Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Automatica 40 (2004) 1017–1023
automatica
www.elsevier.com/locate/automatica

Brief paper

Multi-sensor optimal information fusion Kalman filter☆

Shu-Li Sun*, Zi-Li Deng

Department of Automation, Heilongjiang University, P.O. Box 229, Harbin 150080, People's Republic of China

Received 4 November 2002; received in revised form 27 December 2003; accepted 16 January 2004

Abstract

This paper presents a new multi-sensor optimal information fusion criterion weighted by matrices in the linear minimum variance sense, it is equivalent to the maximum likelihood fusion criterion under the assumption of normal distribution. Based on this optimal fusion criterion, a general multi-sensor optimal information fusion decentralized Kalman filter with a two-layer fusion structure is given for discrete time linear stochastic control systems with multiple sensors and correlated noises. The first fusion layer has a netted parallel structure to determine the cross covariance between every pair of faultless sensors at each time step. The second fusion layer is the fusion center that determines the optimal fusion matrix weights and obtains the optimal fusion filter. Comparing it with the centralized filter, the result shows that the computational burden is reduced, and the precision of the fusion filter is lower than that of the centralized filter when all sensors are faultless, but the fusion filter has fault tolerance and robustness properties when some sensors are faulty. Further, the precision of the fusion filter is higher than that of each local filter. Applying it to a radar tracking system with three sensors demonstrates its effectiveness.

Keywords: Multisensor; Information fusion; Linear minimum variance; Maximum likelihood; Optimal information fusion Kalman filter; Fault tolerance; Radar tracking system

1. Introduction

The information fusion Kalman filtering theory has been studied and widely applied to integrated navigation systems for maneuvering targets, such as airplanes, ships, cars and robots. When multiple sensors measure the states of the same stochastic system, generally we have two different types of methods to process the measured sensor data. The first method is the centralized filter (Willner, Chang, & Dunn, 1976) where all measured sensor data are communicated to a central site for processing. The advantage of this method is that it involves minimal information loss. However, it can result in severe computational overhead due to overloading of the filter with more data than it can handle. Consequently, the overall centralized filter may be unreliable or suffer from poor accuracy and stability when there is severe data fault. The second method is the decentralized filter where the information from local estimators can yield the global optimal or suboptimal state estimate according to certain information fusion criterion. The advantage of this method is that the requirements of communication and memory space at the fusion center are broadened, and the parallel structures would lead to increase in the input data rates. Furthermore, decentralization leads to easy fault detection and isolation. However, the precision of the decentralized filter is generally lower than that of the centralized filter when there is no data fault. Various decentralized and parallel versions of the Kalman filter and their applications have been reported (Kerr, 1987; Hashmipour, Roy, & Laub, 1988), including the federated square-root filter of Carlson (1990). But to some extent, his filter has conservatism because of using the upper bound of the process noise variance matrix instead of process noise variance matrix itself, and assuming the initial estimation errors to be uncorrelated. Roy and Iltis (1991) give a decentralized static filter for the linear system with correlated measurement noises. Kim (1994) gives an optimal fusion filter under the assumption of normal distribution based on the maximum likelihood sense for systems

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Marco Campi under the direction of Editor Tamer Başar.

* Corresponding author. Tel.: +86-0451-86622910.
E-mail address: sunsl@hlju.edu.cn (S.-L. Sun).

0005-1098/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2004.01.014

with multiple sensors, and assumes the process noise to be independent of measurement noises. Saha (1996,1998) discusses the steady-state fusing problem. Deng and Qi (2000) give a multi-sensor fusion criterion weighted by scalars. But the assumption for the estimation errors among the local subsystems to be uncorrelated does not accord with the general case. Qiang and Harris (2001) discuss the functional equivalence of two measurement fusion methods, where the second method requires the measurement matrices to be of identical size.

In this paper, the result of the maximum likelihood fusion criterion under the normal density function, which is presented by Kim (1994), is re-derived as an optimal information fusion criterion weighted by matrices in the linear minimum variance sense. Based on the fusion criterion, an optimal information fusion decentralized filter with fault tolerance and robustness properties is given for discrete time-varying linear stochastic control systems with multiple sensors and correlated noises. It has a two-layer fusion structure. The first fusion layer has a netted parallel structure to determine the cross covariance between every pair of faultless sensors at each time step. The second fusion layer is the fusion center that fuses the estimates and variances of all local subsystems, and the cross covariance among the local subsystems from the first fusion layer to determine the optimal matrix weights and yield the optimal fusion filter.

2. Problem formulation

Consider the discrete time-varying linear stochastic control system with l sensors

$$x(t+1) = \Phi(t)x(t) + B(t)u(t) + \Gamma(t)w(t), \tag{1}$$

$$y_i(t) = H_i(t)x(t) + v_i(t), \quad i = 1,2,\ldots,l, \tag{2}$$

where $x(t) \in R^n$ is the state, $y_i(t) \in R^{m_i}$ is the measurement, $u(t) \in R^{\mathrm{p}}$ is a known control input, $w(t) \in R^{\mathrm{r}}$, $v_i(t) \in R^{m_i}$ are white noises, and $\Phi(t), B(t), \Gamma(t), H_i(t)$ are time-varying matrices with compatible dimensions.

In the following, I_n denotes the $n \times n$ identity matrix, and 0 denotes the zero matrix with compatible dimension.

Assumption 1. $w(t)$ and $v_i(t)$, $i = 1,2,\ldots,l$ are correlated white noises with zero mean and

$$E\left\{\begin{bmatrix} w(t) \\ v_i(t) \end{bmatrix}\begin{bmatrix} w^{\mathrm{T}}(k) & v_i^{\mathrm{T}}(k) \end{bmatrix}\right\} = \begin{bmatrix} Q(t) & S_i(t) \\ S_i^{\mathrm{T}}(t) & R_i(t) \end{bmatrix}\delta_{tk},$$

$$E[v_i(t)v_j^{\mathrm{T}}(k)] = S_{ij}(t)\delta_{tk}, \quad i \neq j, \tag{3}$$

where the symbol E denotes the mathematical expectation, the superscript T denotes the transpose, and δ_{tk} is the Kronecker delta function.

Assumption 2. The initial state $x(0)$ is independent of $w(t)$ and $v_i(t)$, $i = 1,2,\ldots,l$, and

$$Ex(0) = \mu_0, \quad E[(x(0)-\mu_0)(x(0)-\mu_0)^{\mathrm{T}}] = P_0. \tag{4}$$

Our aim is to find the optimal (i.e. linear minimum variance) information fusion Kalman filter $\hat{x}_o(t|t)$ of the state $x(t)$ based on measurements $(y_i(t),\ldots,y_i(1))$, $i=1,2,\ldots,l$, which will satisfy the following performances:

(a) Unbiasedness, namely, $E\hat{x}_o(t|t) = Ex(t)$.
(b) Optimality, namely, to find the optimal matrix weights $\bar{A}_i(t)$, $i = 1,2,\ldots,l$, to minimize the trace of the fusion filtering error variance, i.e. $\mathrm{tr}[P_o(t|t)] = \min\{\mathrm{tr}[P(t|t)]\}$, where the symbol tr denotes the trace of a matrix, $P_o(t|t)$ denotes the variance of the optimal fusion filter with matrix weights and $P(t|t)$ denotes the variance of an arbitrary fusion filter with matrix weights.

3. Optimal information fusion criterion in the linear minimum variance sense

In 1994, Kim provides a maximum likelihood fusion criterion under the assumption of standard normal distribution. Here we will derive the same result in the linear minimum variance sense, where the restrictive assumption of normal distribution is avoided. For simplicity, time t is dropped in the following derivation.

Theorem 1. *Let $\hat{x}_i$, $i = 1,2,\ldots,l$ be unbiased estimators of an n-dimensional stochastic vector x. Let the estimation errors be $\tilde{x}_i = x - \hat{x}_i$, $i=1,2,\ldots,l$. Assume that $\tilde{x}_i$ and $\tilde{x}_j$, $(i \neq j)$ are correlated, and the variance and cross covariance matrices are denoted by P_{ii} (i.e. P_i) and P_{ij}, respectively. Then the optimal fusion (i.e. linear minimum variance) estimator with matrix weights is given as*

$$\hat{x}_o = \bar{A}_1\hat{x}_1 + \bar{A}_2\hat{x}_2 + \cdots + \bar{A}_l\hat{x}_l, \tag{5}$$

where the optimal matrix weights $\bar{A}_i$, $i=1,2,\ldots,l$ are computed by

$$\bar{A} = \Sigma^{-1}e(e^{\mathrm{T}}\Sigma^{-1}e)^{-1}, \tag{6}$$

where $\Sigma = (P_{ij})$, $i,j=1,2,\ldots,l$ is an $nl \times nl$ symmetric positive definite matrix, $\bar{A} = [\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_l]^{\mathrm{T}}$ and $e = [I_n,\ldots,I_n]^{\mathrm{T}}$ are both $nl \times n$ matrices. The corresponding variance of the optimal information fusion estimator is computed by

$$P_o = (e^{\mathrm{T}}\Sigma^{-1}e)^{-1} \tag{7}$$

and we have the conclusion $P_o \leqslant P_i$, $i = 1,2,\ldots,l$.

Proof. Introducing the synthetically unbiased estimator

$$\hat{x} = A_1\hat{x}_1 + A_2\hat{x}_2 + \cdots + A_l\hat{x}_l, \tag{8}$$

where $A_i, i=1,2,\ldots,l$ are arbitrary matrices. From the unbiasedness assumption, we have $E\hat{x}=Ex$, $E\hat{x}_i=Ex$, $i=1,2,\ldots,l$. Taking the expectation of both sides of (8) yields

$$A_1+A_2+\cdots+A_l=I_n. \tag{9}$$

From (8) and (9) we have the fusion estimation error $\tilde{x}=x-\hat{x}=\sum_{i=1}^{l}A_i(x-\hat{x}_i)=\sum_{i=1}^{l}A_i\tilde{x}_i$. Let $A=[A_1,A_2,\ldots,A_l]^{\mathrm{T}}$, so the error variance matrix of the fusion estimator is

$$P=E(\tilde{x}\tilde{x}^{\mathrm{T}})=A^{\mathrm{T}}\Sigma A \tag{10}$$

and the performance index $J=\mathrm{tr}(P)$ becomes

$$J=\mathrm{tr}(A^{\mathrm{T}}\Sigma A). \tag{11}$$

The problem is to find the optimal fusion matrix weights $\bar{A}_i$, $i=1,2,\ldots,l$ under restriction (9) to minimize the performance index (11). Applying the Lagrange multiplier method, we introduce the auxiliary function

$$F=J+2\,\mathrm{tr}[\Lambda(A^{\mathrm{T}}e-I_n)], \tag{12}$$

where $\Lambda=(\lambda_{ij})$ is an $n\times n$ matrix. Set $\partial F/\partial A|_{A=\bar{A}}=0$, and note that $\Sigma^{\mathrm{T}}=\Sigma$, we have

$$\Sigma\bar{A}+e\Lambda=0. \tag{13}$$

Combining (13) with (9) yields the matrix equation as

$$\begin{pmatrix}\Sigma & e\\ e^{\mathrm{T}} & 0\end{pmatrix}\begin{pmatrix}\bar{A}\\ \Lambda\end{pmatrix}=\begin{pmatrix}0\\ I_n\end{pmatrix}, \tag{14}$$

where $\Sigma, e, \bar{A}$ are defined above. Σ is a symmetric positive definite matrix, hence $e^{\mathrm{T}}\Sigma^{-1}e$ is nonsingular. Using the formula of the inverse matrix (Xu, 2001), we have

$$\begin{pmatrix}\bar{A}\\ \Lambda\end{pmatrix}=\begin{pmatrix}\Sigma & e\\ e^{\mathrm{T}} & 0\end{pmatrix}^{-1}\begin{pmatrix}0\\ I_n\end{pmatrix}=\begin{pmatrix}\Sigma^{-1}e(e^{\mathrm{T}}\Sigma^{-1}e)^{-1}\\ -(e^{\mathrm{T}}\Sigma^{-1}e)^{-1}\end{pmatrix} \tag{15}$$

which yields (6). Substituting (6) into (10) yields the optimal information fusion estimation error variance matrix as (7). The following is the derivation of $P_o\leqslant P_i$. In fact, applying Schwartz matrix inequality, we have

$$\begin{aligned}P_o&=(e^{\mathrm{T}}\Sigma^{-1}e)^{-1}=[(\Sigma^{-1/2}e)^{\mathrm{T}}(\Sigma^{1/2}e_i)]^{\mathrm{T}}\\&\quad\times[(\Sigma^{-1/2}e)^{\mathrm{T}}(\Sigma^{-1/2}e)]^{-1}[(\Sigma^{-1/2}e)^{\mathrm{T}}(\Sigma^{1/2}e_i)]\\&\leqslant(\Sigma^{1/2}e_i)^{\mathrm{T}}(\Sigma^{1/2}e_i)=P_i,\end{aligned} \tag{16}$$

where $e_i=[0,\ldots,I_n,\ldots,0]^{\mathrm{T}}$ is an $nl\times n$ matrix whose ith block place is an I_n and others are $n\times n$ block zero matrices. The condition of the equality to hold in (16) is that $P_i=P_{ji}$, $j=1,2,\ldots,l$.

This proof is completed. □

4. Optimal information fusion decentralized Kalman filter with a two-layer fusion structure

Under Assumptions 1 and 2, the ith local sensor subsystem of system (1)–(2) with multiple sensors has the local optimal Kalman filter (Anderson & Moore, 1979)

$$\hat{x}_i(t+1|t+1)=\hat{x}_i(t+1|t)+K_i(t+1)\varepsilon_i(t+1), \tag{17}$$

$$\hat{x}_i(t+1|t)=\bar{\Phi}_i(t)\hat{x}_i(t|t)+B(t)u(t)+J_i(t)y_i(t), \tag{18}$$

$$\varepsilon_i(t+1)=y_i(t+1)-H_i(t+1)\hat{x}_i(t+1|t), \tag{19}$$

$$\begin{aligned}K_i(t+1)&=P_i(t+1|t)H_i^{\mathrm{T}}(t+1)[H_i(t+1)\\&\quad\times P_i(t+1|t)H_i^{\mathrm{T}}(t+1)+R_i(t+1)]^{-1},\end{aligned} \tag{20}$$

$$\begin{aligned}P_i(t+1|t)&=\bar{\Phi}_i(t)P_i(t|t)\bar{\Phi}_i^{\mathrm{T}}(t)+\Gamma(t)\\&\quad\times[Q(t)-S_i(t)R_i^{-1}(t)S_i^{\mathrm{T}}(t)]\Gamma^{\mathrm{T}}(t),\end{aligned} \tag{21}$$

$$P_i(t+1|t+1)=[I_n-K_i(t+1)H_i(t+1)]P_i(t+1|t), \tag{22}$$

$$\hat{x}_i(0|0)=\mu_0,\quad P_i(0|0)=P_0, \tag{23}$$

where $\bar{\Phi}_i(t)=\Phi(t)-J_i(t)H_i(t)$, $J_i(t)=\Gamma(t)S_i(t)R_i^{-1}(t)$. $P_i(t|t)$ and $P_i(t+1|t)$ are the filtering and first-step prediction error variance matrices, respectively, $K_i(t)$ is the filtering gain matrix, and $\varepsilon_i(t)$ is the innovation process, for the ith sensor subsystem, $i=1,2,\ldots,l$.

Since filtering errors of the ith and the jth subsystems are correlated, we have the following Theorem 2.

Theorem 2. *Under Assumptions 1 and 2, the local Kalman filtering error cross covariance between the ith and the jth sensor subsystems has the following recursive form:*

$$\begin{aligned}P_{ij}(t+1|t+1)&=[I_n-K_i(t+1)H_i(t+1)]\\&\quad\times\{\bar{\Phi}_i(t)P_{ij}(t|t)\bar{\Phi}_j^{\mathrm{T}}(t)\\&\quad+\Gamma(t)Q(t)\Gamma^{\mathrm{T}}(t)-J_j(t)R_j(t)J_j^{\mathrm{T}}(t)\\&\quad-J_i(t)R_i(t)J_i^{\mathrm{T}}(t)+J_i(t)S_{ij}(t)J_j^{\mathrm{T}}(t)\\&\quad+\bar{\Phi}_i(t)K_i(t)[S_{ij}(t)J_j^{\mathrm{T}}(t)\end{aligned}$$

$$
\begin{aligned}
&-S_i^{\mathrm{T}}(t)\Gamma^{\mathrm{T}}(t)] + [J_i(t)S_{ij}(t)\\
&-\Gamma(t)S_j(t)]K_j^{\mathrm{T}}(t)\bar{\Phi}_j^{\mathrm{T}}(t)\}\\
&\times[I_n - K_j(t+1)H_j(t+1)]^{\mathrm{T}}\\
&+K_i(t+1)S_{ij}(t+1)K_j^{\mathrm{T}}(t+1), \qquad (24)
\end{aligned}
$$

where $P_{ij}(t|t)$, $i,j=1,2,\dots,l$ $(i\neq j)$ are the filtering error cross covariance matrices between the ith and jth sensor subsystems, and the initial values $P_{ij}(0|0)=P_0$.

Proof. For the ith sensor subsystem, we have the filtering error equation as follows (Anderson & Moore, 1979):

$$
\begin{aligned}
\tilde{x}_i(t+1|t+1) = {}& [I_n - K_i(t+1)H_i(t+1)][\bar{\Phi}_i(t)\tilde{x}_i(t|t)\\
&+\bar{w}_i(t)] - K_i(t+1)v_i(t+1), \qquad (25)
\end{aligned}
$$

where $\bar{w}_i(t)=\Gamma(t)w(t)-J_i(t)v_i(t)$, $\tilde{x}_i(t|t)=x(t)-\hat{x}_i(t|t)$, $i=1,2,\dots,l$. Since the filtering error $\tilde{x}_i(t|t)$ consists of linear combination of $(w(t-1),\dots,w(0),x(0),v_i(t),\dots,v_i(1))$, applying the projection property (Anderson & Moore, 1979), we have $\tilde{x}_i(t|t)\perp v_j(t+1)$, $i,j=1,2,\dots,l$, where $\perp$ denotes orthogonality. Using Assumption 1 and (25) yields the filtering error cross covariance matrix between the ith and the jth sensor subsystems as follows:

$$
\begin{aligned}
P_{ij}(t+1|t+1) = {}& [I_n - K_i(t+1)H_i(t+1)]\\
&\times\{\bar{\Phi}_i(t)P_{ij}(t|t)\bar{\Phi}_j^{\mathrm{T}}(t)\\
&+E[\bar{w}_i(t)\bar{w}_j^{\mathrm{T}}(t)] + \bar{\Phi}_i(t)E[\tilde{x}_i(t|t)\bar{w}_j^{\mathrm{T}}(t)]\\
&+E[\bar{w}_i(t)\tilde{x}_j^{\mathrm{T}}(t|t)]\bar{\Phi}_j^{\mathrm{T}}(t)\}\\
&\times[I_n - K_j(t+1)H_j(t+1)]^{\mathrm{T}}\\
&+K_i(t+1)S_{ij}(t+1)K_j^{\mathrm{T}}(t+1), \qquad (26)
\end{aligned}
$$

where

$$
\begin{aligned}
E[\bar{w}_i(t)\bar{w}_j^{\mathrm{T}}(t)] = {}& \Gamma(t)Q(t)\Gamma^{\mathrm{T}}(t) - \Gamma(t)S_j(t)J_j^{\mathrm{T}}(t)\\
&-J_i(t)S_i^{\mathrm{T}}(t)\Gamma^{\mathrm{T}}(t) + J_i(t)S_{ij}(t)J_j^{\mathrm{T}}(t), \qquad (27)
\end{aligned}
$$

$$
\begin{aligned}
E[\tilde{x}_i(t|t)\bar{w}_j^{\mathrm{T}}(t)] = {}& E\{[(I_n - K_i(t)H_i(t))\\
&\times\tilde{x}_i(t|t-1) - K_i(t)v_i(t)]\\
&\times[\Gamma(t)w(t) - J_j(t)v_j(t)]^{\mathrm{T}}\}\\
= {}& K_i(t)S_{ij}(t)J_j^{\mathrm{T}}(t) - K_i(t)S_i^{\mathrm{T}}(t)\Gamma^{\mathrm{T}}(t). \qquad (28)
\end{aligned}
$$

In (28), $\tilde{x}_i(t|t-1)\perp w(t)$ and $\tilde{x}_i(t|t-1)\perp v_j(t)$ are used. In a similar manner, we have

$$
E[\bar{w}_i(t)\tilde{x}_j^{\mathrm{T}}(t|t)] = J_i(t)S_{ij}(t)K_j^{\mathrm{T}}(t) - \Gamma(t)S_j(t)K_j^{\mathrm{T}}(t). \qquad (29)
$$

Substituting (27)–(29) into (26) yields (24). Using (4) and (23) yields the initial value $P_{ij}(0|0)=P_0$. □

Corollary 1 (Bar-Shalom, 1981). *In Theorem 2, if $S_i(t)=0$, $S_{ij}(t)=0$, the cross covariance $P_{ij}(t+1|t+1)$ is simply given by*

$$
\begin{aligned}
P_{ij}(t+1|t+1) = {}& [I_n - K_i(t+1)H_i(t+1)]\\
&\times[\Phi(t)P_{ij}(t|t)\Phi^{\mathrm{T}}(t) + \Gamma(t)Q(t)\Gamma^{\mathrm{T}}(t)]\\
&\times[I_n - K_j(t+1)H_j(t+1)]^{\mathrm{T}}. \qquad (30)
\end{aligned}
$$

Proof. This follows directly from Theorem 2. □

Based on Theorems 1 and 2, we easily obtain the following corollary.

Corollary 2. *For system (1)–(2) under Assumptions 1 and 2, we have the optimal information fusion decentralized Kalman filter as*

$$
\begin{aligned}
\hat{x}_o(t|t) = {}& \bar{A}_1(t)\hat{x}_1(t|t) + \bar{A}_2(t)\\
&\times\hat{x}_2(t|t) + \cdots + \bar{A}_l(t)\hat{x}_l(t|t), \qquad (31)
\end{aligned}
$$

where $\hat{x}_i(t|t)$, $i=1,2,\dots,l$, are computed by (17)–(23), the optimal matrix weights $\bar{A}_i(t)$, $i=1,2,\dots,l$ are computed by (6) and the optimal fusion variance $P_o(t|t)$ is computed by (7). The filtering error variance $P_i(t|t)$ and cross covariance $P_{ij}(t|t)$, $(i\neq j)$ of the local subsystems are computed by (22) and (24), respectively.

Proof. This follows directly from Theorems 1 and 2. □

The optimal information fusion decentralized filter (31) has a two-layer fusion structure as shown in Fig. 1.

In Fig. 1, every sensor subsystem independently estimates the states, respectively and makes fault detection. The classical approaches for fault detection have WSSR (Willsky, 1976) and U verification (Mehra & Peschon, 1971). If any sensor subsystem is faulty by detection, it is isolated and restored. Otherwise, it is sent to the first fusion layer with a netted parallel structure in which the estimation errors of every pair of sensors are fused to determine the cross covariance between them at each time step, at the same time the estimates and variances are sent to the second fusion layer. The second fusion layer is the final fusion center in which the estimates and variance matrices of all faultless local subsystems, and the cross covariance matrices among the faultless local subsystems from the first fusion layer are fused by Theorem 1 to determine the optimal matrix weights and obtain the optimal fusion filter. On the other hand, after the faulty sensors are restored, they can be re-joined the parallel fusion structure. Since the decentralized structure is used, the computational burden in the fusion center is reduced, and the fault tolerance and reliability is assured.

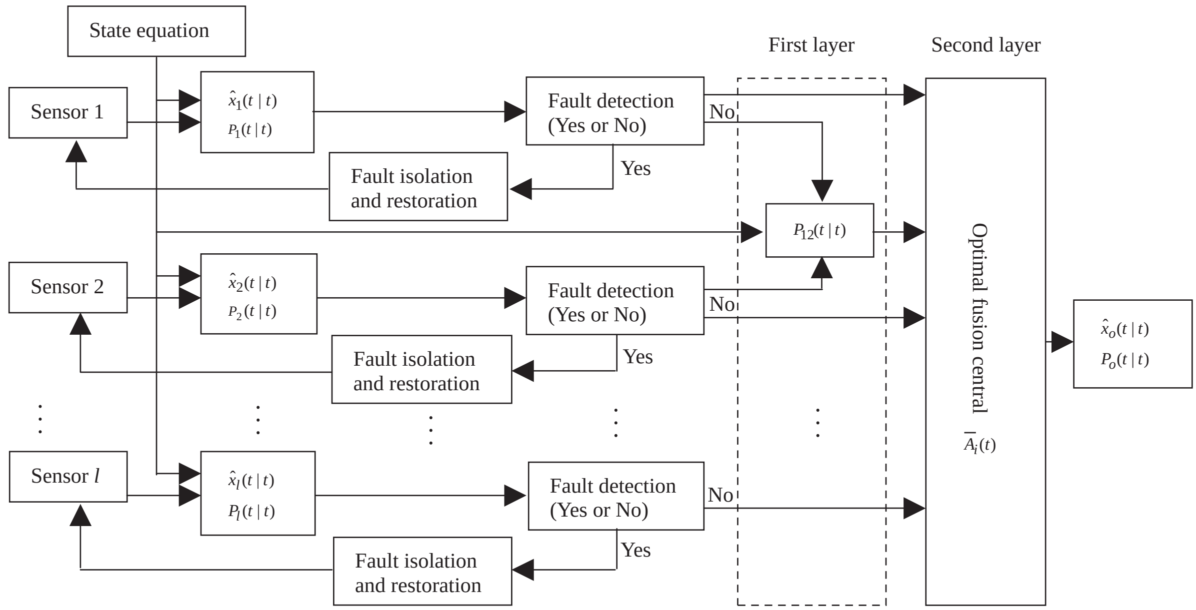

Fig. 1. The optimal information fusion decentralized Kalman filter with a two-layer fusion structure.

5. Simulation example—optimal fusion decentralized Kalman tracking filter

Consider the radar tracking system with three sensors

$$x(t+1)=\begin{bmatrix}1 & T & T^2/2\\ 0 & 1 & T\\ 0 & 0 & 1\end{bmatrix}x(t)+\begin{bmatrix}0\\0\\1\end{bmatrix}w(t), \tag{32}$$

$$y_i(t)=H_ix(t)+v_i(t), \qquad v_i(t)=\alpha_iw(t)+\xi_i(t),$$

$$i=1,2,3, \tag{33}$$

where T is the sampling period. The state is $x(t)=[s(t)\ \dot{s}(t)\ \ddot{s}(t)]^{\mathrm{T}}$, where $s(t)$, $\dot{s}(t)$ and $\ddot{s}(t)$ are the position, velocity and acceleration, respectively, of the target at time t. $y_i(t)$, $i=1,2,3$ are the measurement signals, $v_i(t)$, $i=1,2,3$ are the measurement noises, respectively, of three sensors, which are correlated with Gaussian white noise $w(t)$ with mean zero and variance σ_w^2. The coefficients α_i are constant scalars, and $\xi_i(t)$, $i=1,2,3$ are Gaussian white noises with mean zeros and variance matrices $\sigma_{\xi_i}^2$, and are independent of $w(t)$. Our aim is to find the optimal information fusion decentralized Kalman filter $\hat{x}_o(t|t)$.

In the simulation, setting $T=0.01$; $\sigma_w^2=1$, $\sigma_{\xi_1}^2=5$, $\sigma_{\xi_2}^2=8$, $\sigma_{\xi_3}^2=10$; $\alpha_1=0.5$, $\alpha_2=0.8$, $\alpha_3=0.4$; $H_1=[1,0,0]$, $H_2=[0,1,0]$, $H_3=[0,0,1]$, the initial value $x(0)=0$, $P_0=0.1I_3$, and we take 300 samples.

For every single sensor subsystem, applying (17)–(23), respectively, we can obtain the local optimal Kalman filters $\hat{x}_i(t|t)$ and corresponding variances $P_i(t|t)$, $i=1,2,3$. Applying the two-layer fusion structure in Section 4, we have the optimal information fusion filter $\hat{x}_o(t|t)$ and corresponding variance $P_o(t|t)$. To compare it with the centralized filter, the centralized filter with variance $P_c(t|t)$ is also computed. We select appropriate coordinate to show their differences clearly. From Fig. 2, we see that the precision of the optimal fusion decentralized filter is higher than that of every local filter, but is lower than that of the centralized filter when all sensors are faultless. Though the precision of local filters 1 and 2 is lower, the decentralized filter has better precision and is close to the centralized filter.

To formulate the fault tolerance and robustness properties of the fusion filter, we assume that the first sensor is faulty, so that the measurement equation is that $y_1(t)=H_1x(t)+v_1(t)+f(t)$, where $f(t)$ satisfies $f(t)=0$ $(t<100)$; $f(t)=0.15t$ $(100\leqslant t<200)$; $f(t)=0$ $(t\geqslant 200)$. The values of $f(t)$ show that the first sensor appears faulty at $t=100$ and it is restored at $t=200$. Applying the two-layer fusion structure in Section 4 and WSSR (Willsky, 1976) verification, in Fig. 3, the fusion filter is denoted by dotted lines, the centralized filter is denoted by dash–dot lines, and the real value is denoted by solid line. From the simulation figures, we see that the centralized filter diverges since the first sensor appears faulty at $100\leqslant t<200$. But the decentralized filter can still track the target. It shows that the information fusion decentralized filter with a two-layer fusion structure has better fault tolerance and robustness properties when any sensor is faulty.

6. Conclusions

For the maximum likelihood fusion criterion under the assumption of standard normal distribution presented by

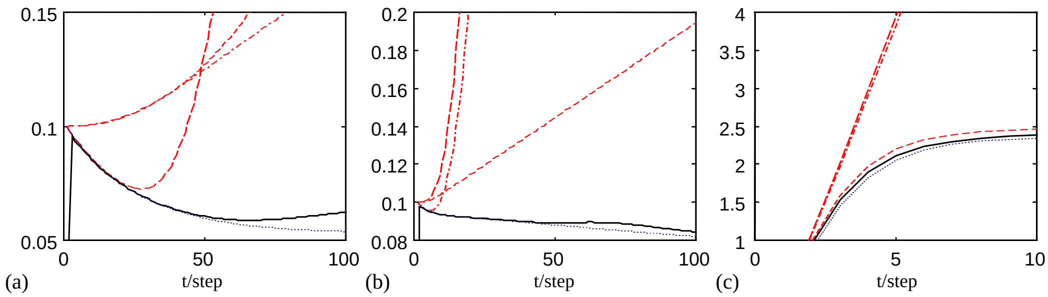

Fig. 2. Comparison of the precision of local filters, decentralized filter and centralized filter when all sensors are faultless: (a) filtering error variance of position $s(t)$; (b) filtering error variance of velocity $\dot{s}(t)$; (c) filtering error variance of acceleration $\ddot{s}(t)$. Local filter 1, – – –; local filter 2, -·-·-·; local filter 3, - - - - -; decentralized filter, —; centralized filter, ...

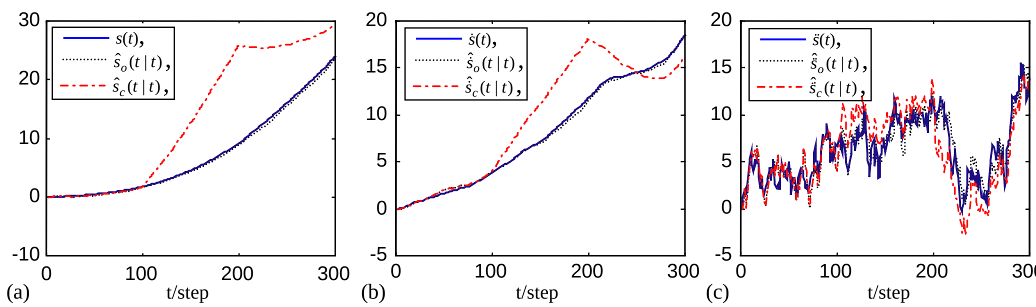

Fig. 3. The tracking performance comparison of the decentralized filter and centralized filter when the first sensor is faulty: (a) position $s(t)$ and Kalman filters $\hat{s}_o(t|t)$, $\hat{s}_c(t|t)$; (b) velocity $\dot{s}(t)$ and Kalman filters $\hat{\dot{s}}_o(t|t)$, $\hat{\dot{s}}_c(t|t)$; (c) acceleration $\ddot{s}(t)$ and Kalman filters $\hat{\ddot{s}}_o(t|t)$, $\hat{\ddot{s}}_c(t|t)$.

Kim (1994), this paper gives a new derivation by using Lagrange multiplier method and new interpretation in the linear minimum variance sense. Based on this fusion criterion, a multi-sensor optimal information fusion decentralized Kalman filter with a two-layer fusion structure is given for discrete time varying linear stochastic control systems with multiple sensors and correlated noises. It has the following properties:

(i) The new derivation of the optimal fusion criterion in the linear minimum variance sense avoids the assumption of normal distribution (Kim, 1994).
(ii) It can solve the optimal fusion problem for systems with multiple sensors and correlated noises.
(iii) The cross covariance between the ith and the jth sensor subsystems is given for systems with correlated noises.
(iv) It avoids the conservatism in the use of the upper bound of the process noise variance instead of process noise variance itself (Carlson, 1990).
(v) It can process the fusion problem when the measurement matrices are of different sizes, thus avoiding the restriction of the measurement matrices to be the same size (Qiang & Harris, 2001).
(vi) Communication of the feedback from the fusion center is avoided.
(vii) A two-layer fusion structure with fault tolerance and robustness properties is given. The netted parallel structure to determine the cross covariance between every pair of sensors is presented.

Acknowledgements

This work was supported by Natural Science Foundation of China under Grant NSFC-60374026. The authors are grateful to the editor and referees for the valuable comments and suggestions.

References

Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering*. Englewood Cliffs, NJ: Prentice-Hall.

Bar-Shalom, Y. (1981). On the track-to-track correlation problem. *IEEE Transactions on Automatic Control*, *AC-26*(2), 571–572.

Carlson, N. A. (1990). Federated square root filter for decentralized parallel processes. *IEEE Transactions on Aerospace and Electronic Systems*, *26*(3), 517–525.

Deng, Z. L., & Qi, R. B. (2000). Multi-sensor information fusion suboptimal steady-state Kalman filter. *Chinese Science Abstracts*, *6*(2), 183–184.

Hashmipour, H. R., Roy, S., & Laub, A. J. (1988). Decentralized structures for parallel Kalman Filtering. *IEEE Transactions on Automatic Control*, *AC-33*(1), 88–93.

Kerr, T. H. (1987). Decentralized filtering and redundancy management. *IEEE Transactions on Aerospace and Electronic Systems*, *23*(6), 83–119.

Kim, K. H. (1994). Development of track to track fusion algorithm. *Proceeding of the American control conference*, Maryland, June (pp. 1037–1041).

Mehra, R. K., & Peschon, J. (1971). An innovation approach to fault detection and diagnosis in dynamic systems. *Automatica*, *7*, 637–640.

Qiang, G., & Harris, C. J. (2001). Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion. *IEEE Transactions on Aerospace and Electronic Systems*, *37*(1), 273–280.

Roy, S., & Iltis, R. A. (1991). Decentralized linear estimation in correlated measurement noise. *IEEE Transactions on Aerospace and Electronic Systems*, *27*(6), 939–941.

Saha, R. K. (1996). Track to track fusion with dissimilar sensors. *IEEE Transactions on Aerospace and Electronic Systems*, *32*(3), 1021–1029.

Saha, R. K. (1998). An efficient algorithm for multisensor track fusion. *IEEE Transactions on Aerospace and Electronic Systems*, *34*(1), 200–210.

Willner, D., Chang, C. B., & Dunn, K. P. (1976). Kalman filter algorithm for a multisensor system. In *Proceedings of IEEE conference on decision and control*, Clearwater, Florida, Dec. (pp. 570–574).

Willsky, A. S. (1976). A survey of design method for failure detection in dynamic systems. *Automatica*, *12*, 601–611.

Xu, N. S. (2001). *Stochastic signal estimation and system control*. Beijing: Beijing Industry University Press.

Shu-Li Sun was born in Heilongjiang, China in 1971. He received his B.Sc. degree in department of mathematics and M.Sc. degree in department of automation from Heilongjiang University in 1996 and 1999, respectively. From 1999, he worked as a faculty member at Heilongjiang University. Currently he is working for his Ph.D. in Harbin Institute of Technology. His main research interests are state estimation, signal processing and information fusion.

Zi-Li Deng was born in Heilongjiang, China in 1938. He graduated in department of mathematics from Heilongjiang University in 1962. Currently he worked as a professor at Heilongjiang University. His main research interests are state estimation, signal processing and information fusion. He has published more than 200 papers and 5 books in these areas.