Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Fault Diagnosis of Machinery based on Support Vector Data Description and D-S Evidence Theory

Bin Chen, Longhua Ma, and Zhaoli Yan

Key Laboratory of Noise and Vibration Research, Institute of Acoustics, Chinese Academy of Sciences. Beijing, 100190, China

PACS: 43.60.-c

ABSTRACT

Condition monitoring and fault diagnosis is essential to the effectiveness and reliability of machinery. To improve the accuracy of fault diagnosis, a novel diagnostic model based on support vector data description (SVDD) and Dempster-Shafer (D-S) evidence theory is proposed. In the method, time and frequency domain fault features are firstly extracted, and used as input vector of single SVDD fault classifier, which is trained according to normal and few faulty data. Then take identifying result of single SVDD classifier at different measuring point around machinery as independent evidence source, and so the evidence set is constructed. Based on unified discernment frame of fault diagnosis, all evidences are aggregated by Dempster's combination rule. Through multi-level inforation fusion, it can make full use of measuring information and resolve the problem of single classifier's misrecognition. Experiment results show that proposed algorithm improves identification precision of fault diagnosis and deal with the contradition between classifiers effectively.

1 INSTRUCTION

Rotating machinery has been widely equipped in some heavy industries. To guarantee its effectiveness and reliability, condition monitoring and fault diagnosis is very necessary. Generally, it usually deploys several sensors at different point around diagnostic machinery to obtain sufficient status information. For large and complex machinery, it may be difficult to establish a mathematical model accurately to identify the machinery's status. However, artifical intelligence techniques can provide an important solution, and have been applied successfully to plant diagnostics, control, identification, and so on.

Fault dianosis can be seen as a problem of pattern recognition. Various intelligent methods, such as artifical neural network (ANN) [1, 2] and support vector machine (SVM) [3-5], have been applied to fault diagnosis of machinery. It requires characterize normal data in a way that distinguishes it from faulty data. However, acquisition of whole abnormal data set of diagnositic machinery would require intentional destruction in several ways. Nevertheless, faulty situations are possible, but expensive, to generate. Above existing classification approaches are not well suited for handling this type of problem, because they seek to develop decision functions using predefined classes including normal and faulty data. Thus, it needs a method that utilizes only normal data. Support vector data description (SVDD) developed by Tax and Duin is very useful for such a one-class classification problem with small or even no faulty data [6]. Its basic idea is to define a boundary around samples with a minimum volume by introducing kernel functions. The nonparametic nature of SVDD would be quite useful when sample distribution is abnormal or when no prior knowledge about the distribution is available. In addition, applying of SVDD to fault diagnosis of machinery permits a flexible decision boundary that uses various nonlinear kernel functions.

For the reason of strong noise or sensor's flaws, single fault SVDD classifier may have the problem of low accuracy and uncertainty of recognition. In addition, multi-sensor data acquisition is adopted for the need of comprehensiveness machinery fault diagnosis. Thus, it need to aggregate multi-channel information to make reasonable decision. Dempster-Shafer (D-S) evidence theory developed by Dempster and Shafer [7], which has more rigorous reasoning process than the probability theory, provides an important way for expression and combination of uncertainty information, and has obtained widespread application in uncertainty reasoning, decision analysis and information fusion [8-10]. In addition, it can construct basic probability assignment function of evidence objectively according to output result of single SVDD classifier.

Based on above considerations, a novel fault diagnosis algorithm based on SVDD and D-S evidence theory is proposed. The organization of paper is as follows. In section 2, we introduce fault diagnosis method in detail. Then numerical examples are given to show the efficiency of the proposed approach in section 3. Conclusions are made in section 4.

2. FAULT DIAGNOSIS ALGORITHM BASED ON SVDD AND EVIDENCE THEORY

2.1 Fault Diagnosis Model

Deploy n sensors at different measuring points around diagnostic machinery. In other ways, we have to aggregate n-channels measuring information to obtain correct fault diagnosis. The structure of fault diagnosis model based on SVDD and D-S evidence theory is showed in Figure 1. As can be seen, the algorithm is made up of four parts: faulty feature extraction, fault classification, evidence combination, and decision making.

The main reasoning process is as follows. Firstly, time or frequency domain fault features are extracted from vibration signals using traditional signal processing method. Collected normal and faulty data from each measuring point are used to train SVDD fault classifer respectively. Then obtained fault features are used as input vector of trained SVDD model to identify. Finally, take classifier as independent evidence source, and the basic probability assignment function of evidence can be constructed according to classifier' output results. All evidences are aggregated based on Dempster's combination rule to make full use of multi-channel identifying information and resolve the contradiction between classifiers. Diagnosis result can be obtained according to *max* BPAF decision rule.

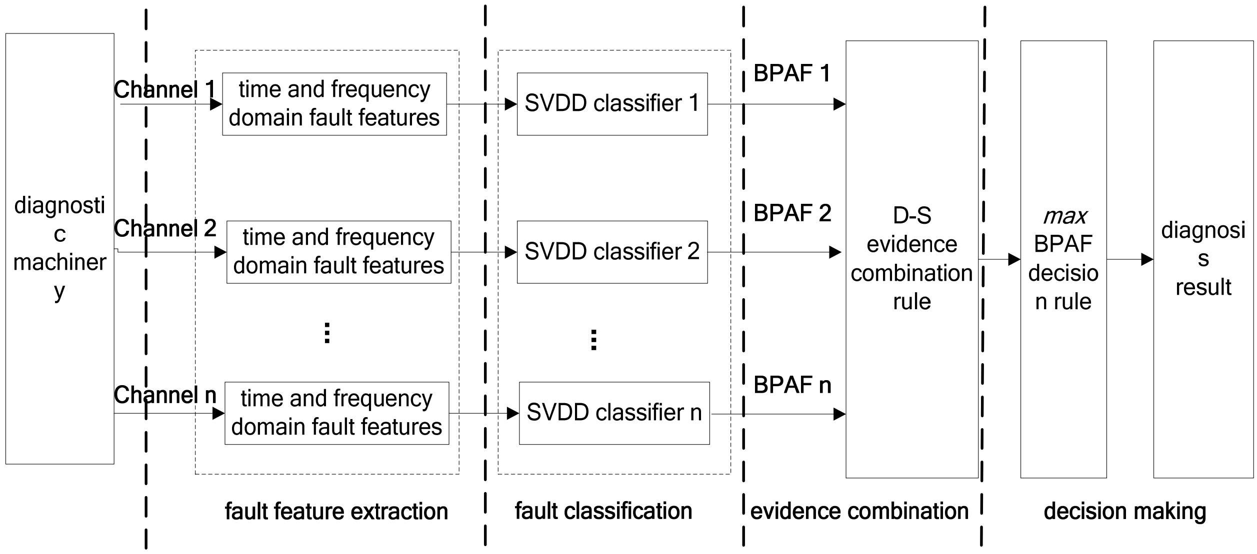

Figure 1. Fault diagnosis algorithm based on SVDD and evidence theory

2.2 Fault Feature Extraction

We can't make sure which feature is sensitive to mechanical fault beforehand. Thus, refering to existing research results, a complete time and frequency domain fault features are extracted from vibration signals.

A) Time-Domain Features

The time-domain vibration signal contains rich status and fault information. Currently, dimensional and dimensionless characteristic parameters are commonly applied. The dimensionless parameters, such as mean (x_{mean}), maximum (x_{max}), average amplitude ($x_{a\text{-}amp}$), root amplitude ($x_{r\text{-}amp}$), mean square root amplitude ($x_{msr\text{-}amp}$), which can be extracted by signal analysis and processing, increase as fault develops. Given vibration signal x (t), the definition of parameters is as follows:

$$\begin{cases} x_{mean} = \dfrac{1}{T}\int_0^T x(t)d_t \\ x_{max} = \max(|x(t)|) \\ x_{a\text{-}amp} = \dfrac{1}{T}\int_0^T |x(t)|d_t \\ x_{r\text{-}amp} = \left[\dfrac{1}{T}\int_0^T |x(t)|^{1/2} d_t\right]^2 \\ x_{msr\text{-}amp} = \sqrt{\dfrac{1}{T}\int_0^T x^2(t)d_t} \end{cases} \quad (1)$$

The dimensionless parameters, such as waveform index ($S_{waveform}$), impulsion index ($S_{impulsion}$), peak index (S_{peak}), tolerance index ($S_{tolerance}$), and kurtosis index ($S_{kurtosis}$), are insensitive to machinery's working conditions including speed or load, but very sensitive to the fault when working conditions change. Their computational methods are as follows:

$$\begin{cases} S_{waveform} = x_{msr\text{-}amp} / x_{a\text{-}amp} \\ S_{impulsion} = x_{max} / x_{a\text{-}amp} \\ S_{peak} = x_{max} / x_{msr\text{-}amp} \\ S_{tolerance} = x_{max} / x_{r\text{-}amp} \\ S_{kurtosis} = \dfrac{1}{T}\int_0^T \left(x(t) - x_{mean}\right)^4 d_t \Big/ \left[\dfrac{1}{T}\int_0^T \left(x(t) - x_{mean}\right)^2 d_t\right]^2 \end{cases} \quad (2)$$

B) Frequency-Domain Features

With the emergence and development of fast Fourier transform (FFT), frequcncy spectrum analysis has been widely used in fault diagnosis of machinery. The amplitude and shape of vibration signal's frequency spectrum are different when machinery works under various operating conditions. Thus, choose frequency domain features as one of vectors in fault feature space.

Given rotating machinery's speed n revolutions per minute, its base frequency (f_{base}) is n/60. Usually, the amplitude spectrum at f_{base}, 2 f_{base} and 3 f_{base} contribute different to machinery's normal and faulty status. In addition, energy spectrum at corresponding feature frequency band is also sensitive to fault classification of machinery.

C) Time-Frequency Domain Wavelet Energy Features

Usually, vibration signal contains some rich nonstationary fault information. Traditional FFT analysis method, which transform only in time or frequency domain completely, can't express signal's time-frequency domain partial feature. Wavelet packet decomposition method, which has high frequency resolution at low-frequency and high time resolution at high-frequency, can obtain such partial feature. For example, to certain signal x(t), its 3-levels decomposed process is as shown in Figure 2. At first level, it decomposes x(t) into $x_1(t)$ and $x_2(t)$ with equal frequency band respectively, and so on.

According to decomposed signals $x_{111}(t)$, $x_{112}(t)$, $x_{121}(t)\cdots$, its corresponding wavelet energy rations coefficient E_{r1}, E_{r2}, $E_{r3}\cdots$ can be extracted, and used as one of vectors in fault feature space.

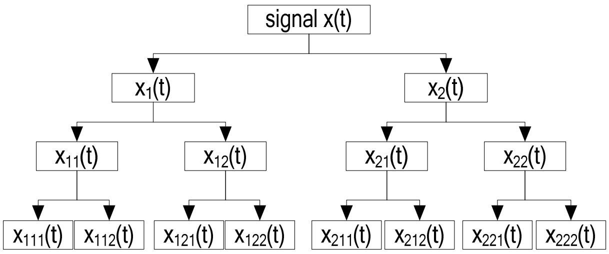

Figure 2. Wavelet decomposition

The above time, frequency, and time-frequency domain features can construct a fault feature space, which will used as input of classifiers to diagnose.

2.3 SVDD Classifier Design

Given a training data set $\{\mathbf{x}_i, i=1,2,\cdots,l\}\subset\Re^n$, the enclosing sphere S is characterized by its center a and radius R, the goal is to minimize the sphere, and keep all training objects inside its boundary, its sketch map in two dimensions is shown in Figure 3.

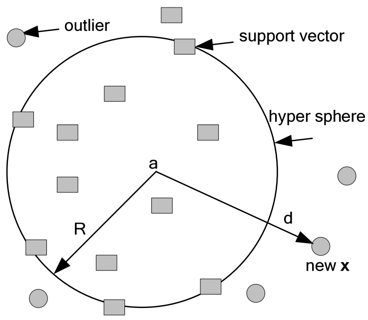

Figure 3. Sketch map of SVDD

The structural error to minimize is:

$$\begin{cases}\min \quad F(R,a)=R^2 \\ s.t \quad \|\mathbf{x}_i-a\|^2\le R^2,\ \forall i\end{cases} \tag{3}$$

To allow for the possibility of some samples falling outside of the sphere, the slack variable ξ_i is introduced to relax the constraints. Then, the sphere can be described as follows:

$$\begin{cases}\min \quad F(R,a)=R^2+C\sum_i\xi_i \\ s.t \quad \|\mathbf{x}_i-a\|^2\le R^2+\xi_i,\quad \xi_i\ge 0\,(i=1,2,\cdots,l)\end{cases} \tag{4}$$

Where, the variable C controls the trade-off between the number of errors and the volume of sphere. The following Lagrangian is constructed:

$$L=R^2+C\sum_i\xi_i-\sum_i\alpha_i(R^2+\xi_i-\|\mathbf{x}_i-a\|^2)-\sum_i\beta_i\xi_i \tag{5}$$

Where， α_i,β_i are the Lagrange multipliers. Setting the partial derivatives of L about R,a,ξ_i to 0, we obtain:

$$\begin{cases}\dfrac{\partial L}{\partial R}=2R(1-\sum_i\alpha_i)=0\Rightarrow\sum_i\alpha_i=1 \\ \dfrac{\partial L}{\partial a}=2\sum_i\alpha_i(\mathbf{x}_i-a)=0\Rightarrow\sum_i\alpha_i\mathbf{x}_i=a \\ \dfrac{\partial L}{\partial\xi_i}=C-\alpha_i-\beta_i=0\Rightarrow C=\alpha_i+\beta_i\end{cases} \tag{6}$$

So, the dual formulation is

$$\begin{cases}\max \quad L(R,a,\alpha_i,\beta_i)=\sum_i\alpha_i(\mathbf{x}_i\cdot\mathbf{x}_i)-\sum_{i,j}\alpha_i\alpha_j(\mathbf{x}_i\cdot\mathbf{x}_i) \\ s.t \quad 0\le\alpha_i\le C,\ \sum_i\alpha_i=1, i=1,2,\cdots,l\end{cases} \tag{7}$$

Objects $\mathbf{x}_i$ with $\alpha_i>0$ are called ***support vectors*** (SVs) of the description. SVs lie on the boundary (if $0<\alpha_i<C$) or outside the boundary (if $\alpha_i=C$) of the sphere. We can see that the center of the sphere is the linear combination of the SVs. As shown in Eq. (7), the problem of SVDD is stated in terms of inner products. For more flexible boundaries, inner products of samples $(\mathbf{x}_i,\mathbf{x}_j)$ can be replaced by a kernel function $K(\mathbf{x}_i,\mathbf{x}_j)$, where $K(\mathbf{x}_i,\mathbf{x}_j)$ satisfies Mercer's theorem [11]. This implicitly maps samples into a nonlinear feature space to obtain a tighter and non-linear boundary. The use of a kernel function allows the computation of dot products in a nonlinear feature space without the use of nonlinear mappings. Commonly used kernel functions include the Gaussian $K(\mathbf{x}_i,\mathbf{x}_j)=\exp(-\|\mathbf{x}_i-\mathbf{x}_j\|^2/2\sigma^2)$ as well as the polynomial $K(\mathbf{x}_i,\mathbf{x}_j)=\left(\mathbf{x}_i\cdot\mathbf{x}_j\right)^d$ functions. Different kernel functions result in different boundaries, and the polynomial function, in general, doesn't produce good results in SVDD [6].

If the distance of new sample $\mathbf{z}$ to the center of sphere is smaller or equal than the radius R, it will be accepted as a target object. Then the state of $\mathbf{z}$ can be judged by:

$$\begin{aligned}f_{SVDD}&=I(\|\mathbf{z}-a\|^2\le R^2) \\ &=I((\mathbf{z}\cdot\mathbf{z})-2\sum_i a_i(\mathbf{z}\cdot x_i)+\sum_{i,j}a_ia_j(x_i\cdot x_i)\le R^2)\end{aligned} \tag{8}$$

Where, $I(A)=\begin{cases}1 & if\ A \text{ is true} \\ 0 & if\ A \text{ is false}\end{cases}$.

2.4 Evidence Combination Model

Every SVDD classifier gives its identifying results. Taking each SVDD classifier as an independent evidence source, the following evidence combination model is established to make valid decision and decrease the recognition uncertainty of single classifier.

A) BPAF Construction of SVDD Fault Classifier

According to evidence theory, the frame of discernment is constituted by all possible mutually exclusive and exhaustive propositions of the identifying objects. Thus, discernment frame of fault diagnosis can be noted by $\Theta=\{A_1,A_2,A_3\cdots\}$, where set elements A_1, A_2 and A_3 denotes normal status, rotor unbalance, rotor friction, etc. The corresponding spheres are $S_i(i=1,2,3,\cdots)$.

Based on identifying result of SVDD, the constructive method of evidence's mass function is as follows. As shown in Figure 4, suppose the relative distance of new sample $\mathbf{z}$ to the center of sphere S_i is d_i, which can be calculated by SVDD algorithm. The basic probability assignment value of set element A_i is $\frac{d_i / R_i}{\sum_i d_i / R_i}$. Thus, the BPAFs of each SVDD are obtained.

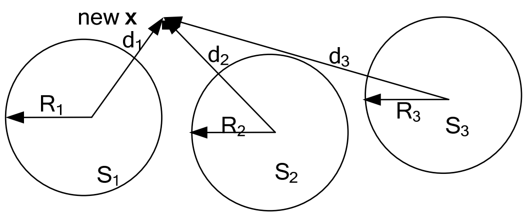

Figure 4. BPAF of evidence

B) Evidence Combination

Because multiple SVDD classifiers provide different assessment of the same frame of discernment according to their own mass functions, combination is a good solution to obtain more relevant information. Demspter gives a useful evidence combination rule, which is defined as follows [7]:

Let m_1 and m_2 are the BPAFs of same discernible frame, the focal elements are $\{A_1, A_2, ..., A_i\}$ and $\{B_1, B_2, ..., B_j\}$ respectively, the combination rule is as follows:

$$m(C) = m_1 \oplus m_2 = \begin{cases} 0 & C = \varnothing \\ \frac{1}{1-K} \sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j) & C \neq \varnothing \end{cases} \quad (9)$$

Where, $\oplus$ called the orthogonal sum, and the conflict degree $k = \sum_{A_i \cap B_j = \varnothing} m_1(A_i) m_2(B_j)$.

C. Decision Rule

Final diagnosis result is obtained according to maximum value of the basic probability assignment functions. $\exists A_1, A_2 \in \Theta$, the BPAFs of A_1 and A_2 are as follows:

$$\begin{cases} m(A_1) = \max\{m(A_j), A_j \subset \Theta\} \\ m(A_2) = \max\{m(A_j), A_j \neq A_1\} \end{cases} \quad (10)$$

Let $\varepsilon_1, \varepsilon_2$ be the threshold value, if it satisfies

$$\begin{cases} m(A_1) - m(A_2) > \varepsilon_1 \\ m(\Theta) < \varepsilon_2 \\ m(A_1) > m(\Theta) \end{cases} \quad (11)$$

Then, element A_1 is the object.

3 EXPERIMENT ANALYSIS

Three class data of rolling machine, including normal, rotor imbalance, and rotor friction, are simulated on ZT-3 rotor vibration simulation test-bed provided by Nanjing DongDa vibration instrument factory. We installed four acceleration sensors at its different location P1, P2, P3 and P4. The data sampling rate is 25.5 KHz. Each class gathers 100 samples, and the data length is 20,000 points. The cross-validation rate between training and testing samples is 0.3. The revolution of rotor is 2000 rpm.

All possible time and frequency-domain statistical features are extracted. From spectrum curve of vibration signal, as shown in Figure 5, we can see that its base frequency f_{base} is 33Hz, which accords with theoretical value (2000/60=33.33 Hz).

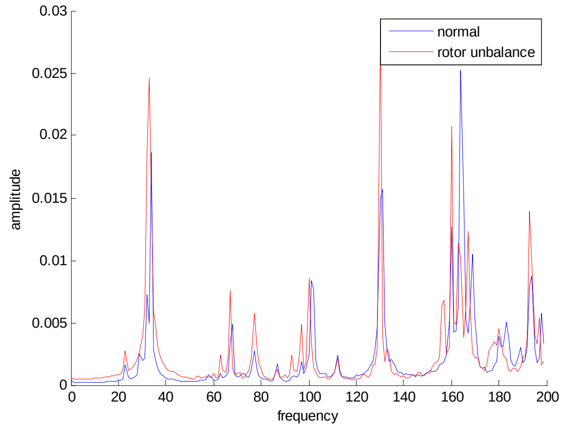

Figure 5. Frequency spectrum

Thus, we choose max amplitude and total energy between frequency band (32-34) Hz, (65-70) Hz, and (128-135) Hz as frequency domain features. In time-frequency domain, we decompose vibration signal into 3 layers using 'db5' wavelet, and obtain 8 wavelet energy coefficients E_{r1}, E_{r2}, E_{r3}, E_{r4}, E_{r5}, E_{r6}, E_{r7}, and Er8 at third level, which are also used as one vector of feature space.

Based on same Gaussian kernel function (set the value of σ to 1), the correct recognition rate of SVDD classifier taking different fault feature as input is shown in Table 1. We can see that the recognition precision of SVDD classifier using all 24 fault features is 91.67%, which is better than single feature. It can satisfy engineering application.

Table 1. Recognition rate of SVDD classifier

Number	Type	Recognition rate
1	mean	55.00%
2	maximum	66.67%
3	average amplitude	60.00%
4	root amplitude	58.33%
5	mean square root amplitude	56.67%
6	waveform index	58.33%
7	impulsion index	55.00%
8	peak index	68.33%
9	tolerance index	66.67%
10	kurtosis index	75.00%
11	maximum amplitude $A_{32\text{-}34}$	63.33%
12	maximum amplitude $A_{65\text{-}70}$	65.00%
13	maximum amplitude $A_{128\text{-}135}$	75.00%
14	totoal energy $E_{32\text{-}34}$	78.33%
15	totoal energy $E_{65\text{-}70}$	73.33%
16	totoal energy $E_{128\text{-}135}$	83.33%
17	wavelet energy coefficients E_{r1}	73.33%
18	wavelet energy coefficients E_{r2}	81.67%
19	wavelet energy coefficients E_{r3}	66.67%
20	wavelet energy coefficients E_{r4}	58.33%
21	wavelet energy coefficients E_{r5}	75.00%
22	wavelet energy coefficients E_{r6}	66.67%
23	wavelet energy coefficients E_{r7}	68.33%
24	wavelet energy coefficients E_{r8}	66.67%
25	all features	91.67%

Suppose the discernment frame is $\Theta = \{A_1, A_2, A_3\}$, and the evidence set contains 4 independent evidences E1, E2, E3, and E4, which are constructed from detected vibration signal at P1, P2, P3 and P4 measuring point. The BPAFs of m1, m2, m3 and m4 are as showed in Table 2, where the element Θ denotes uncertainty. All evidences were fused by Dempster's combination rule, and the result is:

$$\begin{cases} m(A_1) = 0.0016, & m(A_2) = 0.9934 \\ m(A_3) = 0.0050, & m(\Omega) = 0.0000 \end{cases}$$

According to Eq. (11), the decision object is A_2, which indicates rotor unbalanced fault.

Table 2. BPAFs of Evidences

BPAF	A_1	A_2	A_3	Θ
m1(E1)	0.1477	0.7338	0.1003	0.0182
m2(E2)	0.0384	0.4369	0.5069	0.0178
m3(E3)	0.2165	0.6576	0.1182	0.0077
m4(E4)	0.1277	0.7596	0.1020	0.0138

From table 2, we can see that the identifying result of E2 is conflict with E1, E3 and E4 seriously. After evidence aggregation, it can make a reasonable decision and avoid misrecognition of single SVDD classfier 2.

4 CONCLUSION

A novel fault diagnosis algorithm of machinery based on support vector data description and D-S evidence theory is presented. The established single SVDD classifier has high recognition precision of machinery fault. From the combination of multi-SVDD classifiers based on D-S evidence theory, it can make full use of monitoring information and solve the misrecognition problem of single SVDD classifier. Experiment results show that proposed method based on SVDD and evidence theory is feasible and effective.

REFERENCES

1 V. Crupi, E. Guglielmino and G. Milazzo, “Neural network based system for novel fault detection in rotating machinery,” *Journal of Vibration and Control*. 10(8), 1137-1150 (2004).

2 B. Samanta, K.R. Al-Balushi and S.A. Al-Araimi, “Artificial neural networks and genetic algorithm for bearing fault detection,” *Soft Computing*. 10(3) 264-271 (2006).

3 S. Abbasion, A. Rafsanjani, A. Farshidianfar and N. Irani, “Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine,” *Mechanical Systems and SignalProcessing*. 21(7), 2933-2945 (2007).

4 H.J. Shin, D.H. Eom and S.S. Kim, “One class supportvector machines-an application in machine fault detection and classification,” *Computers& Industrial Engineering*. 48(2), 395-408 (2005).

5 A. Widodo and B.S. Yang, “Support vector machine in machine condition monitoring and fault diagnosis,” *Mechanical Systems and Signal Processing*. 21(6), 2560-2574 (2007).

6 D.M.J. TAX and R.T.P.W. DUIN, “Support vector data description,” *Machine Learning*. 54, 45-66 (2004)

7 G. Shafer. *A mathematical theory of evidence* (Princeton University Press, 1976).

8 A. Laha, N.R. Pal, and J. Das, “Land cover classification using fuzzy rules and aggregation of contextual information through evidence theory,” *IEEE Transaction on Geoscience and Remote Sensing*. 44(6), 1633-1641 (2006).

9 H.W. Guo, W.K. Shi, Y. Deng et al, “Evidential conflict and its 3D strategy: discard, discover and disassemble?,” *Systems Engineering and Electronics*, 29(6), 890-898 (2007).

10 B. Chen, R.J. Feng, and J.W. Wan, “A novel fusion method of conflicting evidences for clustering wireless sensor networks,” *Proceedings of Seventh International Symposium on Instrumentation and Control Technology*. 7127, OZ1-OZ7. (2008)

11 B. Scholkopf, A. Smola and K.R. Muller, “Nonlinear component analysis as a kernel eigenvalue problem,” *Neural Computation*. 10(5), 1299-1319 (1998).