Liu, Xiaofeng and Ma, Lin and Zhang, Sheng and Mathew, Joseph (2006) Feature group optimization for machinery fault diagnosis based on fuzzy measures. In Mathew, Joseph and Ma, Lin and Tan, Andy and Anderson, Deryk, Eds. Proceedings 1st World Congress on Engineering Asset Management, pages pp. 352-359, Gold Coast, Australia.

FEATURE GROUP OPTIMIZATION FOR MACHINERY FAULT DIAGNOSIS BASED ON FUZZY MEASURES

Xiaofeng Liu, Lin Ma, Sheng Zhang and Joseph Mathew

CRC for Integrated Engineering Asset Management, *School of Engineering Systems, Queensland University of Technology, Brisbane, QLD 4001, Australia*

With the development of modern multi-sensor based data acquisition technology often used with advanced signal processing techniques, more and more features are being extracted for the purposes of fault diagnostics and prognostics of machinery integrity. Applying multiple features can enhance the condition monitoring capability and improve the fault diagnosis accuracy. However, an excessive number of features also increases the complexity of the data analysis task and often increases the time associated with the analysis process. A method of bringing some efficiency into this process is to choose the most sensitive feature subset instead. Fuzzy measures are helpful in this regard and have the ability to represent the importance and interactions among different criteria. Based on fuzzy measure theory, a feature selection approach for machinery fault diagnosis is presented in this paper. A heuristic least mean square algorithm is adopted to identify the fuzzy measures using training data set. Shapley values with respect to the fuzzy measures are applied as importance indexes to help choose the most sensitive features from a set of features. Interaction indexes with respect to the fuzzy measures are then employed to remove the redundant features. Vibration signals from a rolling element bearing test rig are used to validate the method. The results show that the proposed feature selection approach based on fuzzy measures is effective for fault diagnosis.

Key Word: Fault diagnosis, Feature selection, Fuzzy measures, Importance index; Interaction index

1 INTRODUCTION

The task of condition monitoring and fault diagnosis for modern machinery is becoming increasingly complex as the machines become more automated and complicated. Fortunately, with the development of modern multi-sensor technology and advanced signal processing techniques, the increasing complexity of the tasks of diagnostics and prognostics of these machines can now be managed more meaningfully.

The task of condition monitoring and fault diagnosis usually results in a probabilistic result with a level of uncertainty. According to information fusion theory, information uncertainty can be reduced by applying multi-source information [1]. Hence, condition monitoring and fault diagnosis capability can be enhanced by applying multiple features. However, an excessive number of features on the other hand also increases the difficulties of data analysis and usually escalates maintenance cost. In fact, using all the features for monitoring and diagnosis purposes is also unnecessary as the contribution of each feature varies in a range. Some of them may be of little use to a specific troubleshooting process. Thus, selecting the most valuable feature subset is extremely important.

Various methods have been applied for feature selection purposes. Sub-optimal methods are more widely used than optimal methods. Although sub-optimal methods can generally produce optimal or near-optimal results in most real application cases, they cannot guarantee optimal results. The most frequently used sub-optimal feature selection algorithms are sequential forward search, sequential backward search, plus-L-minus-R and floating search, etc [2].

The exhaustive search is a universal optimal method. It produces an optimal solution by testing all the possible feature combinations. However, as the number of feature combinations increases exponentially with the number of features, this method becomes difficult to implement for high dimension feature space [3] [4] [5]. The only method other than exhaustive search still capable of yielding an optimal result is the branch and bound (BB) based algorithm. However, its application is also conditional. This method is only limited to monotonic criteria. Furthermore, its converging speed is highly dependent on the data sets [3]. To enhance its performance, BB algorithms have been improved by some researchers [6, 7].

Classical set theory and probability theory have been widely applied for dealing with uncertainty problems. Fuzzy set theory and fuzzy measure theory are two more general mathematical methods. Fuzzy methods are more effective than traditional clustering methods in handling fault features which are imprecise, with the boundaries among different failure modes usually being ambiguous in their mapping space. Fuzzy methods describe fault patterns in a non-dichotomous way which is similar to the manner in which human beings process vague information. As an outgrowth of classical measure theory, Fuzzy Measure (FM) and Fuzzy Integral (FI) theory has been applied to pattern recognition [8] [9] [10], image processing [11] [12] [13] and information fusion [14] and have the advantage that they are able to represent importance of criteria and certain interactions among them.

This paper presents a optimal feature selection approach for machinery fault diagnosis based on the fuzzy measure theory. It is suitable for low feature space problems. A heuristic least mean square algorithm (HLMS) is adopted to identify the fuzzy measures from training data set. Shapley values with respect to the fuzzy measures are computed and applied as importance indexes to choose the most sensitive features from a number of features. Interaction indexes with respect to the fuzzy measures are then employed to remove the redundant features. Vibration signals from rolling element bearings are used to validate this method. The results show that the proposed feature selection approach based on fuzzy measures is effective for fault diagnosis and is agreeable with feature subset obtained through other method.

The rest part of this paper is organised as follows. In section 2, the definition of fuzzy measures and related concepts, such as Shapley value and interaction index, are briefly introduced. In Section 3 a heuristic least mean square algorithm for identifying 2-additive fuzzy measures is introduced. A novel approach for fault diagnosis feature optimisation based on fuzzy measures is proposed in section 4. Section 5 presents the results of the experimental evaluation. Discussions based on the results are presented in Section 6, and Section 7 draws the conclusion.

2 FUZZY MEASURE, CHOQUET FUZZY INTEGRAL, IMPORTANCE INDEX AND INTERACTION INDEX

A fuzzy measure on the set X of criteria is a set function

$$\mu : P(X) \rightarrow [0,1]$$

satisfying the following axioms:

1) $\mu(\Phi) = 0, \mu(X) = 1$
2) $A \subset B \subset X$ implies $\mu(A) \leq \mu(B)$

where $X = \{x_1, \ldots, x_n\}$ is the set of criteria, $P(X)$ is the power set of X, i.e. the set of all subsets of X. Here $\mu(A)$ represents the weight of importance of the set of criteria A. Φ denotes the empty set.

The 2-additive fuzzy measure is a special kind of fuzzy measure. It is defined by

$$\mu(K) = \sum_{i=1}^{n} a_i x_i + \sum_{\{i,j\}\subseteq X} a_{ij} x_i x_j \ , \tag{1}$$

where $K \subseteq X$. For any $K \subseteq X$ and $|K| \geq 2$ with x_i=1 if $i \in K$, otherwise x_i=0.

As $\mu_i = a_i$ for all i, the general form for the 2-additive fuzzy measure can be formulated as

$$\mu(K) = \sum_{\{i,j\}\subseteq K} \mu_{ij} - (|K|-2)\sum_{i\in K} \mu_i \ , \tag{2}$$

for any $K \subseteq X$ and $|K| \geq 2$. It can be seen from the expression that the 2-additive fuzzy measure is determined by the coefficients μ_i and μ_{ij}.

The 2-additive fuzzy measures need $n(n+1)/2$ coefficients to be defined. The general number of coefficients required to be defined for k-order additive fuzzy measures is $\sum_{j=1}^{k} \binom{n}{j}$.

Fuzzy measures can describe any of the three interactions between two criteria A and B:

a) Synergetic interaction, which can be represented by

$$\mu(A \cup B) > \mu(A) + \mu(B), \tag{3}$$

b) Inhibitory interaction, which can be represented by

$$\mu(A \cup B) < \mu(A) + \mu(B), \tag{4}$$

c) Non-interaction, which can be represented by

$$\mu(A \cup B) = \mu(A) + \mu(B). \tag{5}$$

Classical probability theory can only be applied to the third situation when there is no interaction between two criteria [15].

The Choquet fuzzy integral of a function $f : X \to [0,1]$ with respect to μ is defined by

$$C_\mu(f(x_1),\ldots,f(x_n)) := \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\mu(A_i), \tag{6}$$

where $X = \{x_1,\ldots,x_n\}$ is the set of criteria; $P(X)$ is the power set of X, i.e. the set of all subsets of X; $\mu(A)$ is the fuzzy measure representing the importance of the set of criteria A, and $A_i = \{x_i, x_{i+1},\ldots, x_n\}$

The importance index (Shapley value) [16] v_i of element i with respect to a fuzzy measure μ is defined by

$$v_i = \sum_{k=0}^{n-1} \gamma_k \sum_{K \subset X \setminus i, |K|=k} [\mu(K \cup \{i\}) - \mu(K)], \tag{7}$$

where $\gamma_k = (n-k-1)!k!/n!$, X is a set of n elements, $K \subset X$.

Importance index (or Shapley value) describes the global importance of each element. It has the property

$$\sum_{i=1}^{n} v_i = 1. \tag{8}$$

It can be seen form the formula that the sum of importance indices of all features is "*1*". The index of the feature *i* is greater than the index of the feature *j* indicates that feature *i* is more important than feature *j* in identifying a given fault.

The importance index demonstrates the contributions of individual features to the recognition of a specific fault class. Based on the importance index, we can determine feature i is more important than feature j if importance index v_i of feature i is greater than importance index v_j of feature j.

The interaction index [17] I_{ji} between two elements i and j with respect to a fuzzy measure μ is defined by

$$I_{ij} = \sum_{k=0}^{n-2} \zeta_k \sum_{K \subset X \setminus \{i,j\}, |K|=k} [\mu(K \cup \{i,j\}) - \mu(K \cup \{i\}) - \mu(K \cup \{j\}) + \mu(K)], \tag{9}$$

where $\xi_k = (n-k-2)!k!/(n-1)!$.

The interaction index describes the degree of redundancy of two features to the recognition of a specific class. Feature i and feature j are complementary for recognition of a class from others if the interaction index I_{ji} between two elements i and j is positive. Both the two features should be employed as their combination will enhance the recognition capability. Feature i and feature j are redundant for the recognition of a class from others if the interaction index I_{ji} between two elements i and j is negative. Choosing of either feature i or feature j will have the same effect on final recognition result. Feature i and feature j are independent for the recognition of a class from others if the interaction index I_{ji} between two elements i and j is zero. In this case, both features have their contributions.

Based on the concept of importance index and interaction index, a feature selection approach for machinery fault diagnosis is proposed. The importance index is used to choose the most important feature set from a number of features. The interaction index is then applied to remove the redundant feature form the set. This approach is expected to optimise the feature group selection for machinery fault diagnosis in different aspects.

3 IDENTIFYING FUZZY MEASURES USING TRAINING DATA SET

As discussed before, the bases of this feature selection method depends on calculation of the importance index and the interaction index with respect to the fuzzy measure μ . This research uses the 2-additive fuzzy measure as it is relatively simple and easy to implement. However, identification of the fuzzy measures is not easy. Here we present a data-driven algorithm for identifying the fuzzy measures.

Identifying fuzzy measures using training data set is an inverse problem of defining fuzzy integrals. Normally the training data set are obtained according to a field expert's knowledge. Here the overall assessment was regarded as Choquet fuzzy integral. The local assessments *f* are based on individual features, see Table 1. In this case, according to the definition of the Choquet integral, identifying the fuzzy integrals means to solve the equation

$$\sum_{i=1}^{n}(f(x_i)-f(x_{i-1}))\mu(A_i)=a_l \quad (l=1,\ldots,\mathrm{m}). \tag{10}$$

Table 1

Training data for identifying fuzzy measures

Data point	Local assessments				Overall assessment
1	$f_1(x_1)$	$f_1(x_2)$	...	$f_1(x_n)$	$C_{\mu 1}$
2	$f_2(x_1)$	$f_2(x_2)$	...	$f_2(x_n)$	$C_{\mu 2}$
⋮	⋮	⋮	⋮	⋮	⋮
m	$f_m(x_1)$	$f_m(x_2)$	...	$f_m(x_n)$	$C_{\mu m}$

A heuristic least mean square algorithm (HLMS) [18] is employed to provide an approximate optimal solution of the equation by minimizing the error between the actual Choquet fuzzy integral and the expected Choquet fuzzy integral output

$$E=[C_\mu(f)-y]^2\,. \tag{11}$$

Conditions such as monotonicity and normalization should also be satisfied when implementing the algorithm. The monotonicity condition can be formulated as

$$\sum_{j\in K}\mu_{ij}-\sum_{j\in K}\mu_j-(n-2)\mu_i\geq 0\,, \tag{12}$$

for $\forall i\in X$, $K\subseteq X\setminus i$, where $|X|=n$.

The normalization condition can be formulated as

$$\sum_{\{i,j\}\subseteq X} \mu_{ij} - (n-2)\sum_{i \in X} \mu_i = 1 . \tag{13}$$

4 FEATURE GROUP OPTIMISATION BASED ON FUZZY MEASURES

The proposed optimisation method for feature group optimization can be implemented in four steps (Figure 1). First we need to obtain the fault knowledge, which is represented by a training data set. Based on the training data set, a specific kind of fuzzy measure is extracted. The interaction index with respect to the fuzzy measures, which is used to remove redundant features from a group of feature set, is then computed. The last step of the approach is to compute the importance index to choose the most important features from a number of features. Sometimes we can change the order of the last two steps.

To obtain the fault knowledge, fuzzy *c*-means (FCM) clustering analysis is employed. It is used by inputting both individual and combinations of fault features to acquire both the local and overall assessment of an object. The overall assessment is used as Choquet fuzzy integrals and the local assessment is used as function f to describe the contributions of individual features to the object.

The core part of the heuristic least mean square algorithm (HLMS) for the 2-additive fuzzy measure identification adopts the gradient descent theory

$$\mu_{new} = \mu_{old} - \alpha \frac{E}{E_{\max}} (x_{n-i} - x_{n-i-1}) , \tag{14}$$

where i denotes the ith element of feature vector in ascending order; $\alpha \in [0,1]$ is a constant, which is known as learning rate; E is the error between actual and expected output.

Once the fuzzy measures are determined, the interaction index of every pair of features and importance index of each feature can be computed. Based on the importance index and interaction index, the less important and redundant features will be eliminated.

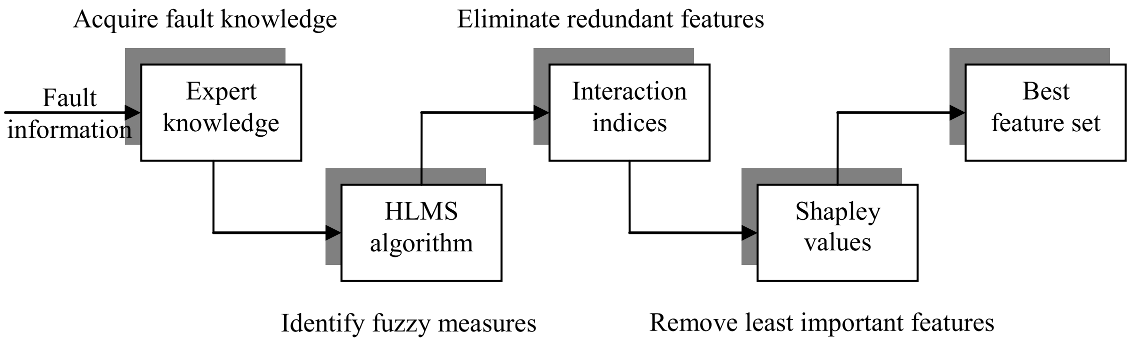

Figure 1 Feature group optimization based on fuzzy measures

5 EXPERIMENT RESULTS

An experiment on rolling element bearing faults was conducted using a machinery fault simulator made by SpectraQuest Inc to evaluate the proposed fault diagnostic method. The data acquisition system included an amplifier PCB 482A20, a filter KROHN-HITE 3202, a NI A/D card Daq6062E with a connector BNC2120, two vibration sensors IMI 608A11 and a Toshiba laptop computer. The software used was LabView. The cutoff frequency was 5k Hz and the sampling rate 20k Hz. Simulated defects included inner race fault, outer race fault, ball fault, looseness of the bearing housing and different fault combinations under different loads and rotating speeds.

Four time-domain features were extracted from the vibration signals, including Root Mean Square (RMS), maximum value, kurtosis and crest factor. The feature values were normalized into range [0, 1] to meet the requirement of the employed fuzzy method input.

For machinery fault diagnosis, although quite a few features are available, their degrees of contribution to the diagnostic process are usually different. The purpose of this research is to keep the minimum number of features while still maintaining the diagnostic accuracy. In this specific scenario, we managed to choose two most suitable features from the group of four to validate our proposed method. Here we used x_1, x_2, x_3, and x_4 to denote RMS, maximum value, kurtosis and crest factor, respectively.

The fault knowledge is normally obtained empirically or according to expert field knowledge. This research used fuzzy *c*-means clustering algorithm as an expert system to acquire the fault knowledge from the original signal. The fault knowledge included local and overall assessment of a specific fault based on different features.

Once the fault knowledge extracted, the 2-additive fuzzy measure was identified. Table 2 shows the identified fuzzy measures for bearing outer race fault. Table 3 shows the Shapley values of each feature for bearing outer race fault.

Table 2

The 2-additive fuzzy measures on the power set of features

Subset	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_4\}$		
Fuzzy measure	0.2975	0.2623	0.3350	0.3268		
Subset	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_1,x_4\}$	$\{x_2,x_3\}$	$\{x_2,x_4\}$	$\{x_3,x_4\}$
Fuzzy measure	0.5680	0.6375	0.5000	0.5784	0.5310	0.5987
Subset	$\{x_1,x_2,x_3\}$	$\{x_1,x_2,x_4\}$	$\{x_1,x_3,x_4\}$	$\{x_2,x_3,x_4\}$		
Fuzzy measure	0.8890	0.7124	0.7769	0.7840		

Table 3

Interaction index

Feature set	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_1,x_4\}$	$\{x_2,x_3\}$	$\{x_2,x_4\}$	$\{x_3,x_4\}$
Interaction index	0.0720	0.0706	-0.0425	0.0187	-0.0837	-0.0532

Table 4

Importance index of different feature

Feature	x_1 (RMS)	x_2 (Maximum)	x_3(Kurtosis)	x_4(Crest factor)
Importance index	0.2493	0.2353	0.3039	0.2115

It can be seen from Table 3 that all the interaction index involving feature x_4 are negative. This indicates that Crest factor is a redundant feature. It can also be seen from Table 4 that the Crest factor plays the least important role in diagnosing bearing fault in this case as the importance index of the Crest factor is the smallest. As such, this feature should be removed from the feature group.

It can be seen from Table 4 that the importance index of the Maximum value is the second smallest. Therefore, Maximum value is the second least important feature in identifying the given fault. As described before, to test our feature selection approach, a feature subset of only two elements was pursued in this case. Although the Maximum value is complementary with other features, it was eliminated from the set.

As a result, the two features of Crest factor and Maximum value were eliminated from the original feature set {*RMS, maximum value, kurtosis, crest factor*} based on important index and interaction index with respect to the 2-additive fuzzy measures. The best feature subset for diagnosing the bearing fault, {*RMS, kurtosis*} was eventually obtained.

6 DISCUSSIONS

Fuzzy measure and fuzzy integral techniques are able to represent importance of criteria and certain interactions among them. Based on this concept, a feature group optimization approach for machinery fault diagnosis is proposed. For rolling element bearing fault diagnosis, the feature subset {*RMS, kurtosis*} was selected from a feature set of four {*RMS, maximum value, kurtosis, crest factor*}. To validate the result, fuzzy *c*-means clustering was used for comparison. Table 5 illustrates the average values of the membership degree of all features of the bearing fault. By comparing Table 4 with Table 5, it can be seen that the two results are in full agreement with each other.

Table 5

Average value of membership degree of features

Feature	RMS	Maximum	Kurtosis	Crest factor
Average value	0.9523	0.9118	0.9524	0.8354

Even from the 2-additive fuzzy measures, some initial signs can be observed. The fuzzy measure for $\{x_1, x_3\}$ is greater than other fuzzy measures of the two-element subset. The fuzzy measure of a three element set which contains x_1 and x_3 is also greater than other fuzzy measures of the three-element subset. The initial assumption was further proven by using the importance index and the interaction index.

The method can improve the efficiency of information utilization, as fuzzy measure and fuzzy integral data fusion techniques are capable of choosing the most sensitive features. As a result, machinery fault diagnostics accuracy can be improved, and machinery reliability, availability and safety can be enhanced. Machinery operating risk and maintenance cost can be reduced. The method can therefore add value to the engineering asset management practice.

7 CONCLUSION

This paper presents a fuzzy measure based feature optimization approach for machinery fault diagnosis. This method can produce an optimal feature set in terms of the importance index and the interaction index. In the case study, a two feature subset {*RMS, kurtosis*} was selected from a feature set of four {*RMS, maximum value, kurtosis, crest factor*} for rolling element bearing fault diagnosis. Experimental results show that the proposed feature selection approach has great potential in efficient and effective fault diagnosis.

8 REFERENCES

1. Xu, Z., et al., Diagnosis Principle and Application for the Fault Information. 2000: China Defence Industry Press.

2. Sun, J., Pattern Recognition. 2002, Changsha: Defence Technology University.

3. Sommol, P. and P. Pudil, Feature selection toolbox. Pattern Recognition, 2002. 35: p. 2749-2759.

4. Pudil, P., J. Novovicova, and P. Somol, Feature selection toolbox software package. Pattern Recognition Letters, 2002. 23: p. 487-492.

5. Schulerud, H. and F. Albregtsen, Many are called, but few are chosen. Feature selection and error estimation in high dimensional spaces. Computer Methods and Programs in Biomedicine, 2004. 73: p. 91-99.

6. Chen, X.-W., An improved branch and bound algorithm for feature selection. Pattern Recognition Letters, 2003. 24: p. 1925-1933.

7. Yu, B. and B. Yuan, A more efficient branch and bound algorithm for feature selection. Pattern Recognition, 1993. 26: p. 883-889.

8. Gader, P.D., M.A. Mohamed, and J.M. Keller, Fusion of handwritten word classifiers. Pattern Recognition Letters, 1996. 17: p. 577-584.

9. Michel, The representation of importance and interaction of features by fuzzy measures. Pattern Recognition Letters, 1996. 17: p. 567-575.

10. Kuncheva, L.I., J.C. Bezdek, and R.P.W. Duin, Decision templates for multiple classifier fusion: an experimental comparison. Pattern Recognition, 2001. 34: p. 299-314.

11. Mirhosseini, A.R. and H. Yan, Human face image recognition: an evidence aggregation approach. Computer Vision and Image Understanding, 1998. 71: p. 213-230.

12. Pham, T.D. and H.Yan, Color image segmentation using fuzzy integral and mountain clustering. Fuzzy Sets and Systems, 1999. 107: p. 121-130.

13. Kwak, K.-C. and W. Pedrycz, Face recognition using fuzzy integral and wavelet decomposition method. IEEE Transactions on Systems, Man, and Cybernetics--Part B: Cybernetics, 2004. 34: p. 1666-1675.

14. Auephanwiriyakul, S., J.M. Keller, and P.D. Gader, Generalized Choquet fuzzy integral fusion. Information Fusion, 2002. 3: p. 69-85.

15. Klir, G.J., Z. Wang, and D. Harmanec, Constructing fuzzy measures in expert systems. Fuzzy Sets and Systems, 1997. 92: p. 251-264.

16. Shapley, L.S., A value for n-person game. Contributions to the theory of games, ed. H.W. Kuhn and A.W. Tucker. Vol. 2. 1953, Princeton, New Jersey: Princeton University Press. 307-317.

17. Murofusion, T. and S. Soneda, Techniques for reading fuzzy measures (3): interaction index. Proceedings of the 9th fuzzy systems symposium, May 1993. Sapporo, Japan.

18. Grabisch, M. A new algorithm for identifying fuzzy measures and its application to pattern recognition. Proceedings of the International Joint Conference of the 4th IEEE International Conference on Fuzzy Systems and the 2nd International Fuzzy Engineering Symposium, 1995. Yokohama, Japan.