Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 52 (2007) 840–854
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

A new fusion formula and its application to continuous-time linear systems with multisensor environment

Vladimir Shin, Georgy Shevlyakov*, Kiseon Kim

School of Information and Mechatronics, Gwangju Institute of Science and Technology, 1 Oryong-Dong Buk-Gu, Gwangju 500-712, South Korea

Available online 14 December 2006

Abstract

The problem of fusion of local estimates is considered. An optimal mean-square linear combination (fusion formula) of an arbitrary number of local vector estimates is derived. The derived result holds for all dynamic systems with measurements. In particular, for scalar uncorrelated local estimates, the fusion formula represents the well-known result in statistics. The fusion formula is applied to fusion of local Kalman estimates in multisensor filtering problem. Examples demonstrate high accuracy of the proposed fusion formula.

Keywords: Kalman filter; Mean-square error; Multisensor; Data fusion

1. Introduction

In recent years, there has been growing interest to fuse multisensor data in order to increase the accuracy of parameter estimates and system states. This interest is motivated by the availability of different types of sensors which use various characteristics of the optical, infrared, and electromagnetic spectra. In many situations, system states or targets are measured by multisensors. The measurements used in estimation are assigned to a common target as a result of the association process (see Bar-Shalom and Li, 1995; Beugnon et al., 2000; Li et al., 2000; Manyika and Durrant-Whyte, 1994; Zhu, 2003). Here the main problem is how to combine local estimates obtained from different types of sensors.

Suppose local scalar unbiased estimates $\hat{x}^{(1)}, \ldots, \hat{x}^{(N)}$ of an unknown random variable x are obtained by independent means. The associated error variances are $P^{(11)}, \ldots, P^{(NN)}$. Then the optimal (in the mean-square sense) linear combination of the local estimates represents the well-known result used in statistics and stochastic control problems (see Gelb, 1974; Kendall and Stuart, 1963; Lewis, 1986)

$$
\begin{aligned}
&\hat{x} = a^{(1)}\hat{x}^{(1)} + \cdots + a^{(N)}\hat{x}^{(N)}, \quad a^{(1)} + \cdots + a^{(N)} = 1,\\
&a^{(i)} = \frac{1}{P^{(ii)}}\left(\frac{1}{P^{(11)}} + \cdots + \frac{1}{P^{(NN)}}\right)^{-1},\\
&P^{(ii)} = E[(x-\hat{x}^{(i)})^2], \quad \hat{x}^{(i)}, a^{(i)} \in \mathbb{R}, \quad i = 1, \ldots, N.
\end{aligned} \tag{1}
$$

* Corresponding author. Tel.: +82 62 970 3123; fax: +82 62 970 2204.
E-mail address: shev@gist.ac.kr (G. Shevlyakov).

0167-9473/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2006.11.031

The optimal mean-square linear combination of two local vector estimates is proposed by Bar-Shalom and Campo (1986)

$$\begin{aligned}
&\hat{x} = A^{(1)}\hat{x}^{(1)} + A^{(2)}\hat{x}^{(2)}, \quad A^{(1)}, A^{(2)} \in \mathbb{R}^{n\times n},\\
&A^{(1)} = (P^{(22)} - P^{(21)})(P^{(11)} + P^{(22)} - P^{(12)} - P^{(21)})^{-1},\\
&A^{(2)} = (P^{(11)} - P^{(12)})(P^{(11)} + P^{(22)} - P^{(12)} - P^{(21)})^{-1},\\
&P^{(ij)} = E[(x - \hat{x}^{(i)})(x - \hat{x}^{(j)})^{\mathrm{T}}], \quad i, j = 1, 2.
\end{aligned} \tag{2}$$

If the two local estimates $\hat{x}^{(1)}$ and $\hat{x}^{(2)}$ are uncorrelated, i.e., $P^{(12)} = P^{(21)} = 0$ in (2), then we have the Millman formulas which are given in Gelb (1974) and Lewis (1986). The generalization of formulas (2) for $N > 2$ is proposed by Shin and Shevlyakov (2005), where linear algebraic equations for matrix weights $A^{(i)}, \ i = 1, \ldots, N$ were derived.

In real-time applications such as estimation and control of large-scale systems with multisensor environment the usage of the matrix weights is restricted by the toughly limited onboard computer memory, data transmission speed, etc. Therefore, in the paper we focus on the fusion formula (1) for an arbitrary number of vector correlated estimates and scalar weights.

The first goal of this paper is derivation of a new fusion formula (FF) for vector correlated local estimates $\hat{x}^{(1)}, \ldots, \hat{x}^{(N)}$. The new FF represents an optimal mean-square linear combination of local vector estimates with scalar weights $a^{(1)}, \ldots, a^{(N)}$ depending on the cross-covariances between local estimates $P^{(ij)} = E[(x - \hat{x}^{(i)})(x - \hat{x}^{(j)})^T], \ i \neq j$.

The second goal is to show how to apply the FF in the linear filtering problem. In this case, we propose a new suboptimal fusion filter (SFF) based on the FF. Examples exhibit high accuracy of the proposed filter.

This paper is organized as follows. In Section 2, we present the FF with some examples. In Section 3, we propose a suboptimal filter to fuse local Kalman estimates in multisensor continuous-time linear systems. In Section 4, the proposed filter is numerically tested. The efficiency of the filter is studied in the case of a damper harmonic oscillator motion and in other examples. It is shown that the proposed filter is only slightly suboptimal as compared with the optimal Kalman filter (KF). Finally, we conclude our results in Section 5.

2. Fusion formula and its particular cases

Suppose that we have local estimates of an unknown random vector $x \in \mathbb{R}^n$,

$$\hat{x}^{(1)}, \ldots, \hat{x}^{(N)} \in \mathbb{R}^n, \tag{3}$$

where $\mathbb{R}^n$ is an n-dimensional Euclidean space. The associated local error cross-covariances are assumed given

$$P^{(ij)} = E[e^{(i)}e^{(j)^{\mathrm{T}}}], \quad e^{(i)} = x - \hat{x}^{(i)}, \ i, j = 1, \ldots, N. \tag{4}$$

Let us consider the FF representing a linear combination of the local estimates (3)

$$\hat{x}^{\mathrm{FF}} = \sum_{i=1}^{N} a^{(i)}\hat{x}^{(i)}, \quad \sum_{i=1}^{N} a^{(i)} = 1, \tag{5}$$

where $a^{(1)}, \ldots, a^{(N)} \in \mathbb{R}$ are scalar unknown weights determined from the mean-square criterion

$$E[\|x - \hat{x}^{\mathrm{FF}}\|^2] = E[(x - \hat{x}^{\mathrm{FF}})^{\mathrm{T}}(x - \hat{x}^{\mathrm{FF}})] \to \min_{a^{(i)}}. \tag{6}$$

The following results completely define the FF and its fusion error covariance:

$$P^{\mathrm{FF}} = E[e^{\mathrm{FF}}(e^{\mathrm{FF}})^{\mathrm{T}}], \quad e^{\mathrm{FF}} = x - \hat{x}^{\mathrm{FF}}. \tag{7}$$

Theorem 1. *Let* $\hat{x}^{(1)}, \ldots, \hat{x}^{(N)} \in \mathbb{R}^n$ *be the local estimates* (3) *of an unknown vector* $x \in \mathbb{R}^n$. *Then the scalar weights* $a^{(1)}, \ldots, a^{(N)} \in \mathbb{R}$ *satisfy the following linear system of equations*:

$$\sum_{i=1}^{N} a^{(i)} \operatorname{tr}(P^{(ij)} + P^{(ji)} - P^{(iN)} - P^{(Ni)}) = 0, \quad j = 1, \ldots, N-1,$$
$$\sum_{i=1}^{N} a^{(i)} = 1, \tag{8}$$

and they can be explicitly written down as

$$a = (\mathbf{1}_N^{\mathrm{T}} \tilde{P}^{-1} \mathbf{1}_N)^{-1} \tilde{P}^{-1} \mathbf{1}_N, \quad a = [a^{(1)} \ldots a^{(N)}]^{\mathrm{T}}, \quad \tilde{P} = [\operatorname{tr}(P^{(ij)})]_{i,j=1}^{N}. \tag{9}$$

In (8), $\operatorname{tr}(A)$ *is the trace of a matrix and* $\mathbf{1}_N = [1 \ldots 1]^{\mathrm{T}} \in \mathbb{R}^N$.

In general, both the results, namely linear equations (8) and expressions (9) are equivalent, being the implicit and explicit forms of the solution, respectively. However, from the computational point of view, when the number of local estimates is large or the covariance matrices are ill conditioned, linear equations (8) may be more preferable than explicit expressions (9).

Corollary 1. *The fusion error covariance* P^{FF} *is given by*

$$P^{\mathrm{FF}} = \sum_{i,j=1}^{N} a^{(i)} a^{(j)} P^{(ij)}. \tag{10}$$

Corollary 2. *If the local estimates* $\hat{x}^{(1)}, \ldots, \hat{x}^{(N)}$ *are unbiased then the fusion estimate* $\hat{x}^{\mathrm{FF}}$ *is also unbiased*, *i.e.*, $E[\hat{x}^{\mathrm{FF}}] = E[x]$.

The proofs of Theorem 1 and corollaries are given in Appendix A. We would like to thank one of the anonymous reviewers for showing us an elegant way of derivation of formula (9).

Further, we consider two examples of application of the obtained results. In the first example, we consider fusion of two correlated and uncorrelated vector local estimates $\hat{x}^{(1)}$ and $\hat{x}^{(2)}$. In the second one, the scalar well-known FF (1) is generalized on the vector case.

2.1. Example 1: fusion of two correlated and uncorrelated vector estimates

In the particular case when $N = 2$, linear equations (8) for the unknown weights $a^{(1)}$ and $a^{(2)}$ take the form

$$\begin{cases} a^{(1)} \operatorname{tr}(2P^{(11)} - P^{(12)} - P^{(21)}) + a^{(2)} \operatorname{tr}(P^{(21)} + P^{(12)} - 2P^{(22)}) = 0, \\ a^{(1)} + a^{(2)} = 1. \end{cases}$$

Solving these equations, we obtain linear fusion of two correlated vector estimates $\hat{x}^{(1)}$ and $\hat{x}^{(2)}$

$$\hat{x}^{\mathrm{FF}} = a^{(1)} \hat{x}^{(1)} + a^{(2)} \hat{x}^{(2)},$$
$$a^{(1)} = \frac{\operatorname{tr}(2P^{(22)} - P^{(12)} - P^{(21)})}{2 \operatorname{tr}(P^{(11)} + P^{(22)} - P^{(12)} - P^{(21)})},$$
$$a^{(2)} = \frac{\operatorname{tr}(2P^{(11)} - P^{(12)} - P^{(21)})}{2 \operatorname{tr}(P^{(11)} + P^{(22)} - P^{(12)} - P^{(21)})}. \tag{11}$$

In Section 4, we apply formulas (11) for fusion of two local estimates of unknown using two independent noisy measurements.

If the estimates $\hat{x}^{(1)}$ and $\hat{x}^{(2)}$ are uncorrelated, i.e., $P^{(12)} = P^{(21)} = 0$ then formulas (11) for weights are reduced to

$$a^{(1)} = \frac{\mathrm{tr}(P^{(22)})}{\mathrm{tr}(P^{(11)} + P^{(22)})}, \quad a^{(2)} = \frac{\mathrm{tr}(P^{(22)})}{\mathrm{tr}(P^{(11)} + P^{(22)})}.$$

2.2. Example 2: fusion of uncorrelated vector estimates

Suppose that the local vector estimates $\hat{x}^{(1)}, \ldots, \hat{x}^{(N)} \in \mathbb{R}^n$ are uncorrelated, i.e., $P^{(ij)} = 0,\ i \neq j$. Then linear system (8) takes the following simple form:

$$\begin{cases} a^{(1)}\,\mathrm{tr}(P^{(11)}) - a^{(N)}\,\mathrm{tr}(P^{(NN)}) = 0, \\ a^{(2)}\,\mathrm{tr}(P^{(22)}) - a^{(N)}\,\mathrm{tr}(P^{(NN)}) = 0, \\ \vdots \\ a^{(N-1)}\,\mathrm{tr}(P^{(N-1,N-1)}) - a^{(N)}\,\mathrm{tr}(P^{(NN)}) = 0, \\ a^{(1)} + \cdots + a^{(N-1)} + a^{N} = 1. \end{cases}$$

The explicit solution of the system and the FF (5) has the form

$$\hat{x}^{\mathrm{FF}} = \sum_{i=1}^{N} a^{(i)} \hat{x}^{(i)}, \quad a^{(i)} = \frac{1}{\mathrm{tr}(P^{(ii)})} \left[\sum_{j=1}^{N} \frac{1}{\mathrm{tr}(P^{(jj)})} \right]^{-1}, \quad i = 1, \ldots, N. \tag{12}$$

In the particular case of the scalar uncorrelated estimates $\hat{x}^{(1)}, \ldots, \hat{x}^{(N)} \in \mathbb{R}$, $\mathrm{tr}(P^{(ii)}) = P^{(ii)}$ and the FF (12) coincides with the well-known result (1).

Remark. The knowledge of the local error cross-covariances $P^{(ij)}$ is needed in many techniques for distributed fusion. The cross-covariance $P^{(ij)}$ is a key quantity for the best fusion estimation (see Berg and Durrant-Whyte, 1994; Hashemipour et al., 1988; Li and Zhang, 2001). In real-life problems, the local error covariances $P^{(ii)},\ i = 1, \ldots, N$, are usually known. For instance, in Kalman filtering, $P^{(ii)}$ is described by the Riccati equation (Bar-Shalom and Li, 1995; Gelb, 1974; Lewis, 1986). However, the local cross-covariances $P^{(ij)},\ i \neq j$ are usually unknown and they should be determined. This paper presents explicit equations for their computation in the linear continuous-time filtering problem.

3. Fusion filtering in multisensor continuous-time linear systems

Many advanced systems now make use of a large number of sensors in practical applications ranging from aerospace and defence, robotics automation systems, to the monitoring and control of process generation plants. Recent developments in integrated sensor network systems have further motivated the search for decentralized signal processing algorithms. An important practical problem in the above systems is to find a fusion estimate to combine the information from various local estimates to produce a global (fusion) estimate. In this section, we propose a new SFF for multisensor continuous-time dynamic systems.

3.1. Problem setting

Consider a continuous-time linear dynamic system with additive white Gaussian noise

$$\dot{x}_t = F_t x_t + G_t v_t, \quad t \geqslant 0, \tag{13}$$

where $x_t \in \mathbb{R}^n$ is a state vector and $v_t \in \mathbb{R}^r$ is a zero-mean Gaussian white system noise with intensity Q_t, i.e., $E[v_t v_t^{\mathrm{T}}] = Q_t \delta_{t-s}$, δ_t is the Dirac delta function.

Suppose that the measurement system involves N sensors

$$
\begin{aligned}
&y_t^{(1)} = H_t^{(1)} x_t + w_t^{(1)}, \quad y_t^{(1)} \in \mathbb{R}^{m_1},\\
&y_t^{(2)} = H_t^{(2)} x_t + w_t^{(2)}, \quad y_t^{(2)} \in \mathbb{R}^{m_2},\\
&\vdots\\
&y_t^{(N)} = H_t^{(N)} x_t + w_t^{(N)}, \quad y_t^{(N)} \in \mathbb{R}^{m_N},
\end{aligned} \tag{14}
$$

where w_t^i represents a zero-mean Gaussian white sensor noise with intensity $R_t^{(i)}$, $i = 1, \ldots, N$. We assume that the initial state $x_0 \sim N(\bar{x}_0, P_0)$ and system and sensor noises $v_t, w_t^{(1)}, \ldots, w_t^{(N)}$ are mutually uncorrelated.

A fundamental problem associated with such systems (13), (14) is that of state estimation of the state x_t from the overall noisy measurements $Y_0^t = \{y_s^{(1)}, y_s^{(2)}, \ldots, y_s^{(N)}, 0 \leqslant s \leqslant t\}$.

3.2. Optimal estimator: Kalman filter

The KF can be used to produce the optimal mean-square state estimate $\hat{x}_t^{\text{opt}}$ based on the overall measurements

$$
Y_t = H_t x_t + w_t, \quad Y_t \in \mathbb{R}^m, \tag{15}
$$

where

$$
Y_t = \begin{bmatrix} y_t^{(1)} \\ \vdots \\ y_t^{(N)} \end{bmatrix}, \quad H_t = \begin{bmatrix} H_t^{(1)} \\ \vdots \\ H_t^{(N)} \end{bmatrix}, \quad w_t = \begin{bmatrix} w_t^{(1)} \\ \vdots \\ w_t^{(N)} \end{bmatrix}, \quad m = m_1 + \cdots + m_N.
$$

The KF equations for the system model (13), (15) are of the following form (Gelb, 1974 and Lewis, 1986):

$$
\begin{aligned}
&\dot{\hat{x}}_t^{\text{opt}} = F_t \hat{x}_t^{\text{opt}} + K_t (Y_t - H_t \hat{x}_t^{\text{opt}}), \quad \hat{x}_0^{\text{opt}} = \bar{x}_0,\\
&K_t = P_t^{\text{opt}} H_t R_t^{-1}, \quad R_t = \operatorname{diag}[R_t^{(1)} \ldots R_t^{(N)}],\\
&\dot{P}_t^{\text{opt}} = F_t P_t^{\text{opt}} + P_t^{\text{opt}} F_t^{\mathrm{T}} - P_t^{\text{opt}} H_t^{\mathrm{T}} R_t^{-1} H_t P_t^{\text{opt}} + G_t Q_t G_t^{\mathrm{T}}, \quad P_0^{\text{opt}} = P_0,\\
&P_t^{\text{opt}} = E[(x_t - x_t^{\text{opt}})(x_t - x_t^{\text{opt}})^{\mathrm{T}}].
\end{aligned} \tag{16}
$$

To compute the state estimate $\hat{x}_t^{\text{opt}}$, the KF requires all the sensor measurements (15) jointly at each time instant t. Unfortunately, because of the limited communication bandwidth or the increased survivability of the system in a poor environment, such as a war situation, every local sensor has to carry out Kalman filtering upon its own measurements $y_t^{(i)}$ first for a local requirement, and then to transmit the processed data—the local state estimates $\hat{x}_t^{(i)}$ to a fusion center. Therefore, the fusion center now needs to fuse all the received local estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$ to yield a globally (fusion) state estimate (see Berg and Durrant-Whyte, 1994; Hashemipour et al., 1988; Zhu, et al., 2001). Next, it is shown that the FF (5) can serve as an alternative to the solution of the filtering problem (13), (14).

3.3. Suboptimal estimator: suboptimal fusion filter

Here we solve the following problem—how to fuse the local Kalman estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$? The proposed filtering algorithm includes two stages: the locally optimal Kalman estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$ computed at the first stage are linearly fused at the second stage based on the FF (5). The proposed filter has a parallel structure and it is suitable for parallel processing of individual sensor measurements. It can also help to minimize the computation time and produce real-time state estimation. Examples of systems containing different types of sensors, demonstrating the accuracy of the proposed filter, are given.

First stage ("Calculation of local Kalman estimates"): According to Eqs. (13) and (14), we have N local dynamic subsystems with the state vector x_t and local (individual) sensor measurement $y_t^{(i)}$:

$$\begin{aligned} &\dot{x}_t = F_t x_t + G_t v_t, \quad t \geqslant 0, \\ &y_t^{(i)} = H_t^{(i)} x_t + w_t^{(i)}, \end{aligned} \tag{17}$$

where the number i of a local subsystem is fixed.

Further, denote a local estimate of the state x_t based on the local sensor measurement $y_t^{(i)}$ by $\hat{x}_t^{(i)}$. To find $\hat{x}_t^{(i)}$, we apply the KF to subsystem (17) and get the local Kalman estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$ with the corresponding local error covariances $P_t^{(11)}, \ldots, P_t^{(NN)}$

$$\begin{aligned} &\dot{\hat{x}}_t^{(i)} = F_t \hat{x}_t^{(i)} + K_t^{(i)} (y_t^{(i)} - H_t \hat{x}_t^{(i)}), \quad \hat{x}_0^{(i)} = \bar{x}_0, \\ &K_t^{(i)} = P_t^{(ii)} H_t^{(i)} (R_t^{(i)})^{-1}, \\ &\dot{P}_t^{(ii)} = F_t P_t^{(ii)} + P_t^{(ii)} F_t^{\mathrm{T}} - P_t^{(ii)} H_t^{(i)^{\mathrm{T}}} (R_t^{(i)})^{-1} H_t^{(i)} P_t^{(ii)} + G_t Q_t G_t^{\mathrm{T}}, \\ &P_0^{(ii)} = P_0, \quad P_t^{(ii)} = E[(x_t - \hat{x}_t^{(i)})(x_t - \hat{x}_t^{(i)})^{\mathrm{T}}]. \end{aligned} \tag{18}$$

Second stage ("Fusion of local Kalman estimates"): To express the final fusion estimate of the state in terms of the local Kalman estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$, we use the FF. From (5) it follows that

$$\hat{x}_t^{\,\mathrm{sub}} = \sum_{i=1}^{N} a_t^{(i)} \hat{x}_t^{(i)}, \quad \sum_{i=1}^{N} a_t^{(i)} = 1. \tag{19}$$

The linear equations (8) for the unknown weights $a_t^{(i)}$ are given by

$$\begin{aligned} &\sum_{i=1}^{N} a_t^{(i)} \operatorname{tr}(P_t^{(ij)} + P_t^{(ji)} - P_t^{(iN)} - P_t^{(Ni)}) = 0, \quad j - 1, \ldots, N-1, \\ &\sum_{i=1}^{N} a_t^{(i)} = 1. \end{aligned} \tag{20}$$

Note that Eqs. (20) depend on the local error covariances $P_t^{(11)}, \ldots, P_t^{(NN)}$ determined by the Riccati equations (18), and the cross-covariances

$$P_t^{(ij)} = E[(x_t - \hat{x}_t^{(i)})(x_t - \hat{x}_t^{(j)})^{\mathrm{T}}], \quad i \neq j$$

that satisfy the following differential equation:

$$\begin{aligned} &\dot{P}_t^{(ij)} = [F_t - P_t^{(ii)} H_t^{(i)^{\mathrm{T}}} (R_t^{(i)})^{-1} H_t^{(i)}] P_t^{(ij)} + G_t Q_t G_t^{\mathrm{T}} + P_t^{(ij)} [F_t - P_t^{(jj)} H_t^{(j)^{\mathrm{T}}} (R_t^{(j)})^{-1} H_t^{(j)}]^{\mathrm{T}}, \\ &P_0^{(ij)} = P_0, \quad i, j = 1, \ldots, N, \ i \neq j. \end{aligned} \tag{21}$$

The derivation of Eqs. (21) is given in Appendix B.

Thus, the local Kalman estimates and covariances $(\hat{x}_t^{(i)}, P_t^{(ii)})$ (see (18)), the local cross-covariances $P_t^{(ij)}, \ i \neq j$ (see (21)), and the fusion equations (19), (20) completely establish the two-stage SFF.

3.4. Discussion

(1) The local Kalman estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$ are separated for different types of sensors, i.e., each estimate $\hat{x}_t^{(i)}$ is found independently of other estimates. Thus, they can be evaluated in parallel for the different sensors $y_t^{(1)}, \ldots, y_t^{(N)}$.

(2) The SFF can be corrected if one of the parallel local estimates $\hat{x}_t^{(i)}$ diverges. In this case, the corresponding weight $a_t^{(i)}$ tends to zero ($a_t^{(i)} \to 0$), thereby indicating that the diverging estimate $\hat{x}_t^{(i)}$ is discarded in the weighted sum (19).

(3) The local error cross-covariances $P_t^{(ij)}$, filter gains $K_t^{(i)}$, and weights $a_t^{(i)}$ can be pre-computed, since they do not depend on the sensor measurements $y_t^{(1)}, \ldots, y_t^{(N)}$ but only on the noise statistics Q_t, $R_t^{(i)}$, and the system matrices and initial conditions $F_t, G_t, H_t^{(i)}, \bar{x}_0, P_0$ which are the part of system model (13), (14). Thus, once the measurement schedule has been settled, the real-time implementation of the SFF requires only the computation of the local Kalman estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$

$$\dot{\hat{x}}_t^{(i)} = F_t \hat{x}_t^{(i)} + K_t^{(i)}(y_t^{(i)} - H_t \hat{x}_t^{(i)}), \quad \hat{x}_0^{(i)} = \bar{x}_0, \ \ i = 1, \ldots, N$$

and the final suboptimal fusion estimate $\hat{x}_t^{\text{sub}}$

$$\hat{x}_t^{\text{sub}} = a_t^{(1)} \hat{x}_t^{(1)} + a_t^{(2)} \hat{x}_t^{(2)} + \cdots + a_t^{(N)} \hat{x}_t^{(N)}.$$

(4) Consider now the relation between the covariances P_t^{opt} and P_t^{sub} of the optimal KF and SFF filter, respectively. The covariance P_t^{opt} in (16) represents an optimal fusion covariance of the Kalman estimate based on overall sensors, i.e., $\hat{x}_t^{\text{opt}} = \hat{x}_t^{\text{opt}}(y_t^{(1)}, \ldots, y_t^{(N)})$. The covariance $P_t^{(ii)}$ in (18) represents an optimal local covariance of the local Kalman estimate based only on the individual sensor $y_t^{(i)}$, i.e., $\hat{x}_t^{(i)} = \hat{x}_t^{(i)}(y_t^{(i)})$. In general, the optimal Kalman estimate $\hat{x}_t^{\text{opt}}$ is more accurate than the local one $\hat{x}_t^{(i)}$, i.e., $\text{tr}(P_t^{\text{opt}}) \leqslant \text{tr}(P_t^{(ii)})$ for all $i = 1, \ldots, N$.

The relation between P_t^{opt} and P_t^{sub} is established by the following result.

Theorem 2. *Let* $P_t^{\text{opt}} = \text{cov}(\widehat{x}_t^{\text{opt}})$ *and* $P_t^{\text{sub}} = \text{cov}(\widehat{x}_t^{\text{sub}})$ *be the covariances of the optimal Kalman estimate* (16) *and suboptimal fusion estimate* (19), *respectively. Then* $P_t^{\text{opt}} \leqslant P_t^{\text{sub}}$.

The proof of Theorem 2 is given in Appendix C.

(5) The implementation of the optimal KF and SFF consists of two stages: off-line and on-line. The off-line stage in the SFF is more complex than the off-line one in the KF, since it requires the computation of the local cross-covariances $P^{(ij)}$, filter gains $K_t^{(i)}$ and weights $a_t^{(i)}$. However, it is not essential since this stage can be pre-computed. The on-line stage (real-time implementation) requires only computation of the local estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$ and fusion estimates $\hat{x}_t^{\text{opt}}$ and $\hat{x}_t^{\text{sub}}$. Therefore, the complexity of on-line stages is similar for the KF and SFF. However, to compute $\hat{x}_t^{\text{opt}}$, the KF requires all sensor measurements jointly at each time instant, whereas the SFF computes $\hat{x}_t^{\text{sub}}$ sequentially.

4. Application of the suboptimal fusion filter

A comparative experimental analysis of the optimal KF and SFF is considered in three examples:

- In Example 3, we estimate the value of an unknown constant scalar from two measurements corrupted by additive white noises. Here, we get precise formulas for mean-square errors (MSEs) for both filters and show that the SFF provides a good estimation accuracy.
- In Example 4, we describe the scalar time-invariant system with N sensor measurements. In this case, we numerically evaluate and compare MSEs depending on the number of sensors.
- In Example 5, the motion governed by Newton's law is considered. Three measurement programs with two sensors are examined.

4.1. Example 3: independent measurements of an unknown scalar

Estimate a random constant $\theta \sim N(m, \sigma^2)$, given two continuous measurements of θ corrupted by Gaussian white noises.

The system and measurement equations describing this situation are

$$\textit{System}:\ \dot{x}_t = 0, \quad t \geqslant 0, \quad x_0 \equiv \theta \sim N(m, \sigma^2). \tag{22}$$

$$\textit{Measurements}:\ y_t^{(1)} = x_t + w_t^{(1)}, \quad y_t^{(2)} = x_t + w_t^{(2)}, \tag{23}$$

where $w_t^{(1)}$ and $w_t^{(2)}$ are the uncorrelated zero-mean white Gaussian noises with intensities r_1 and r_2, respectively.

4.1.1. Optimal estimator: Kalman filter

The KF (16) gives the optimal mean-square estimate $\hat{x}_t^{\,\mathrm{opt}}$ of an unknown $x_t \equiv \theta$ based on the overall sensor measurements $Y_t = [y_t^{(1)}\ \ y_t^{(2)}]^{\mathrm{T}}$. In this case it takes the form:

$$\begin{aligned}
\dot{\hat{x}}_t^{\,\mathrm{opt}} &= P_t^{\,\mathrm{opt}}\left[\frac{1}{r_1}(y_t^{(1)} - \hat{x}_t^{\,\mathrm{opt}}) + \frac{1}{r_2}(y_t^{(2)} - \hat{x}_t^{\,\mathrm{opt}})\right], \quad \hat{x}_0^{\,\mathrm{opt}} = m,\\
\dot{P}_t^{\,\mathrm{opt}} &= -\frac{1}{r}(P_t^{\,\mathrm{opt}})^2, \quad P_t^{\,\mathrm{opt}} = \sigma^2, \quad r = \frac{r_1 r_2}{r_1 + r_2},\\
P_t^{\,\mathrm{opt}} &= E[(\theta - \hat{x}_t^{\,\mathrm{opt}})^2].
\end{aligned}$$

Integrating the Riccati equation for $P_t^{\,\mathrm{opt}}$ by separation of variables, we obtain the exact formula for MSE,

$$P_t^{\,\mathrm{opt}} = \frac{r\sigma^2}{r + \sigma^2 t} = \frac{r_1 r_2 \sigma^2}{r_1 r_2 + (r_1 + r_2)\sigma^2 t}. \tag{24}$$

Together with the optimal KF, we apply the SFF.

4.1.2. Suboptimal estimator: suboptimal fusion filter

First using Eq. (18), we obtain equations for the local Kalman estimates $\hat{x}_t^{(1)}$ and $\hat{x}_t^{(2)}$ and MSEs $P_t^{(11)}$ and $P_t^{(22)}$, i.e.,

$$\begin{aligned}
\dot{\hat{x}}_t^{(i)} &= \frac{1}{r_i} P_t^{(ii)}(y_t^{(i)} - \hat{x}_t^{(i)}), \quad \hat{x}_t^{(i)} = m,\\
\dot{P}_t^{(ii)} &= -\frac{1}{r_i}(P_t^{(ii)})^2, \quad P_0^{(i)} = \sigma^2,\\
P_t^{(ii)} &= E[(\theta - \hat{x}_t^{(i)})^2], \quad i = 1, 2.
\end{aligned} \tag{25}$$

Similar to (24), the solution of the equations for local MSEs $P_t^{(ii)}$ in (25) takes the form:

$$P_t^{(ii)} = \frac{r_i \sigma^2}{r_i + \sigma^2 t}, \quad i = 1, 2. \tag{26}$$

Next, using the scalar version of the FF (11) in Example 1, one can obtain the suboptimal fusion estimate $\hat{x}_t^{\,\mathrm{sub}}$ of the unknown parameter θ as

$$\begin{aligned}
\hat{x}_t^{\,\mathrm{sub}} &= a_t^{(1)}\hat{x}_t^{(1)} + a_t^{(2)}\hat{x}_t^{(2)},\\
a_t^{(1)} &= \frac{P_t^{(22)} - P_t^{(12)}}{P_t^{(11)} + P_t^{(22)} - 2P_t^{(12)}},\\
a_t^{(2)} &= \frac{P_t^{(11)} - P_t^{(12)}}{P_t^{(11)} + P_t^{(22)} - 2P_t^{(12)}},
\end{aligned} \tag{27}$$

where $P_t^{(12)} = E[(\theta - \hat{x}_t^{(1)})(\theta - \hat{x}_t^{(2)})]$ is the local cross-covariance determined by (21) that results in the following differential equation:

$$\dot{P}_t^{(12)} = -\left[\frac{1}{r_1} P_t^{(11)} + \frac{1}{r_2} P_t^{(22)}\right] P_t^{(12)}, \quad P_0^{(12)} = \sigma^2. \tag{28}$$

Substituting (26) for the local MSEs $P_t^{(11)}$ and $P_t^{(22)}$ into (28), we obtain the linear homogeneous differential equation for $P_t^{(12)}$. The solution of (28) by separation of variables is given by

$$P_t^{(12)} = \frac{r_1 r_2 \sigma^2}{(r_1 + \sigma^2 t)(r_2 + \sigma^2 t)}. \tag{29}$$

Finally substituting (26) and (29) into (27) and (10), we obtain precise formulas for the unknown weights $a_t^{(1)}, a_t^{(2)}$, and overall MSE for the SFF, $P_t^{\text{sub}} = E[(\theta - \hat{x}_t^{\text{sub}})^2]$:

$$\begin{aligned} &a_t^{(1)} = r_2/(r_1 + r_2), \quad a_t^{(2)} = r_1/(r_1 + r_2), \\ &P_t^{\text{sub}} = (a_t^{(1)})^2 P_t^{(11)} + 2a_t^{(1)} a_t^{(2)} P_t^{(12)} + (a_t^{(2)})^2 P_t^{(22)} \\ &\quad = r_1 r_2 \sigma^2 (r_1 + r_2 + \sigma t)/[(r_1 + r_2)(r_1 + \sigma^2 t)(r_2 + \sigma^2 t)]. \end{aligned} \tag{30}$$

Comparing the optimal and suboptimal MSEs (24) and (30), we get

$$P_t^{\text{sub}} - P_t^{\text{opt}} = \frac{r_1^2 r_2^2 \sigma^4}{(r_1 + r_2)(r_1 + \sigma^2 t)(r_2 + \sigma^2 t)[r_1 r_2 + (r_1 + r_2)\sigma^2 t]} = \mathrm{O}\left(\frac{1}{t^3}\right).$$

This result shows that the SFF yields good accuracy and certain well-defined convergence properties.

4.2. Example 4: a scalar system with N-sensors

Consider a continuous-time scalar system

$$\dot{x}_t = a x_t + v_t, \quad t \in [0, T], \tag{31}$$

where v_t is a white Gaussian system noise with intensity q, $x_0 \sim N(\bar{x}_0, P_0)$, $a = const$.

The measurement model for the system is

$$y_t^{(i)} = x_t + w_t^{(i)}, \quad i = 1, \ldots, N, \tag{32}$$

where $w_t^{(i)}$ represents a zero-mean Gaussian white sensor noise with intensity $r_i = const$, $i = 1, \ldots, N$.

4.2.1. MSE for the optimal Kalman filter

Let $\hat{x}_t^{\text{opt}}$ be an optimal Kalman estimate of the state x_t based on the overall measurements $Y_t = [y_t^{(1)} \ \cdots \ y_t^{(N)}]^{\mathrm{T}}$ and $P_t^{\text{opt}} = E[(x_t - \hat{x}_t^{\text{opt}})^2]$ be the MSE. By the KF equations (16), P_t^{opt} satisfies the Riccati equation:

$$\dot{P}_t^{\text{opt}} = 2a P_t^{\text{opt}} - \frac{1}{r}(P_t^{\text{opt}})^2 + q, \quad \frac{1}{r} = \frac{1}{r_1} + \cdots + \frac{1}{r_N}, \quad P_t^{\text{opt}} = P_0. \tag{33}$$

Fig. 1. MSE analysis of the KF and SFF for N sensors.

4.2.2. MSE for the suboptimal fusion filter

For the SFF, let $\hat{x}_t^{\text{sub}}$ be a suboptimal estimate of the state x_t, which represents the weighted sum of the local estimates $\hat{x}_t^{(1)}, \ldots, \hat{x}_t^{(N)}$

$$\hat{x}_t^{\text{sub}} = \sum_{i=1}^{N} a_t^{(i)} \hat{x}_t^{(i)}, \qquad \sum_{i=1}^{N} a_t^{(i)} = 1.$$

According to (10), the MSE $P_t^{\text{sub}} = E[(x_t - \hat{x}_t^{\text{sub}})^2]$ is given by

$$P_t^{\text{sub}} = \sum_{i,j=1}^{N} a_t^{(i)} a_t^{(j)} P_t^{(ij)}. \tag{34}$$

Here the local error variances $P_t^{(ii)}$ and cross-covariances $P_t^{(ij)}$ satisfy Eqs. (18) and (21), respectively,

$$\begin{aligned} \dot{P}_t^{(ii)} &= 2a P_t^{(ii)} - \frac{1}{r_i} (P_t^{(ii)})^2 + q, \quad P_0^{(ii)} = P_0, \\ \dot{P}_t^{(ij)} &= 2a P_t^{(ij)} - \left[\frac{P_t^{(ii)}}{r_i} + \frac{P_t^{(jj)}}{r_j} \right] P_t^{(ij)} + q, \\ P_t^{(ij)} &= P_0, \quad i, j = 1, \ldots, N, \ \ i \neq j. \end{aligned} \tag{35}$$

Further, the unknown weights $a_t^{(i)}$, $i = 1, \ldots, N$, can be computed from (9). Finally, Eqs. (34) and (35) determine the overall MSE of the SFF (32). Fig. 1 shows P_t^{opt} and P_t^{sub} for the following parameter values

$$\begin{aligned} &N = 4, \quad q = 0.01, \quad \bar{x}_0 = 0.5, \quad P_0 = 1, \\ &r_1 = 0.2, \quad r_2 = 0.1, \quad r_3 = 0.06, \quad r_4 = 0.04. \end{aligned}$$

From Fig. 1 it can be seen that P_t^{opt} and P_t^{sub} are very close. Also it is clear that estimation accuracy is increasing with the number of sensors N. In the case of a single sensor when $N = 1$, the optimal KF and SFF are identical.

4.3. Example 5: estimation of a damper harmonic oscillator motion

Consider an example of dynamic systems modeled by the second Newton law. For multisensor environment, we examine the following three cases. In the first and second cases, there are two sensors one of which is main and another is reserved (Measurement programs 1 and 2). In the third case, there are two independent sensors observing position and velocity separately (Measurement program 3).

In this example, we verify the SFF using the harmonic oscillator motion in the following system model $\ddot{z} = u_t/M$, where z_t is the position, M is the mass, and u_t is the deterministic input (control). In the canonical form, we have

$$\hat{x}_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ 1/M \end{bmatrix} u_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} z_t \\ \dot{z}_t \end{bmatrix}, \tag{36}$$

where υ_t is a zero-mean Gaussian white system noise with intensity q. This noise has been added for compensation of modeling errors. The initial condition is $x_0 \sim N(\bar{x}_0, P_0)$, where $\bar{x}_0 = [01]^{\mathrm{T}}$, $P_0 = \mathrm{diag}[21]$.

Assume that the measurement model contains two sensors

$$y_t^{(1)} = H^{(1)} x_t + w_t^{(1)}, \quad y_t^{(2)} = H^{(2)} x_t + w_t^{(2)}, \tag{37}$$

where $H^{(1)}$ and $H^{(2)}$ are the 1×2 measurement matrices, $w_t^{(1)}$ and $w_t^{(2)}$ are uncorrelated zero-mean white Gaussian sensor noises with intensities r_1 and r_2, respectively.

The two filters (KF and SFF) for the system model (36), (37) are considered. In this example, there are three measurement programs of measurements being illustrated and compared:

Measurement Program 1: Only the position $x_{1,t} = z_t$ is measured by two different sensors, one of which is main and another is reserved. In this case $H^{(1)} = [1 \;\; 0]$ and $H^{(2)} = [1 \;\; 0]$.

Measurement Program 2: Only the velocity $x_{2,t} = \dot{z}_t$ is measured by two different sensors, one of which is main and another is reserved. Then $H^{(1)} = [0 \;\; 1]$ and $H^{(2)} = [0 \;\; 1]$.

Measurement Program 3: Both the position and velocity are measured. Then $H^{(1)} = [1 \;\; 0]$ and $H^{(2)} = [0 \;\; 1]$.

To study the behavior of the KF and SFF error covariances, set $r_1 = 0.02$, $r_2 = 0.01$ and $q = 1$. The point of interest is the 2×2 MSE matrices in estimation of state components: position$(x_{1,t})$ and velocity $(x_{2,t})$

$$P_t^{\,\mathrm{opt}} = E[e_t^{\,\mathrm{opt}} (e_t^{\,\mathrm{opt}})^{\mathrm{T}}] = \begin{bmatrix} P_{11,t}^{\mathrm{opt}} & P_{12,t}^{\mathrm{opt}} \\ P_{12,t}^{\mathrm{opt}} & P_{22,t}^{\mathrm{opt}} \end{bmatrix}, \quad e_t^{\,\mathrm{opt}} = \begin{bmatrix} x_{1,t} - \hat{x}_{1,t}^{\,\mathrm{opt}} \\ x_{2,t} - \hat{x}_{2,t}^{\,\mathrm{opt}} \end{bmatrix},$$

$$P_t^{\,\mathrm{sub}} = E[e_t^{\,\mathrm{sub}} (e_t^{\,\mathrm{sub}})^{\mathrm{T}}] = \begin{bmatrix} P_{11,t}^{\mathrm{sub}} & P_{12,t}^{\mathrm{sub}} \\ P_{12,t}^{\mathrm{sub}} & P_{22,t}^{\mathrm{sub}} \end{bmatrix}, \quad e_t^{\,\mathrm{sub}} = \begin{bmatrix} x_{1,t} - \hat{x}_{1,t}^{\,\mathrm{sub}} \\ x_{2,t} - \hat{x}_{2,t}^{\,\mathrm{sub}} \end{bmatrix}.$$

These are the quantities shown in Figs. 2 and 3. In Fig. 2 we show the MSE comparisons for "position" ($P_{11,t}^{\mathrm{opt}}$ and $P_{11,t}^{\mathrm{sub}}$) and "velocity" ($P_{22,t}^{\mathrm{opt}}$ and $P_{22,t}^{\mathrm{sub}}$) for Measurement Program 1.

System (36), (37) becomes unobservable for the velocity in Measurement Program 2. In this case, the rank of the observable matrix Θ is equal to 1, i.e.,

$$\mathit{rank}(\Theta) = 1, \quad \Theta = [H^{\mathrm{T}} \;\; F^{\mathrm{T}} H^{\mathrm{T}}], \quad H = \begin{bmatrix} H^{(1)} \\ H^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Fig. 3 shows the MSEs for "position" ($P_{11,t}^{\mathrm{opt}}$ and $P_{11,t}^{\mathrm{sub}}$) and "velocity" ($P_{22,t}^{\mathrm{opt}}$ and $P_{22,t}^{\mathrm{sub}}$) for Measurement Program 3.

From Figs. 2 and 3 it follows that the difference between the optimal MSEs ($P_{ii,t}^{\mathrm{opt}}$, $i = 1, 2$) and the suboptimal one ($P_{ii,t}^{\mathrm{sub}}$, $i = 1, 2$) is negligible for Measurement Programs 1 and 3, especially for steady-state regimes. This means that, for our example, application of the SFF can produce good results in real-time processing requirements. In the case of the unobservability of the pair $\{F, H\}$, Measurement Program 2 must be adjusted.

In Examples 3–5 we observe a relative loss of accuracy of the SFF as compared to the optimal KF.

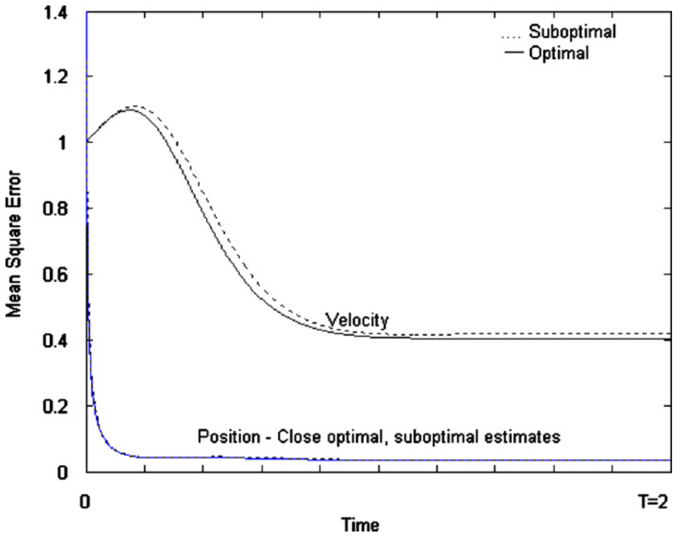

Fig. 2. Program 1: MSE analysis for position and velocity.

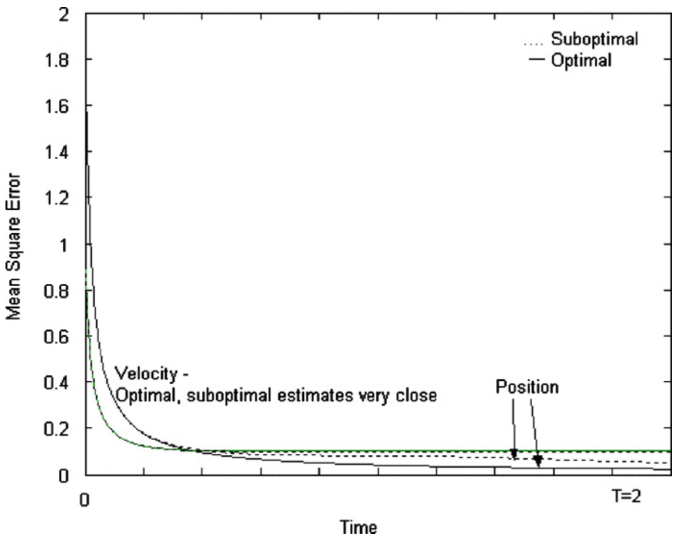

Fig. 3. Program 3: MSE analysis for position and velocity.

5. Conclusion

In this paper the optimal mean-square linear combination of local vector estimates (the fusion formula) is derived. This result is a theoretical basis for the linear estimation fusion. A new suboptimal fusion filter for linear continuous-time dynamic systems with multi-sensor environment based on the FF is proposed. Differential equations for local cross-covariances of filtering errors are derived. These cross-covariances are the key quantities for the best linear fusion. Numerical examples show that the suboptimal filter yields a reasonably good estimation accuracy, especially for the steady-state regime. The obtained fusion filter is slightly suboptimal as compared with the optimal KF.

Acknowledgements

We would like to thank the reviewers for their helpful comments and express our sincere gratitude to Dr. Peter Ruckdeschel for his constructive suggestions in discussion on the accuracy of the optimal and suboptimal estimates.

Appendix A: Proofs of Theorem 1 and corollaries

(i) Using (4) and (5), the mean-square criterion (6) can be rewritten as

$$\begin{aligned} J &= E[(x-\hat{x}^{\mathrm{FF}})^{\mathrm{T}}(x-\hat{x}^{\mathrm{FF}})] = \operatorname{tr}\left\{E[(x-\hat{x}^{\mathrm{FF}})(x-\hat{x}^{\mathrm{FF}})^{\mathrm{T}}]\right\} \\ &= \operatorname{tr}\left\{E\left[\sum_{i,j=1}^{N} a^{(i)}a^{(j)}(x-\hat{x}^{(i)})(x-\hat{x}^{(i)})^{\mathrm{T}}\right]\right\} \\ &= \sum_{i,j=1}^{N} a^{(i)}a^{(j)}\operatorname{tr}(P^{(ij)}) \to \min_{a^{(1)},\ldots,a^{(N)}}, \end{aligned} \tag{A.1}$$

where $P^{(ij)} = E[(x-\hat{x}^{i})(x-\hat{x}^{j})^{\mathrm{T}}]$. So formula (10) of Corollary 1 for the overall fusion error covariance is derived. Next substitute the expression

$$a^{(N)} = 1 - \sum_{j=1}^{N-1} a^{(j)}$$

into (A.1) and obtain that

$$\begin{aligned} J = {} & \sum_{i,j=1}^{N-1} a^{(i)}a^{(j)}\operatorname{tr}(P^{(ij)}) + \left(1-\sum_{j=1}^{N-1} a^{(j)}\right)\sum_{i=1}^{N-1} a^{(i)}\operatorname{tr}(P^{(Ni)}+P^{(iN)}) \\ & + \left(1-\sum_{j=1}^{N-1} a^{(j)}\right)^2 \operatorname{tr}(P^{(NN)}) \to \min_{a^{(1)},\ldots,a^{(N-1)}}. \end{aligned} \tag{A.2}$$

Differentiating criterion (A.2) with respect to $a^{(1)},\ldots,a^{(N-1)}$ and then equating the result to zero, i.e.,

$$\frac{\partial J}{\partial a^{(i)}} = 0, \quad j = 1,\ldots,N-1,$$

we obtain Eq. (8) for the unknown weights $a^{(1)},\ldots,a^{(N-1)}$.

(ii) The derivation of expressions (9) is based on the well-known result in quadratic optimization with restrictions given by (Seber and Lee, 2003)

$$\begin{cases} E[(Y-H\beta)^{\mathrm{T}}(Y-H\beta)] = \min\limits_{\beta} \\ D\beta = d \end{cases} \tag{A.3}$$

from which it follows that

$$\begin{aligned} \hat{\beta} &= \hat{\beta}_0 + [E(H^{\mathrm{T}}H)]^{-1}D^{\mathrm{T}}\left\{D[E(H^{\mathrm{T}}H)]^{-1}D^{\mathrm{T}}\right\}^{-1}(d-D\hat{\beta}_0), \\ \hat{\beta}_0 &= [E(H^{\mathrm{T}}H)]^{-1}E(H^{\mathrm{T}}Y). \end{aligned} \tag{A.4}$$

Putting in (A.3) and (A.4)

$$Y = x, \quad H = [\hat{x}^{(1)}\ldots\hat{x}^{(N)}], \quad \beta = [\hat{a}^{(1)}\ldots\hat{a}^{(N)}], \quad D = [1\ldots1], \quad d = 1,$$

we obtain the explicit expression (9) for scalar weights.

This completes the proof of Theorem 1. □

(iii) If the local estimates $\hat{x}^{(i)}$ are unbiased, i.e., $E[\hat{x}^{(i)}]=E[x],\ i=1,\dots,N$, then we get

$$E[\hat{x}^{\mathrm{FF}}]=\sum_{i=1}^{N}a^{(i)}E[\hat{x}^{(i)}]=\left(\sum_{i=1}^{N}a^{(i)}\right)E[x]=E[x].$$

Corollary 2 is proved.

Appendix B: Derivation of Eq. (21) for cross-covariances

The KF equations (18) yield the following differential equation for the local filtering error $e_t^{(i)}=x_t-\hat{x}_t^{(i)}$:

$$\begin{aligned}\dot{e}_t^{(i)}&=\dot{x}_t-\dot{x}_t^{(i)}=F_tx_t+G_tv_t-F_t\hat{x}_t^{(i)}-K_t^{(i)}(y_t^{(i)}-H_t^{(i)}\hat{x}_t^{(i)})\\&=F_te_t^{(i)}+G_tv_t-K_t^{(i)}(H_t^{(i)}+w_t^{(i)}-H_t^{(i)}\hat{x}_t^{(i)})=(F_t-K_t^{(i)}H_t^{(i)})e_t^{(i)}+G_tv_t-K_t^{(i)}w_t^{(i)}.\end{aligned}$$

Substitute this relation into the Lyapunov equation for the cross-covariance $P_t^{(ij)}=E(e_t^{(i)}e_t^{(j)\mathrm{T}}),\ i\neq j$ (see Lewis, 1986). By virtue of the assumptions that Gaussian white system and sensor noises v_t, $w_t^{(i)}$ and $w_t^{(j)},\ i\neq j$, are mutually uncorrelated, we obtain the linear differential equation for $P_t^{(ij)}$.

This completes the derivation of (21).

Appendix C: Proof of Theorem 2

Consider two filtering errors $\Delta\widehat{x}_t^{\,\mathrm{opt}}=\widehat{x}_t^{\,\mathrm{opt}}-x_t$ and $\Delta\widehat{x}_t^{\,\mathrm{sub}}=\widehat{x}_t^{\,\mathrm{sub}}-x_t$. Since both estimates $\widehat{x}_t^{\,\mathrm{opt}}$ and $\widehat{x}_t^{\,\mathrm{sub}}$ are unbiased, we can write $P_t^{\,\mathrm{opt}}=\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{opt}})$ and $P_t^{\,\mathrm{sub}}=\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{sub}})$.

Next

$$\begin{aligned}0&\leqslant\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{sub}}-\Delta\widehat{x}_t^{\,\mathrm{opt}})\\&=P_t^{\,\mathrm{sub}}-\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{opt}},\Delta\widehat{x}_t^{\,\mathrm{sub}})-\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{sub}},\Delta\widehat{x}_t^{\,\mathrm{opt}})+P_t^{\,\mathrm{opt}}.\end{aligned}\tag{C.1}$$

To prove Theorem 2, it is sufficient to show that $P_t^{\,\mathrm{opt}}=\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{opt}},\Delta\widehat{x}_t^{\,\mathrm{sub}})$ or that

$$\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{sub}}-\Delta\widehat{x}_t^{\,\mathrm{opt}},\Delta\widehat{x}_t^{\,\mathrm{opt}})=0.\tag{C.2}$$

Equality (C.2) is equivalent to

$$E[(\Delta\widehat{x}_t^{\,\mathrm{sub}}-\Delta\widehat{x}_t^{\,\mathrm{opt}})(\Delta\widehat{x}_t^{\,\mathrm{opt}})^{\mathrm{T}}]=0.\tag{C.3}$$

The optimal and suboptimal estimates $\widehat{x}_t^{\,\mathrm{opt}}$ and $\widehat{x}_t^{\,\mathrm{sub}}$ linearly depend on N sensor measurements

$$y_s^{(i)}\in\mathbb{R}^{m_i},\quad 0\leqslant s\leqslant t,\qquad i=1,\dots,N.\tag{C.4}$$

Stack all measurements (C.4) over each other to form one vector $Y_t^{\mathrm{T}}\in\mathbb{R}^{t\cdot m}$, $m=m_1+\cdots+m_N$, and again enlarge it to $\widetilde{Y}_t=[Y_t^{\mathrm{T}},1]\in\mathbb{R}^{v_t}$, $v_t=tm+1$. The linear span $L_t=L_t\{\widetilde{Y}_t\}$ consists of all linear combinations of $\widetilde{Y}_t$. Since $\widehat{x}_t^{\,\mathrm{opt}}$ is optimal among all linear filters, its error $\Delta\widehat{x}_t^{\,\mathrm{opt}}$ is orthogonal to L_t, i.e.,

$$E[(\Delta\widehat{x}_t^{\,\mathrm{opt}})^{\mathrm{T}}(M_t\widetilde{Y}_t)]=0,$$

where $M_t\in\mathbb{R}^{n\times v_t}$ is an arbitrary matrix, and n is the dimension of a state vector $x_t\in\mathbb{R}^n$.

Let e_p, $p=1,\dots,n$, and d_q, $q=1,\dots,v_t$, denote the canonical basis in $\mathbb{R}^n$ and $\mathbb{R}^{v_t}$, respectively. Plugging in $M_t=e_pd_q^{\mathrm{T}}$, $p=1,\dots,n$, $q=1,\dots,v_t$, we can see that (C.5) is equivalent to

$$E[\Delta\widehat{x}_t^{\,\mathrm{opt}}\widetilde{Y}_t^{\mathrm{T}}]=\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{opt}},\widetilde{Y}_t)=0.$$

Then for some particular matrix $\widetilde{M}_t\in\mathbb{R}^{n\times v_t}$, we have

$$\Delta\widehat{x}_t^{\,\mathrm{sub}}-\Delta\widehat{x}_t^{\,\mathrm{opt}}=\widetilde{M}_t\widetilde{Y}_t\in L_t$$

and hence equalities (C.3) and (C.2) hold, i.e.,

$$\mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{opt}}, \Delta\widehat{x}_t^{\,\mathrm{sub}} - \Delta\widehat{x}_t^{\,\mathrm{opt}}) = \mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{opt}}, \widetilde{M}_t \widetilde{Y}_t) = \mathrm{cov}(\Delta\widehat{x}_t^{\,\mathrm{opt}}, \widetilde{Y}_t)\widetilde{M}_t^{\mathrm{T}} = 0.$$

This completes the proof of Theorem 2. □

References

Bar-Shalom, Y., Campo, L., 1986. The effect of the common process noise on the two-sensor fused-track covariance. IEEE Trans. Aerospace Electron. Systems 22 (11), 803–805.

Bar-Shalom, Y., Li, X.R., 1995. Multitarget-multisensor Tracking: Principles and Techniques. YBS Publishing, Storrs.

Berg, T.M., Durrant-Whyte, H.F., 1994. General decentralized Kalman filter. In: Proceedings of the American Control Conference, Maryland, pp. 2273–2274.

Beugnon, C., Singh, T., Llinas, J., Saha, R.K., 2000. Adaptive track fusion in a multisensor environment. In: Proceedings of the International Conference on Information Fusion, vol. 2. Paris, pp. 24–31.

Gelb, A., 1974. Applied Optimal Estimation. MIT Press, Cambridge, MA.

Hashemipour, H.R., Roy, S., Laub, A.J., 1988. Decentralized structures for parallel Kalman filtering. IEEE Trans. Automat. Control 33 (1), 88–94.

Kendall, M.G., Stuart, A., 1963. The Advanced Theory of Statistics. Inference and Relationship, vol. 2. Griffin, London.

Lewis, F.L., 1986. Optimal Estimation with an Introduction to Stochastic Control Theory. Wiley, New York.

Li, X.R., Zhang, P., 2001. Optimal linear estimation fusion—Part III: cross-correlation of local estimation errors. In: Proceedings of the International Conference on Information Fusion, vol. 1. Montreal, pp. 11–18.

Li, X.R., Zhu, Y.M., Han, C.Z., 2000. Unified optimal linear estimation fusion—Part I: unified models and fusion rules. In: Proceedings of the International Conference on Information Fusion, vol. 2. Paris, pp. 10–17.

Manyika, J., Durrant-Whyte, H., 1994. Data Fusion and Sensor Management: a Decentralized Information-theoretic Approach. Ellis Horwood, New York.

Seber, G.A.F., Lee, A.J., 2003. Linear Regression Analysis. Wiley, New Jersey.

Shin, V., Shevlyakov, G., 2005. An optimal mean square combination of estimates with application to filtering problems. Pattern Recognition Image Anal. 15 (2), 327–330.

Zhu, Y.M., 2003. Multisensor Decision and Estimation Fusion. Kluwer Academic, Boston.

Zhu, Y., You, Z., Zhao, J., Zhang, K., Li, X.R., 2001. The optimality for the distributed Kalman filtering fusion with feedback. Automatica 37, 1393–1489.