Suboptimal State Estimation for Uncertain Multisensor Discrete-Time Linear Stochastic Systems

Tyagi Deepak, Georgy Shevlyakov, Kiseon Kim, Vladimir Shin

School of Information and Mechatronics,
Gwangju Institute of Science and Technology, Gwangju, Republic of Korea
E-mail: {deepak, shev, kskim, vishin} @gist.ac.kr
Fax: +82-62-970-2384

Abstract

The problem of recursive estimation for uncertain multisensor linear discrete-time systems is considered. A new suboptimal filtering algorithm is herein proposed. It is based on the fusion formula with scalar weights for an arbitrary number of local Kalman estimates. Each local Kalman estimate is fused by the minimum mean square error criterion. The filter gains and scalar weights do not depend on current observations; therefore the filter can easily be implemented in real-time. Examples demonstrate the efficiency and high-accuracy of the proposed filter.

Key words:
Linear system, multisensor, Kalman filter, uncertainty, fusion formula

1. Introduction

Estimation of the state of a linear uncertain system with multisensor environment is considered. In structure adaptation many methods are available for the adaptation of such kind of systems [1-3]. In this paper we are interested in the Lainiotis-Kalman filter (LKF) that constitutes a partitioning of the original nonlinear filter into a bank of much simpler **L** local Kalman filters (KF's), where each local filter uses its own system model matched with each possible parameter value [1], [4]. The fusion estimate of the state of LKF is given by a weighted sum of local KF's. However, the optimal LKF's weights depend on sensor observations and it is rather difficult to implement the LKF in real-time, given the dimension of state vector and the number of sensors is large.

In [5], [6], it has been proposed to fuse the local KF's by a weighted sum with matrix weights, which do not depend on sensor observations, and therefore can be pre-computed. However, usage of the matrix weights complicates the algorithm of filter design. In this paper we consider uncertain multisensor systems and propose to fuse the LKF's by the use of scalar weights. The new filter can also help to minimize the computation time and produce real-time state estimation, especially for large number of sensors.

The paper is organized as follows. In Section 2, we set up the estimation problem for multisensor linear systems with observation uncertainties. Section 3 gives the optimal filter for the above system based on the LKF for all stacked sensors. In Section 4, we propose the suboptimal filter (SF), which represents a weighted sum of the local KF's with scalar weights depending only on time instance. Each local KF is fused by the minimum mean-square criterion. Section 5 tests the SF numerically. Conclusions are made in Section 6.

2. Problem Setting

Consider the following model of a multisensor system with observation uncertainties:

$$x_{k+1} = F_k x_k + G_k v_k . \qquad (1)$$

$$y_k^{(i)} = \theta^{(i)} H_k^{(i)} x_k + w_k^{(i)} , \quad i = 1,\dots, N , \quad k = 0,1,2,\dots , \qquad (2)$$

where, as usual, $x_k \in \Re^{\mathbf{n}}$ is the state, $v_k \in \Re^{\mathbf{r}} \sim N(0, Q_k)$ is the normally distributed system noise. The system contains N sensors, $y_k^{(i)} \in \Re^{m_i}$ is the observation vector of i^{th} sensor, and $w_k^{(i)} \in \Re^{m_i} \sim N(0, R_k^{(i)})$ is the normally distributed observation error. The system noise v_k and observation errors $w_k^{(1)},\dots,w_k^{(N)}$ are mutually uncorrelated. The initial state x_0 is normal, $x_0 \sim N(\overline{x}_0, P_0)$. Since the estimated state x_k has no superscript "i", all the local filters (sensors) are working on the same state vector.

In a number of applications, there may be a nonzero probability that the observations contain noise only. Therefore we assume, that the unknown parameters $\theta^{(i)},\ i=1,\dots,N$ are taken from the set $\{0,1\}$. The objective is to estimate x_k.

In order to estimate the state of such system optimally, we can use the LKF [1], [4].

3. The Optimal Lainiotis-Kalman Filter

The LKF is based on the Bayesian approach in which the unknown parameter $\theta^{(i)}$ is assumed to be random with *prior* known probabilities

$$p_0^{(i)} = p\left(\theta^{(i)} = \theta_0^{(i)} = 0\right),\quad p_1^{(i)} = p\left(\theta^{(i)} = \theta_1^{(i)} = 1\right),\\ p_0^{(i)} + p_1^{(i)} = 1,\quad i=1,\dots,N. \tag{3}$$

Let us collect all scalar parameters $\theta^{(1)},\dots,\theta^{(N)}$ into vector. Then we obtain the unknown parameter vector $\Theta\in\Re^N$, which takes $L=2^N$ values, i.e.,

$$\Theta = \left[\theta^{(1)}\ \ \dots\ \ \theta^{(N)}\right]^T,\quad \theta^{(i)} = \begin{cases}\theta_0^{(i)} = 0,\\ \theta_1^{(i)} = 1.\end{cases} \tag{4}$$

With the above as preamble, we can rewrite the multisensor system model (1), (2) in the following form

$$x_{k+1} = F_k x_k + G_k v_k,\\ y_k = \widetilde{H}_k(\Theta) x_k + w_k, \tag{5}$$

where

$$Y_k = \begin{bmatrix} y_k^{(1)} \\ \vdots \\ y_k^{(N)} \end{bmatrix} \in \Re^m,\quad \widetilde{H}_k = \begin{bmatrix} \theta^{(1)} H_k^{(1)} \\ \vdots \\ \theta^{(N)} H_k^{(N)} \end{bmatrix} \in \Re^{m\times n}, \tag{6}$$

$$w_k = \begin{bmatrix} w_k^{(1)} \\ \vdots \\ w_k^{(N)} \end{bmatrix} \in \Re^m,\quad m = m_1 + \dots + m_N.$$

Given the parameter vector Θ belongs to the discrete space (4), i.e., $\Theta=\Theta_i,\ i=1,\dots,2^N$, the optimal mean-square state estimate $\hat{x}_k^{opt}$ represents the weighted sum of the local Kalman estimates

$$\hat{x}_k^{(i)} \equiv \hat{x}_k(\Theta_i),\quad i=1,\dots,L=2^N \tag{7}$$

matched to the linear system (5), (6) at fixed

$$\Theta = \Theta_i = \left[\theta_{i_1}^{(1)}\ \ \dots\ \ \theta_{i_N}^{(N)}\right]^T,\quad i_1,\dots,i_N = 0,1. \tag{8}$$

We have

$$\hat{x}_k^{opt} = \sum_{i=1}^{L} \tilde{c}_k^{(i)} \hat{x}_k^{(i)}, \tag{9}$$

where $\hat{x}_k^{(i)}$ represents the local Kalman estimate determined by the standard KF equations [1], [7]:

$$\begin{aligned}
&\hat{x}_k^{(i)} = F_k \hat{x}_{k-1}^{(i)} + K_k^{(i)}\left[y_k - \widetilde{H}_k^{(i)} F_k \hat{x}_{k-1}^{(i)}\right],\quad \hat{x}_0^{(i)} = \bar{x}_0,\\
&M_k^{(i)} = F_{k-1} P_{k-1}^{(ii)} F_{k-1}^T + G_{k-1} Q_{k-1} G_{k-1}^T,\quad P_0^{(ii)} = P_0,\\
&K_k^{(i)} = M_k^{(i)} \widetilde{H}_k^{(i)^T}\left[\widetilde{H}_k^{(i)} M_k^{(i)} \widetilde{H}_k^{(i)^T} + R_k\right]^{-1},\\
&P_k^{(ii)} = \left(I_n - K_k^{(i)} \widetilde{H}_k^{(i)}\right) M_k^{(i)},\quad \widetilde{H}_k^{(i)} = \theta^{(i)} H_k^{(i)},\\
&R_k = \mathrm{diag}\left[R_k^{(1)}\ \ \cdots\ \ R_k^{(N)}\right],\quad i=1,\dots,L,
\end{aligned} \tag{10}$$

and the weights

$$\tilde{c}_k^{(i)} = p\left(\Theta_i \middle| Y_k\right),\quad Y_k = \{y_0,\dots,y_k\},\quad i=1,\dots,2^N \tag{11}$$

represent *a posteriori* probabilities of Θ_i given Y_k, in their turn defined by the recursive Bayesian formula [1], [6]. As already stated above, the LKF (7)-(11) is effective for solving problems of low dimensions, since it requires calculations of a large number of *a posteriori* probabilities $p\left(\Theta_i \middle| Y_k\right),\ i=1,\dots,2^N$ in real-time.

In this paper we devise an alternative SF for the system (1), (2). This filter does not require calculations of a *posteriori* probabilities $p\left(\Theta_i \middle| Y_k\right)$ at each time instance $k>0$. The obtained suboptimal filtering algorithm reduces the computational burden and on-line computational requirements considerably.

4. The Suboptimal Filter

Identical to the optimal LKF, the SF represents the state estimate as a weighted sum of the local KF's (7), however, the weights depend only on time instances $k > 0$ and are independent of current observations y_k. Consequently, giving an opportunity to design the SF with minimal complexity, which can be easily implemented in real-time, especially in high dimension problems unlike optimal LKF. According to this idea, we have

$$\hat{x}_k^{sub} = \sum_{i=1}^{L} c_k^{(i)} \hat{x}_k^{(i)}, \qquad \sum_{i=1}^{L} c_k^{(i)} = 1, \tag{12}$$

where $c_k^{(1)},\dots,c_k^{(N)}$ are the scalar weights depending only on time instance k and determined by the mean-square criterion

$$J = E\left\| x_k - \hat{x}_k^{sub} \right\|^2 \to \min_{c_k^{(i)}}. \tag{13}$$

Theorem. (i) *The weights* $c_k^{(1)},\dots,c_k^{(L)}$ *satisfy the linear algebraic equations*

$$\begin{gathered}\sum_{i=1}^{L} c_k^{(i)} \operatorname{tr}\left(P_k^{(ij)} + P_k^{(ji)} - P_k^{(iL)} - P_k^{(Li)}\right) = 0,\\ j = 1,\dots,L-1, \quad c_k^{(1)} + \dots + c_k^{(L)} = 1.\end{gathered} \tag{14}$$

(ii) *The overall error covariance*

$$P_k^{sub} = E\left(\left(x_k - \hat{x}_k^{sub}\right)\left(x_k - \hat{x}_k^{sub}\right)^T\right) \tag{15}$$

is given by

$$\begin{gathered}P_k^{sub} = \sum_{i,j=1}^{L} c_k^{(i)} c_k^{(j)} P_k^{(ij)}, \quad P_k^{(ij)} = E\left(\tilde{x}_k^{(i)} \tilde{x}_k^{(j)^T}\right),\\ \tilde{x}_k^{(i)} = x_k - \hat{x}_k^{(i)}, \quad i = 1,\dots,L.\end{gathered} \tag{16}$$

In (13), $\operatorname{tr}(A)$ is the trace of a matrix A.

The proof of Theorem is given in Appendix A.

Note that formulas (14)-(16) depend on the local error covariances $P_k^{(ii)}$ determined by the Riccati equations (10), and the local cross-covariances $P_k^{(ij)}$, $i \neq j$, which satisfy the following recursive equation:

$$\begin{gathered}P_k^{(ij)} = (I_n - K_k^{(i)} \tilde{H}_k^{(i)})\\ \times (F_{k-1} P_{k-1}^{(ij)} F_{k-1}^T + G_{k-1} Q_{k-1} G_{k-1}^T)(I_n - K_k^{(j)} \tilde{H}_k^{(j)})^T,\\ P_0^{(ij)} = P_0, \quad i,j = 1,\dots,L; \quad i \neq j,\end{gathered} \tag{17}$$

where $K_k^{(i)}$ *stands for the local Kalman gains* (10).

The derivation of (17) is given in Appendix B.

Thus, the local Kalman estimates and covariances $\hat{x}_k^{(i)}, P_k^{(ii)}$ (see (10)), the local cross-covariances $P_k^{(ij)}$, $i \neq j$ (see (17)), and the fusion equations (14) completely establish the suboptimal filter.

Remark 1. It is very easy to note that the local Kalman gains $K_k^{(i)}$, the local error cross-covariances $P_k^{(ij)}$, and the weights $c_k^{(i)}$ can be pre-computed, as they do not depend on the present observations Y_k, but only on the noise statistics and system matrices, and on the values Θ_i of the parameter Θ, which are the part of system model (1),(2), (4).

Therefore, once the observation schedule is settled, the real-time implementation of the SF requires only the computation of the local Kalman estimates $\hat{x}_k^{(1)},\dots,\hat{x}_k^{(L)}$ and the final fusion suboptimal estimate $\hat{x}_k^{sub}$.

Remark 2. Since Θ takes a finite number of values (3), (8), the local Kalman estimates (10) are separated for each value of $i = 1,\dots,L$. Each local estimate $\hat{x}_k^{(i)}$ is found independently of the other estimates $\hat{x}_k^{(j)}$, $j = 1,\dots,L;\ j \neq i$, and therefore they can be evaluated in parallel.

5. Examples

Example 1. Consider the scalar linear system

$$x_{k+1} = a x_k + v_k, \quad k = 0,1,2,\dots, \tag{18}$$

where $a = \text{const}$, $v_k \sim N(0,q)$, $x_0 \sim N(\bar{x}_0, \sigma^2)$.

The observation model contains three sensors:

$$y_k^{(i)} = \theta^{(i)} x_k + w_k^{(i)}, \quad i = 1,2,3, \tag{19}$$

where $w_k^{(i)} \sim N(0, r_i)$, $i = 1,2,3$, and the unknown parameters $\theta^{(i)}$, $i = 1,2,3$ take only two values with equal prior probabilities, i.e.,

$$\theta^{(i)} = \begin{cases} \theta_0^{(i)} = 0, & p\left(\theta_0^{(i)}\right) = 0.5, \\ \theta_1^{(i)} = 1, & p\left(\theta_1^{(i)}\right) = 0.5. \end{cases} \tag{20}$$

Formulas (19) and (20) represent an observation model which takes two modes, which are,

$\theta^{(i)} = 1$ (signal-present) and $\theta^{(i)} = 0$ (signal-absent). Here the vector parameter $\Theta = \left[\theta^{(1)} \quad \theta^{(2)} \quad \theta^{(3)}\right]^T$ takes $L = 8$ values, $\Theta = \Theta_i$ as given below:

$$\begin{aligned} &\Theta_1 = [1 \quad 1 \quad 1]^T, \quad \Theta_2 = [1 \quad 1 \quad 0]^T, \\ &\Theta_3 = [1 \quad 0 \quad 1]^T, \quad \Theta_4 = [1 \quad 0 \quad 0]^T, \\ &\Theta_5 = [0 \quad 1 \quad 0]^T, \quad \Theta_6 = [0 \quad 1 \quad 1]^T, \\ &\Theta_7 = [0 \quad 0 \quad 1]^T, \quad \Theta_8 = [0 \quad 0 \quad 0]^T. \end{aligned} \tag{21}$$

The model parameters, noises statistics, and initial conditions are set to

$$\begin{aligned} &a = 0.7, \quad q = 0.025, \quad \bar{x}_0 = 10, \quad \sigma^2 = 2, \\ &r_1 = 1, \quad r_2 = 2, \quad r_3 = 3. \end{aligned} \tag{22}$$

Here we compare two filters: the optimal LKF

$$\hat{x}_k^{opt} = \sum_{i=1}^{8} \tilde{c}_k^{(i)} \hat{x}_k^{(i)}, \quad \hat{x}_k^{(i)} = \hat{x}_k(\Theta_i) \tag{23}$$

and SF

$$\hat{x}_k^{sub} = \sum_{i=1}^{8} c_k^{(i)} \hat{x}_k^{(i)}. \tag{24}$$

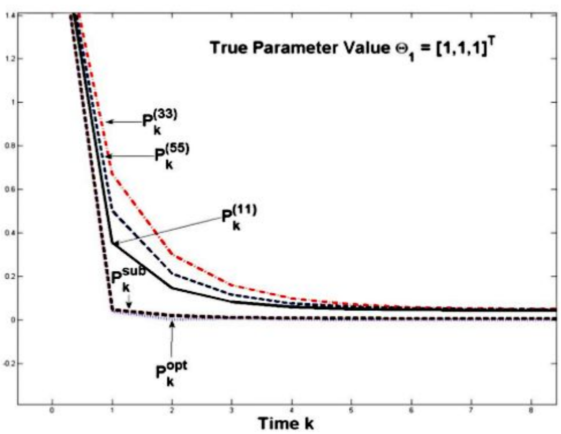

Fig.1 Comparison of MSEs for LKF and SF

The system (18)-(20) is simulated for all values of the parameter (21). Figs. 1-3 present the time histories of the LKF and SF characteristics for the first case, $\Theta = \Theta_1$. Such time histories are perfect analogy for the other cases. Fig.1 shows the overall mean-square errors (MSE's) $P_k^{opt} = E\left(\left(x_k - \hat{x}_k^{opt}\right)^2\right)$ $P_k^{sub} = E\left(\left(x_k - \hat{x}_k^{sub}\right)^2\right)$ and $P_k^{(ii)} = E\left(\left(x_k - \hat{x}_k^{(i)}\right)^2\right)$ for the different values of the vector parameter Θ, but among them only the value $\Theta = \Theta_1$ is the true value in (19).

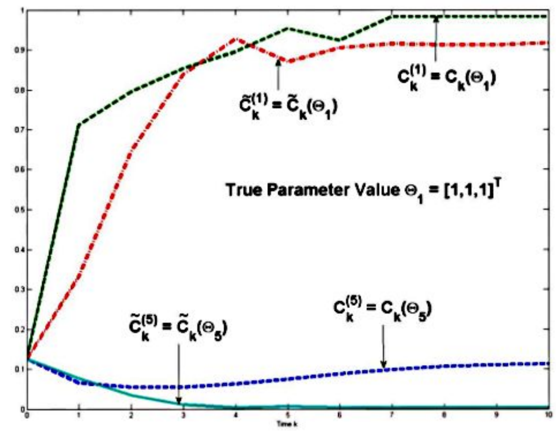

Fig. 2 Comparison of the weights $\tilde{c}_k^{(i)}$ and $c_k^{(i)}$

From Fig.1 it can be seen that P_k^{opt} and P_k^{sub} are very close, it follows from the fact that optimal and suboptimal weights corresponding to the true value $\Theta = \Theta_1$ are very close to each other, i.e., $c_k^{(1)} \approx \tilde{c}_k^{(1)}$ (see Figs.1 and 2). In Fig.3, the comparisons of the optimal and suboptimal estimates show us that performance of the SF is quite similar to the optimal one. This proves that SF is a good alternative for LKF.

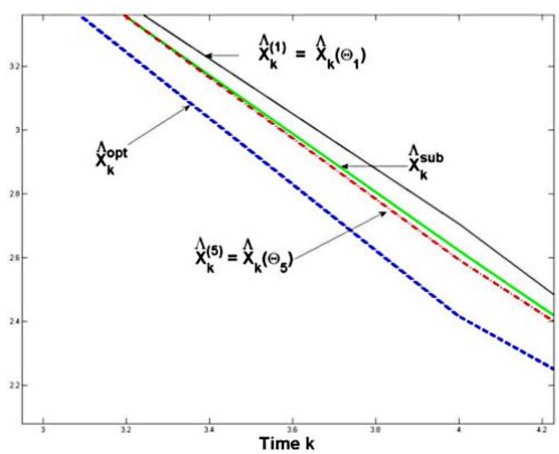

Fig.3 Optimal and suboptimal estimates

Example 2. Consider the 2-dimensional system

$$x_{k+1} = \begin{bmatrix} 1 & 0 \\ 0.9 & 0.4 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 0 \end{bmatrix} v_k, \quad k = 0,1,2,\ldots, \tag{25}$$

where

$$x_k = [x_{1,k} \quad x_{2,k}]^T, \quad v_k \sim N(0, q), \quad x_0 \sim N(\bar{x}_0, P_0).$$

We are measuring the position $x_{1,k}$ and velocity $x_{2,k}$ using two sensors,

$$y_k^{(1)} = \theta^{(1)} x_{1,k} + w_k^{(1)}, \quad y_k^{(2)} = \theta^{(2)} x_{2,k} + w_k^{(2)}, \qquad (26)$$

where $w_k^{(i)} \sim N(0, r), \quad \theta^{(i)} \in \{0,1\}, \quad i = 1,2.$

In this case, the vector parameter $\Theta = [\theta^{(1)} \quad \theta^{(2)}]^T$ takes $L = 4$ values with equal prior probabilities, $\Theta = \Theta_i$, as given below:

$$\Theta_1 = [1 \quad 1]^T, \quad \Theta_2 = [1 \quad 0]^T, \quad \Theta_3 = [0 \quad 1]^T, \quad \Theta_4 = [0 \quad 0]^T. \qquad (27)$$

We compare the optimal LKF and SF. The model parameters, noise statistics, and initial conditions are set to

$$q = 0.2, \quad r = 1, \quad \bar{x}_0 = [10.0 \quad 0.0]^T, \quad P_0 = \mathrm{diag}[2.0 \quad 3.0]. \qquad (28)$$

Figs. 4-6 present the time histories of the filter characteristics for the true value of $\Theta = \Theta_1$. This time history is similar for the other values of Θ.

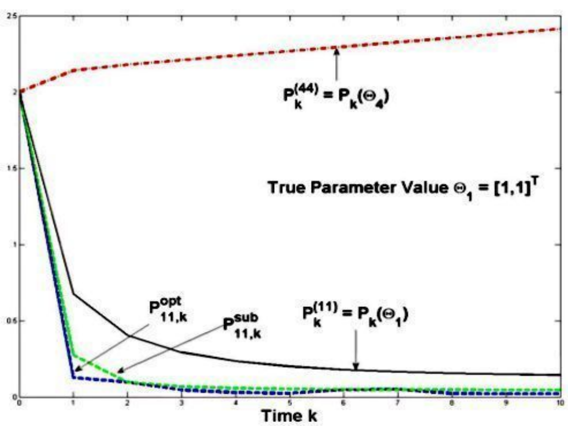

Fig.4 Comparison of MSEs for position $x_{1,k}$

In Figs. 4 and 5, we show the overall optimal and suboptimal MSE's for the position $P_{11,k}^{opt} = E\left((x_{1,k} - \hat{x}_{1,k}^{opt})^2\right)$ and velocity $P_{22,k}^{sub} = E\left((x_{2,k} - \hat{x}_{2,k}^{sub})^2\right)$ respectively, where $P_k^{(ii)} = E\left((x_{i,k} - \hat{x}_k^{(i)})^2\right)$.

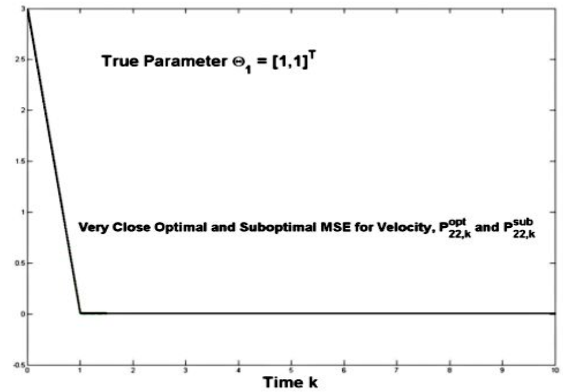

Fig.5 Comparison of MSEs for $x_{2,k}$

From Figs. 4 and 5 we observe that the difference between the optimal MSE ($P_{ii,k}^{opt}$, $i = 1,2$) and suboptimal MSE ($P_{ii,k}^{sub}$, $i = 1,2$) is negligible. It is worth noting in Fig. 4 that the local MSE $P_k^{(44)}$ corresponding to Θ_4, the noise measurement, is very bad because the optimal and suboptimal weights $\tilde{c}_k^{(4)}$, $c_k^{(4)}$ corresponding to Θ_4 are very small (see Fig. 6). Fig. 4 also shows that in the steady state regime local MSE $P_k^{(11)}$, corresponding to $\Theta = \Theta_1$ is very close to optimal MSE P_k^{opt}.

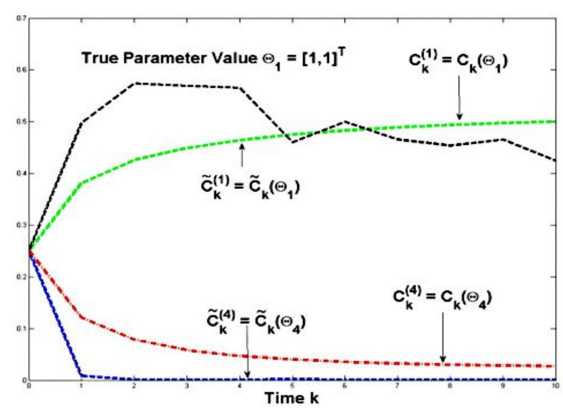

Fig.6 Comparison of the weights $\tilde{c}_k^{(i)}$ and $c_k^{(i)}$

6. Conclusion

In this paper, we have designed a new SF for uncertain multisensor linear discrete-time systems. This filter represents a linear combination of the local KF's with scalar weights depending only on time instance. Each local KF is fused by the minimum mean-square criterion. The proposed filter has a parallel structure and as a result of that is suitable for parallel processing. Simulation results demonstrate the relative loss of accuracy of the SF as compared to the optimal KF.

References

[1] W. Watanabe, Adaptive Estimation and Control: Partitioning Approach, Prentice-Hall, New York, 1991.

[2] P.D. Hanlon, and P.S. Maybeck, "Interrelationship of single-filter and multiple-model adaptive algorithms", IEEE Trans. Aerospace and Electronic Systems, 34(3), 1998, 934-946.

[3] F. Wang, and V. Balakrishnan, "Robust adaptive Kalman filters for linear time-varying systems with stochastic parametric uncertsinties", Proc. Conference Decision & Control, San Diego, California, USA, 1999, 440-444.

[4] D.G. Lainiotis, Partitioned linear estimation algorithm: discrete case, IEEE Trans. Automation Control, 20(3), 1975, 255-257.

[5] D. Choi, V. Shin, J. Ahn, and B. Ahn, "Suboptimal Adaptive Filter for Discrete-Time Stochastic Systems", IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences, vol. E88-A, No. 3, 2005, 620-625.

[6] V. Shin, J. Ahn, and D. Kim, "New Algorithm for Recursive Estimation in Linear Discrete-Time Systems with Unknown Parameters", Inter. Journal of Control, Automation, and Systems, 4(4), 2006, 456-465.

[7] F.L. Lewis, Optimal Estimation with an Introduction to Stochastic Control Theory, John Wiley & Sons, New, York, 1986.

Appendix A: Proof

Using (12), the criterion (13) can be rewritten as

$$\begin{aligned} J &= \mathrm{tr}E\left(\left(x_k - \hat{x}_k^{sub}\right)\left(x_k - \hat{x}_k^{sub}\right)^T\right) \\ &= \mathrm{tr}E\left(\left(\sum_{i=1}^{L} c_k^{(i)}\left(x_k - \hat{x}_k^{(i)}\right)\right)\left(\sum_{j=1}^{L} c_k^{(j)}\left(x_k - \hat{x}_k^{(j)}\right)^T\right)\right) \\ &= \sum_{i,j=1}^{L} \mathrm{tr}\left(c_k^{(i)} c_k^{(j)} E\left(x_k - \hat{x}_k^{(i)}\right)\left(x_k - \hat{x}_k^{(j)}\right)^T\right) \\ &= \mathrm{tr}\left\{\sum_{i,j=1}^{L} c_k^{(i)} c_k^{(j)} P_k^{(ij)}\right\} \to \min_{c_k^{(i)}}. \end{aligned} \tag{A.1}$$

Formula (A.1) gives the overall covariance (16).

Next substituting $c_k^{(L)} = 1 - \left(c_k^{(1)} + \cdots + c_k^{(L-1)}\right)$ into (A.1), we obtain

$$\begin{aligned} J &= \sum_{i,j=1}^{L-1} c_k^{(i)} c_k^{(j)} \mathrm{tr}\left(P_k^{(ij)}\right) + \left(1 - \sum_{h=1}^{L-1} c_k^{(h)}\right)\mathrm{tr}\left(P_k^{(LL)}\right) \\ &+ \left(1 - \sum_{h=1}^{L-1} c_k^{(h)}\right)\sum_{i=1}^{L-1} c_k^{(i)} \mathrm{tr}\left(P_k^{(hL)} + P_k^{(Lh)}\right) \to \min_{c_k^{(i)}}. \end{aligned} \tag{A.2}$$

Differentiating each summand of the criterion J in (A.2) with respect to $c_k^{(1)}, \ldots, c_k^{(L-1)}$ and then setting the result to zero, we obtain the linear algebraic equations (14) for the unknown weights $c_k^{(1)}, \ldots, c_k^{(L)}$.

This completes the proof of Theorem.

Appendix B: Derivation of equation (17)

The derivation of equations (17) is based on the recursive equations for the state x_k and estimate $\hat{x}_k^{(i)}$. Using equations (1), (2), and (10), we obtain recursive equations for the local error $\tilde{x}_k^{(i)} = x_k - \hat{x}_k^{(i)}$, i.e,

$$\begin{aligned} \tilde{x}_k^{(i)} &= F_{k-1} x_{k-1} + G_{k-1} v_{k-1} - F_{k-1} \hat{x}_{k-1}^{(i)} \\ &- K_k^{(i)}\left[y_k - \tilde{H}_k^{(i)} F_{k-1} \hat{x}_{k-1}^{(i)}\right] = F_{k-1} \tilde{x}_{k-1}^{(i)} \\ &- K_k^{(i)}\left[\tilde{H}_k^{(i)}\left(F_{k-1} x_{k-1} + G_{k-1} v_{k-1}\right) + w_k^{(i)} \right. \\ &\left. - \tilde{H}_k^{(i)} F_{k-1} \hat{x}_{k-1}^{(i)}\right] + G_{k-1} v_{k-1} \\ &= \left(I_n - K_k^{(i)} \tilde{H}_k^{(i)}\right) F_{k-1} \tilde{x}_{k-1}^{(i)} \\ &+ \left(I_n - K_k^{(i)} \tilde{H}_k^{(i)}\right) G_{k-1} v_{k-1} - K_k^{(i)} w_k^{(i)}. \end{aligned} \tag{B.1}$$

According to the assumptions that the errors $\tilde{x}_{k-1}^{(i)}$, and white noises v_{k-1} and $w_k^{(i)}$ are mutually uncorrelated, equation (B.1) yields recursive equations (17) for the cross-covariances $P_k^{(ij)} = E\left(\tilde{x}_k^{(i)} \tilde{x}_k^{(j)^T}\right)$, $i \neq j$.