TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 04/24 pp262-265
Volume 9, Number 3, June 2004

Multisensor Data Fusion for Automotive Engine Fault Diagnosis*

WANG Yunsong (王 松)**, CHU Fulei (褚福磊),
HE Yongyong (何永勇), GUO Dan (郭 丹)

State Key Laboratory of Tribology, Department of Precision Instruments and Mechanology, Tsinghua University, Beijing 100084, China

Abstract: This paper describes mainly a decision-level data fusion technique for fault diagnosis for electronically controlled engines. Experiments on a SANTANA AJR engine show that the data fusion method provides good engine fault diagnosis. In data fusion methods, the data level fusion has small data preprocessing loads and high accuracy, but requires commensurate sensor data and has poor operational performance. The decision-level fusion based on Dempster-Shafer evidence theory can process noncommensurate data and has robust operational performance, reduces ambiguity, increases confidence, and improves system reliability, but has low fusion accuracy and high data preprocessing cost. The feature-level fusion provides good compromise between the above two methods, which becomes gradually mature. In addition, acquiring raw data is a precondition to perform data fusion, so the system for signal acquisition and processing for an automotive engine test is also designed by the virtual instrument technology.

Key words: engine; fault diagnosis; data fusion; Dempster-Shafer evidence theory

Introduction

In recent years, multisensor data fusion has received significant attention for both military and nonmilitary applications. Data fusion techniques combine data from multiple sensors and related information from associated databases to achieve improved accuracies, and then more specific inferences could be achieved by the use of a single sensor alone.

Automobiles are complex machines with integrated mechanical, electronic, and hydraulic technology, so automotive engine diagnosis is a process during which data of multisource sensors monitoring automotive operating conditions are acquired, combined, and applied. From the diagnostics aspect, single source information is always fuzzy, imprecise, and incomplete. Only multisource data acquired and combined provide more reliable and accurate test and diagnosis for automobiles, so we can apply data fusion technology to automotive fault diagnosis.

1 Signal Preprocessing and Feature Extraction

1.1 Signal acquisition and processing based on virtual instrument technology

Acquiring raw data is a precondition to perform data fusion. We designed a system for signal acquisition and processing for an automotive engine test by

Received: 2003-11-07; revised: 2003-12-22; selected from Proceedings of the Symposium on Frontiers and Challenges of Mechanical Science and Technology

* Supported by the Trans-Century Training Programme Foundation for the Talents by the Ministry of Education, China and Shandong Natural Science Foundation, China (No.Y2002F17)

** To whom correspondence should be addressed.
E-mail: syw992308@163.com; Tel: 86-10-62783968

employing virtual instrument technology. The system hardware consists of sensors, signal conditioning circuits, PCI-6023E data acquisition board, and personal computer. The software was programmed by means of a virtual instrument developing workbench—LabVIEW. The system acquires real time data, displays historical and transient data, implements data postprocessing of digital filtering, frequency spectrum analysis, and waveform drawing.

1.2 Feature selection and extraction

In engine fault diagnosis the virtual instrument system acquires and processes data from sensors and actuators to obtain features reflecting engine operating condition. Some of the features are redundant. This brings much computing load to data postprocessing, and as a result reduces diagnosis efficiency. Hence, the number of features must be reduced for data fusion.

The simplest feature-selecting method is on the basis of expert knowledge and experience. Unsteady idling is a commonly-found malfunction resulting from fuel injection valve blockage, poor ignition or intake manifold vacuum leakage for automotive electronically controlled engine. We can choose the four electronic waveforms of oxygen sensors, ignition system, injection fuel valve, and vacuum sensor to diagnose the fault. The information of the sensor and actuator data is embodied in the waveform structure which is signified by extracted features. For example, the oxygen waveform can completely be expressed by the four features of maximum, minimum, and average voltages, and permissible response time to air-fuel mixture from rich to lean; ignition waveform expressed by puncture, spark, and minimum voltages, dwell angle, spark duration, and residual voltage oscillation times in ignition coil; fuel injection waveform expressed by maximum and minimum voltage, and fueling duration; and vacuum waveform expressed by maximum, minimum, and average voltages.

The feature vectors after feature selection have multiple dimensions and can be reduced by the Karhunen-Loeve transform. Thus a few main components of linear independence are used to represent features. For instance, we can choose the first-order and second-order main components as features of the oxygen waveform.

2 Application of the Decision-Level Data Fusion Method

Observational data may be combined, or fused, at a variety of levels from the raw data (or observation) level to a state vector level, or at the decision level. Raw sensor data can be directly combined if the sensor data are commensurate (i.e., if the sensors are measuring the same physical phenomena such as two temperature sensors). Conversely, if the sensor data are noncommensurate, then the data must be fused at a feature/state vector level or decision level. Feature level fusion involves the extraction of representative features from sensor data. In feature level fusion, features are extracted from multiple sensor observations, and combined into a single concatenated feature vector which is input pattern recognition approach based on neural networks method. Decision level fusion involves fusion of sensor information, after each sensor has made a preliminary determination of fault diagnosis. An example of decision level fusion methods is Dempster-Shafer (D-S) evidence theory, which is mainly described in this paper.

D-S evidence theory offers an interesting tool to combine data from multisensors. It is used extensively for target tracking, automated identification of targets, and limited automated reasoning applications, but scarcely for fault diagnosis. In this section we will study the application of D-S evidence theory to fault diagnosis for SANTANA AJR engine.

D-S evidence theory introduces a concept of discernment frame, i.e., a set of mutually exclusive propositions. The correlated propositions are assumed to be expressed as subsets of the frame Θ, namely, a finite set. Basic probability assignment, or mass function, over Θ is

$m(\varnothing)=0$, $\sum_{A\subseteq 2^{\Theta}} m(A)=1$, when the mapping is

$$2^{\Theta}\rightarrow[0,1] \tag{1}$$

where $\varnothing$ is null set; A is a subset; $A\subseteq\Theta$.

Subset B of set Θ is called the focal element of a mass function. The belief measure $\mathrm{Bel}(B)$ and plausibility measure $\mathrm{Pl}(B)$ can be computed using:

$$\mathrm{Bel}(B)=\sum_{A\subseteq B} m(A) \tag{2}$$

$$\mathrm{Pl}(B)=\sum_{B\cap A\neq\varnothing} m(A) \tag{3}$$

The belief measure Bel(B) quantifies strength of the belief that event A occurs. The functions m, Bel, and Pl are derived from the concept of lower and upper bounds for a set of compatible probability distributions. In addition, D-S evidence theory allows the fusion of several sources using Dempster's combination operator. It is defined like orthogonal sum (commutative and associative) in terms of the following equation,

$$m(B) = m_1(B) \oplus m_2(B) \oplus \cdots \oplus m_M(B) \tag{4}$$

For two sources, the aggregation of evidence for a hypothesis $B \subseteq \Theta$ can be expressed as follows:

$$m(B) = \frac{1}{Q} \sum_{A \cap C = B} m_i(A) \cdot m_j(C) \tag{5}$$

where $Q = 1 - \sum_{A \cap C = \varnothing} m_i(A) \cdot m_j(C)$.

Now study the application of D-S evidence theory to engine fault diagnosis. Suppose that there are N mutually independent fault types to be classified, and M sensors to monitor automotive engine conditions. Use the vector $\boldsymbol{H}_j=\{h_{j0}, h_{j1}, \cdots, h_{jn}\}$ ($j = 1, 2, \cdots, N$) to denote features of the j-th fault, and $\boldsymbol{S}_k=\{s_{k0}, s_{k1}, \cdots, s_{kn}\}$ ($k = 1, 2, \cdots, M$) to denote features from measurement data of the k-th sensor, where n is the number of elements in the discernment frame. The smaller the degree of relation between the two vectors, the larger the probability of the j-th fault assigned by the k-th sensor. The degree of relation between the two vectors can be expressed by Euclidean distance:

$$d_{kj} = d(\boldsymbol{S}_k, \boldsymbol{H}_j) = [\sum_{i=1}^{n} (s_{ki} - h_{ji})^2]^{1/2},$$
$$k = 1, 2, \cdots, M, \quad j=1,2,\cdots, N \tag{6}$$

where s_{ki} is the i-th component of the vector $\boldsymbol{S}_k$, h_{ji} is the i-th component of the vector $\boldsymbol{H}_j$.

Using Eq. (6) , we obtain the distance matrix $\boldsymbol{D}$,

$$\boldsymbol{D} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1N} \\ d_{21} & d_{22} & \cdots & d_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ d_{M1} & d_{M2} & \cdots & d_{MN} \end{bmatrix} \tag{7}$$

After substituting d_{kj} for $p_{kj} = 1/d_{kj}$ and normalizing every row of the matrix $\boldsymbol{D}$, we can obtain another matrix $\boldsymbol{P}$,

$$\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MN} \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \vdots \\ \boldsymbol{P}_M \end{bmatrix} \tag{8}$$

The elements $p_{i1}, p_{i2}, \cdots, p_{iN}$ ($i = 1, 2, \cdots, M$) in the rows of the matrix $\boldsymbol{P}$ can be considered as the values of mass function which the i-th sensor assigns for all faults.

During D-S evidence reasoning, we can define a mass function over the fault discernment frame Θ on the basis of the information one evidence provides. Thus a belief structure can be considered by reflection of the relevant evidence over discernment frame. For instance, a vacuous structure whose mass function is zero, i.e., $m(\varnothing) = 0$, implies that the relevant evidence provides nothing of information; certainty structure whose mass function is one, i.e., $m(A) = 1$, confirms existence of the fault A. Information capacity provided for fault discernment frame by a belief structure can be measured by the evidence entropy as follows:

$$E_m = \prod_{A \subset \Theta} [\mathrm{Pl}(A)]^{m(A)} = \prod_{m(A) \neq 0} [\mathrm{Pl}(A)]^{m(A)} \tag{9}$$

The bigger or closer to one the evidence entropy E_m is, the more information the evidence provides. So we can say that the evidence entropy reflects actually decision confidence of combined evidences.

Through a great deal of experiments on the SANTANA AJR engine, we obtain three standard feature vectors:

(1) $\boldsymbol{H}_0$={–823.9472, –442.2903, 3.2895, 5.1562, 35.3872, 7.7889, –53.0390}, when the engine is in good condition;

(2) $\boldsymbol{H}_1$={536.5645, –649.1083, 3.9637, –7.9034, –32.2867, –97.1227, –21.0975}, when unsteady idling results from fuel injection valve blockage;

(3) $\boldsymbol{H}_2$={90.1887, –601.0368, 5.5516, –5.2591, –32.3023, –80.4196, –13.5560}, when unsteady idling results from intake manifold vacuum leakage.

The goal of the data fusion technique in fault diagnosis is to properly discern every proposition in the discernment frame, namely, to correctly detect and classify faults. Here, suppose that the fault discernment frame is $\boldsymbol{\theta} = \{\boldsymbol{H}_0, \boldsymbol{H}_1, \boldsymbol{H}_2\}$. Feature vectors from the observational measured data are obtained: 1) for the oxygen sensor, $\boldsymbol{O}$={530.3708, –650.9331}; 2) for the ignition system, $\boldsymbol{D}$={3.9587, –7.9640}; 3) for the fuel injection valve, $\boldsymbol{P}$={–35.9027}; and 4) for the intake manifold vacuum sensor, $\boldsymbol{Z}$={–97.0850, –22.1008}.

We use D-S evidence theory to deduce the condition of AJR engine from the vectors. The feature vectors from the four sensors were combined at the decision level to get the final fusion results as follows:

$$\mathrm{Bel}(\boldsymbol{H}_0) = m(\boldsymbol{H}_0) = 0.0276;$$

$$\mathrm{Bel}(\boldsymbol{H}_1) = m(\boldsymbol{H}_1) = 0.9539;$$

$$\mathrm{Bel}(\boldsymbol{H}_2) = m(\boldsymbol{H}_2) = 0.0185.$$

According to the above results, we conclude that there may be a failure of fuel injection valve blockage. This conclusion is consistent with the practical experiment.

We can compute the evidence entropy from Eq. (9). The change of the evidence entropy during the data fusion process is shown in Fig. 1. Along with proceeding of data fusion, the evidence entropy increases. This implies that the diagnosis decision confidence improves. Multisensor data fusion has significant advantages over single sensor data, decreases uncertainty of fault diagnosis, and improves classification ability of diagnosis system for fault patterns.

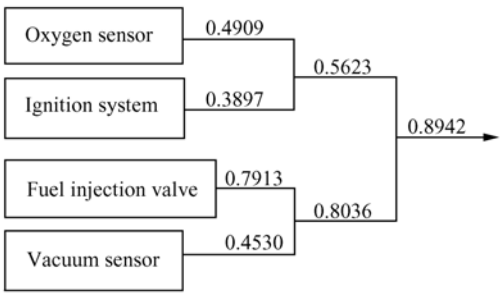

Fig. 1 Change of evidence entropy during data fusion process

3 Conclusions

In data fusion methods, data-level fusion has a small data preprocessing load and high accuracy, but requires commensurate sensor data and has poor operational performance. The decision-level fusion method can process noncommensurate data and has robust operational performance, reduces ambiguity, increases confidence, and improves system reliability, but has low fusion accuracy and high data-preprocessing costs. Feature-level fusion based on neural networks improves the fault recognition rate for training samples and the fault recognition performance for test samples, and uses information from multiple sources to supplement other sensors, thus increasing system reliability and diagnosis accuracy. The feature-level fusion method provides good compromise between the other two methods, which becomes gradually mature. Experiments show that the decision-level data fusion method can be used for engine fault diagnosis.

The selection among these methods for a particular application is a system engineering problem which depends upon issues such as the available communications bandwidth, sensor characteristics, computational resources available, and other issues. There is no one universal method which is applicable to all situations or applications.

Acknowledgements

The experimental data in the present work have been supplied by the Automobile Laboratory of Shandong University of Technology. The authors wish to thank Yang Zhongyu for the help of the experiment test.

References

[1] Ampazis N, Perantonis S J, Taylor J G. Dynamics of multiplayer networks in the vicinity of temporary minima. *Neural Networks*, 1999, **12**(1): 43-58.

[2] Hall D L. An introduction to multisensor data fusion. *Proceedings of the IEEE,* 1997, **85**(1): 6-23.

[3] Ganonlay A. Multiple platforms sensor integration model: MULSIM computer program. ADA 079951. *Naval Research Laboratory,*1999, **31**(12): 1-62.

[4] Tenney R. Detection with distributed sensors. *IEEE Trans. On AES*, 1981, **17**(1): 98-101.

[5] Ronald R Y. Entopy and specificity in a mathematical theory of evidence. *Gen. Systems,* 1983, **9**(2): 249-260.

[6] Stephane J. Contribution to multisensor fusion formalization. *Robotics and Autonomous Systems*, 1994, **13**(2): 56-60.

[7] Varshney P K. Distributed Detection and Information Fusion. New York: Springer-Verlag, 1996.

[8] Philip R D. A new failure detection approach and its application to GPS autonomous integrity monitoring. *IEEE Transactions on AES*, 1995, **31**(1): 499-506.

[9] Philip L, Bogler H. Shafer-Dempeter reasoning with applications to multisensor target identification systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 1997, **17**(6): 901-930.

[10] Isermann R. Process fault detection based on modeling and estimation methods—A survey. *Automatica*, 1984, **20**(2): 47-54.

[11] Carnault J. Third generation antiskid: Safety diagnostive features of the ECU. SAE890093.