# Multi-source Information Fusion Based on Data Driven

Xin Zhang[1], Li Yang[1, 2], Yan Zhang[1], Jinxue Sui[1]
[1]Shandong Institute of Business and Technology / School of Information and Electronic Engineering, YanTai, China
suijinxue@163.com
[2]Department of Control Engineering / Naval Aeronautical Engineering University, YanTai, China,
suijinxue@163.com

***Abstract*—Take data driven method as the theoretical basis, study multi-source information fusion technology. Using online and off-line data of the fusion system, does not rely on system's mathematical model, has avoided question about system modeling by mechanism. Uses principal component analysis method, rough set theory, Support Vector Machine(SVM) and so on, three method fusions and supplementary, through information processing and feature extraction to system's data-in, catches the most important information to lower dimensional space, realizes knowledge reduction. From data level, characteristic level, decision-making three levels realize information fusion. Through the fusion example to the data which the fire surveys to confirming, it indicated that this method reduced computational complexity, reduced information loss in the fusion process, and enhanced the fusion accuracy.**

***Index Terms*—Information fusion, Data driven, Principal component analytic method, rough set theory, Support Vector Machine (SVM))**

## I. Introduction

The Multi-source information fusions are one new research field that has very broad prospect, widely application in the multi-sensor information processing. It uses many kinds of type sensor's different characteristic, gain system different attribute information from the multi-position comprehensively, from the time and space to carry on examination, relevance, estimate and synthesis and so on multi-aspect various processing to information which comes from the multiple source, rational dealing multiple source information which various sensors provide has redundancy, complementarity and cooperativity, in order to obtain the precise condition and status estimation that as well as complete, prompt situation appraisal and treat assessment[1].

The Multi-source information fusions have fault tolerance, complementarity, timeliness and so on characteristics, can increase the measuring dimension and confidence, improve system's detective performance and survivability, expand space and time coverage area, improve system reliability and maintainability, enhances the system fault tolerance and robustness, achieved complement each other's advantages, resource sharing, raises the utilization rate of resources in the system [2].

## II. Question Statement

The commonly used multi-sensor information fusion method may divide into Bayesian Inference, Weighted Average Method, Maximun Likelihood, Dempster-Shafer Inference, Kalman Filter, Cluster Analysis, Fuzzy Logic, Neural Networks, Wavelet Theory, Rough Set Theory, Support Vector Machines and so on methods[3]. But these fusion algorithms are proposed mostly according to the concrete question that can obtain the most superior effect to the specific domain's question. Therefore, the existing fusion algorithms have their certain application.

The Multi-source information fusion technology as the emerging technology, has not formed the basic theory frame and the generalized fusion algorithm, because the information will be affected by many kinds of factors in the gathering process, therefore information's uncertainty, incompletion and redundancy universal exist; Multi-sensor's observed result exists the associated ambiguity; The information fusion level is insufficient, is unable to realize fusion in the true sense [4].

When the input/output mechanism model of the fusion system is inaccurate, very large uncertainty, or too complex, high order, strong nonlinear, difficult to analyze and design, perhaps when it is very difficult to establish the mechanism model, is very difficult to realize the high-grade fusions of the multi-source information using the traditional fusion method based on the model, needs gain mass data by using the information technology fully, establishes the processing method based on data driven [5]. The data driven includes the data driven thought and the data driven control, its thought refers to realizing the data predict, appraisal, dispatch, monitoring, diagnosis, decision-making and optimization of the system and so on each kind of expectation function by using the online and off-line data of the controlled system [6].

Realizes the fusion inference with the data driven method, uses the online and off-line data of the fusion system, does not rely on system's mathematical model, has avoided system mechanism modeling question[7]. Its goal reduces influence that the multi-sensor observation data is affected by the environment condition and the sensor itself characteristic, realizes the data conversion, reduces the information loss, controls and the complexity


doi:10.4304/jsw.6.6.1125-1132

of the related computation in the fusion process, maintains the data uniformity, reduces the data associated ambiguity; Realizes the information fusion from the data level, the characteristic level and the decision-making level; Carries on the fusion optimization, improves the fusion system's robustness and stability. Uses the theoretical research, optimized processing and the example confirmation method to study, carries on the appraisal through the fusion example, and further enhances its result precisely reliable.

## III. THE MULTI-SOURCE INFORMATION FUSION BASED ON DATA DRIVEN

### *A. Overall fusion thought*

Fuses the existing multi-source information fusion method, based on the data driven thought, through analysis to multiple source information which comes from sensor, classifies the redundant information, the supplementary information, the coordination information, carries on information processing and feature extraction, catches the most important information to the lower dimensional space, realizes the knowledge reduction. Uses the principal element analytic method, the Rough Set Theory and the support vector machines separately, from the data level of the information fusion, the characteristic level, the decision-making level three levels realize multiple source information fusion based on the data driven, and carries on the examination to the fusion result through the test sample. Fusion system principle diagram is shown in Figure 1.

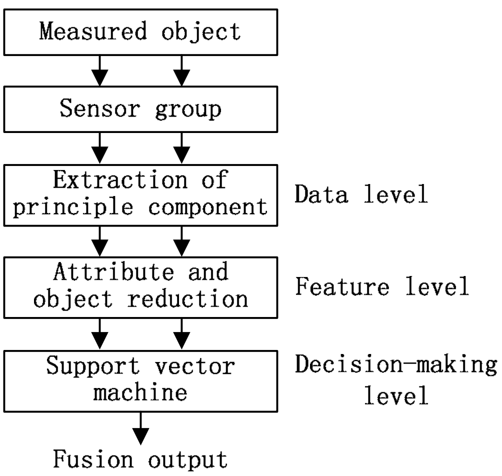



Figure 1. The Multi-source information fusing system functional block diagram based on data driven

### *B. Data level information fusion*

Uses the principal Component analytic method (PCA) to process information of the data level. The principal Component analytic method is one multi-dimensional statistical method that transforms many correlated variables about object of study as the minority non-correlated variable, it defines the primary and secondary status of data change direction according to eigenvalue of the stochastic matrix's covariance matrix, separates the information, eliminate the partial noises that contains in the multivariable time series and the redundant information, thus obtains various ranks principal element variable according to the order, assigns the number of the principal element according to information how many which needs to retain. The highly related primary information will project to the low-dimensional space, reduces the dimension of the multivariable time series, causes to the variance of the various vector components in the low-dimensional space to be biggest, retains the original useful information, reflects the relevance among the different variable, reduces the relativity, eliminates the false influence which the dimension difference brings and solves the unusual sampling value and the missing data question, causes each variable be in the 'equality' status, achieves the optimization of the feature extraction. The projection subspace has reflected the main change of the procedure variable, but the residual space mainly has reflected the process noise and disturbance and so on. The data after pretreatment projects separately to the principal element subspace and the residual subspace through PCA, carries on the statistical analysis and establishes the corresponding statistics to carry on the supposition examination[8][9].

The step that the principal element analysis method realizes the data level fusion to be as follows:

*1) primitive sample standardization processing*

Suppose there is n sensor to carry on measurement to some object $X$ , each sensor measured data number is m, then the sampled data matrix is:

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix} \quad (1)$$

Carries on standardized processing to the sampled data matrix's various variables, eliminates the influence of each measured dimension. The corresponding standardized matrix is:

$$Y = \left(\frac{X_{ij} - \overline{X}_j}{\sigma_j}\right)_{m\times n} \quad (2)$$

Where $\overline{X}_j$ and $\sigma_j$ is the sample mean and the sample standard deviation of various targets respectively, $i$ =1,2,…, m; $j$ =1,2,…, n

*2) Calculates correlation matrix $R$ and solves its eigenvalue and the eigenvector*

Obtains the covariance of the principal element $Z$ according to data matrix Y after standardized:

$$\mathrm{cov}(Z) = u'\mathrm{cov}(Y)u = u\sum u' = \Lambda \quad (3)$$

where $\sum$ is the covariance matrix of the matrix Y after standardization, $\Lambda$ is the diagonal matrix,

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

Random constant vector $u$ is the p order orthogonal matrix, that is $u'u = I$ .Let

$$u = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{p1} \\ l_{12} & l_{22} & \cdots & l_{p2} \\ \cdots & \cdots & \cdots & \cdots \\ l_{1p} & l_{2p} & \cdots & l_{pp} \end{bmatrix}$$

Replaces the covariance matrix by random variable's correlation matrix to calculate the principal element, reduces the bit drop-out as far as possible which the dimensionality reduction and the feature extraction brings, causes it intently to reflect the change information which contains in the original variable, and effective separation system information and noise. Because data matrix $Y$ is the standardized matrix, therefore may use correlation matrix $R$ to replace the covariance matrix, therefore type (3) may be expressed as:

$$uRu' = \Lambda \tag{4}$$

Using $u'$ left multiply by (4),

$$Ru' = u'\Lambda \tag{5}$$

May obtain by conversion:

$$|\lambda I - R| = 0 \tag{6}$$

The solved principal element variance $\lambda_i$ ( $i$ =1,2,…, n) is n root of the above equation, $\lambda$ is eigenvalue of the correlation matrix.

*3) Selects the principal element*

In order to describe the information content size quantificationally which the principal element provides, defines variance contribution rate ci of the principal element variable $Z_i$ and first k principal element accumulative variance contribution rate $\eta_k$ as follows:

$$c_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}, \eta_k = \sum_{i=1}^{k} c_i = \frac{\sum_{j=1}^{k} \lambda_j}{\sum_{j=1}^{n} \lambda_j} \tag{7}$$

The type (7) may use to weigh the first k principal element share of the total information content. As a general rule, may take k principal element, causes its accumulation variance contribution rate $\eta_k \geqslant 85\%$ (may determine this value according to system's request).

*4) Calculates the principal element as this level outputs*

Because R is the positive definite matrix, therefore its characteristic root is the nonnegative real number, ranged the correlation matrix R characteristic root according to size $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$, its corresponding eigen vector is $l_1, l_2, \cdots, l_n$ . Therefore the principal element might write:

$$\begin{aligned} Z_1 &= l_1^T x = l_{11}x_1 + l_{12}x_2 + \ldots + l_{1n}x_n \\ Z_2 &= l_2^T x = l_{21}x_1 + l_{22}x_2 + \ldots + l_{2n}x_n \\ &\cdots \\ Z_n &= l_n^T x = l_{n1}x_1 + l_{n2}x_2 + \ldots + l_{nn}x_n \end{aligned} \tag{8}$$

When the accumulation contribution rate of the current k principal element $\eta_k \geqslant 85\%$, indicated that definitely may replace the original n initial condition attribute with the first k principal element not to lose too many information, reduced the input standard vector. Calculates various principal elements value which needs, various principal elements' weight obtains directly from the primary data of the information computation which assigns through the technical progress factor of the corresponding principal element's variance, obtains the initial fusion result of the data level through the weighting, thus forms the output sample sets of the data level information fusion, as the comparative input of the policy-making level fusion.

### *C. Feature level information fusion*

Takes the data level fusion output's result(Principal element variable which principal element analysis obtains) as this level input, uses the rough set theory, through the undistinguished relations and the undistinguished type to determine the approximate territory of the given question, on the premise of maintaining the classified ability invariable, through the knowledge reduction, derives the classifying rule of the concept, defines attribute dependence and importance, depend upon the historical data to determine that weights of the various essential factors in the multiple source information, may obtain are more objective and scientific weight. Under the premise of not losing the information, by simple expressing decision system's conclusion attribute to condition attribute set dependence and connection, solves its importance in decision system. Realizes the attribute selective preference using the attribute reduction, then carries on the object reduction, rejects the redundant ingredient of the characteristic data, extracts independent essential characteristic. Therefore, does not need any subjective apriori information outside the data set to be able to carry on the objective description and processing to the uncertainty, can effectively analyze and process the imprecise, inconsistent and incomplete information. The attribute reduction is carries on the important pre-processing process of the data mining and the rule extraction, may determine that the duty related attribute set, deletes attributes which does not constitute influence, thus reduces search space and raises efficiency.

*1) Rough sets theory* [10]-[12]

The rough sets theory was proposed by Polish mathematician Z.Pawlak in 1982 that was one kind of new mathematical tool that processes fuzzy and inaccurate question. In the rough sets theory the knowledge may be regard as the classified ability for domain of discourse $U$ .

Supposes $U$ is the domain of discourse, R is one equivalent relation of U. If $X \subset U$ , then defines $R^*(X)$ for the upper approximate sets of X, the lower approximate sets $BN_R(X)$ and the boundary $BN_R(X)$ respectively is:

$$R^*(X) = \{x \in U, R(x) \cap X \neq \Phi\} \tag{9}$$

$$R_*(X) = \{x \in U, R(x) \subseteq X\} \tag{10}$$

$$BN_R(X) = R^*(X) - R_*(X) \tag{11}$$

where $\Phi$ is the null set.

$R_*(X)$ is the set that all can certainly belong to $X$ element regarding knowledge $R$ , $R^*(X)$ is the set that possibly belongs to X element regarding knowledge $R$ , $BN_R(X)$ is the set that already not only cannot belong to $X$ , but also cannot belong regarding knowledge $R$ . When $X$ can express the certain R basic category, called that $X$ is $R$ definable, otherwise called that $X$ is $R$ indefinable. $R$ definable set may be called as the R precise set, but $R$ indefinable set can be called as the R rough set. If $R_*(X)$ and $R^*(X)$ is not equal, then $X$ is rough.

*2) Knowledge expression system and decision-making table*

The rough set in fact is taking the objective world as a definition knowledge expresses system in the abstract, the expression is $S = \langle U, A, V, f \rangle$ , $U$ is the non-null finite set of the object, expresses universe of discourse; A is the non-null finite set, expresses attribute; $V$ is the value domain set of attribute; $f$ is the information function (information, function).

Attribute set A also is divided frequently two sets $C$ and $D$ , $D$ has an attribute generally; If $A = C \bigcup D$ , $C \bigcap D = \phi$ , C is called the condition attribute, D is called the decision-making attribute. $V = \bigcup(V_a)$ , $V_a$ is definition domain of the attribute a, $a \in A$. $f$ : $U \times A \to V$, $f(X_i, A_j) \in V_j$ , it entrusts with an information value for each object's each attribute.

Knowledge expression system is decision-making table that has the condition attribute and the decision-making attribute, calls $T = \langle U, A, C, D \rangle$. The equivalence class of the relation $ind(C)$ and $ind(D)$ respectively are the condition class and the decision-making class.

Regarding each $x \in U$ and $a \in C \bigcup D$ , defines function:

$$d_x : A \to V, d_x(a) = a(x) \tag{12}$$

Function $d_x$ is called the decision rule in decision-making table $T$ ; $x$ is the identification of the decision rule $d_x$ , namely the set $U$ element does not express any actual thing in the policy-making table, is only the decision rule identifier.

*3) Attribute dependences and importance*

In the rough set theory, does not use the supposition information in advance, but merely using data in table to compute all attribute important degree.

Suppose there is two attribute sets $C_1$ and $C_2$ , then defines $C_1$ 's degree of dependence on $C_2$ is:

$$\gamma_{C_1}(C_2) = \frac{pos_{C_1}(C_2)}{|U|} \tag{13}$$

Let $a \in C \bigcup D$ , C is the condition attribute set, D is the decision-making attribute set, then defines the attribute importance of $a$ is:

$$\gamma_C(D) - \gamma_{C-|a|}(D) \tag{14}$$

The type (14) indicates the influence on classification decision-making after removing $a$ from $C$ .

*4) Attribute reduction and core*

Reduction: let $R$ be an equivalent relation, and $r \in R$, when $ind(R) = ind(R - \{r\})$ , named that $R$ may reduce to be $R - \{r\}$ , if it may not go a step further reduction, then said that it is a $R$ reduction, named it red $(R)$ .

Core: $R$ all reduction's intersection is called core of $R$ , names it $core(R)$ , that is:

$$core(R) = \bigcap red(R) \tag{15}$$

The core contains in all reduction bunch, the core cannot eliminate knowledge characteristic part set.

One of rough set theory's key contents is the knowledge reduction, namely on the condition that maintains knowledge library classification ability invariable, deletes non-correlated or unimportant knowledge. Attribute is no less important in decision-making table, some attributes are the redundancy.

If there is a subset of properties $Q \subseteq C$ , the classified ability regarding the decision-making attribute D will be invariable, and $Q$ is independent on $D$ , then $Q$ is called $D$ relative reduction (RelativeDeduct) of $C$ , and $Q_i (i = 1,2,\cdots,n)$ each reduction maintains classification ability to be invariable to the decision-making attribute, each relative reduction may construct one decision-making table in the classified ability commends which is same as the original decision-making table, therefore, their output result must be consistent.

The relative reduction's computation may use discernibility function to solve, in the canonical form of discrimination function, all conjunction form in the minimum canonical of the discernibility function is $D$ relative reduction of $C$.

The attribute reduction's method is examining whether exsits $r_R(Q)=1$ according to the rough set theory, judges the line in the information table is determined only by which conditions, sloves the $Q$ core of $R$, deletes the unnecessary condition attribute and the repeated message, obtains decision-making table which the condition attribute simplifies. The simplified decision-making table has the function before simplification, but the simplified decision-making table has the less condition attribute. The simplified decision-making table is a 'incomplete' decision-making table, it only contains condition attribute that is necessary for decision-making. The decision-making table simplification of a knowledge representation system is not the only one, namely the minimal solution of the question is not the only one, therefore may carry on the optimization to the question solution according to certain requests.

*5) The step that realizes the characteristic level fusion using the rough set theory*

- The sample information which gathers according to the condition attribute and the conclusion attribute establishes an information table, carries on the discretization to the input information.
- Utilizes the rough set theory to carry on the reduction to the primary data, uses the rough set to carry on the attribute reduction first, and realizes the attribute optimal, then carries on the object reduction, eliminates the noise (inconsistent object) and the redundant object in the sample
- Solves the core value table.
- We gain the simplified form of the information table by the core value table.
- Compiles the corresponding smallest rule, choose optimal character subset, obtains the quickest fusion algorithm.

### *D. Decision-making Level Information Fusion*

Takes the characteristic level fusion output's characteristic variable as the input, produces the training sample. Uses support vector machine(SVM) to carry on the information fusion. SVM was one kind of new machine learning algorithm that was proposed by Vapnik et al. in 1995 [12] based on the statistical learning theory at the first time. It causes the classified error of the unknown sample to be minimum through building optimal hyperplane; Maps the input data into the feature space to realize the classification and approximation of function, is the realization to the VC dimension theory and the structure risk minimum principle to statistical learning under the limited sample condition [13].

Takes the characteristic level fusion output's characteristic variable as the input, produces training sample. Uses SVM to carry on the information fusion. SVM is one kind intelligent study mechanism to be established aims at the machine learning question which specially under the small sample condition, it was one kind of new machine learning algorithm [12] that Vapnik et al. in 1995 proposed for the first time in the statistics theory of learning's foundation, did not need excessively many apriori knowledges and expert knowledge, during minimum experience risk, minimum the upper boundary of the confidence interval, thus obtained the stronger generalization, has avoided the overfitting question effectively.

SVM causes the classified error of the unknown sample to be smallest through structuring most superior planoid; Through maps the data-in into the feature space to realize the classification and approximation of function, is realizing the minimum principle of the VC-dimension theory and the structure risk to the statistic study under the limited sample condition [13]. Chooses the core function, through revising its parameter unceasingly, carries on the training study with the training sample, determines the related parameter of the accuracy error and the core function, uses the interior point algorithms to solve the regression function, guaranteed that the extremal solution is the global optimal solution; Carries on the test with the test sample, achieves the system accuracy, thus realizes the policy-making level fusion.

At present support vector machines method commonly is used in classification and regression analysis. Its essential method is representing the entire sample set with the minority support vector, through predetermined nonlinear mapping, the input vector x will map in a higher dimentional space, choose the core function which satisfies the Mercer theorem to substitute for the dot product operation of the higher dimentional feature space, carries on the linear regression in the higher dimentional feature space, thus obtains the original spatial nonlinear regression effect. The optimal function $f(x)$ which obtains through the optimization problems namely for decision function.

The core function's form mainly has linear, the polynomial, the radial basis core function (RBKF), two perceptrons and so on. The choice of the core function must satisfy the Mercer theorem. The regression measure function may use the ε-insensit function which is proposed by Vapnik.

Let the system input include n sample, each sample $X_i$ is composed of m data, produces the training sample data matrix (omitted).

$$X=\begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix} \tag{16}$$

In matrix, each row comes from the identical sensor, altogether m sensors. These m sensors gather one of the k kind of information. Therefore the training sample set may be split k training samples set. That is:

$x_{ij} \in L_T$

Where $i$ =1, 2, …, $n$ ; $j$ =1, 2, …, $m$ ; $L_T \in \{1,2,\cdots,k\}$ , is one kind of information.

Using the support vector to k group of data to find the optimal hyperplane separately, may obtain the optimized question of k quadratic programming plans, its maximized function is:

$$\omega(\alpha) = \sum_{i=1}^{n} \alpha_i - \tfrac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (17)$$

Constraint condition: $0 \le \alpha_i \le C$ (C is penalty factor) and

$$\sum_{i=1}^{n} \alpha_i y_j = 0 \quad (18)$$

As well as hyperplane coefficient vector:

$$\omega = \sum_{i=1}^{n} \alpha_i y_i x_i \quad (19)$$

Through solving type (17) ~ (19) to confirm hyperplane coefficient, obtains the optimal solution and k classified functions. Choosing a components of $\alpha$ which is bigger than 0, and its corresponding sample $(x_j, y_j)$ --- namely the support vector corresponding sample, calculates $b = y - \sum_{i=-}^{n} y_i \alpha_i K(x_i, y_i)$ according to it, so obtains the decision function is:

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i k(x_i, y_i) + b \quad (20)$$

The steps that support vector machines realize the decision-making level fusion is:

- Takes the characteristic level's output as this level input, determined the number of input $X$ , produces the training sample $\{(x_i, y_i | i = 1,2,\cdots,k)\}$, xi is the input value, yi is the predicted value.
- Chooses satisfies Mercer the theorem core function $K(x_i, y_i)$ : This paper chooses the radial basis core function $K(x_i, x_j) = e^{\frac{-\|x_i - x_j\|^2}{2\sigma^2}}$ , the width $\sigma^2$ is the same to all cores, assigns by the user.
- Determines the core function and the related parameter. Ensure the accuracy error and the related parameter $\alpha, \omega, b$ .
- Solves the decision function $f(x)$ .
- Carries on the test decision function with the test sample, if can satisfy the system required precision, uses this decision function as the multi-sensor information fusion application system model, realizes the information fusion. Otherwise, repeat (3), (4), until satisfies the system accuracy.
- For the identical test procedure, may increases system's fusion test precision through adjusting$\varepsilon$and core function $K(x_i, x_j)$ and the related parameter, increases system's fusion test precision. Regarding the linearly nonseparable case, Cortes and Vapnik have introduced the soft edge optimal hyperplane concept in 1995, through introducing the slack variable and the penalty factor to be solved.

*E. Example validation*

Selects fire detection as an object, carries on multiple source information fusion example confirmation based on the data driven.

The fire early character state is unstable and has the different manifestation, like the slow smoldering, in the fire process, follows the solid state high temperature product, the gaseous state combustion product, the flame, the combustion sound and so on. Because the fire event is very accidental, the observed data is very little, therefore the fire signal is the signal which beforehand unknown or cannot determine. The environmental variation like climate, the humidity, the dust, the electronic noise and the other artificial activities possibly cause detect the change of the signal, therefore the fire detection is one kind of non-structure question.

The main signal of the fire detection have the gas, the smog, the flame, the combustion sound, the temperature and so on, as a general rule, the content of CO is extremely low in the air, often when the burning will occur, only cause CO content in the air sharp rise, when the fire occurs, simultaneously be accompanied by the temperature and the flame increasing with the smog density, therefore using examines the CO content, temperature and smog density change completes fire's detection.

From Chinese standard open fire SH4, standard smoldering fire SHI and the typical unwanted signal curve under the kitchen environment selects the tentative data, the temperature of Chinese standard open fire SH4 and the standard smoldering fire SH1, the smog and the CO fire parameter variation along with the time, like figure 2, 3 and 4 are shown [14 -15].

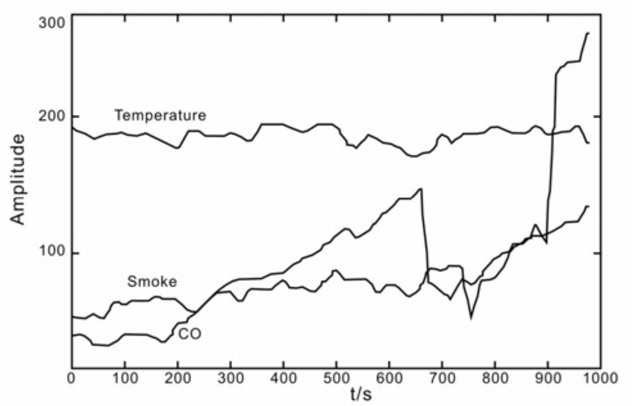



Figure 2. Standard open fire SH4 parameter change figure

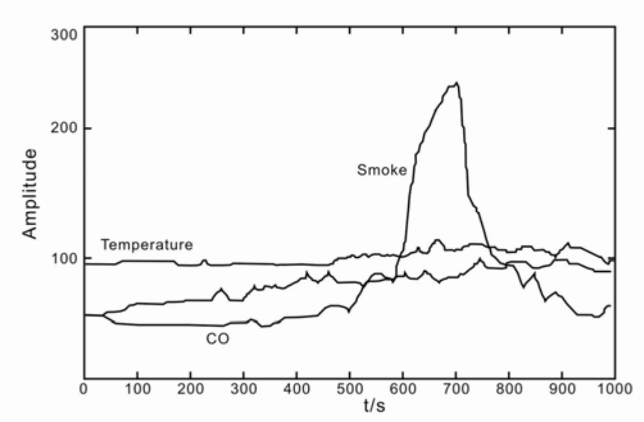



Figure 3. Standard smoldering fire SHI parameter change figure

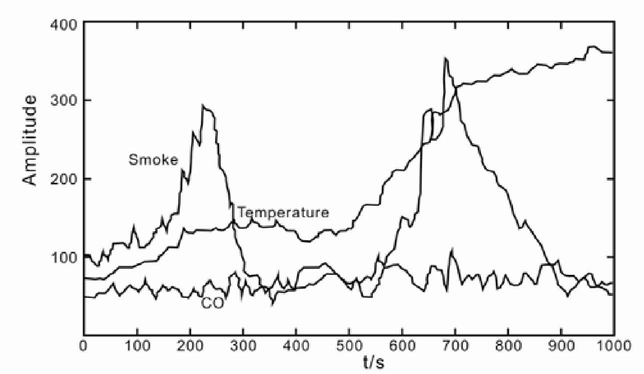



Figure 4. Typical unwanted signal curve under the kitchen environment

From figure 2, 3, 4, sampling 100 groups of data, namely 4 sensors, each sensor detects 100 data. Selects 70 groups as the system input sample, 30 groups as the test sample. Carries on the information fusion according to the multiple source information fusion step which above gives.

First carries on normalized processing to the sampled data, eliminates the metrical data dimension influence; Next uses the principal element analytic method to carry on the data level the information processing, obtains various ranks principal element variable according to the order, calculates various principal elements value which needs, assigns the principal element according to the principal element accumulation technical progress factor the number is 16, forms the data level information fusion the output sample collection; Once more, take the data level fusion output's result as this level input, uses the rough collection theory, carries on the discretization to the infed information, establishes the information table, carries on the attribute reduction, the object reduction, calculates the nuclear value table, extracts the information table the simplified form, the extraction most superior character subset, forms the characteristic level information fusion the output sample collection; Finally, takes by the characteristic level fusion output's character subset this level inputs, produces training sample, uses the support vector machines law to carry on the policy-making level information fusion, choose radial direction base nuclear function, determines the accuracy error and the related parameter $a$ , $\omega$ , $b$ , compute decision function $f(x)$ , through uses Matlab SVM the Toolbox training simulation, obtains the fusion output result finally. Tests the sample with 30 groups to carry on the examination to the fusion result, reports mistakenly 1 group, the accuracy reaches 96.7%.

## VI. CONCLUSION

This paper take data-driven method as theoretical principle, uses the principal component analytic method, the rough set theory, the support vector machine and so on many kinds of method unions and supplementary, realizes th Multi-source information fusions from the data level, the characteristic level, the decision-making level three levels. May carries on the confirmation through the system simulation and the actual system fusion. Has provided the new way for the establishing multiple source information fusion basic theory frame and the generalized fusion algorithm.

The Multi-source information fusion technology based on the data driven method take full advantage of the online and off-line data of the fusion system, does not rely on system's mathematical model, has avoided the system mechanism modeling question. Fully using redundancy of the fusion system, supplementary and synergistic multiple source information, reduces the multi-sensor data influence on the environment condition and the itself characteristic, reduced the computation complexity, reduced in the fusion process information loss, rejected characteristic data redundant ingredient, reduced the data association ambiguity, enhanced the fusion accuracy.

The multiple source information fusion technology based on data driven has fault tolerance, complementarity, timeliness characteristics and so on, can increase the dimension and the confidence of the survey, expand space and time coverage area, improve system reliability and maintainability, achieve in the system sharing complementary advantages, the resource sharing, to raise the resources utilization. Can reduce the system total uncertainty, increases the precision, reduces ambiguity to the object model, enhances the system correct decision-making ability, compensate sole information source inaccuracy and the measuring range limitation, obviously enhances the speed of the information processing, reduces the cost of the information acquisition.

## ACKNOWLEDGMENT

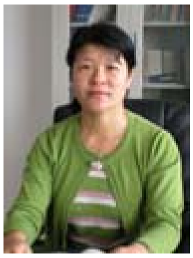
This work is supported by NSF ZR2010FL021 of Shandong Province.

## REFERENCES

[1] Waltz E, Llinas J.Multisensor Data Fusion[M]. Boston: Artech House,1990.

[2] KAM M,RORRES C,CHANG W,et al. Performance and geometric interpretation for decision fusion with memory[J].IEEE Transactions on system, man, cybernetic,1999,29(1):52~62.

[3] HE You, GUAN Xin, WANG Guo-hong. Survey on the Progress and Prospect of Multisensor Information Fusion[J]. Journal of Astronautics,2005,26(4):524-529
[4] Jian Xiaogang, Jia Hongsheng, Shi Laide. Advances on multi-sensor information fusion technologies[J]. Chinese Journal of Construction Machinery,2009,7(2):227-232
[5] Chai Tianyou, Liu Derong. control. Decision-making, dispatcher and failure diagnosis Based on data [J]. Acta Automatica Sinica, 2009,35(6): Ⅱ-Ⅲ
[6] HOU Zhong-Sheng, XU Jian-Xin. On Data-driven Control Theory: the State of the Art and Perspective[J].Acta Automatica Sinica,2009,35(6): 650-667.
[7] GAO Chuan-Hou, JIAN Ling, CHEN Ji-Ming, SUN You-Xian. Data-driven Modeling and Predictive Algorithm for Complex Blast Furnace Ironmaking Process[J], Acta Automatica Sinica, 2009,35(6): 721-730
[8] J E Jackson. A user's guide to principal components[M]. New York:Wiley-Interscience,1991.
[9] Wen Cheng-Lin, Hu Jing, Wang Tian-Zhen, Chen Zhi-Guo. Relative PCA with applications of data compression and fault diagnosis. Acta Automatica Sinica,2008,34(9):1128-1139.
[10] Pawlak Z. Rough sets [J]. International Journalof Information and ComputerScience, 1982, 11(5): 341~356.
[11] Pawlak Z. Rough Sets Theoretical Aspectsof Reasoning about Data[M]. Dordrecht: Kluwer Academic Publishers, 1991. 68~162.
[12] Vapnik V. The nature of statistical learning theory[M]. New York:Springer-Verlag, 1995.
[13] ZHAO Chun-ling, HUANG Kun, CHEN Sen-fa. Research on an Optimal Fusion Method of Multi-source Information[J].Information and Control, 2005,34(1): 66-70.
[14] People's Republic of China national standards. Point type thermal fire detector specification and experimental technique (GB4716-84). Beijing: Chinese Standard Publishing house, 1993, 69-47.
[15] Wu Longbiao, Yuan Hongyong. Fire detection and control engineering. Hefei: University of Science and Technology of China Publishing house, 1999, 20-25



**Xin Zhang**, professor, in July, 1985, graduated from the Harbin University of Science and Technology and obtained the bachelor's degree; in April, 1988 graduates from the Harbin university of science and technology, and obtains the technology master's degree.

At present she works in Shandong Institute of Business and Technology, School of Information and Electronic Engineering, YanTai, China. She is engaged in for a long time in the embedded system application, intelligent control and intelligent system, information fusion, and published more than 20 related papers.