Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Aerospace Science and Technology 8 (2004) 57–62
Aerospace Science and Technology
www.elsevier.com/locate/aescte

Multi-sensor optimal information fusion Kalman filters with applications

Shu-li Sun [a,b,*]

[a] *Deep Space Exploration Research Center, Harbin Institute of Technology, Harbin 150001, China*
[b] *Department of Automation, Heilongjiang University, Harbin 150080, China*

Received 19 November 2002; accepted 6 August 2003

Abstract

Using Kalman filtering theory, a new multi-sensor optimal information fusion algorithm weighted by matrices is presented in the linear minimum variance sense, which is equivalent to the maximum likelihood fusion algorithm under the assumption of normal distributions. The algorithm considers the correlation among local estimation errors, and it involves the inverse of certain matrix with high dimension. Another two new multi-sensor suboptimal information fusion algorithms weighted by vectors and weighted by scalars are given for reducing the computational burden and increasing the real-time property. Based on these fusion algorithms, the multi-sensor optimal and suboptimal information fusion Kalman filters with two-layer fusion structures are given. The simulation researches of the comparisons among them as well as the centralized filter in a radar tracking system with three sensors show their effectiveness.

Keywords: Multi-sensor; Optimal information fusion; Distributed Kalman filter; Cross-covariance

1. Introduction

The information fusion Kalman filtering theory, which yields the combination of information fusion theory and Kalman filtering theory, has been widely applied in integrated navigation systems for maneuvering targets, for example airplane, ship, automobile and robot etc. When multiple sensors measure the state for 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 [9] where all measured sensor data are communicated to the central site for processing. The advantage of this method is that there is a minimal information loss. However, it can result in severe computational problem due to overloading 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 distributed filter where the local estimators from all sensors can yield the global optimal or suboptimal state estimator according to certain information fusion criterion. The advantages of this method are that the requirement of memory space to the fusion center is broadened, and the parallel structures can increase the input data rates, furthermore, the decentralization makes for easy fault detection and isolation. Hence the distributed filter is widely studied. However, the precision of the distributed filter is generally lower than that of the centralized filter when there is not data fault. Recently, Various distributed and parallel versions of the Kalman filter and applications have been reported in [3–8]. Hashemipour et al. [6] give a parallel Kalman filtering structure for multisensor networks amenable to parallel processing. Carlson [3] presents the famous federated square root filter that assumes the initial estimation error cross-covariance matrices among the local subsystems to be zero, i.e. the local estimation errors among the local subsystems are uncorrelated at the initial time, which doesn't accord with the general case. To some extent, it has the conservatism because of using the upper bound of the process noise variance matrix instead of the process noise variance matrix itself. Deng and Qi [5] give a fusion algorithm weighted by scalars for the systems with multiple sensors, but the assumption for state estimation errors among the local subsystems to be uncorrelated doesn't accord with the general case. Kim [7] and H. Chen et al. [4] respectively give the multi-sensor optimal information fusion estimator in the maximum likelihood sense under the assumption of normal distributions. H. Chen et al. [4] indicate that it is also the weighted least squares estimate under the assumption of normal distributions. Qiang Gan and Chris J. Harris [8] dis-

* Corresponding author.
E-mail address: sunsl@hlju.edu.cn (S.-l. Sun).

1270-9638/$ – see front matter
doi:10.1016/j.ast.2003.08.003

cuss the functional equivalence of two measurement fusion methods, where the first method is the centralized filter, but the second method combines the multisensor data based on a minimum-mean-square-error criterion, which requires the measurement matrices to be of identical size.

This paper presents a new multi-sensor optimal information fusion algorithm weighted by matrices in the linear minimum variance sense. Under the assumption of normal distributions, it is equivalent to the maximum likelihood fusion estimate [4,7]. However, here it has the generality since the assumption of normal distributions is avoided. Another two optimal fusion algorithms weighted by vectors and weighted by scalars in the linear minimum variance sense are also given for reducing the computational burden, but they are suboptimal compared with one weighted by matrices. Based on these fusion algorithms, the corresponding fusion filters with two-layer fusion structures are given. Since the distributed structure is used, the fusion estimator has a better reliability.

The paper is organized as follows: first the problem formulation and preliminary lemmas are given in Section 2. Section 3 presents three new optimal information fusion algorithms. The optimal fusion Kalman filters with two-layer fusion structures are given in Section 4. The comparison researches among the new fusion algorithms as well as the centralized filter in a radar tracking system with three sensors are shown in Section 5. Finally, the conclusions are given.

2. Problem formulation and lemmas

Consider the discrete linear stochastic system with multiple sensors

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

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

where $x(t) \in R^n$ is the state, $y_i(t) \in R^{m_i}$, $i = 1, 2, \ldots, l$, are the measurements, Φ, Γ, H_i are the constant matrices with compatible dimensions, and l is the number of sensors.

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

Assumption 1. $w(t) \in R^r$ and $v_i(t) \in R^{m_i}$, $i = 1, 2, \ldots, l$, are the independent white noises with zero mean and variance matrices Q and R_i, respectively, i.e.

$$\mathrm{E}\big[v_i(t)v_j^{\mathrm{T}}(k)\big] = 0, \quad i \neq j;\ \forall t, k,$$

$$\mathrm{E}\left\{\begin{bmatrix} w(t) \\ v_i(t) \end{bmatrix}\begin{bmatrix} w^{\mathrm{T}}(k) & v_i^{\mathrm{T}}(k) \end{bmatrix}\right\} = \begin{bmatrix} Q & 0 \\ 0 & R_i \end{bmatrix}\delta_{tk}, \tag{3}$$

where E is the expectation, the superscript T denotes the transpose, 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

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

Here we assume that all sensors are faultless. The problem is to find the optimal (i.e. linear minimum variance) information fusion distributed Kalman filter $\hat{x}_0(t \mid t)$ of the state $x(t)$ based on the measurements $(y_i(t), \ldots, y_i(1))$, $i = 1, 2, \ldots, l$, such that

(a) $\hat{x}_0(t \mid t) = \bar{A}_1(t)\hat{x}_1(t \mid t) + \cdots + \bar{A}_l(t)\hat{x}_l(t \mid t)$, where $\bar{A}_i(t)$, $i = 1, 2, \ldots, l$, are the weights, and $\hat{x}_i(t \mid t)$, $i = 1, 2, \ldots, l$, are the local filters.
(b) Unbiasedness, namely, $\mathrm{E}[\hat{x}_0(t \mid t)] = \mathrm{E}[x(t)]$.
(c) Optimality, namely, to minimize the trace of the error variance matrix of the fusion estimator, i.e. $\min\{\mathrm{tr}(P(t \mid t))\}$, where $P(t \mid t)$ denotes the variance of arbitrary a fusion filter, and the symbol "tr" denotes the trace of a matrix.

Lemma 1. *Under the Assumptions* 1 *and* 2, *for the ith sensor subsystem of the system* (1)–(2) *with multiple sensors, we have the local optimal Kalman filters* [1]

$$\hat{x}_i(t+1 \mid t+1) = \hat{x}_i(t+1 \mid t) + K_i(t+1)\varepsilon_i(t+1), \tag{5}$$

$$\hat{x}_i(t+1 \mid t) = \Phi\hat{x}_i(t \mid t), \tag{6}$$

$$\varepsilon_i(t+1) = y_i(t+1) - H_i\hat{x}_i(t+1 \mid t), \tag{7}$$

$$K_i(t+1) = P_i(t+1 \mid t)H_i^{\mathrm{T}} \times \big[H_i P_i(t+1 \mid t)H_i^{\mathrm{T}} + R_i\big]^{-1}, \tag{8}$$

$$P_i(t+1 \mid t) = \Phi P_i(t \mid t)\Phi^{\mathrm{T}} + \Gamma Q \Gamma^{\mathrm{T}}, \tag{9}$$

$$P_i(t+1 \mid t+1) = [I_n - K_i(t+1)H_i]P_i(t+1 \mid t), \tag{10}$$

$$\hat{x}_i(0 \mid 0) = \mu_0, \quad P_i(0 \mid 0) = P_0, \tag{11}$$

where $P_i(t \mid t)$ and $P_i(t+1 \mid t)$ are the filtering and first-step prediction error variances respectively, $K_i(t)$ is the filtering gain matrix, and $\varepsilon_i(t)$ is the innovation process, for the ith sensor subsystem, $i = 1, 2, \ldots, l$.

Lemma 2. *Under the Assumptions* 1 *and* 2, *the local filtering error cross-covariance between the ith and the jth sensor subsystems for the system* (1)–(2) *with multiple sensors has the following recursive form* [2]

$$P_{ij}(t+1 \mid t+1) = \big[I_n - K_i(t+1)H_i\big], \big[\Phi P_{ij}(t \mid t)\Phi^{\mathrm{T}} + \Gamma Q \Gamma^{\mathrm{T}}\big]\big[I_n - K_j(t+1)H_j\big]^{\mathrm{T}}, \tag{12}$$

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

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

In 1994, Kim [7] gave a maximum likelihood fusion estimate under the assumption of normal distributions. Now we will give the same result based on the linear minimum variance sense, where the assumption of normal distributions is avoided. For simplicity, time t is dropped in the derivation.

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

$$\hat{x}_0 = \bar{A}_1\hat{x}_1 + \bar{A}_2\hat{x}_2 + \cdots + \bar{A}_l\hat{x}_l \tag{13}$$

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

$$\bar{A} = \Sigma^{-1}e\big(e^{\mathrm{T}}\Sigma^{-1}e\big)^{-1} \tag{14}$$

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

$$P_0 = \big(e^{\mathrm{T}}\Sigma^{-1}e\big)^{-1} \tag{15}$$

and we have the relation $\mathrm{tr}(P_0)\leqslant \mathrm{tr}(P_{ii}),\ i=1,2,\ldots,l$.

Proof. Introducing the synthetically unbiased estimator

$$\hat{x} = A_1\hat{x}_1 + A_2\hat{x}_2 + \cdots + A_l\hat{x}_l \tag{16}$$

where $A_i,\ i=1,2,\ldots,l$, are arbitrary matrices. Let the variance matrix of error $\tilde{x}=x-\hat{x}$ be P. From the unbiased assumption, we have $\mathrm{E}\hat{x}=\mathrm{E}x$, $\mathrm{E}\hat{x}_i=\mathrm{E}x,\ i=1,2,\ldots,l$. Taking the expectation in (16) yields

$$A_1 + A_2 + \cdots + A_l = I_n. \tag{17}$$

From (16) and (17) we have the fusion estimation error

$$\tilde{x} = x-\hat{x} = \sum_{i=1}^{l} A_i(x-\hat{x}_i) = \sum_{i=1}^{l} A_i\tilde{x}_i \tag{18}$$

so the error variance matrix of the fusion estimator is

$$P = \mathrm{E}\big(\tilde{x}\tilde{x}^{\mathrm{T}}\big) = \sum_{i,j=1}^{l} A_i P_{ij} A_j^{\mathrm{T}} \tag{19}$$

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

$$J = \sum_{i,j=1}^{l} \mathrm{tr}\big(A_i P_{ij} A_j^{\mathrm{T}}\big). \tag{20}$$

Our goal is to find the optimal weighting matrices $\bar{A}_i,\ i=1,2,\ldots,l$, under the restriction (17), to minimize the performance index (20). Applying LaGrange multiplier method, we introduce an auxiliary function

$$F = J + \sum_{j=1}^{n}\left[\lambda_j^{\mathrm{T}}\left(\sum_{i=1}^{l} A_i - I_n\right)e_j\right] \tag{21}$$

where $\lambda_j=[\lambda_{1j},\ldots,\lambda_{nj}]^{\mathrm{T}}$ is a n-dimension vector, $e_j=[0,\ldots,0,1,0,\ldots,0]^{\mathrm{T}}$ is a n-dimension vector where the jth element is 1 and others are 0.

Setting $\partial F/\partial A_i|_{A_i=\bar{A}_i}=0,\ i=1,2,\ldots,l$, we have

$$P_{i1}\bar{A}_1^{\mathrm{T}} + P_{i2}\bar{A}_2^{\mathrm{T}} + \cdots + P_{il}\bar{A}_l^{\mathrm{T}} + \frac{1}{2}\Lambda^{\mathrm{T}} = 0 \tag{22}$$

where $\Lambda=(\lambda_{ij})_{n\times n}$. Setting $U=\frac{1}{2}\Lambda^{\mathrm{T}}$ and combining (22) with (17) yield the matrix equation as

$$\begin{pmatrix}\Sigma & e\\ e^{\mathrm{T}} & 0\end{pmatrix}\begin{pmatrix}\bar{A}\\ U\end{pmatrix} = \begin{pmatrix}0\\ I_n\end{pmatrix}, \tag{23}$$

where $\Sigma, e, \bar{A}$ are defined in front. Σ is a symmetric positive definite matrix since the estimation errors of the local subsystems are not completely same generally, hence $e^{\mathrm{T}}\Sigma^{-1}e$ is nonsingular. Using the formula of the inverse matrix [11], we have

$$\begin{pmatrix}\bar{A}\\ U\end{pmatrix} = \begin{pmatrix}\Sigma & e\\ e^{\mathrm{T}} & 0\end{pmatrix}^{-1}\begin{pmatrix}0\\ I_n\end{pmatrix} = \begin{pmatrix}\Sigma^{-1}e(e^{\mathrm{T}}\Sigma^{-1}e)^{-1}\\ -(e^{\mathrm{T}}\Sigma^{-1}e)^{-1}\end{pmatrix}. \tag{24}$$

Using (24) yields (14) and substituting (14) into (19) yields the error variance matrix of the optimal fusion estimator as (15).

Since $\partial^2 F/(\partial A)^2|_{A=\bar{A}} = \Sigma > 0$, $\bar{A}$ is the optimal solution to minimize the performance index $J=\mathrm{tr}(P)$.

Setting $A_i=I_n,\ A_j=0,\ j=1,2,\ldots,l,\ j\neq i$, in (19) yields $\mathrm{tr}(P_0)\leqslant\mathrm{tr}(P_{ii})$. □

Corollary 1. *If the estimation error $\tilde{x}_i$ and $\tilde{x}_j,\ i\neq j$, are uncorrelated, then the optimal matrix weights $\bar{A}_i$ are*

$$\bar{A}_i = \left(\sum_{j=1}^{l} P_{jj}^{-1}\right)^{-1} P_{ii}^{-1},\quad i=1,2,\ldots,l. \tag{25}$$

The optimal information fusion estimator is

$$\hat{x}_0 = \left(\sum_{j=1}^{l} P_{jj}^{-1}\right)^{-1}\left(\sum_{i=1}^{l} P_{ii}^{-1}\hat{x}_i\right) \tag{26}$$

and the error variance of the optimal fusion estimator is

$$P_0 = \left(\sum_{j=1}^{l} P_{jj}^{-1}\right)^{-1}. \tag{27}$$

Proof. From the uncorrelation of estimation errors $\tilde{x}_i$ and $\tilde{x}_j,\ i\neq j$, we have $P_{ij}=0$, so $\Sigma=\mathrm{diag}(P_{ii})$ is a block diagonal matrix in Theorem 1, then using (14), (13) yields (25) and (26), using (15) yields (27). □

From Theorem 1, we see that Σ^{-1} will be a computational burden since Σ is a $nl\times nl$ matrix if nl is larger. For reducing the computational burden and increasing the real-time property, we will give another two fusion algorithms weighted by vectors (i.e. weighted by diagonal matrices) and weighted by scalars.

Theorem 2. *Under the assumptions of Theorem* 1, *we have the optimal fusion estimator weighted by vectors as* (13) *with the diagonal matrix weights $\bar{A}_i,\ i=1,2,\ldots,l$,* *computed by* (14) *with the* $nl \times nl$ *sparse matrix* $\Sigma = (\overline{P}_{ij})_{nl\times nl}$, $i, j = 1, 2, \ldots, l$. $\overline{P}_{ij}$, $i, j = 1, 2, \ldots, l$, *are the diagonal matrices consisting of the diagonal elements of the covariance matrices* P_{ij}. *The error variance of the optimal fusion estimator with vector weights is computed by*

$$P_0 = \left(e^{\mathrm{T}}\Sigma^{-1}e\right)^{-1}e^{\mathrm{T}}\Sigma^{-1}(P_{ij})_{nl\times nl}\Sigma^{-1}e\left(e^{\mathrm{T}}\Sigma^{-1}e\right)^{-1}, \tag{28}$$

and we still have the relation $\mathrm{tr}(P_0) \leqslant \mathrm{tr}(P_{ii})$, $i = 1, 2, \ldots, l$.

Corollary 2. *If the estimation error* $\tilde{x}_i$ *and* $\tilde{x}_j$, $i \neq j$, *are uncorrelated, the optimal diagonal matrix weights* $\bar{A}_i$ *are*

$$\bar{A}_i = \left(\sum_{j=1}^{l}\overline{P}_{jj}^{-1}\right)^{-1}\overline{P}_{ii}^{-1}, \quad i = 1, 2, \ldots, l. \tag{29}$$

The optimal fusion estimator weighted by vectors is

$$\hat{x}_0 = \left(\sum_{j=1}^{l}\overline{P}_{jj}^{-1}\right)^{-1}\sum_{i=1}^{l}\overline{P}_{ii}^{-1}\hat{x}_i \tag{30}$$

and the error variance of the optimal fusion estimator is

$$P_0 = \left(\sum_{j=1}^{l}\overline{P}_{jj}^{-1}\right)^{-1}\left(\sum_{i=1}^{l}\overline{P}_{ii}^{-1}P_{ii}\overline{P}_{ii}^{-1}\right) \times \left(\sum_{j=1}^{l}\overline{P}_{jj}^{-1}\right)^{-1}. \tag{31}$$

Theorem 3. *Under the assumptions of Theorem* 1, *we have the optimal fusion estimator weighted by scalars as* (13) *with the scalar weights* $\bar{A}_i$, $i = 1, 2, \ldots, l$, *computed by* (14) *with the* $l \times l$ *positive definite matrix* $\Sigma = (\mathrm{tr}(P_{ij}))_{l\times l}$, $i, j = 1, 2, \ldots, l$, *and the* l*-dimension vector* $e = [1, 1, \ldots, 1]^{\mathrm{T}}$. *The error variance of the optimal fusion estimator with scalar weights is computed by*

$$P_0 = [\bar{A}_1 I_n, \ldots, \bar{A}_l I_n](P_{ij})_{nl\times nl}[\bar{A}_1 I_n, \ldots, \bar{A}_l I_n]^{\mathrm{T}}, \tag{32}$$

and we still have the relation $\mathrm{tr}(P_0) \leqslant \mathrm{tr}(P_{ii})$, $i = 1, 2, \ldots, l$.

Corollary 3. *If the estimation errors* $\tilde{x}_i$ *and* $\tilde{x}_j$, $i \neq j$, *are uncorrelated, the optimal scalar weights* $\bar{A}_i$ *are*

$$\bar{A}_i = \left(\sum_{j=1}^{l}\frac{1}{\mathrm{tr}\,P_{jj}}\right)^{-1}\frac{1}{\mathrm{tr}\,P_{ii}}, \quad i = 1, 2, \ldots, l. \tag{33}$$

The optimal fusion estimator weighted by scalars is

$$\hat{x}_0 = \left(\sum_{j=1}^{l}\frac{1}{\mathrm{tr}\,P_{jj}}\right)^{-1}\sum_{i=1}^{l}\frac{1}{\mathrm{tr}\,P_{ii}}\hat{x}_i \tag{34}$$

and the error variance of the optimal fusion estimator is

$$P_0 = \left(\sum_{j=1}^{l}\frac{1}{\mathrm{tr}\,P_{jj}}\right)^{-2}\sum_{i=1}^{l}\left(\frac{1}{\mathrm{tr}\,P_{ii}}\right)^{2}P_{ii}. \tag{35}$$

Theorems 2 and 3 can be proved similar to Theorem 1. Theorem 2 gives an optimal information fusion algorithm weighted by vectors based on the linear minimum variance sense. It is a suboptimal fusion algorithm compared to one weighted by general matrices in Theorem 1. However, since the $nl \times nl$ matrix Σ in Theorem 2 is a sparse matrix, its inverse is also a sparse matrix, particularly, when the local estimation errors $\tilde{x}_i$ and $\tilde{x}_j$, $i \neq j$, are uncorrelated, since $\overline{P}_{ii}$, $i = 1, 2, \ldots, l$, are diagonal matrices, so $\overline{P}_{ii}^{-1}$ can easily be computed. Hence the computational burden can be reduced compared to Theorem 1. In a similar manner, Theorem 3 also gives an optimal information fusion algorithm weighted by scalars based on the linear minimum variance sense. It is also a suboptimal fusion algorithm compared to Theorem 1. However, since Σ in Theorem 3 is a $l \times l$ matrix. Particularly, when the local estimation errors are uncorrelated, it only involves the computation of scalars. Hence the computational burden can be reduced obviously.

4. Optimal information fusion distributed filters with a two-layer fusion structure

Theorem 4. *Under the Assumptions* 1 *and* 2, *we have the optimal fusion in a unified form for distributed Kalman filters, which are respectively weighted by matrices, vectors and scalars for the system* (1)–(2) *with multiple sensors as*

$$\hat{x}_0(t \mid t) = \bar{A}_1(t)\hat{x}_1(t \mid t) + \cdots + \bar{A}_l(t)\hat{x}_l(t \mid t), \tag{36}$$

where $\hat{x}_i(t \mid t)$, $i = 1, 2, \ldots, l$, *are computed by Lemma* 1, *and the optimal weights* $\bar{A}_i(t)$, $i = 1, 2, \ldots, l$, *are computed by* (14). *The minimal fusion variance* $P_0(t \mid t)$ *is computed by* (15) *or* (28) *or* (32). *The local filtering error variance* $P_{ii}(t \mid t)$ (*i.e.* $P_i(t \mid t)$) *and cross-covariance* $P_{ij}(t \mid t)$, $i \neq j$, *of the local subsystems are computed by Lemmas* 1 *and* 2, *respectively.*

Proof. From Theorems 1–3 we obtain Theorem 4. □

The information fusion filters in Theorem 4 all have the following two-layer fusion structures (Fig. 1).

In Fig. 1, the first fusion layer has the netted parallel structure in which the estimation errors of any two sensors are fused to determine the cross-covariance matrix between them. The second fusion layer is the fusion center in which the estimates and error variance matrices of all local subsystems, and the cross-covariance matrices among the local subsystems from the first fusion layer are fused to determine the optimal weights and obtain the optimal fusion filters.

Remark. In Theorem 4, we assume all sensors to be faultless. If some sensors are fault, they will be isolated and restored by fault detection [10]. The rest sensors continue the fusion. We can implement fault detection by adding the element for the detection in every subsystem.

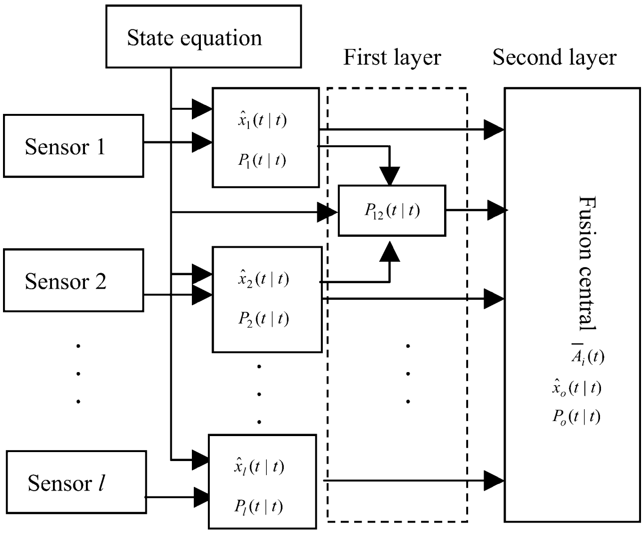

Fig. 1. The two-layer fusion structure of distributed fusion filter.

5. Simulation researches

Consider the radar tracking system with three sensors

$$x(t+1)=\begin{bmatrix}1 & T & T^2/2\\ 0 & 1 & T\\ 0 & 0 & 1\end{bmatrix}x(t)+\begin{bmatrix}0\\0\\1\end{bmatrix}w(t), \tag{37}$$

$$y_i(t)=H_i x(t)+v_i(t),\quad i=1,2,3, \tag{38}$$

where T is the sampling period. The state $x(t)=[s(t),\dot{s}(t),\ddot{s}(t)]^{\mathrm{T}}$ where $s(t)$, $\dot{s}(t)$ and $\ddot{s}(t)$ are the position, velocity and acceleration respectively of the target at time tT. $y_i(t)$, $i=1,2,3$, are the measurement signals. $v_i(t)$, $i=1,2,3$, are the measurement noises of three sensors respectively with zero mean and variances $\sigma_{v_i}^2$, and are independent of Gaussian white noise $w(t)$ with zero mean and variance σ_w^2. Our goal is to find the optimal fusion distributed Kalman filter $\hat{x}_0(t\mid t)$.

In the simulation, setting $T=0.01$ s, $\sigma_w^2=1$, $\sigma_{v_1}^2=8$, $\sigma_{v_2}^2=15$, $\sigma_{v_3}^2=20$, the initial values $x(0)=0$, $P_0=0.1I_3$, we take 200 sampling data.

Respectively, (37) and (38) can bring three local subsystems, and the corresponding measurement matrices are $H_1=[1,0,0]$, $H_2=[0,1,0]$ and $H_3=[0,0,1]$. For every single sensor subsystem 1–3, applying Lemma 1, respectively we can obtain the local optimal Kalman filters $\hat{x}_i(t\mid t)$ and variance matrices $P_i(t\mid t)$, $i=1,2,3$. Applying Theorem 4, we respectively obtain the optimal information fusion Kalman filters $\hat{x}_0^m(t\mid t)$ weighted by matrices, $\hat{x}_0^v(t\mid t)$ weighted by vectors and $\hat{x}_0^s(t\mid t)$ weighted by scalars. The corresponding variance matrices are respectively $P_0^m(t\mid t)$, $P_0^v(t\mid t)$ and $P_0^s(t\mid t)$. To compare with the centralized filter where the observations are federated in one vector observation, the centralized filter $\hat{x}_c(t\mid t)$ with the variance matrix $P_c(t\mid t)$ is also given. We compute the traces of their filtering error variance matrices and list performance indexes at some time steps in Table 1. From Table 1 we have the conclusion of $\mathrm{tr}(P_c(t\mid t))\leqslant \mathrm{tr}(P_0^m(t\mid t))\leqslant \mathrm{tr}(P_0^v(t\mid t))\leqslant \mathrm{tr}(P_0^s(t\mid t))\leqslant \mathrm{tr}(P_i(t\mid t))$, $i=1,2,3$, when all sensors are faultless. We see that the precision of the fusion filter with matrix weights approximates to the centralized filter, but it requires the inverse of a 9×9 matrix at each time step; the precision of the fusion filter with vector weights is lower than that of one with matrix weights, but it requires the inverse of a 9×9 sparse matrix at each time step; while the precision of the fusion filter with scalar weights is lower than that of one with vector weights, is higher than that of any local sensor subsystem, but it only requires the inverse of a 3×3 matrix at each time step. The precision of the fusion filters is dropped with reduction of the computational burden.

6. Conclusions

This paper presents three new multi-sensor optimal information fusion algorithms, which are respectively weighted by matrices, vectors and scalars in the linear minimum variance sense. Based on three fusion algorithms, the optimal information fusion distributed Kalman filters with two-layer fusion structures are given for discrete linear stochastic system with multiple sensors. They have the following properties:

(i) The algorithms can handle the optimal information fusion problems for systems with multiple sensors when the estimation errors of the local subsystems are correlated.
(ii) The optimal fusion criterion weighted by matrices in the linear minimum variance sense avoids the assump-

Table 1
The performance indexes for the filters

Index	tT				
	10	50	100	150	200
$\mathrm{tr}(P_1(t\mid t))$	9.3118	50.3013	62.4860	68.8420	69.2998
$\mathrm{tr}(P_2(t\mid t))$	9.2955	27.3613	29.0053	29.1654	29.2429
$\mathrm{tr}(P_3(t\mid t))$	4.0863	4.3164	4.5403	4.9083	5.4704
$\mathrm{tr}(P_0^s(t\mid t))$	4.0862	4.3090	4.5268	4.8757	5.3900
$\mathrm{tr}(P_0^v(t\mid t))$	4.0685	4.2576	4.3330	4.4329	4.5282
$\mathrm{tr}(P_0^m(t\mid t))$	4.0683	4.2145	4.2389	4.2437	4.2497
$\mathrm{tr}(P_c(t\mid t))$	4.0683	4.2074	4.2235	4.2283	4.2293

tion of normal distributions in [4,7]. It has the generality.

(iii) They avoid the conservative way to use the upper bound of the process noise variance matrix instead of the process noise variance matrix itself and to assume the initial estimation errors among the local subsystems to be uncorrected in [3].

(iv) They can process the optimal fusion problem that the measurement matrices are different sizes, and avoid the restriction of the measurement matrices to be identical size in [8].

(v) A two-layer fusion structure is given. The netted parallel structure is presented to determine the cross-covariance matrix between any two sensors.

Acknowledgements

The author is indebted to the reviewers for the valuable comments and suggestions.

References

[1] B.D.O. Anderson, J.B. Moore, Optimal Filtering, Prentice-Hall, Englewood Cliffs, NJ, 1979.

[2] Y. Bar-Shalom, On the track-to-track correlation problem, IEEE Trans. Automat. Contr. 26 (2) (1981) 571–572.

[3] N.A. Carlson, Federated square root filter for decentralized parallel processes, IEEE Trans. Aerospace Electron. Syst. 26 (3) (1990) 517–525.

[4] H. Chen, T. Kirubarajan, Y. Bar-Shalom, Performance limits of track-to-track fusion vs. centralized estimation: theory and application, in: Proc. 4th Internat. Conf. on Information Fusion, Montreal, Quebec, 2001.

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

[6] H.R. Hashmipour, S. Roy, A.J. Laub, Decentralized structures for parallel Kalman filtering, IEEE Trans. Automat. Contr. 33 (1) (1988) 88–93.

[7] K.H. Kim, Development of track to track fusion algorithm, in: Proc. Amer. Contr. Conf., MD, 1994.

[8] G. Qiang, C.J. Harris, Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion, IEEE Trans. Aerospace Electron. Syst. 37 (1) (2001) 273–280.

[9] D. Willner, C.B. Chang, K.P. Dunn, Kalman filter algorithm for a multisensor system, in: Proc IEEE Conf. Decision and Control, 1976.

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

[11] N.S. Xu, Stochastic Signal Estimate and System Control, Beijing Industry Univ. Press, Beijing, 2001.