1

State Optimal Estimation for Nonstandard Multi-sensor Information Fusion System

Jiongqi Wang and Haiyin Zhou
National University of Defense Technology
China

1. Introduction

In the field of information fusion, state estimation is necessary[1-3]. The traditional state estimation is a process to use statistics principle to estimate the target dynamical (or static) state by using of measuring information including error from single measure system. However, single measure system can't give enough information to satisfy the system requirement for target control, and is bad for the precision and solidity of state estimation. Therefore, developing and researching information fusion estimation theory and method is the only way to obtain state estimation with high precision and solidity.
The traditional estimation method for target state (parameter) can be traced back to the age of Gauss; in 1975, Gauss presented least square estimation (LSE), which is then applied in orbit determination for space target. In the end of 1950s, Kalman presented a linear filter method, which is widely applied in target state estimation and can be taken as the recursion of LSE[4]. At present, these two algorithms are the common algorithms in multi-sensor state fusion estimation, which are respectively called as batch processing fusion algorithm and sequential fusion algorithm.
The classical LSE is unbiased, consistent and effective as well as simple algorithm and easy operation when being applied in standard multi sensor information fusion system (which is the character with linear system state equation and measuring equation, uncorrelated plus noise with zero mean)[5]. However, because of the difference of measuring principle and character of sensor as well as measuring environment, in actual application, some non-standard multi-sensor information fusion systems are often required to be treated, which mainly are as follows:

1) Each system error, mixing error and random disturbed factor as well as each nonlinear factor, uncertain factor (color noise) existing in multi sensor measuring information[6];

2) Uncertain and nonlinear factors existing in multi sensor fusion system model, which is expressed in two aspects: one is much stronger sense, uncertain and nonlinear factors in model structure and another is time-change and uncertain factor in model parameter[7];

3) Relativity between system noise and measuring noise in dynamical system or relativity among sub-filter estimation as well as uncertain for system parameter and unknown covariance information[8-9].

Ignoring the above situations, the optimal estimation results cannot be obtained if still using the traditional batch processing fusion algorithm or sequential fusion algorithm. So to research the optimal fusion estimation algorithm for non standard multi-sensor system with the above situations is very essential and significative[10].

In the next three sections, the research work in this chapter focuses on non-standard multi-sensor information fusion system respectively with nonlinear, uncertain and correlated factor in actual multi-sensor system and then presents the corresponding resolution methods.

Firstly, the modeling method based on semi-parameter modeling is researched to solve state fusion estimation in nonstandard multi-sensor fusion system to eliminate and solve the nonlinear mixing error and uncertain factor existing in multi-sensor information and moreover to realize the optimal fusion estimation for the state.

Secondly, a multi-model fusion estimation methods respectively based on multi-model adaptive estimation and interacting multiple model fusion are researched to deal with nonlinear and time-change factors existing in multi-sensor fusion system and moreover to realize the optimal fusion estimation for the state.

Thirdly, self-adaptive optimal fusion estimation for non-standard multi-sensor dynamical system is researched. Self-adaptive fusion estimation strategy is introduced to solve local dependency and system parameter uncertainty existed in multi-sensor dynamical system and moreover to realize the optimal fusion estimation for the state.

2. Information Fusion Estimation of Nonstandard Multisensor Based on Semi parametric Modeling

From the perspective of parameter modeling, any system models generally consist of two parts: deterministic model (It means that the physical model and the corresponding parameters are determined) and non-deterministic model (It means that the physical models are determined but some parameter uncertainty, or physical models and parameters are not fully identified). In general case, the practical problems of information fusion can be described approximately by means of parametric modeling, then to establish the compact convergence of information processing model. Namely, the part of the systematic error of measurement can be deduced or weaken through the establishment of the classic parametric regression model, but it cannot inhibit mixed errors not caused by parametric modeling and uncertainty errors and other factors. Strictly speaking, the data-processing method of classical parametric regression cannot fundamentally solve the problem of uncertainty factors[11]. Yet it is precisely multi-sensor measurement information in the mixed errors and uncertainties that have a direct impact on the accuracy indicated by the model of multi-sensor fusion system, then in turn will affect the state estimation accuracy to be estimated and computational efficiency. So, it is one of the most important parts to research and resolve such error factors of uncertainty, and to establish a reasonable estimation method under the state fusion estimation.

As for this problem, there are a large number of studies to obtain good results at present. For instance, systematic error parameter model suitable for the engineering background is established to deal with the system error in measurement information. Extended-dimensional vector is employed to directly turn systematic error into the problem of the state fusion estimation under the standard form[12]. However, due to the increase of the number of parameters to be estimated, the treatment not only lowered the integration of estimation accuracy, but also increased the complexity of the computation of the matrix inversion. In addition, robust estimation theory and its research are designed to the problem of the incomplete computing of the abnormal value and the condition of systems affected by the large deviation[13]. A first order Gauss - Markov process is used to analyze and handle the random noise in measurement information. However, most of these treatments and researches are based on artificial experience and strong hypothesis, which are sometimes so contrary to the actual situation that they can doubt the feasibility and credibility of the state fusion estimation.

The main reason for the failure of the solution of the above-mentioned problems is that there is no suitable uncertainty modeling method or a suitable mathematical model to describe the non-linear mixed-error factors in the multi-sensor measurement information[14].

Parts of the linear model (or called) semi-parameter model can be used as a suitable mathematical model to describe the non-linear mixed-error factors in the measurement information [15]. Semi-parametric model have both parametric and non-parametric components. Its advantages are that it focused on the main part of (i.e. the parameter component) the information but without neglecting the role of the interference terms (non-parametric component). Semi-parametric model is a set of tools for solving practical problems with a broad application prospects. On the one hand, it solves problems which are difficult for only parameter model or non-parametric model alone to solve, thus enhancing the adaptability of the model; on the other, it overcome the issue of excessive loss of information by the non-parametric method and describe practical problems closer to the real and made fuller use of the information provided by data to eliminate or weaken the impact of the state fusion estimation accuracy caused by non-linear factors more effectively.

This section attempts to introduce the idea of semi-parametric modeling into the fusion state estimation theory of the non-standard multi-sensor. It establishes non-standard multi-sensor fusion state estimation model based on semi-parametric regression and its corresponding parameters and non-parametric algorithm. At the same time of determining the unknown parameters, it can also distinguish between nonlinear factors and uncertainties or between system error and accidental error so as to enhance the state fusion estimation accuracy.

2.1 State Fusion Estimation Based on Mutual Iteration Semi-parametric Regression

In terms of the optimal state fusion estimation of the multi-sensor fusion system integration, its main jobs are to determine the "measurement information" and the state of mapping relationship to be assessed, to reveal statistical characteristics of measurement errors, and then to reach the result to be consistent with the optimal state fusion of the project scene. The mapping matrix is determined by specific engineering and the model established by the physical background, having a clear expression generally. Therefore, the core task of the multi-sensor consists in the statistical characteristics of the measurement error analysis. But in practice, the differences of sensor measuring principle and its properties often touch upon the existence of the observing system error and the non-standard multi-sensor data fusion system under the influence of nonlinear uncertain elements. Among them, the errors in constant-value system or parameterized system are rather special but conventional system error. For these systems, it is easy to deal with[12]. But in fact, some systematic errors, non-linear uncertainties in particular, which occur in the multi-sensor information fusion

system, are difficult to be completely expressed by parameters. In the first place, there are many factors which effect the value-taken of nonlinearity but all of these cannot be considered when establishing mathematical models. Secondly, some relative simple functional relations are chosen to substitute for functional relation between those factors and their parameters, so the established functional model are often said to be the approximate expression of the practical problems, that is to say, there is the existence of the model representation for error. When the error value of the model is a small amount, there is nothing much influence on the result of the assessment of the general state of this system if omitting it. But when the error value of the model is comparatively large, the neglect of it will exert a strong influence and lead to the wrong conclusion. Therefore, we main focused on the refinement issues of the state fusion estimation model under the condition of the non-linear uncertain factors (those non-linear uncertainties which are not parameterized fully), introducing semi-parametric regression analysis to establish non-standard multi-sensor information fusion estimation theory based on semi-parametric regression and its corresponding fusion estimation algorithm.
(1) Semi-parametric Regression Model
Assuming a unified model of linear integration of standard multi-sensor fusion system is:

$$\boldsymbol{Y}^N = \boldsymbol{HX} + v^N$$

Where, Y^N is the observation vector, X the state vector of the fusion to be estimated, v^N observation error, H the mapping matrix between metrical information and the state fusion to be estimated. In this model, v^N is supposed to be white noise of the zero mean. That is to say, except observation error, the observation vector Y^N is completely used as the function of status to be assessed. However, if the model is not accurate, with nonlinear uncertainties, the above formula cannot be strictly established and should be replaced by:

$$\boldsymbol{Y}^N = \boldsymbol{H}^N \boldsymbol{X} + \boldsymbol{S}^N + v^N \tag{2.1}$$

Where, $S^N(t)$ is the amount of model error which describes an unknown function relationship, it is the function of a certain variables t.
Currently, there are three methods for using semi-parametric model to estimate the error with nonlinear factor model in theory, including the estimation of part of the linear model of approximation parameterization, the estimation of part of the linear model of regularization matrix compensation, and part of the two-stage linear model estimation[16]. But the process of its solution implies that the algorithm realization is comparative complex, and that the accuracy of estimation depends on the cognition of characteristics of non-parametric component as well as the choice of basis function. Taking the estimation of part of the linear model of regularization matrix compensation for instance, the programming of key factors like regular matrix and smoothing factor are highly hypothetical, including some elements presumed in advance, furthermore, the solution process is very complex. If there is something error or something that cannot meet the model requirement in the solution of smoothing factor α and regularization matrix R_s, it will directly lead to unsolvable result to the semi-parametric fusion model. Here, we propose an algorithm based on the state fusion estimation of mutual-iteration semi-parametric regression, by the compensation for the error of the non-standard multi-sensor fusion model and the spectrum feature analysis to non-linear uncertainties, through aliasing frequency estimation method of

decoupling to define the best fitting model, thus establishing the algorithm between the model compensation for the state fusion estimation model and the state fusion estimation of mutual iteration semi-parametric regression, isolating from non-linear uncertainties and eliminating the influence on its accuracy of the state fusion estimation.

(2) The basis function of nonlinear uncertainties is expressed as a method for decoupling parameter estimation of the aliasing frequency.

According to the signal processing theory, in the actual data processing, model errors and random errors under the influence of the real signal, non-linear uncertainties are often at different frequency ranges. Frequency components which are included in the error of measurement model are higher than the true signal frequency, but lower than random errors, so it can be called sub-low-frequency error[17-18]. It is difficult for classical least squares estimation method to distinguish between non-linear model error and random errors. However, the error characteristics of the measurement model can be extracted from the residual error in multi-sensor observation. Namely, it is possible to improve the state estimation accuracy if model error of residual error (caused mainly by the non-linear uncertainties) can be separated from random noise and the impact of model error deducted in each process of iterative solution.

On consideration that nonlinear factors S in semi-parametric model can be fitted as the following polynomial modulation function forms:

$$S(t) = \sum_{m=0}^{M-1} \left(\sum_{i=0}^{N_m - 1} a_i^{(m)} t^i \right) \cdot \exp\{ j2\pi f_m t \} \overset{\text{def}}{=} \sum_{m=0}^{M-1} b_m(t) \cdot \exp\{ j2\pi f_m t \} \tag{2.2}$$

Where, f_m is the frequency item of non-linear uncertainties, $b_m(t)$ the amplitude envelope of each component signal, $a_k^{(m)}$ polynomial coefficients corresponding to envelope function. From Equation (2.2), $S(t)$ is a multi-component amplitude-harmonic signal. It is complicated to directly use maximum likelihood estimation method to distinguish the frequency parameters of various components and amplitude parameter but apply the combination of matching pursuit and the ESPRIT method of basis pursuit to decouple parameter estimation.

Firstly, recording $y_0(t) = S(t)$, the method of ESPRIT [19] to estimate the characteristic roots closest to the unit circle from $y_0(t)$ is used to estimate frequency of corresponding harmonic components $\hat{\lambda}$. Without loss of generality, if the estimation corresponded to f_0, that is, $\hat{f}_0 = (1/2\pi) \cdot \mathrm{Arg}\{\hat{\lambda}\}$, according to this, the original signal frequency is shifted to frequency to get

$$\tilde{y}_0(t) = y_0(t) \cdot \exp\{ -j2\pi \hat{f}_0 t \} \tag{2.3}$$

The baseband signal is obtained from low pass and filter of the shifted signal $\tilde{y}_0(t)$. Namely, it can be used as an estimate of amplitude envelope $b_0(t)$.

Noting: $\hat{b}_0(t) = \mathrm{LPF}[\tilde{y}_0(t)]$, among them, $\mathrm{LPF}[\cdot]$ refers to low-pass filter. The observation model of amplitude envelope is deduced from Formula (2.2):

$$\hat{b}_0(t) = \sum_{i=0}^{N_0-1} a_i^{(0)} t^i + \varepsilon(t) \tag{2.4}$$

The corresponding coefficient $\hat{a}_i^{(0)}$ is estimated by Least Square, which is also used to reconstruct the corresponding signal components.

$$\overline{b}_0(t) = \sum_{i=0}^{N_0-1} \hat{a}_i^{(0)} t^i \cdot \exp\{j2\pi\hat{f}_0 t\}$$

To move forward a step, the reconstruction of the harmonic component of amplitude modulation is subtracted from $y_0(t)$, then we can obtain residual signal:

$$y_1(t) = y_0(t) - \overline{b}_0(t) \tag{2.5}$$

The residual signal is used as a new observing signal to repeat the above processes to get parameter estimates of multi-component signals, that is $\{\hat{f}_k, \hat{a}_i^{(k)}\}$, $i = 0,1,\cdots,N_k - 1,\ k = 0,1,\cdots,M-1$. The stop condition of iterative algorithm can be represented as residual control criterion and the order selection of other models.

(3) Steps of how to calculate mutual iterative state estimation

By means of the basis function to nonlinear uncertainties and the estimation method of decoupling parameter of corresponding aliasing frequency, nonlinear uncertainties can be extracted by fitting method, establishing multi-sensor fusion system model. The optimal fusion estimate of the state $\boldsymbol{X}$ to be estimated can be determined by the mutual iteration method of the following linear and nonlinear factors. If the degree of the Monte-Carlo simulation test is L, the implementation algorithm will be as following.

Step1: For the obtaining multi-sensor fusion system, least squares estimation fusion is used to get $\boldsymbol{X}_j$ in the known observation sequence $Y_{1j}, Y_{2j}, \cdots, Y_{Nj}$ $(j = 1,2,\cdots,L)$;

Setp2: Computing observation residuals $Y'_{1j}, Y'_{2j}, \cdots, Y'_{Nj}$ in multi-sensor fusion system;

Setp3: Examining whether the observation residual family $\{Y'_{i1}, Y'_{i2}, \cdots, Y'_{iL} \mid i = 1,2,\cdots,N\}$ is white noise series, if it is, turn to Step5, if not, turning to Step4;

Step4: With the method for aliasing frequency estimation, nonlinear uncertainties vector $\boldsymbol{S}^N = \{S_1, S_2, \cdots, S_N\}$ can be determined. That is to say, S_i should satisfy the following conditions:

$$\begin{cases} Y'_{i1} = S_i + v_{i1} \\ Y'_{i2} = S_i + v_{i2} \\ \cdots\cdots\cdots \\ Y'_{iL} = S_i + v_{iL} \end{cases}, i = 1, 2, \cdots, N$$

Where, the white noise series is $v_{i1}, v_{i2}, \cdots, v_{iL}$, $Y_{1j}, Y_{2j}, \cdots, Y_{Nj}$ is replaced by $Y_{1j} - S_1, Y_{2j} - S_j, \cdots, Y_{Nj} - S_N$, and then turn to Step1.

Step5: The exported value of $\hat{\boldsymbol{X}} = \frac{1}{L}\sum_{j=1}^{L}\boldsymbol{X}_j$ is the optimal fusion estimate of the state to be estimated.

The above algorithm is all dependent on iteration. We will keep estimating and fitting the value of the nonlinear uncertainty vector S^N . Simultaneously, it is also a process of being close to the true value of a state to be assessed. When approaching the true state values, observation residuals equaled to Gaussian white noise series. This method is in essence an improvement to iterative least squares estimation of Gauss-Newton by the use of two-layer iterative correction. Step4 (a process of fitting nonlinear uncertainties) is a critical part.

2.2 Analysis of Fusion Accuracy

Two theorems will be given in the following. Comparing to the classical least square algorithm, the state fusion estimation accuracy based on mutual iteration semi-parametric regression will be analyzed in theory to draw a corresponding conclusion.

Theorem 2.2: On the condition of nonlinear uncertain factors, the estimation for $\hat{\boldsymbol{X}}$ is $\hat{\boldsymbol{X}}_{\mathrm{BCS}}$ which is called unbiased estimation, and $\hat{\boldsymbol{X}}$ is obtained from the state fusion estimation based on mutual iteration semi-parametric regression, while with classical weighted least squares, the estimate value $\hat{\boldsymbol{X}}_{\mathrm{WLSE}}$ is biased estimate.

Demonstration: Under the influence of factors of the nonlinear uncertain error, the state fusion estimation based on semi-parametric regression from the generalized unified fusion model (2.1) is deduced as:

$$\hat{\boldsymbol{X}}_{\mathrm{BCS}} = \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{Y} - \hat{\boldsymbol{S}}) \tag{2.6}$$

And $\hat{\boldsymbol{S}}$ is function fitted values of nonlinear uncertain error vector, then its expectation is:

$$\mathrm{E}[\hat{\boldsymbol{X}}_{\mathrm{BCS}}] = \mathrm{E}[(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{Y} - \hat{\boldsymbol{S}})] = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X} = \boldsymbol{X} \tag{2.7}$$

$\hat{\boldsymbol{X}}_{\mathrm{BCS}}$ is the unbiased estimation of $\boldsymbol{X}$. The estimated value $\hat{\boldsymbol{X}}_{\mathrm{WLSE}}$ is computed by the method of weighted least squares estimation fusion. That is:

$$\hat{\boldsymbol{X}}_{\mathrm{WLSE}} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{Y} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{H}\boldsymbol{X} + \hat{\boldsymbol{S}}) \tag{2.8}$$

Its expectation is:

$$\mathrm{E}[\hat{\boldsymbol{X}}_{\mathrm{WLSE}}]=\mathrm{E}[(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{HX}+\hat{\boldsymbol{S}})]=\boldsymbol{X}+(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\hat{\boldsymbol{S}} \quad (2.9)$$

The following relationship formula is from Formula (2.6) and (2.8):

$$\hat{\boldsymbol{X}}_{\mathrm{WLSE}}=\hat{\boldsymbol{X}}_{\mathrm{BCS}}+(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\hat{\boldsymbol{S}} \quad (2.10)$$

Theorem 2.3: On the condition of nonlinear error factors, the valuation accuracy of $\hat{\boldsymbol{X}}$ which is based on the state fusion estimation of the mutual iteration semi-parametric regression ranked above the valuation accuracy which is based on the method of weighted least squares estimation fusion.

Demonstration: The estimation accuracy of semi-parametric state fusion is supposed to be $\mathrm{Cov}[\hat{\boldsymbol{X}}_{\mathrm{BCS}}]$, so:

$$\mathrm{Cov}[\hat{\boldsymbol{X}}_{\mathrm{BCS}}]=\mathrm{E}[(\hat{\boldsymbol{X}}_{\mathrm{BCS}}-\boldsymbol{X})(\hat{\boldsymbol{X}}_{\mathrm{BCS}}-\boldsymbol{X})^{\mathrm{T}}]=(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1} \quad (2.11)$$

However, the valuation accuracy $\mathrm{Cov}[\hat{\boldsymbol{X}}_{\mathrm{WLSE}}]$ obtained by the method of weighted least squares estimation fusion.

$$\begin{aligned}\mathrm{Cov}[\hat{\boldsymbol{X}}_{\mathrm{WLSE}}]&=\mathrm{E}[(\hat{\boldsymbol{X}}_{\mathrm{WLSE}}-\boldsymbol{X})(\hat{\boldsymbol{X}}_{\mathrm{WLSE}}-\boldsymbol{X})^{\mathrm{T}}]\\&=\mathrm{E}[(\hat{\boldsymbol{X}}_{\mathrm{BCS}}+\boldsymbol{P}-\boldsymbol{X})(\hat{\boldsymbol{X}}_{\mathrm{BCS}}+\boldsymbol{P}-\boldsymbol{X})^{\mathrm{T}}]=(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}+\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\end{aligned} \quad (2.12)$$

Where, $\boldsymbol{P}=(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\hat{\boldsymbol{S}}$, obviously $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}>0$, the estimation accuracy of $\hat{\boldsymbol{X}}$ based on the state fusion estimation of the mutual iteration semi-parametric regression is superior to the estimation accuracy obtained by the method of weighted least squares estimation fusion.

2.3 Numerical Examples and Analysis

In order to verify these conclusions, the experiment of numerical simulation is conducted on the basis of the method of the state fusion estimation of the mutual iteration semi-parametric.

On consideration of the state fusion estimation of the constant-value $x=10$, the fusion system, which consists of three sensors, is used to the conduction of the state fusion estimation. The measurement equation of non-standard multi-sensor fusion is: $y_i=x+b_i+v_i$, $i=1,2,3$; where v_i noise which is zero mean, and each variance is the Gaussian-noise of $R_1=1, R_2=2$ and $R_3=3$. Simultaneously, non-linear error component $b_i(i=1,2,3)$ is something related to cycle colored noise of the number of Monte-Carlo simulation L, each amplitude is $b_1=0.5, b_2=1$ and $b_3=1.5$. The simulation times $L=100$. The estimate values and estimated variance of the state to be estimated are obtained from the method of the classical least squares estimation and the state fusion estimation of the mutual iteration semi-parametric given in the Table 2.1. Comparing the simulation results by the two methods, the fusion estimation accuracy is relatively low by the use of the least squares due to the influence of nonlinearity error. And it can also be predicted that with the increase of non-linear error factors, the estimation accuracy is bound to reduce more and more. But the method for the state fusion estimation of the mutual iteration semi-parametric can separate white noise series in observation noise from non-linear error factors, canceling its influence to state fusion estimation accuracy by fitting estimates. If there is nonlinearity error, the state estimator, which is obtained by the method for the state fusion estimation of the iteration semi-parametric, will be the optimal estimation of true value.

Fusion Algorithm	State Fusion Estimation	Fusion Estimation Variance
Method of Weighted Least Squares Estimation Fusion.	11.084	1.957
Method of State Fusion Estimation of Iteration Semi-parametric	10.339	1.434

Table 2.1. Comparison of Estimation Result between the two Fusion Algorithms

3. Nonstandard Multisensor Information Fusion Estimate Based on Multi-model Fusion

In recent years, it becomes a hot research topic to establish the parametric / semi-parametric model in the control of a complex nonlinear model, which has also been a great application, but there are so few tactics which are used in actual projects. The main reason for this problem is due to the difficulties of establishing an accurate model for complex non-linear parameters and the uncertainty of the actual system to a degree. These uncertainties sometimes are performed within the system, sometimes manifests in the system outside. The designer can not exactly describe the structure and parameters of the mathematical model of the controlled object in advance within the system. As the influence to the system from external environment, it can be equivalent to be expressed by many disturbances, which are unpredictable but might be deterministic or even random. Furthermore, some other measurement noise logged in the system from the feedback loop of the different measurement, and these random disturbances and noise statistics are always unknown. In this case, for dynamic parameters of the model which is from doing experiments on the process of parametric modeling, it is hard for the accuracy and adaptability expressed by the test model, which is even a known model structure, to estimate parameters and their status in the real-time constraints conditions.

Multi-model fusion processing is a common method for dealing with a complex nonlinear system[20-21], using multi-model to approach dynamic performance of the system, completing real-time adjustment to model parameter and noise parameter which is related to the system, and programming multiple model estimator based on multiple model. This estimator avoided the complexity of the direct model due to the reason that it can achieve better estimation to its accuracy, complex tracking speed and stability. Compared with the single model algorithm, multi-model fusion has the following advantages: it can refine the modeling by appropriate expansion model; it can improve the transient effect effectively; the estimation will be the optimal one in the sense of mean square error after assumptions are met; the algorithm with parallel structure will be conducive to parallel computing.

Obtaining the state optimal fusion estimate is the processing of using multi-model to approach dynamic performance of the system at first, then realizing the disposure of multi-model multi-sensor fusion to the controlled object tracking measurement, This is the problem of the multi-model fusion estimation in essence[22]. The basic idea of it is to map the uncertainty of the parameter space (or model space) to model set. Based on each model parallel estimator, the state estimation of the system is the optimal fusion of estimation obtained by each estimator corresponding to the model. As it is very difficult to analyze this system, one of these methods is to use a linear stochastic control system to denote the original nonlinear system approximately and to employ the treatment of thinking linear regression model to solve the nonlinear problem which should be solved by uncertain control systems [23]. The fusion approach diagram is shown in Fig. 3.1.

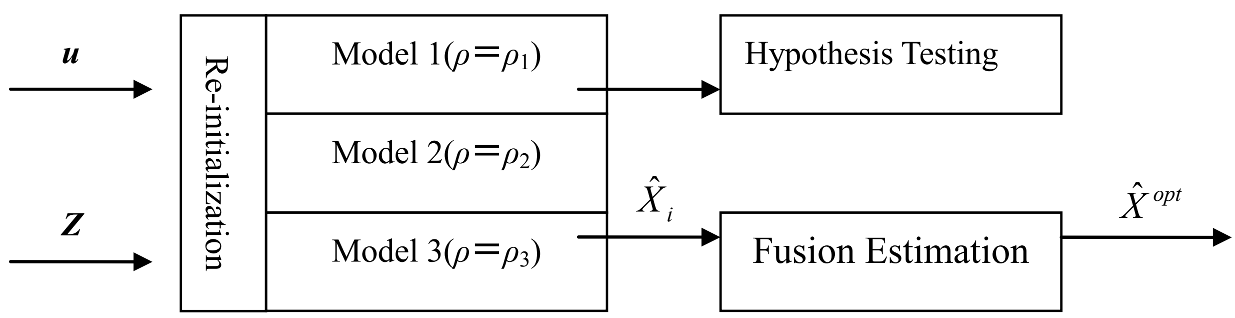


Fig. 3.1. Multi-model Fusion Estimation Approach Principle

Where, since different operational modes of stochastic uncertain systems worked with a group of parallel estimator, the input of each estimator will be the control input u and metrical information Z in a system, while the output of each estimator will be each one based on output residuals and state estimation X_i in a single model. According to the residual information, a hypothesis testing principle is used for programming model weight of an estimator corresponding to each model to reflect the situation that the probability of a model-taken at the determining time in a system. And the overall system state estimation is the weighted average value of the state estimation of each estimator.

3.1 Basic Principles of Multi-model Fusion

The description of multi-model fusion problem can be summed up: if the mathematical model of the object and the disturbance cannot be fully determined, multiple models will be designed as control sequence to approach the process of complex nonlinear time-varying in a system so as to make specified performance approaching as much as possible and keep it best.

The following nonlinear system will be given:

$$\begin{cases} \boldsymbol{X}(k+1) = F(\boldsymbol{X}(k), \boldsymbol{\theta}(k)) \\ \boldsymbol{Z}(k) = G(\boldsymbol{X}(k), \boldsymbol{\theta}(k)) \end{cases} \tag{3.1}$$

Where, $X(k) \in R^n$ is supposed to be the system state vector, $Z(k) \in R^m$ being the system output vector, F, G being nonlinear functions, $\theta(k)$ being the vector of uncertain parameters.

(1) Model Design

Without loss of generality, the system output space is supposed to be Υ, then some outputs $Z_1 \le \cdots \le Z_N$ can be chosen from Υ and get a corresponding equilibrium $(X_i, \theta_i, Z_i), i = 1, \cdots, N$. The linearization expansion of the system at each equilibrium point can get some linear model $\sum_i$ from the original nonlinear system, and they constituted linear multi-model representation of the original system. Now the parameter $\theta \in \{\theta_1, \theta_2, \cdots, \theta_N\}$ can choose some discrete values. Thus the following model set can be obtained:

$$\Omega = \{M_i \mid i = 1, 2, \cdots, N\} \tag{3.2}$$

Where, M_i is related to the parameter θ. In a broad sense, M_i can express plant model and also feedback matrix of different states and the different local area where the error fall on. Also defined a collection of design-based estimator Ω:

$$\mathrm{E} = \{E_i \mid i = 1, 2, \cdots, N\} \tag{3.3}$$

Where, E_i is supposed to be designed based estimator M_i.

Based on the above analysis, the linear multi-model of the nonlinear systems (3.1) can be described as follows:

$$\begin{cases} \boldsymbol{X}(k+1) = \boldsymbol{\Phi}(\theta_i, k)\boldsymbol{X}(k) + \boldsymbol{C}(\theta_i, k)\boldsymbol{u}(k) + \boldsymbol{\Gamma}(\theta_i, k)w(k) \\ \boldsymbol{Z}_i(k) = \boldsymbol{H}(\theta_i, k)\boldsymbol{X}(k) + v_i(k) \end{cases} \quad i = 1, 2, \cdots, N \tag{3.4}$$

Where, $\boldsymbol{\Phi}(\boldsymbol{\theta}, k), \boldsymbol{C}(\boldsymbol{\theta}, k), \boldsymbol{\Gamma}(\boldsymbol{\theta}, k)$ are the system matrixes, $\boldsymbol{u}(k)$ being the control vector of the system, $\boldsymbol{H}(\boldsymbol{\theta}, k)$ being the mapping matrix, $w(k)$ being the n dimensional system noise sequence, $v(k)$ being the m dimensional system noise sequence. The meanings of other symbol are the same as those in Equation (3.1). Here, the multi-model fusion refers to use some linear stochastic control systems given in Equation (3.4) to solve nonlinear problems in Equation (3.1).

(2) Selection of Estimator

This is the second most important aspect, namely, choosing some estimators that can reasonably describe nonlinear systems to complete the process of the state fusion estimation.

(3) Rules and Model Fusion

In order to generate the global optimal fusion estimation, fusion rules can be fallen into three patterns:

1) Soft Decision or No Decision: At any k moment, global estimates are obtained from the estimation $\hat{\boldsymbol{X}}_{ik}$ $(i = 1, 2, \cdots, N)$ based on all estimators instead of the mandatory use of the estimator to estimate the value. It is claimed to be the mainstream multi-model fusion

method. If the conditional mean of the system state is considered as estimation, global estimates will be the sum of the probability weighted of estimated value of all estimators. That is:

$$\hat{\boldsymbol{X}}_{k|k} = \mathrm{E}\left(\boldsymbol{X}_k \mid \boldsymbol{Z}^k\right) = \sum_{i=1}^{N} \hat{\boldsymbol{X}}_{ik} P(M_{ik} \mid \boldsymbol{Z}^k) \tag{3.5}$$

2) Hard Decision: The approximation of the obtained global estimates is always from the estimated value of some estimators. The principle of selection of these estimators is the model maximum possible matching with the current model and the final state estimation will be mandatory. If only one model is to be selected in all models by maximum probability, consider its estimated value as the global one.
3) Random Decision: Global estimates are determined approximately based on some of the randomly selected sequence of the estimated model. The first fusion mode is the main method of multi-model fusion estimation. With the approximation of the nonlinear system and the improvement for system fault tolerance, the tendency of the multi-model fusion will be: estimating designing real-time adaptive weighted factor and realizing the adaptive control between models.
In reality, according to different model structures and fusion methods, multi-model fusion algorithm can be divided into two categories: (1) fixed multi-model (FMM); (2) interacting multiple model (IMM)[24-25]. The latter is designed for overcoming the shortcomings of the former. It can expand the system to the new mode without changing the structure of the system, but requires some prior information of a probability density function and the condition that the switching between the various models should satisfy the Markov process.
Related closely to the fixed structure MM algorithms, there is a virtually ignored question: the performance of MM Estimator is heavily dependent on the use of the model set. There is a dilemma here: more models should be increased to improve the estimation accuracy, but the use of too many models will not only increase computation, but reduce the estimator's performance.
There are two ways out of this dilemma: 1) Designing a better model set (But so far the available theoretical results are still very limited); 2) using the variable model set.
It will be discussed Multi-model Adaptive Estimation (MMAE) and Interactive Multiple Model in Multi-model estimation method in a later paper.

3.2 Multi-model Adaptive Estimation

(1) The Fusion Architecture in MMAE
Multiple model adaptive estimators consisted of a parallel Kalman filter bank and hypothesis testing algorithm. Each library has a special filter system model, the independent vector parameters ($a_i, i = 1, 2, \cdots, N$) are used to describe its inherent Kalman filter model. Each Kalman filter model formed the current system state estimation $\hat{X}_i$ according to the independent unit under its own model and input vector, then using the estimate of the formation of the predictive value of the measurement vector, considering the residual error obtained by subtracting this value to the actual measurement vector Z as the similar levels of instruction between the filter model and the real system model. The smaller the residual error is the more matching between filter model and the real system model. Assuming the

residual error is used to calculate the conditional probability p_i in the conditions of the actual measured values and the actual vector parameter a by test algorithm. The conditional probability is used to weigh the correctness of each Kalman filter state estimate. The probability weighted average being from the state estimation, formed the mixed state estimation of the actual system $\hat{X}_{MMAE}$. Multiple model adaptive estimators are shown in Fig. 3.2.

(2) The Filtering Algorithm in MMAE

Step1 Parallel Filtering Equation

The Kalman filter of the $i\ (i=1,2,\cdots,N)$ linear model is:

$$\begin{cases} \boldsymbol{X}_i(t_k)=\boldsymbol{\Phi}_i\boldsymbol{X}_i(t_{k-1})+\boldsymbol{C}_i\boldsymbol{u}(t_{k-1})+\boldsymbol{\Gamma}_i w_i(t_{k-1}) \\ \boldsymbol{Z}_i(t_k)=\boldsymbol{H}_i\boldsymbol{X}_i(t_k)+v_i(t_k) \end{cases} \tag{3.6}$$

The symbols have the same meaning as those of Formula (3.4). In addition, systematic noise $w_i(t_k)$ and observation noise $v_i(t_k)$ are both zero mean white noise, and for all k, j, satisfying:

$$\left.\begin{aligned} &\mathrm{E}[w_i(t_k)]=0 \\ &\mathrm{E}[v_i(t_k)]=0 \\ &\mathrm{E}[w_i(t_k)w_i^T(t_j)]=\boldsymbol{Q}_i\delta_{k,j} \\ &\mathrm{E}[v_i(t_k)v_i^T(t_j)]=\boldsymbol{R}_i\delta_{k,j} \\ &\mathrm{E}[w_i(t_k)v_i^T(t_j)]=0 \end{aligned}\right\} \tag{3.7}$$

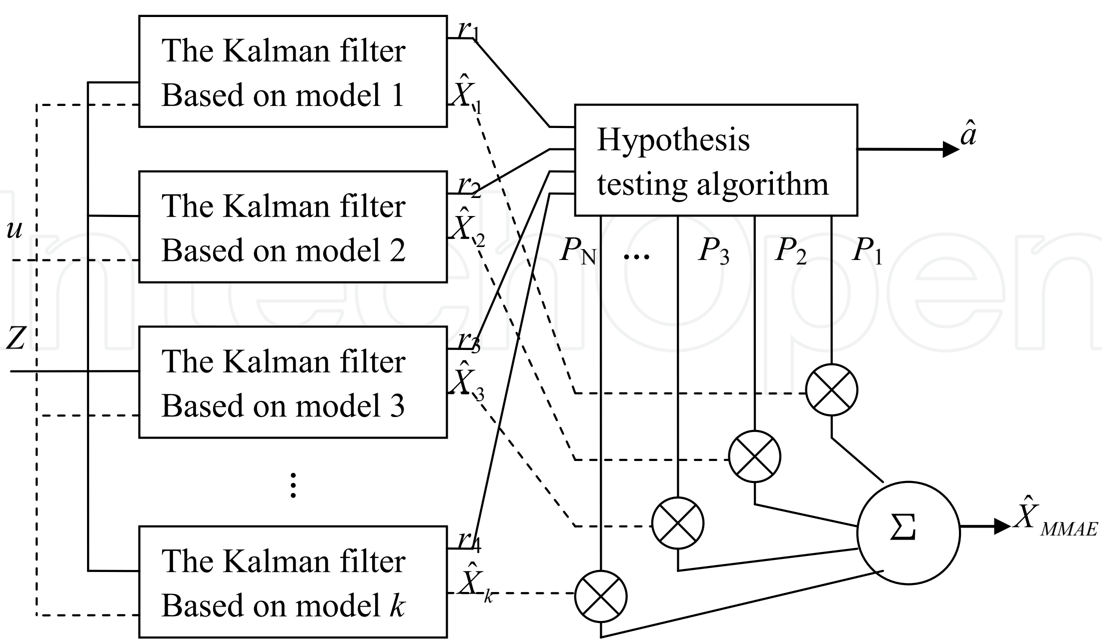


Fig. 3.2 Structure Diagram of Multiple-Model Adaptive Estimators

The Kalman filter algorithm use the above model to determine the optimum time to update the prediction and measurement of Kalman filter state estimation, optimum estimate update equation and state estimation error covariance matrix. Based on Kalman filter model, the update time equation of the Kalman filter state estimation is as follows:

$$\begin{cases} \hat{\boldsymbol{X}}_i(k/k-1) = \boldsymbol{\Phi}_i \hat{\boldsymbol{X}}_i(k-1/k-1) + \boldsymbol{C}_i \boldsymbol{u}(k-1) \\ \hat{\boldsymbol{Z}}_i(k/k-1) = \boldsymbol{H}_i \hat{\boldsymbol{X}}_i(k/k-1) \end{cases} \tag{3.8}$$

The update time equation of the state estimation error covariance matrix is:

$$\boldsymbol{P}_i\left(k/k-1\right) = \boldsymbol{\Phi}_i \boldsymbol{P}_i\left(k-1/k-1\right)\boldsymbol{\Phi}_i^{\mathrm{T}} + \boldsymbol{\Gamma}_i \boldsymbol{Q}_i \boldsymbol{\Gamma}_i^{\mathrm{T}} \tag{3.9}$$

The Kalman filter state estimation can achieve the measurement update by the following formula:

$$\hat{\boldsymbol{X}}_i(k/k) = \hat{\boldsymbol{X}}_i(k/k-1) + \boldsymbol{K}_i(k)\boldsymbol{r}_i(k) \tag{3.10}$$

And the gain of Kalman is:

$$\boldsymbol{K}_i\left(k\right) = \boldsymbol{P}_i\left(k/k-1\right)\boldsymbol{H}_i^{\mathrm{T}} \boldsymbol{A}_i(k)^{-1} \tag{3.11}$$

The O-C residual vector referred to the deviation by subtracting the measured value $Z_i(k)$ to the Kalman estimation based on previous measurements $Z_i(k/k-1)$, and that is:

$$\boldsymbol{r}_i(k) = \boldsymbol{Z}_i(k) - \boldsymbol{H}_i \hat{\boldsymbol{X}}_i(k/k-1) \tag{3.12}$$

Its variance matrix is:

$$\boldsymbol{A}_i(k) = \boldsymbol{H}_i \boldsymbol{P}_i\left(k/k-1\right)\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i \tag{3.13}$$

And the update equation of the state estimate covariance matrix is:

$$\boldsymbol{P}_i\left(k/k\right) = [\mathrm{I} - \boldsymbol{K}_i\left(k\right)\boldsymbol{H}_i]\boldsymbol{P}_i\left(k/k-1\right) \tag{3.14}$$

Step2 Solving of Model Probability

It can obtain the new residual income of single linear model at any moment through the calculation of each parallel filter system of local filtering equation. At this moment, on the basis of the residual information and a hypothesis test principle, the model probability, corresponding to each estimator model, is designed reflect real-time system model in determining the time to take the possibility of a size. The representation of the probability of two models will be given as:

1) The representation of the model probability based on statistical properties of residuals

It is known to all: If the Single Kalman model and the system model phase are matching, the residual is the Gaussian white noise of the sequence zero-mean, and the variance matrix can be obtained by Formula (3.13). Therefore, the conditional probability density function under the condition of the measured values $\boldsymbol{Z}(t_k)$ of the $i\ (i = 1, 2, \cdots, N)$ filter model at the kth moment is:

$$f_{\boldsymbol{Z}(t_k)|\boldsymbol{H}_i, \boldsymbol{Z}(t_{k-1})}(\boldsymbol{Z}(t_k) \mid \boldsymbol{H}_i, \boldsymbol{Z}(t_{k-1})) = \frac{1}{(2\pi)^{m/2} \mid \boldsymbol{A}_i \mid^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{r}_i^{\mathrm{T}}(k)\boldsymbol{A}_i^{-1}\boldsymbol{r}_i(k)\right\} \tag{3.15}$$

Defining the following objective function:

$$J_i(k) = p(\boldsymbol{\theta}_i \mid \boldsymbol{Z}_k) = p_i(t_k) = \Pr\{\boldsymbol{H} = \boldsymbol{H}_i \mid \boldsymbol{Z}(t_k) = \boldsymbol{Z}_k\} \tag{3.16}$$

And there will be the following recurrence relations:

$$p_i(t_k) = f_{\boldsymbol{Z}(t_k)|\boldsymbol{H}_i,\boldsymbol{Z}(t_{k-1})}(\boldsymbol{Z}(t_k) \mid \boldsymbol{H}_i, \boldsymbol{Z}(t_{k-1})) \cdot p_i(t_{k-1}) \tag{3.17}$$

For the normalized of the above objective function, if:

$$J_i(k) = p_i(t_k) = \frac{f_{\boldsymbol{Z}(t_k)|\boldsymbol{H}_i,\boldsymbol{Z}(t_{k-1})}(\boldsymbol{Z}(t_k) \mid \boldsymbol{H}_i, \boldsymbol{Z}(t_{k-1})) \cdot p_i(t_{k-1})}{\sum_{j=1}^{N} f_{\boldsymbol{Z}(t_k)|\boldsymbol{H}_j,\boldsymbol{Z}(t_{k-1})}(\boldsymbol{Z}(t_k) \mid \boldsymbol{H}_j, \boldsymbol{Z}(t_{k-1})) \cdot p_j(t_{k-1})} \tag{3.18}$$

The representation of the model probability based on statistical properties of residuals will be obtained.

2) The representation of the model probability based on normalized residuals

From the preceding analysis, it shows that O-C residual error $r_i(k)$ meant that the error between the actual output at the k time and the output of the ith model, so the residual can be used directly to define the following performance index function:

$$J_i(k) = \omega(\theta_i \mid \boldsymbol{Z}^k) = \frac{S(k) - r_i^2(k)}{(N-1)S(k)} \tag{3.19}$$

Where, the model weight of the ith estimator will be $S(k) = \sum_{i=1}^{N} r_i^2(k)$, $\omega(\theta_i \mid \boldsymbol{Z}^k)$, which is the weighted value in the actual model. The more accurate the ith estimator is, the smaller the corresponding residuals will be. Therefore, the greater the model weight of the estimator is, the smaller other corresponding models will be.

It does not involve the statistical distribution residuals in this model probabilistic representation, but the calculation is relatively simple.

Step3 Optimal Fusion Estimate

The optimal fusion estimate of the state is the product integration of the local estimates corresponding to local parallel linear model and their corresponding performance index function. That is:

$$\hat{\boldsymbol{X}}_k^{\text{opt}} = \sum_{i=1}^{N} J_i(k) \hat{\boldsymbol{X}}_i(k \mid k) \tag{3.20}$$

There are the following forms in the covariance matrix:

$$\boldsymbol{P}^{\text{opt}}(k \mid k) = \sum_{i=1}^{N} J_i(k) \{\boldsymbol{P}_i(k/k) + [\hat{\boldsymbol{X}}_i(k \mid k) - \hat{\boldsymbol{X}}_k^{\text{opt}}][\hat{\boldsymbol{X}}_i(k \mid k) - \hat{\boldsymbol{X}}_k^{\text{opt}}]^{\text{T}}\} \tag{3.21}$$

In addition, the estimates of the actual model parameters at the kth moment will be:

$$\hat{\boldsymbol{\theta}}_k^{\text{opt}} = \sum_{i=1}^{N} J_i(k) \boldsymbol{\theta}_i \tag{3.22}$$

3.3 Interacting Multiple Model Algorithm

The American Scholar Blom was the first one who proposed IMM algorithm in 1984. There are the following advantages in the interacting multiple model algorithm. In the first place, IMM is the optimum estimate after the completeness and the exclusive condition are

satisfied in the model. Secondly, IMM can expand the new operation model of the estimated system without changing the structure of the system. Furthermore, the amount of computation in IMM is moderate, having advantages of nonlinear filtering.
(1) The Fusion Architecture of IMM
Assuming a certain system can be described as the following state equation and measurement equation:

$$\begin{cases} X(k+1) = \Phi(k, m(k))X(k) + w(k, m(k)) \\ Z(k) = H(k, m(k))X(k) + v(k, m(k)) \end{cases} \tag{3.23}$$

Where, $X(k)$ is the system state vector, $\Phi(k, m(k))$ being the state transition matrix;

$w(k, m(k))$ is a mean zero, the variance being the Gaussian white noise $Q(k, m(k))$;
$Z(k)$ is the measurement vector, $H(k, m(k))$ being the observation matrix;
$v(k, m(k))$ is a mean zero, the variance being the Gaussian white noise $R(k, m(k))$;
And there is no relation between $w(k, m(k))$ and $v(k, m(k))$.
Where, $m(k)$ means an effective mode at t_k sampling time. At t_k time, the effective representation of m_i is $m_i(k) = \{m(k) = m_i\}$. All possible system mode set is $M = \{m_1, m_2, \cdots, m_N\}$. The systematic pattern sequence is assumed to be first-order Markov Chain, then the transition probability from $m_i(k+1)$ to $m_j(k)$ will be:

$$P\{m_i(k+1) \mid m_j(k)\} = \pi_{ji}, \ \ m_i, m_j \in M \tag{3.24}$$

And

$$\sum_{i=1}^{N} \pi_{ji} = 1 \quad j = 1, 2, \cdots, N \tag{3.25}$$

When received measurement information, the actual transition probability between models is the maximum posterior probability based on the above π_{ji} and measurement set $\{Z^k\}$.
The core of the interacting multiple model algorithms can modify the filter's input/output by using the actual transition probability in the above. The schematic figure of inter-influencing multiple model algorithms will be given in Fig. 3.3.

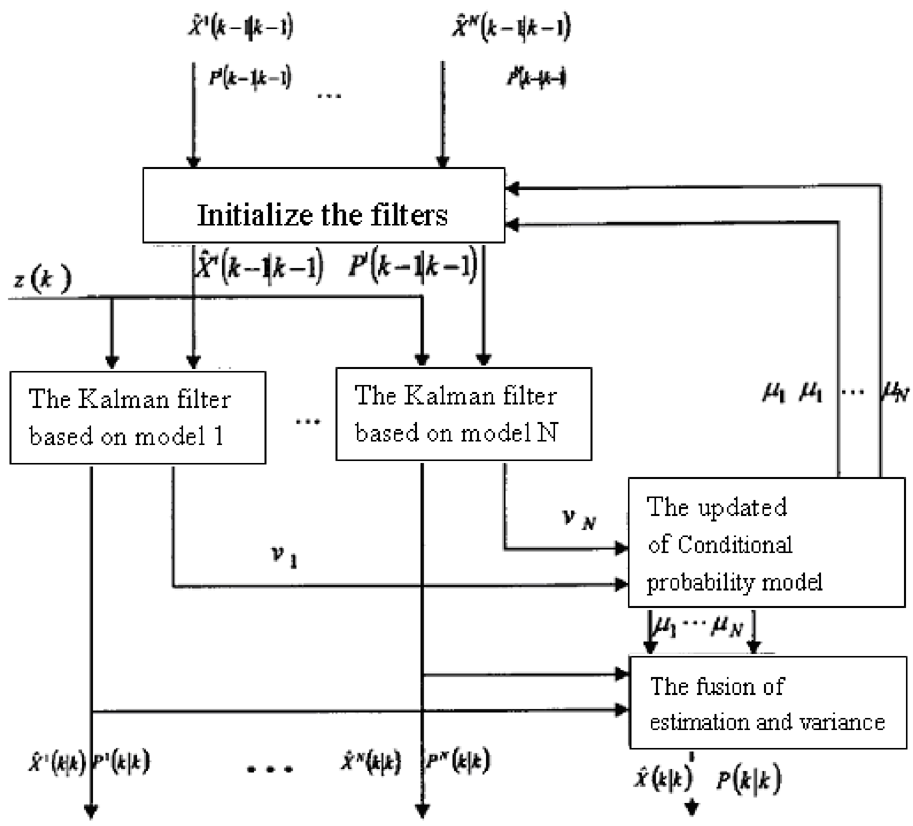


Fig. 3.3 Algorithm Flow of the Interacting Multiple Model

(2) The Filtering Algorithm for IMM

The interacting multiple model expanded the state of the conditional mean along the model space to do the Bayes probability. It is the optimum estimate under the condition of target motion model set covering model and the model of mutual independence. The interacting multiple model algorithm is a recursive algorithm: the model number is supposed to be limited, and each algorithm included 4-step in a cycle: input interaction, filter calculation, the updated for model probability and output interaction.

Step1 Input interaction

Input interaction is the most typical step of the interacting multiple model algorithm, using all state and model conditional probability obtained at last circulation as the computation and input state of each filtering model and the input state error covariance matrix.

That is:

$$\hat{X}_{0i}(k-1\,|\,k-1)=\sum_{j=1}^{N}\hat{X}_{j}(k-1\,|\,k-1)\mu_{j|i}(k-1\,|\,k-1) \tag{3.26}$$

$$P_{0i}(k-1\,|\,k-1)=\sum_{j=1}^{N}\mu_{j|i}(k-1\,|\,k-1)\{P_{j}(k-1\,|\,k-1)+a\cdot a^{T}\} \tag{3.27}$$

And

$$a=[\hat{X}_{j}(k-1\,|\,k-1)-\hat{X}_{0i}(k-1\,|\,k-1)] \tag{3.28}$$

The predicted probability of the model $\mu_{j|i}$ is:

$$\mu_{j|i}(k-1|k-1) = P\{m_j(k-1)|m_i(k), Z^{k-1}\} = \frac{1}{c_i}\pi_{ji}\mu_j(k-1) \tag{3.29}$$

Where

$$c_i = \sum_{j=1}^{N}\pi_{ji}\mu_j(k-1) \tag{3.30}$$

$\mu_i(k)$ means the probability of model m_i at the kth time,

And that is: $\mu_i(k) = P\{m_i(k)|Z^k\}$.

Step2 Filter Calculation

Each filter will do the Kalman filtering after obtaining measurement data collection $Z(k)$ signal. What the filter of each model outputted are the mode estimation, covariance matrix, the residual covariance matrix of the Kalman filter and the updated state vector. Kalman filter equations of the ith model at the kth time will be introduced below.

The state and covariance prediction of the ith model at the kth time is:

$$X_i(k|k-1) = \Phi_i\hat{X}_{0i}(k-1|k-1) \tag{3.31}$$

$$P_i(k|k-1) = \Phi_i P_{0i}(k-1|k-1)\Phi_i^T + Q_i \tag{3.32}$$

The residual vector of the Kalman Filter is the difference between measured values and the Kalman filter estimates of the previous measured values. That is:

$$v_i(k) = Z(k) - H_i\hat{X}_i(k|k-1) \tag{3.33}$$

And $Z(k)$ is the measurement data collection at the kth time.

The residual covariance matrix of the Kalman filter is:

$$S_i(k) = H_i P_i(k|k-1)H_i^T + R_i \tag{3.34}$$

The gain matrix of the Kalman filter is:

$$K_i(k) = P_i(k|k-1)H_i^T S_i^{-1} \tag{3.35}$$

The updated state equation of the Kalman filter is:

$$\hat{X}_i(k|k) = \hat{X}_i(k|k-1) + K_i v_i \tag{3.36}$$

The state covariance updated equation of the Kalman filter is:

$$P_i(k|k) = (I - K_i H_i)P_i(k|k-1) \tag{3.37}$$

Step3 Updated for Model Probability

Model probability provides information for working of a model at any time, which is given by Bayes Theorem. The updated equation of the specific model probability is:

$$\mu_i(k) = P\{m_i(k)|Z^k\} = \frac{1}{c}\Lambda_i(k)\sum_{j=1}^{N}\pi_{ji}\mu_j(k-1) \tag{3.38}$$

Where, $c = P\{Z(k)|Z^{k-1}\} = \sum_{i=1}^{N}\Lambda_i(k)\cdot c_i$

So, $\Lambda_i(k)$ is the likelihood function for model m_i the kth time, the likelihood value will be calculated by the residual error and the updated amount of covariance. That is:

$$\Lambda_i(k) = N[v_i(k):0, S_i(k)] = |2\pi S_i|^{-1/2} \exp\{-\frac{1}{2} v_i^T S_i^{-1} v_i\} \tag{3.39}$$

Step4 Output Fusion

The final state of the output is obtained by weighting and combining all sub-model state estimation, namely, by the product of the state estimation of each model and model probability.

$$\hat{X}(k|k) = \sum_{i=1}^{N} \hat{X}_i(k|k) \cdot \mu_i(k) \tag{3.40}$$

Simultaneously, the estimated covariance matrix is:

$$P(k|k) = \sum_{i=1}^{N} \mu_i(k)\{P_i(k|k) + b \cdot b^T\} \tag{3.41}$$

And

$$b = [\hat{X}_i(k|k) - \hat{X}(k|k)] \tag{3.42}$$

As will be readily seen, when IMM estimation is taking into historical information of mode at the kth time, it also mixes the previous estimated information in the beginning of each circulation to avoid the shortcoming that the complexity of the optimal estimation will present an exponential growth with time. It is the main aspect that can distinguish interacting multiple model algorithm from other non-interacting multiple model estimation.

4. Nonstandard Multi-sensor Information Fusion Based on Local Filtering Estimate Decoupling

The algorithm of the state fusion estimation of dynamic multi-sensor system is related to the fusion structure. There commonly are: centralization, distribution and hybrid[26-27]. Each fusion structure has its own particular advantages and disadvantages. For instance, problems as heavy computational burden and poor tolerance are existed in the centralization, but all the raw sensor measurements are used without loss, so the fusion result is the optimal one. In regard to the distribution, it adopts two-level information processing to use a primary filter and several local filters replace the original single centralized fusion model. In the first stage, each local filter processed the information of each corresponding subsystem measurement in parallel; then in the second stage, the primary filter will filter the local state of each local filter to improve the computational efficiency and error tolerance of the system. However, the distributed fusion estimation always assumes that the local estimates obtained from each sensor are independent of each other and that the local covariance is diagonal to achieve the decoupling of the estimated state of each sensor, which is the basis for the distributed optimal algorithm. In the multi-sensor system, state estimates of the corresponding local filter in each subsystem are often related. In view of the relevant local filter, the distributed fusion filter is needed to transform in order to achieve the global optimal estimates to make the local filtering estimate irrelevant in the actual operation.

The distributed joint filter (FKF, Federal Kalman Filter) was proposed by an American Scholar N.A. Carlson in 1988 concerning with a special form of distributed fusion. It has been considered as a new information fusion method which is only directed towards the synthesis of the estimated information of sub-filter. The sub-filter is also a parallel structure and each filter adopted the Kalman filter algorithm to deal with its own sensor measurements. In order to make the structure of the master filter and the accuracy of the centralized fusion estimation similar, the feature which distinguished the combined filter from the general distributed filter is that the combined filter applied variance upper bound technique and information distribution principle to eliminate the correlation estimates of the sub-filter in each sensor, and distributed the global state estimate information and noise information of the system to each sub-filter without changing the form of sub-filter algorithm. Therefore, it has the advantages of more simple in algorithm, better fault tolerance and easy to implement, etc. When information distribution factor determined the performance of the combined filter, the selection rules became the focus of recent research and debate[28]. Under the present circumstances, it is the main objective and research direction in this field to search for and design "information distribution" which will be simple, effective and self-adaptive.

4.1 Analysis and Decoupling for the Relevance of the Combined Filter

The system description will be given as:

$$\boldsymbol{X}(k+1)=\boldsymbol{\Phi}(k+1,k)\boldsymbol{X}(k)+\boldsymbol{\Gamma}(k+1,k)w(k) \tag{4.1}$$

$$\boldsymbol{Z}_i(k+1)=\boldsymbol{H}_i(k+1)\boldsymbol{X}_i(k+1)+v_i(k+1) \qquad i=1,2,\cdots,N \tag{4.2}$$

Where, $\boldsymbol{X}(k+1)\in \mathrm{R}^n$ is the system state vector at the $k+1$ time, $\boldsymbol{\Phi}(k+1,k)\in \mathrm{R}^{n\times n}$ being the state transition matrix of the system, $\boldsymbol{\Gamma}(k+1,k)$ being the process noise distribution matrix, $\boldsymbol{Z}_i(k+1)\in \mathrm{R}^m$ $(i=1,2,\cdots,N)$ being the measurements of the i sensor at the $k+1$ time, and $\boldsymbol{H}_i(k+1)$ being the mapping matrix of the ith sensor at the $\boldsymbol{H}_i(k+1)$ time. Assume $\mathrm{E}[w(k)]=0$, $\mathrm{E}[w(k)w^{\mathrm{T}}(j)]=\boldsymbol{Q}(k)\delta_{kj}$, $\mathrm{E}[v_i(k)]=0$, and $\mathrm{E}[v_i(k)v_i^{\mathrm{T}}(j)]=\boldsymbol{R}_i(k)\delta_{kj}$.

Theorem 4.1: In the multi-sensor information fusion system described by Equation (4.1) and (4.2), if local estimates are unrelated, the global optimal fusion estimate of the state $\hat{\boldsymbol{X}}_g$ can have the following general formulas:

$$\begin{cases} \hat{\boldsymbol{X}}_g = \boldsymbol{P}_g \sum_{i=1}^{N} \boldsymbol{P}_i^{-1}\hat{\boldsymbol{X}}_i = \boldsymbol{P}_g\boldsymbol{P}_1^{-1}\hat{\boldsymbol{X}}_1 + \boldsymbol{P}_g\boldsymbol{P}_2^{-1}\hat{\boldsymbol{X}}_2 + \cdots + \boldsymbol{P}_g\boldsymbol{P}_N^{-1}\hat{\boldsymbol{X}}_N \\ \boldsymbol{P}_g = (\sum_{i=1}^{N}\boldsymbol{P}_i^{-1})^{-1} = (\boldsymbol{P}_1^{-1}+\boldsymbol{P}_2^{-1}+\cdots+\boldsymbol{P}_N^{-1})^{-1} \end{cases} \tag{4.3}$$

Where, $\hat{\boldsymbol{X}}_i, \boldsymbol{P}_i$ $i=1,2,\cdots,N$ are respectively referred as the local estimates of the subsystem and the corresponding estimated covariance matrix.

Supposing $\hat{X}_g(k|k), P_g(k|k)$ are the optimal estimates and the covariance matrix of the combined Kalman filter (the fusion center), $\hat{X}_i(k|k), P_i(k|k)$ being the estimate and the covariance matrix of the i sub-filter, $\hat{X}_m(k|k), P_m(k|k)$ being the estimate and the covariance matrix of the Master Filter, and if there is no feedback from the fusion center to the sub-filter, when the Master Filter completed the fusion process at k time, there will be $\hat{X}_m(k|k) = \hat{X}(k|k)$, $P_m(k|k) = P(k|k)$. The forecast for the main filter will be (Because of no measurements, the Master Filter only had time updates, but no measurement updates.):

$$\begin{cases} \hat{X}_m(k+1|k) = \Phi_k \hat{X}(k|k) \\ P_m(k+1|k) = \Phi(k) P(k|k) \Phi^{\mathrm{T}}(k) + \Gamma(k) Q(k) \Gamma^{\mathrm{T}}(k) \end{cases} \tag{4.4}$$

Where, the meanings of $\Phi(k)$, $\Gamma(k)$ and $Q(k)$ are the same as those above. As the ith sub-filter has both time updates and measurement updates, it should have:

$$\begin{aligned} \hat{X}_i(k+1|k+1) &= \hat{X}_i(k+1|k) + K_i(k+1)(Z_i(k+1) - H_i(k+1)\hat{X}_i(k+1|k)) \\ &= \Phi(k)\hat{X}_i(k|k) + K_i(k+1)(Z_i(k+1) - H_i(k+1)\Phi(k)\hat{X}_i(k|k)) \end{aligned} \tag{4.5}$$

Accordingly,

$$\begin{aligned} &\tilde{X}_i(k+1|k+1) = X(k+1|k+1) - \hat{X}_i(k+1|k+1) \\ &= \Phi(k)X(k|k) + \Gamma(k)w(k) - \Phi(k)\hat{X}_i(k|k) \\ &\quad - K_i(k+1)[H_i(k+1)(\Phi(k)X(k|k) + \Gamma(k)w(k)) + v_i(k+1) - H_i(k+1)\Phi(k)\hat{X}_i(k|k)] \\ &= (\mathrm{I} - K_i(k+1)H_i(k+1))\Phi(k)\tilde{X}_i(k|k) \\ &\quad + (\mathrm{I} - K_i(k+1)H_i(k+1))\Gamma(k)w(k) - K_i(k+1)v_i(k+1) \end{aligned} \tag{4.6}$$

Then we can get the covariance of the local sub-filters i and j at the $k+1$ th time:

$$\begin{aligned} &P_{i,j}(k+1) = \mathrm{Cov}(\tilde{X}_i(k+1|k+1), \tilde{X}_j(k+1|k+1)) \\ &= (\mathrm{I} - K_i(k+1)H_i(k+1))\Phi(k)P_{i,j}(k)\Phi^{\mathrm{T}}(k)(\mathrm{I} - K_j(k+1)H_j(k+1))^{\mathrm{T}} \\ &\quad + (\mathrm{I} - K_i(k+1)H_i(k+1))\Gamma(k)Q(k)\Gamma^{\mathrm{T}}(k)(\mathrm{I} - K_j(k+1)H_j(k+1))^{\mathrm{T}} \\ &= (\mathrm{I} - K_i(k+1)H_i(k+1))(\Phi(k)P_{i,j}(k)\Phi^{\mathrm{T}}(k) + \Gamma(k)Q(k)\Gamma^{\mathrm{T}}(k))(\mathrm{I} - K_j(k+1)H_j(k+1))^{\mathrm{T}} \end{aligned} \tag{4.7}$$

There is no measurement in the master filter, so the time updates is also the measurement updates:

$$\begin{cases} \hat{X}_m(k+1|k+1) = \hat{X}_m(k+1|k) = \Phi(k)\hat{X}(k|k) \\ \tilde{X}_m(k+1|k+1) = X(k+1|k+1) - \hat{X}_m(k+1|k+1) \\ \qquad = \Phi(k)X(k|k) + \Gamma(k)w(k) - \Phi(k)\hat{X}(k|k) = \Phi(k)\tilde{X}(k|k) + \Gamma(k)w(k) \end{cases} \tag{4.8}$$

Therefore, the covariance of any sub-filter i and the Master Filter m at the $(k+1)$ th time will be:

$$\begin{aligned}\boldsymbol{P}_{i,m}(k+1) &= \mathrm{Cov}(\tilde{\boldsymbol{X}}_i(k+1\mid k+1), \tilde{\boldsymbol{X}}_m(k+1\mid k+1)) \\ &= (\mathrm{I}-\boldsymbol{K}_i(k+1)\boldsymbol{H}_i(k+1))\boldsymbol{\Phi}(k)\boldsymbol{P}_{i,m}(k)\boldsymbol{\Phi}^{\mathrm{T}}(k) \\ &\quad + (\mathrm{I}-\boldsymbol{K}_i(k+1)\boldsymbol{H}_i(k+1))\boldsymbol{\Gamma}(k)\boldsymbol{Q}(k)\boldsymbol{\Gamma}^{\mathrm{T}}(k)\end{aligned} \tag{4.9}$$

As can be seen, only on the condition of both $\boldsymbol{Q}(k)=0$ and $\boldsymbol{P}_{i,j}(k)=0$, the filtering errors between each sub-filter and the Master Filter at $(k+1)$ time are not related to each other. While in the usual case, both constraint conditions are hard to establish.
In addition, supposing:

$$\begin{aligned}\boldsymbol{B}_i(k+1) &= (\mathrm{I}-\boldsymbol{K}_i(k+1)\boldsymbol{H}_i(k+1))\boldsymbol{\Phi}(k), \boldsymbol{C}_i(k+1) \\ &= (\mathrm{I}-\boldsymbol{K}_i(k+1)\boldsymbol{H}_i(k+1))\boldsymbol{\Gamma}(k), \quad (i=1,2,\cdots,N)\end{aligned} \tag{4.10}$$

And:

$$\begin{aligned}
&\begin{bmatrix} \boldsymbol{P}_{1,1}(k+1) & \cdots & \boldsymbol{P}_{1,N}(k+1) & \boldsymbol{P}_{1,m}(k+1) \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{P}_{N,1}(k+1) & \cdots & \boldsymbol{P}_{N,N}(k+1) & \boldsymbol{P}_{N,m}(k+1) \\ \boldsymbol{P}_{m,1}(k+1) & \cdots & \boldsymbol{P}_{m,N}(k+1) & \boldsymbol{P}_{m,m}(k+1) \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{B}_1(k+1)\boldsymbol{P}_{1,1}(k)(\boldsymbol{B}_1(k+1))^{\mathrm{T}} & \cdots & \boldsymbol{B}_1(k+1)\boldsymbol{P}_{1,N}(k)(\boldsymbol{B}_N(k+1))^{\mathrm{T}} & \boldsymbol{B}_1(k+1)\boldsymbol{P}_{1,m}(k)\boldsymbol{\Phi}^{\mathrm{T}}(k) \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{B}_N(k+1)\boldsymbol{P}_{N,1}(k)(\boldsymbol{B}_1(k+1))^{\mathrm{T}} & \cdots & \boldsymbol{B}_N(k+1)\boldsymbol{P}_{N,N}(k)(\boldsymbol{B}_N(k+1))^{\mathrm{T}} & \boldsymbol{B}_N(k+1)\boldsymbol{P}_{N,m}(k)\boldsymbol{\Phi}^{\mathrm{T}}(k) \\ \boldsymbol{\Phi}(k)\boldsymbol{P}_{m,1}(k)(\boldsymbol{B}_1(k+1))^{\mathrm{T}} & \cdots & \boldsymbol{\Phi}(k)\boldsymbol{P}_{m,N}(k)(\boldsymbol{B}_N(k+1))^{\mathrm{T}} & \boldsymbol{\Phi}(k)\boldsymbol{P}_m(k)\boldsymbol{\Phi}^{\mathrm{T}}(k) \end{bmatrix} \\
&+ \begin{bmatrix} \boldsymbol{C}_1(k+1)\boldsymbol{Q}(k)(\boldsymbol{C}_1(k+1))^{\mathrm{T}} & \cdots & \boldsymbol{C}_1(k+1)\boldsymbol{Q}(k)(\boldsymbol{C}_N(k+1))^{\mathrm{T}} & \boldsymbol{C}_1(k+1)\boldsymbol{Q}(k)\boldsymbol{\Gamma}^{\mathrm{T}}(k) \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{C}_N(k+1)\boldsymbol{Q}(k)(\boldsymbol{C}_1(k+1))^{\mathrm{T}} & \cdots & \boldsymbol{C}_N(k+1)\boldsymbol{Q}(k)(\boldsymbol{C}_N(k+1))^{\mathrm{T}} & \boldsymbol{C}_N(k+1)\boldsymbol{Q}(k)\boldsymbol{\Gamma}^{\mathrm{T}}(k) \\ \boldsymbol{\Gamma}(k)\boldsymbol{Q}(k)(\boldsymbol{C}_1(k+1))^{\mathrm{T}} & \cdots & \boldsymbol{\Gamma}(k)\boldsymbol{Q}(k)(\boldsymbol{C}_N(k+1))^{\mathrm{T}} & \boldsymbol{\Gamma}(k)\boldsymbol{Q}(k)\boldsymbol{\Gamma}^{\mathrm{T}}(k) \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{B}_1(k+1) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \boldsymbol{B}_N(k+1) & 0 \\ 0 & \cdots & 0 & \boldsymbol{\Phi}(k) \end{bmatrix} \begin{bmatrix} \boldsymbol{P}_{1,1}(k) & \cdots & \boldsymbol{P}_{1,N}(k) & \boldsymbol{P}_{1,m}(k) \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{P}_{N,1}(k) & \cdots & \boldsymbol{P}_{N,N}(k) & \boldsymbol{P}_{N,m}(k) \\ \boldsymbol{P}_{m,1}(k) & \cdots & \boldsymbol{P}_{m,N}(k) & \boldsymbol{P}_{m,m}(k) \end{bmatrix} \begin{bmatrix} \boldsymbol{B}_1^{\mathrm{T}}(k+1) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \boldsymbol{B}_N^{\mathrm{T}}(k+1) & 0 \\ 0 & \cdots & 0 & \boldsymbol{\Phi}^{\mathrm{T}}(k) \end{bmatrix} \\
&+ \begin{bmatrix} \boldsymbol{C}_1(k+1) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \boldsymbol{C}_N(k+1) & 0 \\ 0 & \cdots & 0 & \boldsymbol{\Gamma}(k) \end{bmatrix} \begin{bmatrix} \boldsymbol{Q}(k) & \cdots & \boldsymbol{Q}(k) & \boldsymbol{Q}(k) \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{Q}(k) & \cdots & \boldsymbol{Q}(k) & \boldsymbol{Q}(k) \\ \boldsymbol{Q}(k) & \cdots & \boldsymbol{Q}(k) & \boldsymbol{Q}(k) \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_1^{\mathrm{T}}(k+1) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \boldsymbol{C}_N^{\mathrm{T}}(k+1) & 0 \\ 0 & \cdots & 0 & \boldsymbol{\Gamma}^{\mathrm{T}}(k) \end{bmatrix}
\end{aligned} \tag{4.11}$$

As can be seen, due to the influence of the common process noise $w(k)$, even if $\boldsymbol{P}_{i,j}(k)=0$, there cannot get $\boldsymbol{P}_{i,j}(k+1)=0$. At this time, "variance upper-bound" technology can be used to eliminate this correlation. Known by the matrix theory[29], there are upper-bound existed in the phalanx being composed of $\boldsymbol{Q}(k)$ from the Formula (4.11).

$$\begin{bmatrix} \boldsymbol{Q}(k) & \cdots & \boldsymbol{Q}(k) & \boldsymbol{Q}(k) \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{Q}(k) & \cdots & \boldsymbol{Q}(k) & \boldsymbol{Q}(k) \\ \boldsymbol{Q}(k) & \cdots & \boldsymbol{Q}(k) & \boldsymbol{Q}(k) \end{bmatrix} \le \begin{bmatrix} \beta_1^{-1}\boldsymbol{Q}(k) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \beta_N^{-1}\boldsymbol{Q}(k) & 0 \\ 0 & \cdots & 0 & \beta_m^{-1}\boldsymbol{Q}(k) \end{bmatrix} \tag{4.12}$$

And: $\beta_1+\beta_2+\cdots+\beta_N+\beta_m=1, \quad 0\le\beta_i\le 1$

As can be seen, the positive definite of the upper-bound in Formula (4.12) is stronger than that of original matrix. That is to say, the difference between the upper-bound matrix and the original matrix is positive semi-definite.

A similar upper-bound can also be set in the initial state covariance $\boldsymbol{P}_0$. That is:

$$\begin{bmatrix} \boldsymbol{P}_{1,1}(0) & \cdots & \boldsymbol{P}_{1,N}(0) & \boldsymbol{P}_{1,m}(0) \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{P}_{N,1}(0) & \cdots & \boldsymbol{P}_{N,N}(0) & \boldsymbol{P}_{N,m}(0) \\ \boldsymbol{P}_{m,1}(0) & \cdots & \boldsymbol{P}_{m,N}(0) & \boldsymbol{P}_{m,m}(0) \end{bmatrix} \le \begin{bmatrix} \beta_1^{-1}\boldsymbol{P}_{1,1}(0) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \beta_N^{-1}\boldsymbol{P}_{N,N}(0) & 0 \\ 0 & \cdots & 0 & \beta_m^{-1}\boldsymbol{P}_{m,m}(0) \end{bmatrix} \tag{4.13}$$

It also can be seen from this, there is no related items in the right side of the Formula (4.13). Namely, if enlarge the initial covariance of the master filter and each sub-filter, the correlation of the initial covariance errors of the mater filter and each sub-filter. Then, it can be known from Formula (4.7) and (4.9).

$$\boldsymbol{P}_{i,j}(k)=0 \quad (i\ne j, i,j=1,2,\cdots,N,m).$$

It can be got the following by substituting Formula (4.12) and (4.13) into Formula (4.11):

$$
\begin{bmatrix}
P_{1,1}(k+1) & \cdots & P_{1,N}(k+1) & P_{1,m}(k+1) \\
\vdots & \ddots & \vdots & \vdots \\
P_{N,1}(k+1) & \cdots & P_{N,N}(k+1) & P_{N,m}(k+1) \\
P_{m,1}(k+1) & \cdots & P_{m,N}(k+1) & P_{m,m}(k+1)
\end{bmatrix}
$$

$$
\leq \begin{bmatrix}
B_1(k+1) & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & B_N(k+1) & 0 \\
0 & \cdots & 0 & \Phi(k)
\end{bmatrix}
\begin{bmatrix}
P_{1,1}(k) & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & P_{N,N}(k) & 0 \\
0 & \cdots & 0 & P_{m,m}(k)
\end{bmatrix}
\begin{bmatrix}
B_1^{\mathrm{T}}(k+1) & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & B_N^{\mathrm{T}}(k+1) & 0 \\
0 & \cdots & 0 & \Phi^{\mathrm{T}}(k)
\end{bmatrix}
$$

$$
+ \begin{bmatrix}
C_1(k+1) & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & C_N(k+1) & 0 \\
0 & \cdots & 0 & \Gamma(k)
\end{bmatrix}
\begin{bmatrix}
\beta_1^{-1}Q(k) & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & \beta_N^{-1}Q(k) & 0 \\
0 & \cdots & 0 & \beta_m^{-1}Q(k)
\end{bmatrix}
\begin{bmatrix}
C_1^{\mathrm{T}}(k+1) & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & C_N^{\mathrm{T}}(k+1) & 0 \\
0 & \cdots & 0 & \Gamma^{\mathrm{T}}(k)
\end{bmatrix}
\quad (4.14)
$$

If taken the equal sign, that is, achieved the de-correlation of local estimates, on the one hand, the global optimal fusion estimate can be realized by Theorem 4.1 , but on the other, the initial covariance matrix and process noise covariance of the sub-filter themselves can enlarged by β_i^{-1} times. What's more, the filter results of every local filter will not be optimal.

4.2 Structure and Performance Analysis of the Combined Filter

The combined filter is a 2-level filter. The characteristic to distinguish from the traditional distributed filters is the use of information distribution to realize information share of every sub-filter. Information fusion structure of the combined filter is shown in Fig. 4.1.

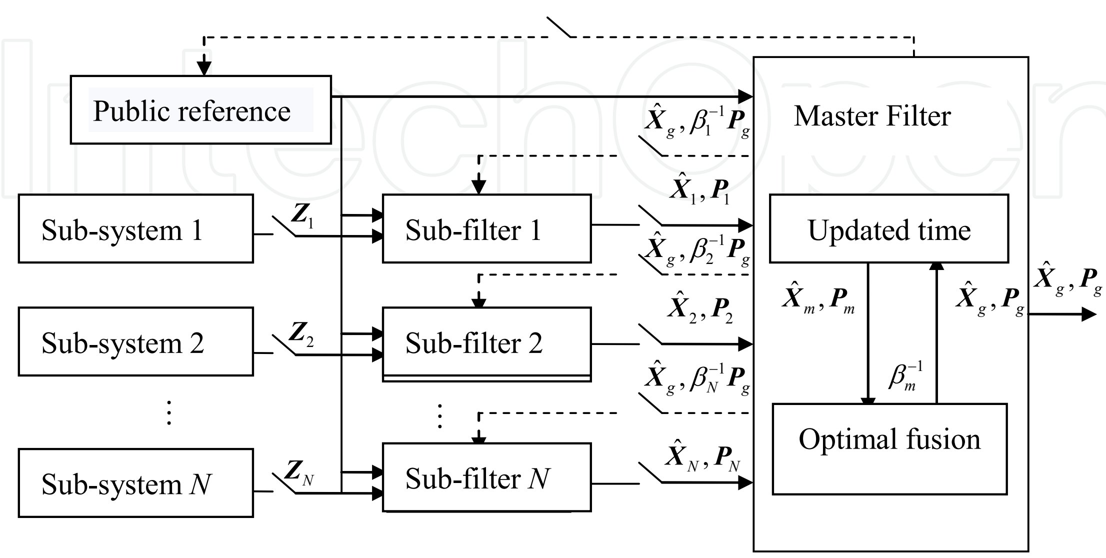


Fig. 4.1 Structure Indication of the Combined Filter

From the filter structure shown in the Fig. 4.1, the fusion process for the combined filter can be divided into the following four steps.

Step1 Given initial value and information distribution: The initial value of the global state in the initial moment is supposed to be $\boldsymbol{X}_0$, the covariance to be $\boldsymbol{Q}_0$, the state estimate vector of the local filter, the system covariance matrix and the state vector covariance matrix separately, respectively to be $\hat{\boldsymbol{X}}_i, \boldsymbol{Q}_i, \boldsymbol{P}_i, \quad i = 1, \cdots, N$, and the corresponding master filter to be $\hat{\boldsymbol{X}}_m, \boldsymbol{Q}_m, \boldsymbol{P}_m$.The information is distributed through the information distribution factor by the following rules in the sub-filter and the master filter.

$$\begin{cases} \boldsymbol{Q}_g^{-1}(k) = \boldsymbol{Q}_1^{-1}(k) + \boldsymbol{Q}_2^{-1}(k) + \cdots + \boldsymbol{Q}_N^{-1}(k) + \boldsymbol{Q}_m^{-1}(k) \quad \boldsymbol{Q}_i^{-1}(k) = \beta_i \boldsymbol{Q}_g^{-1}(k) \\ \boldsymbol{P}_g^{-1}(k|k) = \boldsymbol{P}_1^{-1}(k|k) + \boldsymbol{P}_2^{-1}(k|k) + \cdots + \boldsymbol{P}_N^{-1}(k|k) + \boldsymbol{P}_m^{-1}(k|k) \quad \boldsymbol{P}_i^{-1}(k|k) = \beta_i \boldsymbol{P}_g^{-1}(k|k) \\ \hat{\boldsymbol{X}}_i(k|k) = \hat{\boldsymbol{X}}_g(k|k) \qquad i = 1, 2, \cdots, N, m \end{cases} \tag{4.15}$$

Where, β_i should meet the requirements of information conservation principles:

$$\beta_1 + \beta_2 + \cdots + \beta_N + \beta_m = 1 \quad 0 \le \beta_i \le 1$$

Step2 the time to update the information: The process of updating time conducted independently, the updated time algorithm is shown as follows:

$$\begin{cases} \hat{\boldsymbol{X}}_i(k+1|k) = \boldsymbol{\Phi}(k+1|k)\hat{\boldsymbol{X}}_i(k|k) \qquad i = 1, 2, \cdots, N, m \\ \boldsymbol{P}_i(k+1|k) = \boldsymbol{\Phi}(k+1|k)\boldsymbol{P}_i(k|k)\boldsymbol{\Phi}^{\mathrm{T}}(k+1|k) + \boldsymbol{\Gamma}(k+1|k)\boldsymbol{Q}_i(k)\boldsymbol{\Gamma}^{\mathrm{T}}(k+1|k) \end{cases} \tag{4.16}$$

Step3 Measurement update: As the master filter does not measure, there is no measurement update in the Master Filter. The measurement update only occurs in each local sub-filter, and can work by the following formula:

$$\begin{cases} \boldsymbol{P}_i^{-1}(k+1|k+1)\hat{\boldsymbol{X}}_i(k+1|k+1) = \boldsymbol{P}_i^{-1}(k+1|k)\hat{\boldsymbol{X}}_i(k+1|k) + \boldsymbol{H}_i^{\mathrm{T}}(k+1)\boldsymbol{R}_i^{-1}(k+1)\boldsymbol{Z}_i(k+1) \\ \boldsymbol{P}_i^{-1}(k+1|k+1) = \boldsymbol{P}_i^{-1}(k+1|k) + \boldsymbol{H}_i^{\mathrm{T}}(k+1)\boldsymbol{R}_i^{-1}(k+1)\boldsymbol{H}_i(k+1) \qquad i = 1, 2, \cdots, N \end{cases} \tag{4.17}$$

Step4 the optimal information fusion: The amount of information of the state equation and the amount of information of the process equation can be apportioned by the information distribution to eliminate the correlation among sub-filters. Then the core algorithm of the combined filter can be fused to the local information of every local filter to get the state optimal estimates.

$$\begin{cases} \hat{\boldsymbol{X}}_g(k|k) = \boldsymbol{P}_g(k|k)\sum_{i=1}^{N,m}\boldsymbol{P}_i^{-1}(k|k)\hat{\boldsymbol{X}}_i(k|k) \\ \boldsymbol{P}_g(k|k) = (\sum_{i=1}^{N,m}\boldsymbol{P}_i^{-1}(k|k))^{-1} = (\boldsymbol{P}_1^{-1}(k|k) + \boldsymbol{P}_2^{-1}(k|k) + \cdots + \boldsymbol{P}_N^{-1}(k|k) + \boldsymbol{P}_m^{-1}(k|k))^{-1} \end{cases} \tag{4.18}$$

It can achieve the goal to complete the workflow of the combined filter after the processes of information distribution, the updated time, the updated measurement and information fusion. Obviously, as the variance upper-bound technique is adopted to remove the

correlation between sub-filters and the master filter and between the various sub-filters in the local filter and to enlarge the initial covariance matrix and the process noise covariance of each sub-filter by β_i^{-1} times, the filter results of each local filter will not be optimal. But some information lost by the variance upper-bound technique can be re-synthesized in the final fusion process to get the global optimal solution for the equation.
In the above analysis for the structure of state fusion estimation, it is known that centralized fusion structure is the optimal fusion estimation for the system state in the minimum variance. While in the combined filter, the optimal fusion algorithm is used to deal with local filtering estimate to synthesize global state estimate. Due to the application of variance upper-bound technique, local filtering is turned into being suboptimal, the global filter after its synthesis becomes global optimal, i.e. the fact that the equivalence issue between the combined filtering process and the centralized fusion filtering process. To sum up, as can be seen from the above analysis, the algorithm of combined filtering process is greatly simplified by the use of variance upper-bound technique. It is worth pointing out that the use of variance upper-bound technique made local estimates suboptimum but the global estimate after the fusion of local estimates is optimal, i.e. combined filtering model is equivalent to centralized filtering model in the estimated accuracy.

4.3 Adaptive Determination of Information Distribution Factor

By the analysis of the estimation performance of combined filter, it is known that the information distribution principle not only eliminates the correlation between sub-filters as brought from public baseline information to make the filtering of every sub-filter conducted themselves independently, but also makes global estimates of information fusion optimal. This is also the key technology of the fusion algorithm of combined filter. Despite it is in this case, different information distribution principles can be guaranteed to obtain different structures and different characteristics (fault-tolerance, precision and amount of calculation) of combined filter. Therefore, there have been many research literatures on the selection of information distribution factor of combined filter in recent years. In the traditional structure of the combined filter, when distributed information to the subsystem, their distribution factors are predetermined and kept unchanged to make it difficult to reflect the dynamic nature of subsystem for information fusion. Therefore, it will be the main objective and research direction to find and design the principle of information distribution which will be simple, effective and dynamic fitness, and practical. Its aim is that the overall performance of the combined filter will keep close to the optimal performance of the local system in the filtering process, namely, a large information distribution factors can be existed in high precision sub-system, while smaller factors existed in lower precision sub-system to get smaller to reduce its overall accuracy of estimated loss. Method for determining adaptive information allocation factors can better reflect the diversification of estimation accuracy in subsystem and reduce the impact of the subsystem failure or precision degradation but improve the overall estimation accuracy and the adaptability and fault tolerance of the whole system. But it held contradictory views given in Literature [28] to determine information distribution factor formula as the above held view. It argued that global optimal estimation accuracy had nothing to do with the information distribution factor values when statistical characteristics of noise are known, so there is no need for adaptive determination.

Combined with above findings in the literature, on determining rules for information distribution factor, we should consider from two aspects.
1) Under circumstances of meeting conditions required in Kalman filtering such as exact statistical properties of noise, it is known from filter performance analysis in Section 4.2 that: if the value of the information distribution factor can satisfy information on conservation principles, the combined filter will be the global optimal one. In other words, the global optimal estimation accuracy is unrelated to the value of information distribution factors, which will influence estimation accuracy of a sub-filter yet. As is known in the information distribution process, process information obtained from each sub-filter is $\beta_i \boldsymbol{Q}_g^{-1}, \beta_i \boldsymbol{P}_g^{-1}$, Kalman filter can automatically use different weights according to the merits of the quality of information: the smaller the value of β_i is, the lower process message weight will be, so the accuracy of sub-filters is dependent on the accuracy of measuring information; on the contrary, the accuracy of sub-filters is dependent on the accuracy of process information.
2) Under circumstances of not knowing statistical properties of noise or the failure of a subsystem, global estimates obviously loss the optimality and degrade the accuracy, and it is necessary to introduce the determination mode of adaptive information distribution factor. Information distribution factor will be adaptive dynamically determined by the sub-filter accuracy to overcome the loss of accuracy caused by fault subsystem to remain the relatively high accuracy in global estimates. In determining adaptive information distribution factor, it should be considered that less precision sub-filter will allocate factor with smaller information to make the overall output of the combined filtering model had better fusion performance, or to obtain higher estimation accuracy and fault tolerance.
In Kalman filter, the trace of error covariance matrix $\boldsymbol{P}$ includes the corresponding estimate vector or its linear combination of variance. The estimated accuracy can be reflected in filter answered to the estimate vector or its linear combination through the analysis for the trace of $\boldsymbol{P}$. So there will be the following definition:
Definition 4.1: The estimation accuracy of attenuation factor of the i th local filter is:

$$EDOP_i \stackrel{\Delta}{=} \sqrt{\mathrm{tr}(\boldsymbol{P}_i \boldsymbol{P}_i^{\mathrm{T}})} \tag{4.19}$$

Where, the definition of $EDOP_i$ (Estimation Dilution of Precision) is attenuation factor estimation accuracy, meaning the measurement of estimation error covariance matrix in i local filter, $\mathrm{tr}(\bullet)$ meaning the demand for computing trace function of the matrix.
When introduced attenuation factor estimation accuracy $EDOP_i$, in fact, it is said to use the measurement of norm characterization $\boldsymbol{P}_i$ in $\boldsymbol{P}_i$ matrix: the bigger the matrix norm is, the corresponding estimated covariance matrix will be larger, so the filtering effect is poorer; and vice versa.
According to the definition of attenuation factor estimation accuracy, take the computing formula of information distribution factor in the combined filtering process as follows:

$$\beta_i = \frac{EDOP_i}{EDOP_1 + EDOP_2 + \cdots + EDOP_N + EDOP_m} \tag{4.20}$$

Obviously, β_i can satisfy information on conservation principles and possess a very intuitive physical sense, namely, the line reflects the estimated performance of sub-filters to improve the fusion performance of the global filter by adjusting the proportion of the local estimates information in the global estimates information. Especially when the performance degradation of a subsystem makes its local estimation error covariance matrix such a singular huge increase that its adaptive information distribution can make the combined filter participating of strong robustness and fault tolerance.

5. Summary

The chapter focuses on non-standard multi-sensor information fusion system with each kind of nonlinear, uncertain and correlated factor, which is widely popular in actual application, because of the difference of measuring principle and character of sensor as well as measuring environment.

Aiming at the above non-standard factors, three resolution schemes based on semi-parameter modeling, multi model fusion and self-adaptive estimation are relatively advanced, and moreover, the corresponding fusion estimation model and algorithm are presented.

(1) By introducing semi-parameter regression analysis concept to non-standard multi-sensor state fusion estimation theory, the relational fusion estimation model and parameter-non-parameter solution algorithm are established; the process is to separate model error brought by nonlinear and uncertainty factors with semi-parameter modeling method and then weakens the influence to the state fusion estimation precision; besides, the conclusion is proved in theory that the state estimation obtained in this algorithm is the optimal fusion estimation.

(2) Two multi-model fusion estimation methods respectively based on multi-model adaptive estimation and interacting multiple model fusion are researched to deal with nonlinear and time-change factors existing in multi-sensor fusion system and moreover to realize the optimal fusion estimation for the state.

(3) Self-adaptive fusion estimation strategy is introduced to solve local dependency and system parameter uncertainty existed in multi-sensor dynamical system and moreover to realize the optimal fusion estimation for the state. The fusion model for federal filter and its optimality are researched; the fusion algorithms respectively in relevant or irrelevant for each sub-filter are presented; the structure and algorithm scheme for federal filter are designed; moreover, its estimation performance was also analyzed, which was influenced by information allocation factors greatly. So the selection method of information allocation factors was discussed, in this chapter, which was dynamically and self-adaptively determined according to the eigenvalue square decomposition of the covariance matrix.

6. Reference

Hall L D, Llinas J. Handbook of Multisensor Data Fusion. Bcoa Raton, FL, USA: CRC Press, 2001

Bedworth M, O'Brien J. the Omnibus Model: A New Model of Data Fusion. IEEE Transactions on Aerospace and Electronic System, 2000, 15(4): 30-36

Heintz, F., Doherty, P. A Knowledge Processing Middleware Framework and its Relation to the JDL Data Fusion Model. Proceedings of the 7th International Conference on Information Fusion, 2005, pp: 1592-1599

Llinas J, Waltz E. Multisensor Data Fusion. Norwood, MA: Artech House, 1990

X. R. Li, Yunmin Zhu, Chongzhao Han. Unified Optimal Linear Estimation Fusion-Part I: Unified Models and Fusion Rules. Proc. 2000 International Conf. on Information Fusion, July 2000

Jiongqi Wang, Haiyin Zhou, Deyong Zhao, el. State Optimal Estimation with Nonstandard Multi-sensor Information Fusion. System Engineering and Electronics, 2008, 30(8): 1415-1420

Kennet A, Mayback P S. Multiple Model Adaptive Estimation with Filter Pawning. IEEE Transaction on Aerospace Electron System, 2002, 38(3): 755-768

Bar-shalom, Y., Campo, L. The Effect of The Common Process Noise on the Two-sensor Fused-track Covariance. IEEE Transaction on Aerospace and Electronic Systems, 1986, Vol.22: 803-805

Morariu, V. I, Camps, O. I. Modeling Correspondences for Multi Camera Tracking Using Nonlinear Manifold Learning and Target Dynamics. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June, 2006, pp: 545-552

Stephen C, Stubberud, Kathleen. A, et al. Data Association for Multisensor Types Using Fuzzy Logic. IEEE Transaction on Instrumentation and Measurement, 2006, 55(6): 2292-2303

Hammerand, D. C. ; Oden, J. T. ; Prudhomme, S. ; Kuczma, M. S. Modeling Error and Adaptivity in Nonlinear Continuum System, NTIS No: DE2001-780285/XAB

Crassidis. J Letal.A. Real-time Error Filter and State Estimator.AIAA-943550.1994:92-102

Flammini, A, Marioli, D. et al. Robust Estimation of Magnetic Barkhausen Noise Based on a Numerical Approach. IEEE Transaction on Instrumentation and Measurement, 2002, 16(8): 1283-1288

Donoho D. L., Elad M. On the Stability of the Basis Pursuit in the Presence of Noise. http: //www-stat.stanford.edu/-donoho/reports.html

Sun H Y, Wu Y. Semi-parametric Regression and Model Refining. Geospatial Information Science, 2002, 4(5): 10-13

Green P.J., Silverman B.W. Nonparametric Regression and Generalized Linear Models. London: CHAPMAN and HALL, 1994

Petros Maragos, FangKuo Sun. Measuring the Fractal Dimension of Signals: Morphological Covers and Iterative Optimization. IEEE Trans. On Signal Processing, 1998(1): 108~121

G, Sugihara, R.M.May. Nonlinear Forecasting as a Way of Distinguishing Chaos From Measurement Error in Time Series, Nature, 1990, 344: 734-741

Roy R, Paulraj A, kailath T. ESPRIT--Estimation of Signal Parameters Via Rotational Invariance Technique. IEEE Transaction Acoustics, Speech, Signal Processing, 1989, 37:984-98

Aufderheide B, Prasad V, Bequettre B W. A Compassion of Fundamental Model-based and Multi Model Predictive Control. Proceeding of IEEE 40th Conference on Decision and Control, 2001: 4863-4868

Aufderheide B, Bequette B W. A Variably Tuned Multiple Model Predictive Controller Based on Minimal Process Knowledge. Proceedings of the IEEE American Control Conference, 2001, 3490-3495

X. Rong Li, Jikov, Vesselin P. A Survey of Maneuvering Target Tracking-Part V: Multiple-Model Methods. Proceeding of SPIE Conference on Signal and Data Proceeding of Small Targets, San Diego, CA, USA, 2003

T.M. Berg, et al. General Decentralized Kalman filters. Proceedings of the American Control Conference, Mayland, June, 1994, pp.2273-2274

Nahin P J, Pokoski Jl. NCTR Plus Sensor Fusion of Equals IFNN. IEEE Transaction on AES, 1980, Vol. AES-16, No.3, pp.320-337

Bar-Shalom Y, Blom H A. The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients. IEEE Transaction on Aut. Con, 1988, AC-33: 780-783

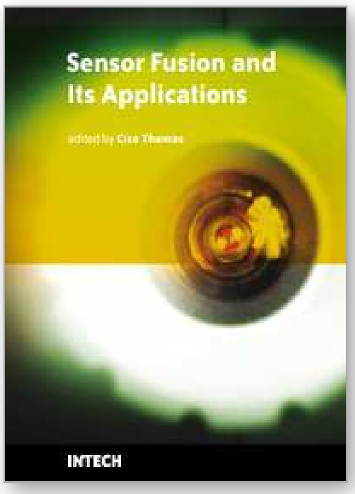
X.Rong Li, Vesselin P. Jilkov. A Survey of Maneuvering Target Tracking-Part I: Dynamic Models. IEEE Transaction on Aerospace and Electronic Systems, 2003, 39(4): 1333-1361

Huimin Chen, Thiaglingam Kirubarjan, Yaakov Bar-Shalom. Track-to-track Fusion Versus Centralized Estimation: Theory and Application. IEEE Transactions on AES, 2003, 39(2): 386-411

F.M.Ham. Observability, Eigenvalues and Kalman Filtering. IEEE Transactions on Aerospace and Electronic Systems, 1982, 19(2): 156-164

Xianda, Zhang. Matrix Analysis and Application. Tsinghua University Press, 2004, Beijing

X. Rong Li. Information Fusion for Estimation and Decision. International Workshop on Data Fusion in 2002, Beijing, China



Sensor Fusion and its Applications

Edited by Ciza Thomas

ISBN 978-953-307-101-5
Hard cover, 488 pages
Publisher Sciyo
Published online 16, August, 2010
Published in print edition August, 2010