calculate the correlation function. Consider

$$z^k = \frac{\left(G_L^k + G_U^k\right)}{2} \tag{21}$$

$$K_{nk} = \sum_{i=1}^{n} \left[\frac{\left|x_{nk}^j - z_{jk}\right| - \left(G_U^{jk} - G_L^{jk}\right)/2}{\left|\left(G_U^{jk} - G_L^{jk}\right)/2\right|} + 1 \right]. \tag{22}$$

FIGURE 2: The engine of Nissan Cefiro 2.0.

Step 4. Find min (K_{nk}). If K_{nt} is bigger than k, then the data does not belong to any group set.

Step 5. Continue until recognizing is finished. If it is not finished, go to Step 2.

4. Fault Diagnosis of Car Engine

The tested object of this research is the engine of the Nissan Cefiro 2.0. As shown in Figure 2, the engine temperature is working between 80 and 95°C, and the base configuration of the engine has about 1.0 mm spark-plug gap. At the time of the experiment, the car was in the parked gear when the engine was in either a normal condition or the fault tests condition.

4.1. The Tested Configuration. The main specifications of the engine are shown in Table 1. The back pressure was received by a digital storage oscilloscope by using a pressure transmitter; the exhaust temperature was received by a temperature sensor. The exhaust component was received by an exhaust gas analyzer. The experimental structure is shown in Figure 3. The engine signals were all delivered to the diagnosis system by the sensors, and the detailed records of signals were easily designed by LabView 8.5 software [13]. Typical screenshots of the fault diagnostic software are shown in Figures 4 and 5. Here, the fault types are divided into 16 types (including no fault), and there are 8 characteristics for the input data. The components of engine exhaust include HC (ppm), CO (%), and CO_2 (%), and this study separately installed the temperature sensor in $T1$, $T2$, $T3$, and $T4$ so that we could promptly monitor the temperature for every exhaust position in order to speculate on engine faults. The exhaust pressures relative to the multiple frequencies of the engine's rotational speed are shown in Table 2. The processed data was sent to the next stage for fault diagnosis.

4.2. Testing Results and Discussion. In this paper, 208 sets of tested data were used according to [14] to test the practicability of the proposed method. In the training stage, there were 160 sets of training data, as shown in Table 3. The other data (48 sets) was used to test patterns. The input data of a fault diagnosis system will unavoidably contain some uncertainties

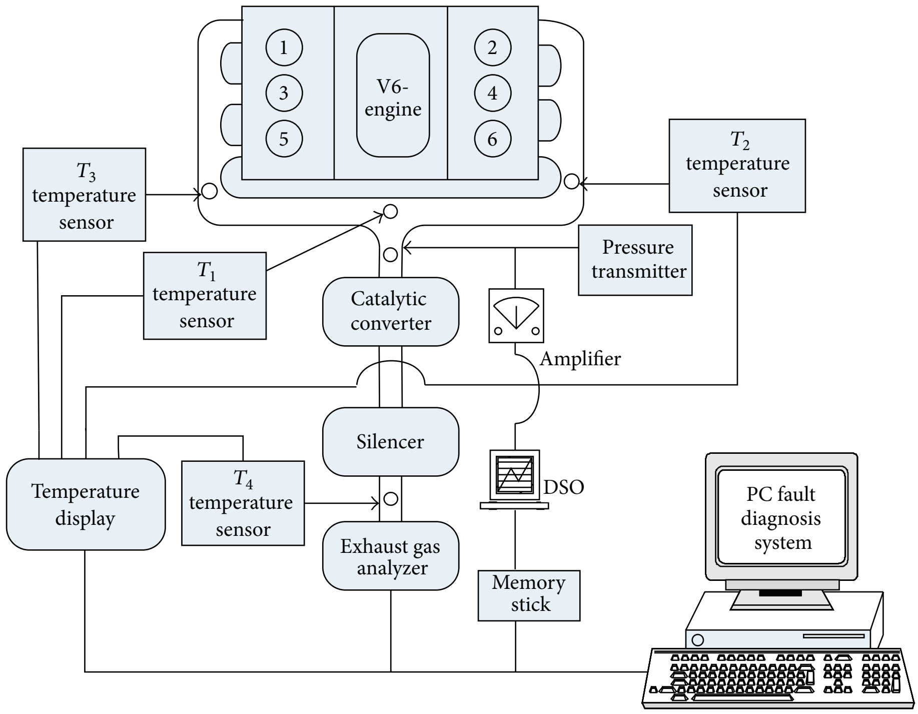

Figure 3: The experimental structure.

Table 1: The main specifications of the car engine.

Engine from	Fuel supply style	The cylinder arranging	Valve train	Displacement	Compression ratio	Maximum horse power	Maximum torque
Six-cylinder four-stroke petrol engine	Electronic ignition and engine-control computer	V6	DOHC 24V	1995 cc	9.5 : 1	150 ps/6400 rpm	19.3 kg-m/4000 rpm

Table 2: Fault types relative to the multiple frequency of the engine rotational speed.

Items	Fault types	Characteristic frequency 9.5th~12th multiple frequencies
1	No fault	Nothing
2	Spark plug is too large to 2.0 mm	Nothing
3	Spark plug is too small to 0.2 mm	12
4	No fuel injection in the 1st cylinder	9.5, 10.5, 12
5	No fuel injection in the 2nd cylinder	9.5, 12
6	No fuel injection in the 3rd cylinder	9.5, 10.5, 12
7	No fuel injection in the 4th cylinder	9.5, 12
8	No fuel injection in the 5th cylinder	9.5, 10.5, 12
9	No fuel injection in the 6th cylinder	9.5, 12
10	No ignition in the 6th cylinder	9.5, 10.5, 12
11	No ignition in the 5th cylinder	9.5, 12
12	No ignition in the 4th cylinder	9.5, 10.5, 12
13	No ignition in the 3rd cylinder	9.5, 12
14	No ignition in the 2nd cylinder	9.5, 10.5, 12
15	No ignition in the 1st cylinder	9.5, 12
16	Oxygen sensor malfunction	12

Table 3: The engine fault data and fault types (partial samples).

Number	Temperature				Exhaust			10.5 multiple frequencies	Fault number	Fault types
	T_1	T_2	T_3	T_4	CO (%)	HC (ppm)	CO_2 (%)			
11	207	288	213	82	0.11	40	14.21	2	1	Normal
66	313	258	240	152	0.03	0	12.89	2	5	The second cylinder not fuel injection
89	276	221	245	84	0.18	50	12.82	2	7	The fourth cylinder not fuel injection
102	305	279	233	95	2.11	100	12	1	8	The fifth cylinder not fuel injection
127	392	263	246	116	2.9	100	13.53	2	10	The sixth cylinder not ignition cylinder
152	372	237	254	115	2.89	180	13.47	2	12	The fourth cylinder not ignition cylinder
175	364	257	237	111	2.94	370	13.12	2	14	The second cylinder not ignition cylinder
195	260	301	233	131	0.06	50	14.49	2	16	Oxygen sensor malfunction

Table 4: Diagnosis performances of methods compared.

Method	Training time	Accuracy rate (%)
Proposed method	1000	98%
K-means method	N/A	85%
MNN-I (8-8-16)	1000	62%
MNN-II (8-10-16)	1000	80%
MNN-III (8-15-16)	1000	95%

Table 5: Diagnosis performances of proposed method.

Noise percentage (%)	Accuracy rate (%)
±0%	98%
±5%	95%
±10%	87%
±15%	77%
±20%	65%

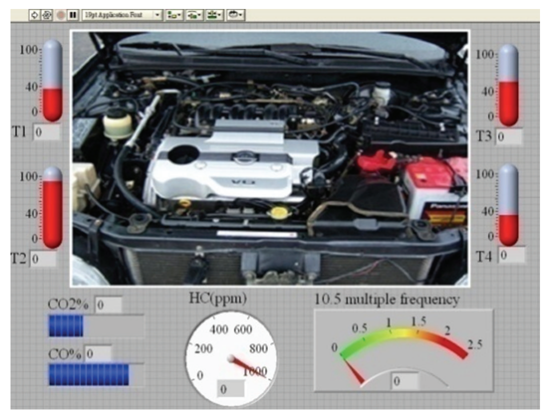

Figure 4: The LabView recording of the fault diagnosis system.

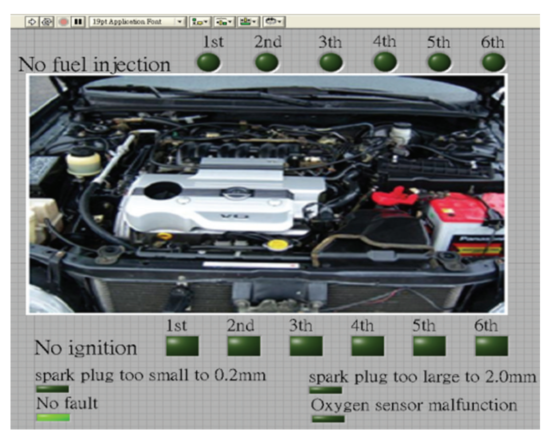

Figure 5: The fault diagnosis system.

and noise. The sources of error include environmental noise, transducers, and human mistakes, all of which can lead to data uncertainties. To take the noise and uncertainties into account, 1,800 sets of testing data were created by adding ±5% to ±15% of random, uniformly distributed error to the training data in order to appraise the fault-tolerant abilities of the proposed method. To take noise and uncertainties into account, 48 sets of testing data were created by adding ±5% to ±20% of random, uniformly distributed errors to the training data, in order to appraise the fault-tolerant abilities of the proposed method. The reason for this is that the input data of an engine system will contain some noise and uncertainties. Table 4 shows the recognition results of different methods. When using the multilayer neural network (MNN) and k-means-based methods to diagnose the faults of the engine, the maximum accuracy was 95% for the MNN-based method and 85% in the k-means-based method. The accuracy of the proposed diagnostic method is 98%, which is quite high and better than the other methods.

The test results using different numbers of added errors are shown in Table 5. Usually, the error-containing data degrades the recognition capabilities in proportion to the number of errors added. This table shows that these methods all bear remarkable tolerance to the errors contained in the data. The proposed method shows good tolerance to added errors and has high accuracy rate of 65% in the extreme case where the errors were ±20%.

5. Conclusions

This paper presents a novel fault diagnosis method for car engines based on GA and extension theory. The calculation of the proposed recognized method is fast and very simple. It can be easily implemented by PC software. When a diagnosed data is input into the proposed diagnosis system, the proposed recognized method will output the possibility of all fault types. It provides useful information to engine fault diagnosis and maintenance. Test results show that the proposed method does not only diagnose the main fault types but can also detect useful information about future trends and multifault analysis. Moreover, the proposed method has a significantly higher degree of diagnosis accuracy than current methods and shows good tolerance to added errors.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] B. Das, "Fuzzy logic-based fault-type identification in unbalanced radial power distribution system," *IEEE Transactions on Power Delivery*, vol. 21, no. 1, pp. 278–285, 2006.

[2] B. Kerezsi and I. Howard, "Vibration fault detection of large turbogenerators using neural networks," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 1, pp. 121–126, December 1995.

[3] M. B. I. Reaz, F. Choong, M. S. Sulaiman, F. Mohd-Yasin, and M. Kamada, "Expert system for power quality disturbance classifier," *IEEE Transactions on Power Delivery*, vol. 22, no. 3, pp. 1979–1988, 2007.

[4] S.-F. Hwang and R.-S. He, "A hybrid real-parameter genetic algorithm for function optimization," *Advanced Engineering Informatics*, vol. 20, no. 1, pp. 7–21, 2006.

[5] M. H. Wang and P. Y. Chen, "Using extension theory to design a low-cost and high-accurate personal recognition system," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 952568, 12 pages, 2013.

[6] F. Yang and C. Zhang, "Research of an integration model combining rough set with extension theory for fault diagnosis," in *Proceedings of the 2nd International Conference on Mechanic Automation and Control Engineering (MACE '11)*, pp. 5131–5134, July 2011.

[7] X. Yuan, "A new extension theory-based production operation method in industrial process," *Chinese Journal of Chemical Engineering*, vol. 21, no. 1, pp. 44–54, 2013.

[8] Z. Y. Lei and W. Chen, "Safety evaluation of fire protection system of high-rise building based on matter-element extension method," *Journal of Yangtze University*, vol. 6, no. 2, 2009.

[9] G. Renner and A. Ekárt, "Genetic algorithms in computer aided design," *Computer Aided Design*, vol. 35, no. 8, pp. 709–726, 2003.

[10] J. C. Tovar and W. Yu, "Fuzzy neural modeling via clustering and support vector machines," in *Proceedings of the 16th IEEE International Conference on Control Applications (CCA '07)*, pp. 24–29, October 2007.

[11] S. J. Aguilar-Ruiz, R. Giráldez, and J. C. Riquelme, "Natural encoding for evolutionary supervised learning," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 4, pp. 466–479, 2007.

[12] W. Cai, "The extension set and incompatibility problem," *Journal of Scientific Exploration*, vol. 1, no. 1, pp. 81–93, 1983.

[13] W. Cai, "Extension theory and its application," *Journal of Scientific Exploration*, vol. 44, no. 7, pp. 673–682, 1999.

[14] Y. Y. Chen, *An artificial neural network model for the diagnosis of automotive engine [thesis]*, National Chung Hsing University, 2006.