www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

DATA FUSION ALGORITHM OF FAULT DIAGNOSIS CONSIDERING SENSOR MEASUREMENT UNCERTAINTY

Xu Xiaobin[1], Zhou Zhe[1,2], Wen Chenglin[1]

[1]Insititute of Systems Science and Control Engineering, School of Automation
Hangzhou Dianzi University
Hangzhou, China

[2]State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control
Zhejiang University
Hangzhou, China

Emails: xuxiaobin1980@163.com; zhouzhe@zju.edu.cn; wencl@hdu.edu.cn

Submitted: Aug. 30, 2012 ***Accepted: Jan. 11, 2013*** ***Published: Feb. 20, 2013***

Abstract- This paper presents data fusion algorithm of fault diagnosis considering sensor measurement uncertainty. Random-fuzzy variables (RFV) are used to model testing patterns (TPs) and fault template patterns (FTPs) respectively according to on-line sensor monitoring data and typical historical sensor data reflecting every fault mode. A similarity measure is given to calculate matching degree between a TP and each FTP in fault database such that Basic Probability Assignment (BPA) can be obtained by normalizing matching degree. Several BPAs provided by many sensor sources are fused by Dempster's rule of combination. A diagnosis decision-making can be done according to the fusion results. Finally,

the diagnosis examples of machine rotor system with vibration sensors show that the proposed method can enhance accuracy and reliability of data fusion-based diagnosis system.

Index terms***:*** **sensor data fusion, fault diagnosis, random-fuzzy variable, similarity measure, Dempster-Shafer evidence theory.**

I. INTRODUCTION

In modern industry, fault diagnosis plays an important role in accident prevention, human safety, maintenance decision-making, and cost minimization. For critical equipment, on-line condition monitoring and on-line fault diagnosis are often needed. There is no single sensor that can reliably obtain all the information required for fault diagnosis. With development of sensor technology, a great deal of monitoring data can be obtained by sensor instruments mounted on equipment. Generally, these sensor data are uncertain because of some unavoidable factors including the random disturbances in measurement environment and system errors of sensor instrument, etc [1]. New challenges have arisen with regard to making more reasonable inferences based on multi-source information with uncertainty.

Dempster-Shafer (DS) evidence theory can combine multi-source information to reduce the uncertainty and yield more accurate diagnosis results than any single-source information [2]. In the framework of DS evidence theory, fault modes are modeled as elements in frame of discernment. The Basic Probability Assignment (BPA) function can be considered as the matching degrees between an on-line testing pattern and each fault template pattern in fault database. These BPAs can be fused by Dempster's rule of combination. A diagnosis decision-making can be done according to the fusion results.

Testing pattern (TP) and fault template pattern (FTP) can be extracted respectively from on-line monitoring data and typical historical data reflecting every fault mode. Hence, how to mode TP and FTP based on uncertain data, and obtain BPA according to relation between TP and FTP are two keys in fault diagnosis based on DS evidence theory.

International Electro-technical Commission (IEC) established a Guide to the expression of uncertainty in measurement. It suggested that statistical approaches can be used for the expression and estimation uncertainty of sensor data [3]. However, statistical approaches are only suitable for data with randomness. Reference [4] also presented that TP and TFP can be all modeled as Gaussian membership functions. BPA can be calculated by fuzzy max-min operator.

Nevertheless, in practice, sensor data have not only randomness caused by disturbances in environment but also non-randomness, i.e., unknown systematic uncertainty, derived from systematic error of sensor itself. Single statistical method or fuzzy method cannot comprehensively deal with two different uncertainties [5].

Alessandro Ferrero presented that fuzzy variable and random variable can be used to respectively describe unknown systematic uncertainty and random uncertainty, and both kinds of variables can be integrated as “random-fuzzy variable (RFV)” to model sensor data completely and naturally [5].

This paper presents data fusion algorithm of fault diagnosis considering random uncertainty and unknown systematic uncertainty of sensor data. Random-fuzzy variables (RFV) are used to model testing pattern (TP) and fault template pattern (FTP) respectively according to on-line monitoring data and typical historical data. A similarity measure fit for RFVs is given to calculate matching degree between a TP and each FTP in fault database such that BPA can be obtained by normalizing matching degree. Several BPAs provided by many sensor sources are fused by Dempster’s rule of combination. A diagnosis decision-making can be done according to the fusion results. The fault diagnosis structure is shown in Figure 1. Finally, the diagnosis examples of machine rotor system with a vibration displacement sensor (VDS) and a vibration acceleration sensor (VAS) show that the proposed method can enhance accuracy and reliability of fusion-based diagnosis system.

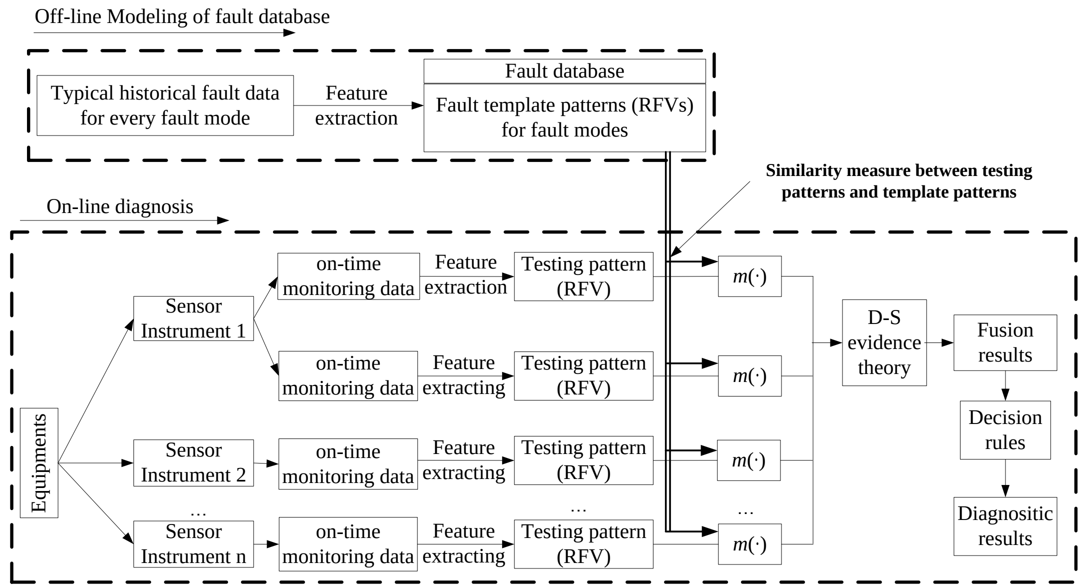

Figure 1. Fault diagnostic structure

II. THE DEMPSTER-SHAFER EVIDENCE THEORY

Let Θ be a finite nonempty set of mutually exclusive alternatives. It is called the frame of discernment including all possible propositions. In fault diagnosis, $\Theta=\{Fj|\ j=1, 2, \dots, N\}$ is a set of fault propositions, where Fj is the proposition that fault Fj happens. Let 2Θ be the power set of Θ and contain all subsets of Θ. The Basic Probability Assignment (BPA) is a mapping from 2Θ to [0, 1] defined by m: $2\Theta\rightarrow[0,1]$, which satisfies the following conditions: (1) $m(\varnothing)=0$; (2) For any $A \in 2^{\Theta}$, $\sum_A m(A)=1$. If $m(A)>0$, A is called focal element[4,6]. In this paper, only single faults are considered, namely, the form of BPAs are $\{m(F1), \dots, m(FN), m(\Theta)\}$, where $m(Fj)$ means the possibility that fault Fj occurs and $m(\Theta)$ means the degree of ignorance.

Two BPAs coming from independent sources can be combined by Dempster's rule of combination, $m=m_1 \oplus m_2$, defined as follows [4, 6]:

$$m(C)=\begin{cases}\dfrac{\sum_{C_1\cap C_2=C} m_1(C_1)m_2(C_2)}{1-k}, & C\neq\varnothing \\ 0 & C=\varnothing\end{cases} \qquad (1)$$

where $k=\sum_{C_1\cap C_2\neq\varnothing} m_1(C_1)m_2(C_2)$ is called the degree of conflict.

III. THE METHOD OF FAULT FEATURE EXTRACTION BASED ON RANDOM-FUZZY VARIABLE

When the uncertainty of sensor data is mainly affected by randomness, statistical approach is able to express and estimate the uncertainty, since statistical theory is the most known and used mathematical tool to deal with statistical distributions of data. But it fails when sensor data are simultaneously affected by non-randomness, i.e., unknown systematic uncertainty. On the other hand, Fuzzy variables are able to express and deal with sensor data affected by unknown systematic uncertainty [5]. However, it can hardly handle randomness.

Although fuzzy variable and random variable are defined in a different way and they obey different mathematics, with respect to Theory of Uncertainty, fuzzy variable and random variable should not be considered competitive but rather complementary. In practice, uncertainties arising in the sensor measurement processes consist of randomness and unknown systematic uncertainty, so both kinds of variables are all needed [7].

The random-fuzzy variable can effectively express the contributions of different effects (random and unknown systematic) to uncertainties of sensor data. In this section, the definition of RFV is briefly introduced, and then it will be used to model testing pattern and fault template pattern.

a. The definition of Random-Fuzzy Variables

RFV is a particular kind of type-Ⅱ fuzzy variables, whose α-cut B_α is confidence intervals of type Ⅱ [7, 8],

$$B_\alpha = [[b_1^\alpha, b_2^\alpha], [b_3^\alpha, b_4^\alpha]] \ , \ \forall \alpha \in [0,1] \qquad (2)$$

and obeys the following constraints [7, 8]:

1) $b_1^\alpha \le b_2^\alpha \le b_3^\alpha \le b_4^\alpha$ for $\forall \alpha \in [0,1]$;
2) For all $\alpha \in [0,1]$, the corresponding sequences of type 1 external confidence interval $[b_1^\alpha, b_4^\alpha]$ and inner confidence interval $[b_2^\alpha, b_3^\alpha]$ can generate respective normal and convex membership function;
3) For $\forall \alpha, \alpha' \in [0,1]$

$$\alpha' > \alpha \Rightarrow \begin{cases} [b_1^{\alpha'}, b_3^{\alpha'}] \subset [b_1^\alpha, b_3^\alpha] \\ [b_2^{\alpha'}, b_4^{\alpha'}] \subset [b_2^\alpha, b_4^\alpha] \end{cases}$$

4) $[b_2^{\alpha=1}, b_3^{\alpha=1}] \equiv [b_1^{\alpha=1}, b_4^{\alpha=1}]$.

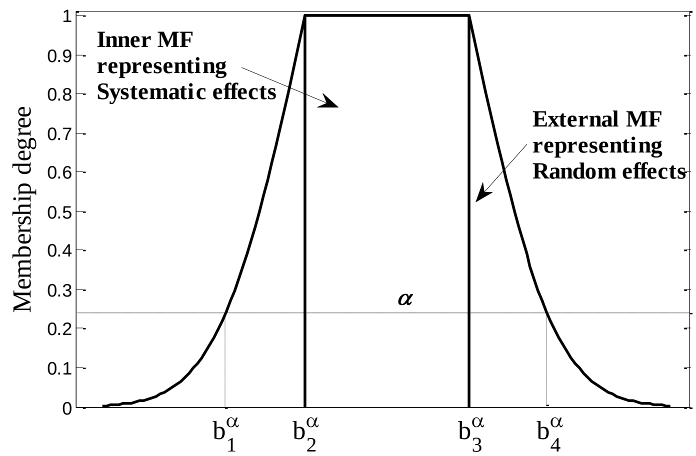

Figure 2. Example of RFV

Figure 2 gives an example of RFV. For a sensor datum, the confidence level that it is within α-cut B_α is p=1-α, for instance, α=0.25, p=1-0.25=0.75. So B_α can be called a confidence interval.

In fact, RFV is obtained by combining the inner and external membership function according to their α-cuts for all possible $\alpha \in [0,1]$.

When RFV is adopted to express uncertainties, the widths of the closed intervals $[b_1^\alpha, b_2^\alpha]$ and $[b_3^\alpha, b_4^\alpha]$ in (1) reflect randomness contribution to whole uncertainties. On the other hand, the closed internal interval $[b_2^\alpha, b_3^\alpha]$ in (1) is a type 1 confidence interval and its width reflects the contribution of unknown systematic error to whole uncertainties.

It can be concluded that a RFV can perfectly deal with uncertainties of sensor data because it can not only model randomness and unknown systematic error, but also distinguish their different contributions to whole uncertainties by using a unified form.

b. Model Fault Template Pattern and Testing Pattern as RFVs

Here, Testing pattern (TP) and fault template pattern (FTP) are all modeled as RFVs respectively according to on-line monitoring data and typical historical data reflecting every fault mode. All RFVs of FTPs compose fault database.

b.i RFV model of fault template pattern

The detailed steps of determining RFV are as follows

1) Model the external membership function μ_1 for randomness contribution

Suppose x is a variable denoting one of some fault features, which may result from different faults or fault modes. When a certain fault was simulated or really occurred in past monitoring process, x was observed by the corresponding sensor instrument, n data of x was recorded, generally $n \geq 200$. Thus, statistical histogram can be constructed using these typical historical data. Gaussian probability density function (pdf) of x is obtained through interpolation fitting.

There is a case in point, for fault diagnosis of a machine rotor system (the detail will be given in section 5), let x be the amplitude of foundational frequency (f_v) of vibration acceleration (unit m/s^2), which is observed by vibration displacement sensor. Suppose 200 data under fault of "*rotor unbalance*" are recorded. The fitted pdf $p(x)$ is shown in Figure 3.

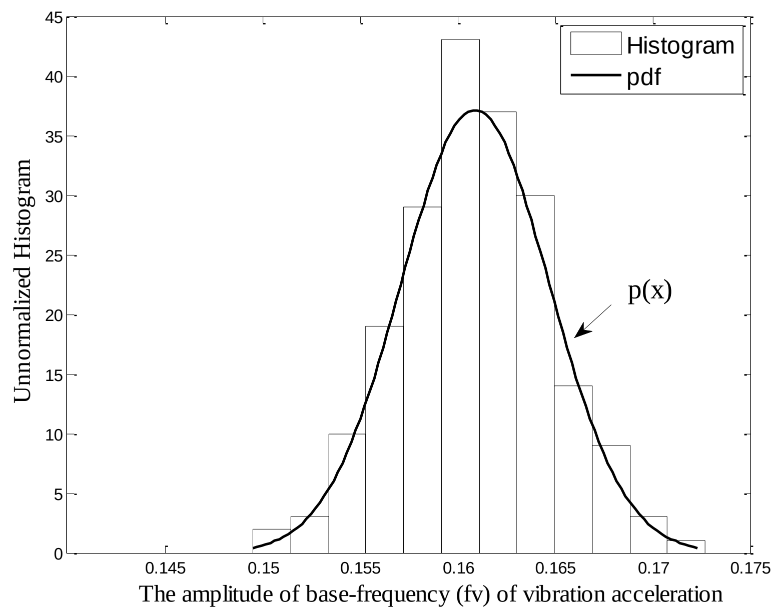

Figure 3. The fitted Gaussian pdf

The pdf $p(x)$ needs to be transformed into a possibility density function, i.e., an external membership function. The specifics are shown in the following ways [7, 9].

Firstly, the value x_p corresponding to the peak value of pdf is determined, and its membership degree in the external membership function is set 1. Let x_L and x_R be the bounds of x's distribution. Since $p(x)$ is normal distribution, so $[x_L, x_R]=[x_p-3\sigma, x_p+3\sigma]$, where x_p is mean value, σ is standard deviation.

Secondly, for interval [x_L, x_p], let $x_{L_i} \in [x_L, x_p]$, i=1,2,…,M and $x_{L_i} = x_L + i \times (x_p - x_L)/(M+1)$. For interval [$x_p$,$x_R$], let $x_{R_i} \in [x_p, x_R]$, i=1,2,…,M and $x_{R_i} = x_p + (M-i+1) \times (x_R - x_p)/(M+1)$. The distribution of these points in [x_L, x_R] is showed in Figure 4. The following M+2 nested intervals can be generated $[x_p, x_p] \subseteq [x_{L_M}, x_{R_M}] \subseteq [x_{L_{M-1}}, x_{R_{M-1}}] \subseteq ... \subseteq [x_{L_1}, x_{R_1}] \subseteq [x_L, x_R]$

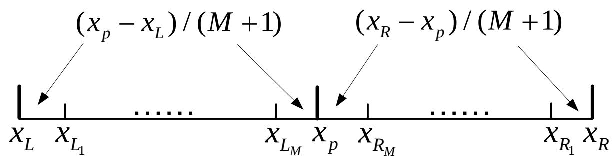

Figure 4. The distribution of points on interval [x_l, x_r]

Obviously, these intervals are the different α-cuts of the external membership function according to the corresponding values of α. The greater is M, the higher is the resolution of the desired fuzzy variable. The value of M is chosen according to the required resolution. It is known that the confidence level about the generated interval $[x_{L_k}, x_{R_k}]$ is

$$\lambda_k = \int_{x_{L_k}}^{x_{R_k}} p(x)dx \qquad (3)$$

and the corresponding α_k=1-λ_k.

In this example, M=100, x_p=0.1607 m/s^2, x_l =0.1496 m/s^2, x_r =0.1718 m/s^2 σ=0.0037 m/s^2, using M+2 generated nested intervals, $p(x)$ can be transformed into the external membership function μ_1 as shown in Figure 5.

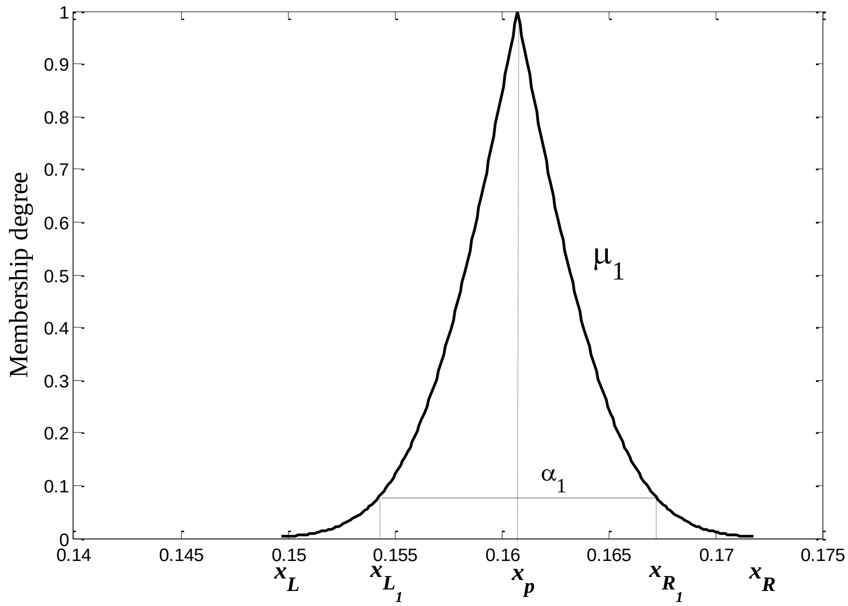

Figure 5. The external membership function transformed from pdf $p(x)$

2) Model the inner membership function μ_2 for non-randomness contribution

The unknown uncertainty (non-randomness) is caused by systematic error of sensor instruments, which can be described by inner membership function. Generally, systematic error is provided by manufacturer, e.g., for the vibration displacement sensor mentioned above, the manufacturer represents systematic error specification as $x_p \pm y\%$, where y=0.2% describes sensor accuracy. In this case, a rectangular membership function $\mu_{2,}$ can be used to model the systematic error as shown in Figure 6.

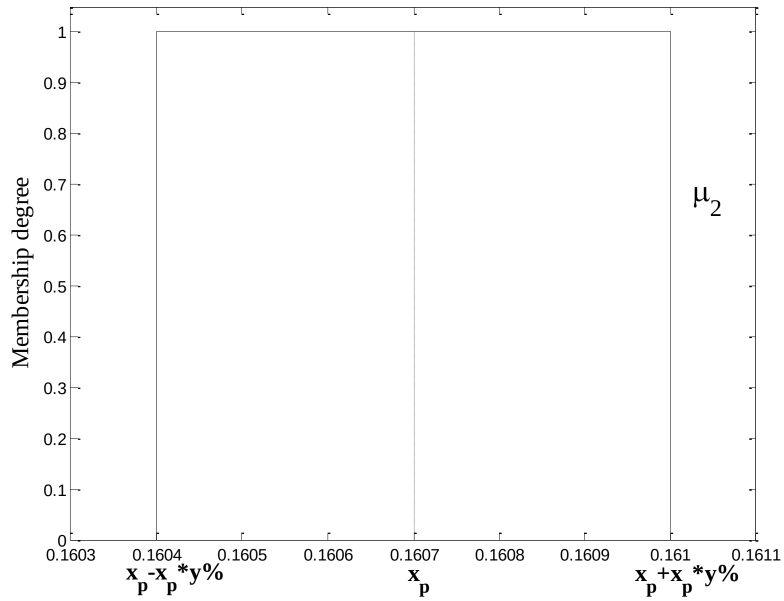

Figure 6. Rectangular membership function

If the corrected data from instruments or experts' empirical knowledge can be obtained, then the rectangular membership function can be revised and its shape may be no longer rectangular generally [10].

3) Model RFV of fault template pattern through combining μ_1 and μ_2

The final RFV can be determined by combining μ_1 and μ_2. At the same level α, the generated α-cuts of μ_1 and μ_2 are $[x_{E_L}^{\alpha}, x_{E_R}^{\alpha}]$ and $[x_{I_L}^{\alpha}, x_{I_R}^{\alpha}]$ respectively, where "E" and "I" denote "external membership function" and "inner membership function" respectively. Interval $[x_{E_L}^{\alpha}, x_{E_R}^{\alpha}]$ is divided into the two intervals $[x_{E_L}^{\alpha}, x_p^{\alpha}]$ and $[x_p^{\alpha}, x_{E_R}^{\alpha}]$. The α-cuts $X_\alpha = \{x_a^{\alpha}, x_b^{\alpha}, x_c^{\alpha}, x_d^{\alpha}\}$ of final RFV is defined by

$$\begin{aligned} x_b^{\alpha} &= x_{I_L}^{\alpha} \\ x_c^{\alpha} &= x_{I_R}^{\alpha} \\ x_a^{\alpha} &= x_b^{\alpha} - (x_p - x_{E_L}^{\alpha}) \\ x_d^{\alpha} &= x_c^{\alpha} + (x_{E_R}^{\alpha} - x_p) \end{aligned} \tag{4}$$

Therefore, A RFV can be represented using a M×5 matrix, whose row is $\{x_a^{\alpha}, x_b^{\alpha}, x_c^{\alpha}, x_d^{\alpha}, \alpha\}$. Obviously, the number of calculations required will increase with the numbers of α-cuts. In real applications, generally, the number of row usually is set 100, thus a RFV is represented by a 100×5 matrix [9]. Due to this simplicity, RFV could be conveniently realized through computer programming.

b.ii RFV model of testing patterns

During on-line monitoring and diagnosis, considering the affects of randomness and systematic error, it is unreliable to judge faults by single monitoring datum unless sudden accidents happen, because generation of fault is a gradual change process. Therefore, it is assumed that, at interval $\triangle t$, equipment runs stably and at least m (m≥60) monitoring data need to be collected by sensor instrument. Therefore, the method in subsection 3.2.1 can also be used to construct RFV of testing patterns. Comparing with the single monitoring datum or the mean value of several monitoring data, it is able to objectively reflect uncertainties involved in measurements of fault features in time interval Δt.

IV. DETERMINE BPA BASED ON SIMILARITY MEAUSRE BETWEEN RFVs

a. Similarity measure between RFVs

In fault diagnosis, it is necessary to calculate matching degree between a testing pattern (TP) and each fault template pattern (FTP) in fault database, and then matching degree will be converted into BPA. If the matching degree between a PT with a certain FTP is maximal, that is to say, the fault about this FTP happens most likely. In this section, a similarity measure fit for RFVs is given to calculate matching degree such that corresponding BPA can be obtained by normalizing matching degree.

Definition 1 For random-fuzzy variables A and B, the similarity measure between them is defined as

$$S(A,B) = e^{-d(A,B)} \qquad (5)$$

where

$$d(A,B) = \sum_{i=1}^{n} \alpha_i \cdot \frac{\left|a_{E_{Li}} - b_{E_{Li}}\right| + \left|a_{E_{Ri}} - b_{E_{Ri}}\right|}{2} \qquad (6)$$

$[a_{E_{Li}}, a_{E_{Ri}}],[b_{E_{Li}}, b_{E_{Ri}}],\ i=1,\ldots,M$ represent external confidence interval of α_i-cuts of RFV A and B, respectively.

Remark 1: The external membership function includes inner membership function of RFV, so it is enough that the external confidence interval is used to measure similarity between RFVs. Actually, RFV are composed by a set of its external and inner confidence intervals and the corresponding value α. Therefore, firstly, $d(A,B)$ measures the difference between them, and then, the negative exponential function maps $d(A,B)$ into interval [0,1] such that $S(A,B)$ satisfies the following conditions of similarity measure

C1. $0 \le S(A, B) \le 1$

Proof. Because $d(A,B) \ge 0$, so $0 \le e^{-d(A,B)} \le 1$, according to (5), we have $0 \le S(A,B) \le 1$. □

C2. $S(A, B) = 1 \Leftrightarrow A = B$

Proof. **Necessity:** if $A=B$, then $d(A,B)=0$. $e^{-d(A,B)}=1$, so $S(A,B)=1$; **Sufficiency:** $S(A,B)=1$ means $d(A,B)=0$, namely, $a_{E_{Li}} = b_{E_{Li}}, a_{E_{Ri}} = b_{E_{Ri}}$, so $A=B$ □

C3. $S(A, B) = S(B, A)$

Proof. Because

$$d(A,B) = \sum_{i=1}^{n} \alpha_i \cdot \frac{\left|a_{E_{Li}} - b_{E_{Li}}\right| + \left|a_{E_{Ri}} - b_{E_{Ri}}\right|}{2} = \sum_{i=1}^{n} \alpha_i \cdot \frac{\left|b_{E_{Li}} - a_{E_{Li}}\right| + \left|b_{E_{Ri}} - a_{E_{Ri}}\right|}{2} = d(B,A)$$

so we have $S(A,B)=e^{-d(A,B)}= e^{-d(B, A)}=S(B,A)$. □

b. Calculate BPA using the similarity measure between RFVs

In this subsection, an example illustrates that how to determine BPA based on similarity degree, and the method in [4] is also given for comparisons in following section.

Suppose machine rotor system have 3 fault modes F_j, j=1, 2, 3. The amplitude of foundational frequency (f_v) of vibration acceleration is still selected as fault feature. For this feature, FTP of every fault is constructed by using off-line method in 3.2.1. Specially, TP of F_1 is constructed by using on-line method in 3.2.2. The every FTP and TP are all modeled as corresponding RFVs, denoted as respectively A_j and B, j=1, 2, 3, shown in Figure 7.

100×3 matrices are used to represent RFVs of A_j and B. Elements of every row of the matrix are two endpoints of external confidence interval of α_i-cuts of RFV and α_i itself, where i=1… M.

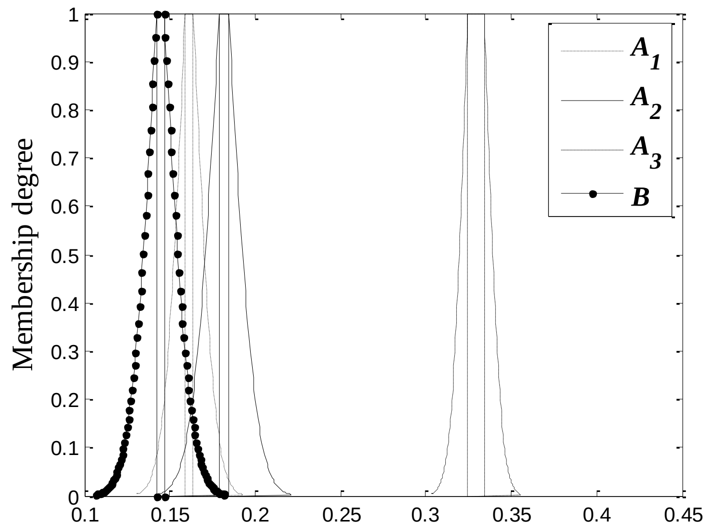

Figure 7. TP and Three FTPs

$$A_1 = \begin{bmatrix} 0.1302 & 0.1927 & 0.0027 \\ 0.1308 & 0.1921 & 0.0033 \\ \vdots & \vdots & \vdots \\ 0.1587 & 0.1645 & 0.9512 \\ 0.1590 & 0.1639 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0.1423 & 0.2212 & 0.0027 \\ 0.1431 & 0.2204 & 0.0033 \\ \vdots & \vdots & \vdots \\ 0.1783 & 0.1852 & 0.9512 \\ 0.1790 & 0.1845 & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0.3029 & 0.3559 & 0.0027 \\ 0.3034 & 0.3554 & 0.0033 \\ \vdots & \vdots & \vdots \\ 0.3240 & 0.3348 & 0.9512 \\ 0.3245 & 0.3343 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0.1069 & 0.1822 & 0.0027 \\ 0.1076 & 0.1815 & 0.0033 \\ \vdots & \vdots & \vdots \\ 0.1416 & 0.1474 & 0.9512 \\ 0.1424 & 0.1467 & 1 \end{bmatrix}$$

The matching degree between TP B and each FTP is

$$\begin{cases} \rho(F_j) = S(A_j, B) \\ \rho(\Theta) = \prod_{j=1}^{3} (1 - \rho(F_j)) \end{cases} \quad j = 1, 2, 3 \qquad (7)$$

The corresponding BPA can be obtained by normalizing matching degree

$$\begin{cases} m(F_j) = \dfrac{\rho(F_j)}{\sum_{j=1}^{3} \rho(F_j) + \rho(\Theta)} \\ m(\Theta) = \dfrac{\rho(\Theta)}{\sum_{j=1}^{3} \rho(F_j) + \rho(\Theta)} \end{cases} \tag{8}$$

In this example, $m(F_1)$=0.5111, $m(F_2)$=0.3884, $m(F_3)$=0.0527, $m(\Theta)$=0.0478.

In [4], FTPs for F_j, j=1, 2, 3, and TP are all modeled by Gaussian membership functions. FTPs are described as

$$\mu_{F_j}(x) = \begin{cases} \exp\left(-\dfrac{(x-M_{a,j})^2}{2\sigma_{a,j}^2}\right) & x < M_{a,j} \\ 1 & M_{a,j} \le x \le M_{b,j} \\ \exp\left(-\dfrac{(x-M_{b,j})^2}{2\sigma_{b,j}^2}\right) & x > M_{b,j} \end{cases} \tag{9}$$

where $M_{a,\,j}$ and $M_{b,\,j}$ are min and max mean value of n/5 typical historical data (n typical historical data are divided into 5 subgroup).

The TP for a certain fault is calculated as follows:

$$\mu_o(x) = \exp\left(-\frac{(x-M_o)^2}{2\sigma_o^2}\right) \tag{10}$$

where, M_o is the mean value of m on-line monitoring data.

The same typical historical and monitoring data are used to construct FTPs and PT about f_v. Thus, the parameters in (9) and (10) are as follow

$$M_{a,1} = 0.1596\,, \sigma_{a,1} = 0.01 \qquad M_{a,2} = 0.1696\,, \sigma_{a,2} = 0.0031$$
$$M_{b,1} = 0.1644\,, \sigma_{b,1} = 0.02 \qquad M_{b,1} = 0.1932\,, \sigma_{b,2} = 0.0138$$

$$M_{a,3} = 0.3247\,, \sigma_{a,3} = 0.0023 \qquad M_o = 0.1445\,, \sigma_o = 0.0118$$
$$M_{b,3} = 0.3387\,, \sigma_{b,3} = 0.0074$$

The membership functions of FTPs $\mu_{F_j}(x)$ of F_j, j=1, 2, 3 and TP $\mu_o(x)$ of F_1 about fault feature f_v are shown in Figure 8.

From Figure 8, we can see that the ordinate value of crossed points between TP and FTPs, namely, the matching degrees. It is calculated as follows:

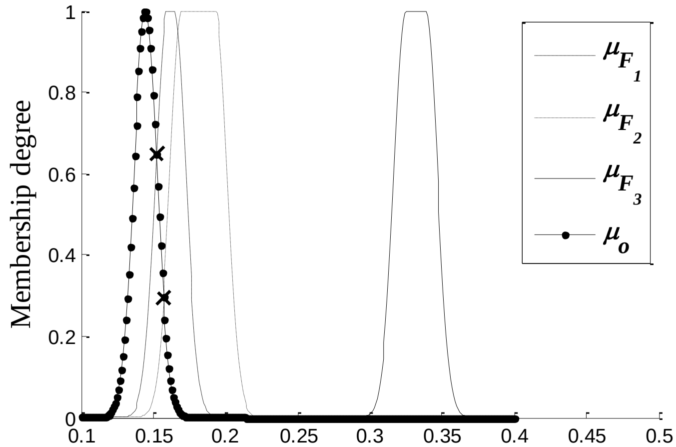

Figure 8. Membership functions of fault templates and tested model

$$\begin{cases} \rho(F_j) = \min_x \max(\mu_{F_j}(x), \mu_o(x)) \\ \rho(\Theta) = 1 - \max(\rho(F_j)) \end{cases} \quad j = 1,2,3 \qquad (11)$$

Similarly, by using (8), the corresponding BPA can be obtained as

$m(F_1)$=0.4926, $m(F_2)$=0.2237, $m(F_3)$=0.0000, $m(\Theta)$=0.2837.

Comparing with the BPA obtained by (11), our BPA supports F_1 more definitely, i.e., 0.5111>0.4926. The reason for this predominance is that RFV can model randomness and systemic error more comprehensively. In next section, the fault diagnosis experiments of machine rotor system will further illustrate this predominance in data fusion-based decision-making.

V. FAULT DIAGNOSIS EXPERIMENTS OF MACHINE ROTOR SYSTEM

The proposed method is applied in ZHS-2 machine rotor system shown in Figure 9 [11]. A vibration displacement sensor (VDS) and a vibration acceleration sensor (VAS) are installed on the bracket of rotor to collect vibration signals in both horizontal and vertical directions. The collected vibration signals are inputted into HG-8902 data collector, and then processed by signal conditioning circuits, finally the processed signals is inputted into laptop. The fault features can be acquired by HG-8902 data analysis software (under environment of Labview).The structure diagram of data collection system is shown in Figure 10, it is obvious that, besides of the systematic errors of vibration sensors, signal conditioning circuit and A/D convertor are also have conversion errors. The accumulation of these errors composes the systematic errors of the whole sensor instrument. It can be generally found in specifications of instrument provided by manufacturer. For VDS and VAS instruments in this experiment, their accuracies are all *y*= 0.02%.

Three typical fault of this rotor system are *rotor unbalance*、*rotor misalignment* and *motor bracket loosening*. The analyses to the large amount of experimental data show that abnormal vibration caused by faults will lead to increasing or decreasing of amplitudes of vibration frequencies. Therefore, the amplitude of frequency $k \times f_v$ ($k \times f_v$ means k times the frequency of f_v, k=1, 2, 3) of vibration acceleration and the average amplitude of vibration displacement (AAVD) are selected as fault features [11]. The decision-making is made based on fusion results of several BPAs obtained from every fault feature.

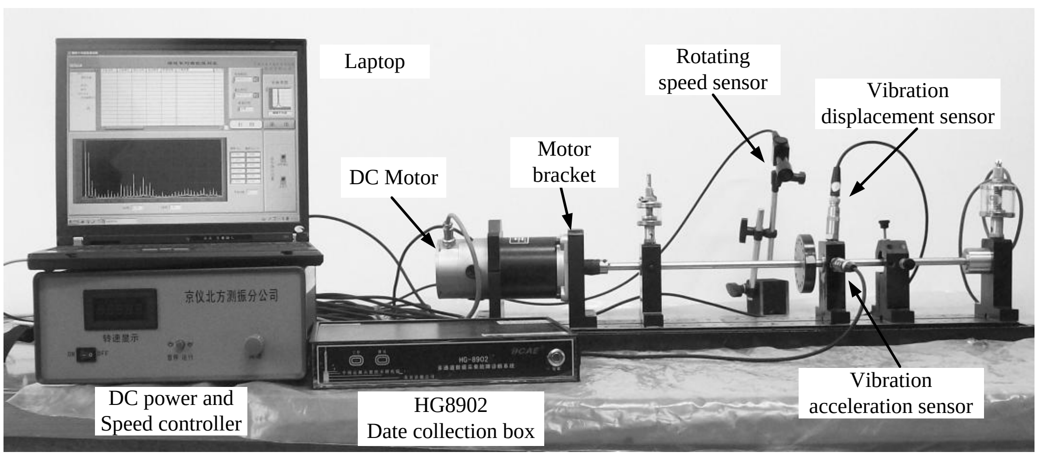

Figure 9. Diagram of experiment set-up

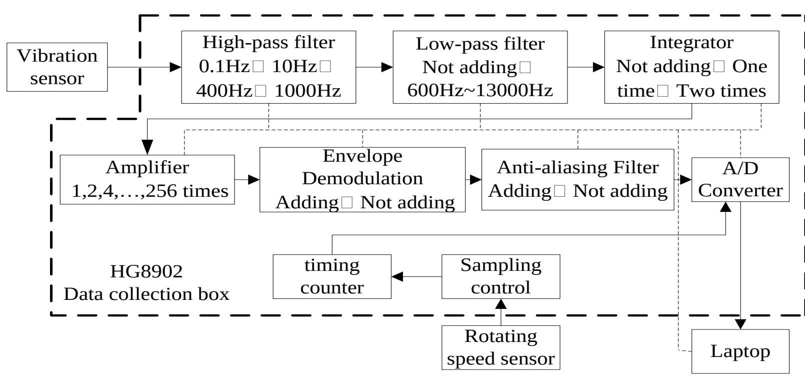

Figure 10. HG8902 data collection system

The specific steps are as follow:

1) Determining the frame of discernment

The frame of discernment is $\Theta = \{F_1, F_2, F_3\}$, where, F_1=*rotor unbalance*,F_2=*rotor misalignment*, and F_3=*motor bracket loosening*.

2) Off-line constructing the fault template patterns as FTP database

12 RFVs of FTPs can be modeled according to the method in section 3.2.1, denoted as A_{ij}, where i=4 represents the number of fault features, j=3 represents the number of fault modes.

3) On-line diagnosing

When a certain fault happens during on-line monitoring process, for 4 fault features, RFVs of TPs can be modeled according to the method in section 3.2.2, denoted B_i, where i=4 represents the number of fault features.

The matching degrees between TP B_i and each FTP A_{ij} is calculated by (7), Total 4 BPAs about 4 fault features can be got by normalizing the corresponding matching degree.

Dempster's rule of combination is used to fuse these i pieces of BPAs. Fault diagnosis is made based on fusion results. Generally, the determined fault mode needs to satisfy following conditions [1, 4], namely, the decision rules:

1) The BPA of the determined fault mode must be maximal. its value should be larger than a given threshold γ, here γ is set 0.6 empirically;
2) $m(\Theta)$ must be smaller than a given threshold Δ, here Δ is chosen as 0.3 empirically;
3) The difference between BPA of the determined fault type and BPA of other fault modes must be larger than a given threshold ξ, here ξ is set 0.15 empirically.

The experimental results are listed in table 1. The diagnosis result based on single BPA and fusion results of 4 BPAs are all given. From this table, it can be seen that we cannot correctly diagnose fault only based on single BPA obtained from the corresponding fault feature. But according to the fusion results of all BPAs, we can correctly judge faults, namely, the diagnosis results are identical with the current fault state of rotor system. Moreover, comparing with the fusion results by using method in [4] (listed in table 1), the proposed method can correctly diagnose fault while the method in [4] cannot, e.g., for F_2 in table 1. For the diagnosis of F_1 and F_3, the both methods can all make correct decisions. However, the BPA of the determined fault obtained by the proposed method is large than that in [4]. The larger the BPA is, the more confident the decision-making is.

Furthermore, in order to further validate the effectiveness of the proposed method, the procedures of on-line fault simulations (step 3) is repeated for 100 times. The experimental results are given in table 2. The values outside parentheses are the times of correct diagnosis for both methods, and the values in parentheses are the times that one method is more confident than counterpart. From the table we can see that, for F_1 and F_3, the both methods can all effective in 100 times of repeated experiments. However, the proposed method is more confident than method in [4] except for twice of diagnostic experiments of F_3. For F_2, the method in [4] hardly makes correct decisions, but the proposed method can diagnose fault F_2 in the entire repeated experiments.

Table 1: The experimental results

Fault mode			**BPA**				**Diagnosis results**
			$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(\Theta)$	
rotor unbalance F_1	Fault features	f_v	0.5396	0.3630	0.0223	0.0751	*Uncertain*
rotor unbalance F_1	Fault features	$2f_v$	0.5323	0.0177	0.0129	0.4372	*Uncertain*
rotor unbalance F_1	Fault features	$3f_v$	0.4912	0.0563	0.4362	0.0164	*Uncertain*
rotor unbalance F_1	Fault features	AAVD	0.0136	0.1029	0.3429	0.5405	*Uncertain*
rotor unbalance F_1	RFV fusion		**0.8631**	0.0466	0.0889	0.0015	$\boldsymbol{F_1}$
rotor unbalance F_1	Method in [5]		**0.6779**	0.1842	0.0842	0.0537	$\boldsymbol{F_1}$
rotor misalignment F_2	Fault features	f_v	0.5468	0.3679	0.0226	0.0628	*Uncertain*
rotor misalignment F_2	Fault features	$2f_v$	0.0191	0.5665	0.4128	0.0015	*Uncertain*
rotor misalignment F_2	Fault features	$3f_v$	0.1890	0.2796	0.2852	0.2462	*Uncertain*
rotor misalignment F_2	Fault features	AAVD	0.1409	0.3298	0.0001	0.5293	*Uncertain*
rotor misalignment F_2	RFV fusion		0.0294	**0.8904**	0.0801	0.0001	$\boldsymbol{F_2}$
rotor misalignment F_2	Method in [5]		0.0002	**0.5975**	0.4022	0.0001	***Uncertain***
motor bracket loosening F_3	Fault features	f_v	0.0657	0.0976	0.5138	0.3229	*Uncertain*
motor bracket loosening F_3	Fault features	$2f_v$	0.0122	0.3669	0.5039	0.1171	*Uncertain*
motor bracket loosening F_3	Fault features	$3f_v$	0.4547	0.0603	0.4676	0.0174	*Uncertain*
motor bracket loosening F_3	Fault features	AAVD	0.2647	0.2111	0.0001	0.5241	***Uncertain***
motor bracket loosening F_3	RFV fusion		0.1135	0.0698	**0.8146**	0.0021	$\boldsymbol{F_3}$
motor bracket loosening F_3	Method in [5]		0.1713	0.0254	**0.7803**	0.0230	$\boldsymbol{F_3}$

Table 2 Compared results of repeated experiments

	rotor unbalance	***rotor misalignment***	***motor bracket loosening***
The proposed method	100 (100)	100 (100)	100 (95)
Method in [5]	100 (0)	22 (0)	100 (5)

In conclusion, RFV considers the overall uncertainties of sensor measurement. Moreover, the new similarity measure is appropriate to RFV so that more precise matching results can be obtained, and then the fusion results are also more reasonable.

VI. CONCLUSIONS

In this paper, random-fuzzy variable, which can represent randomness and unknown systematic error simultaneously, is used to express and deal with the uncertainty in the procedure of sensor measurement. A new similarity measure appropriated for RFVs is presented for matching FTP and TP extracted from sensor data with uncertainty. The matching degree can be transformed into BPA in evidence theory by normalization. Then, Dempster's rule of combination is used to fuse several BPAs provided by many sensor instruments. The fault diagnostic experiments of machine rotor system illustrate that the proposed method outperforms the method in [4] which never considers the overall uncertainties. In a word, the experimental results show that the proposed method can enhance accuracy and reliability of data fusion-based diagnosis system.

ACKNOWLEDGMENTS

This research is supported by National Science Foundation of China (61004070, 61104019, 60934009, 61034006), The Zhejiang Province research Program Project of commonweal technology application (2012C21025).

REFERENCES

[1] X. Fan, M. Zuo. "Fault diagnosis of machines based on DS evidence theory. Part 2: Application of the improved DS evidence theory in gearbox fault diagnosis", Pattern Recognition Letters, Vol. 27, No. 5, 2006, pp. 377-385.

[2] X. B. Xu, H. S. Feng, Z. Wang, C. L. Wen. "An information fusion method of fault diagnosis based on interval basic probability assignment", The Chinese Journal of Electronics, Vol. 20, No. 2, 2011, pp. 255-260.

[3] IEC-ISO, *Guide to the Expression of Uncertainty in Measurement*, 1992.

[4] X. B. Xu, C. L. Wen, Y. C. Wang. "Information fusion algorithm of fault diagnosis based on random set metrics of fuzzy fault features", Journal of Electronics & Information Technology, Vol. 31, No. 7, 2009, pp. 1635-1640.

[5] A. Ferrero and S. Salicone. "The random-fuzzy variables: a new approach to the expression of uncertainty in measurement", IEEE Transactions on Instrumentations and Measurement, Vol. 53, No. 5, 2004, pp. 1370-1377.

[6] G. Shafer. *A Mathematical Theory of Evidence*, New Jersey: Princeton University Press, Princeton, 1976.

[7] S. Salicone. *Measurement Uncertainty: an Approach via the Mathematical Theory of Evidence,* Springer series in reliability engineering, Springer-Verlag, New York, 2007.

[8] A. Ferrero and S. Salicone. "Modeling and processing measurement uncertainty within the theory of evidence: Mathematics of random–fuzzy variables", IEEE Transactions on Instrumentations and Measurement, Vol. 56, No. 3, 2007, pp. 704–716.

[9] A. Ferrero and S. Salicone. "A method based on random-fuzzy variables for online estimation of the measurement uncertainty of DSP-based instruments", IEEE Transactions on Instrumentations and Measurement, Vol. 53, No. 5, 2004, pp. 1362-1369.

[10] A. Ferrero, S. Salicone. "The construction of random-fuzzy variable from the available relevant metrological information", IEEE Transactions on Instrumentations and Measurement, Vol. 58, No. 2, 2009, pp. 365-374.

[11] L. Ma. Equipment fault diagnosis technique and its application, Chinese academy of carrier rocket technology, China: Beijing Jinghang Corporation, Technical Report, 1998, pp.34-58.