Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Optimal Distributed Fusion Filtering with the Mean and Variance Information of the State *

Jiabing Sun, Chengjin Zhang, and Runtao Yang
School of Mechanical, Electrical & Information Engineering
Shandong University
Weihai, 264209, Shandong Province, China
{*sunjb, cjzhang, yrt*}*@sdu.edu.cn*

Xiaoxia Sun †
School of Applied Technology
China University of Labor Relations
Beijing, 100048, China
sunxxciir@163.com

***Abstract*— The distributed optimal (linear minimum variance) fusion filtering problem for multi-sensor discrete-time stochastic systems is considered. Based on the projection theory, a new fusion algorithm is presented. The mean and variance information of the state can be computed off-line. The local filters are Kalman filters using local measurement data. After obtaining the local filtering error variances and cross covariances, using the new fusion algorithm, the distributed optimal fusion filter is derived. Commonly, this fusion filter is better than (at least, as good as) the fusion filter weighted by matrices. A numerical example shows the effectiveness of the new fusion algorithm.**

***Index Terms*— Distributed estimation, Information fusion, Multi-sensor systems,**

I. Introduction

The filtering problem for stochastic systems is a classical problem. In recent years, the fusion filtering problem for stochastic systems has attracted a lot of interest([1], [2], [3], [4]). There are two methods for information fusion filtering. The first one is the centralized method. In this method, the fusion center collects all raw measurement data, and give the globally optimal (linear minimum variance) estimate by directly combining local measurements. But it is computationally expensive and has low robustness and less flexibility if some of the sensors break down. The second one is the distributed (decentralized) fusion method. In this method, each sensor gives local estimate of the state by utilizing its own measurement data. Instead of the raw measurement data, all local estimates are send to the fusion center. In the fusion center, the local estimates are used to give an estimate. In comparison with the centralized method, the distributed method has less a little lower accuracy. But it also has smaller computational burden. And the fault diagnosis and isolation become easier in this method.

In recent years, the distributed fusion filters have been widely investigated (see, e.g., [5], [6], [7], [8], [9]). Reference [9] gives the matrix-weighting optimal fusion rule in the linear minimum variance sense. This fusion rule is widely use in distributed fusion filtering([10], [11]). In the matrix-weighting fusion rule, the weighted matrices is obtained by solving a constraint optimization problem, not by the projection theory. In this fusion rule, the sum of the weighted matrices must be equal to an identity matrix. This condition is a sufficient condition for the the unbiasedness of the fused estimator, when the mean of the estimated state is unknown. But it is unnecessary, when the mean of estimated state is known. In this paper, we will show that the matrix-weighting fusion rule is suboptimal, not optimal, given the mean and the variance information of the estimated state.

In this paper, the distributed optimal fusion estimation problems for a class of multi-sensor systems is considered. Based on the projection theory, we give a new distributed fusion algorithm. In special cases, our algorithm is identical to the matrix-weighting fusion rule. Commonly, it is better than the latter. In Sect. II, the problem formulation is introduced. In Sect. III, some preliminaries and the new fusion algorithm are introduced. In Sect. IV, the mean and variance of the state, the local filters and the fusion filter are presented. In Sect. V, a simulation example is shown, where three local filters and two fusion filters are compared. Finally, the paper is concluded in Sect. VI.

II. Problem Formulation

Consider the following linear discrete-time multi-sensor stochastic system:

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

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

where $x(t) \in R^n$ is the state, $y_i(t) \in R^{m_i}$ is the measurement of the ith sensor, $w(t)$ and $v_i(t)$ are white noises, Φ and H_i are constant matrices with suitable dimensions, L is the number of sensors.

In this paper, we assume system (1) satisfies the following assumptions.

Assumption 1 : $w(t)$ and $v_i(t)$ are zero mean white noises,

*This work is supported by Natural Science Foundation of China under Grant #61673245 to C. Zhang.

†The corresponding author.

and satisfy

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

$$E[v_i(t)v_j^T(l)] = R_{ij}\delta_{tl}, \quad i \neq j. \tag{2}$$

Assumption 2: The initial state $x(0)$, is independent of $w(t)$ and $v_i(t)$, and satisfies

$$E\left[x(0)\right] = \mu_0, E\left[(x(0) - \mu_0)(x(0) - \mu_0)^T\right] = P_0. \tag{3}$$

Remark 1: For ease of presentation, we assume the matrices Φ, H_i, Q, S_i, R_i, and R_{ij} are constant. Indeed, these matrices can be time-varying. When they are time-varying, the following results of this paper still hold.

In this paper, our aim is to find the distributed optimal (linear minimum variance) fusion filter $\hat{x}_0(t|t)$, based on local filters $\hat{x}_i(t|t), \; i = 1, 2, \ldots, L.$

III. Preliminaries

First, for ease of reference, we give some results about the linear minimum variance estimation theory.

Suppose $x \in R^n, y \in R^m$, are random vectors. And $Ex = \bar{x}, Ey = \bar{y}, C_x, C_{xy}$, and C_y are all known.

Lemma 1 (Projection theory) [12] : Using data y, the linear minimum variance estimation of x is given by

$$\hat{x} = \bar{x} + C_{xy}C_y^{-1}(y - \bar{y}). \tag{4}$$

Define $\tilde{x} = x - \hat{x}$, then $E\tilde{x} = 0, \tilde{x} \perp \hat{x}$. The estimation error variance is

$$C_{\tilde{x}} = C_x - C_{xy}C_y^{-1}C_{yx}. \tag{5}$$

When C_y is singular, C_y^{-1} could be replaced by C_y^{+}, i.e., the Moore-Penrose inverse of C_y.

Furthermore, we obtain the covariances in the following lemma.

Lemma 2: Define $C_{\tilde{x}} = P$, then

$$C_{x\tilde{x}} = P, \tag{6}$$

$$C_{x\hat{x}} = C_x - P, \tag{7}$$

$$C_{\hat{x}} = C_x - P. \tag{8}$$

Proof: First,

$$C_{x\tilde{x}} = E[(x - \bar{x})\tilde{x}^T] = E[x\tilde{x}^T] = E[(\hat{x} + \tilde{x})\tilde{x}^T] = P, \tag{9}$$

where $x = \tilde{x} + \hat{x}$ and $\tilde{x} \perp \hat{x}$ are used.

Second,

$$\begin{aligned} C_{x\hat{x}} &= E[(x - \bar{x})(\hat{x} - \bar{x})^T] \\ &= E[(x - \bar{x})(x - \bar{x} - \tilde{x})^T] \\ &= C_x - E[(x - \bar{x})\tilde{x}^T] \\ &= C_x - C_{x\tilde{x}} \\ &= C_x - P. \end{aligned} \tag{10}$$

Third,

$$\begin{aligned} C_{\hat{x}} &= E[(\hat{x} - \bar{x})(\hat{x} - \bar{x})^T] \\ &= E[(x - \bar{x} - \tilde{x})(x - \bar{x} - \tilde{x})^T] \\ &= C_x - C_{x\tilde{x}} - C_{\tilde{x}x} + P \\ &= C_x - P. \end{aligned} \tag{11}$$

This ends the proof.

Now, we give the following fusion algorithm, based on Lemma 1 and lemma 2.

Lemma 3 : Suppose $x \in R^n$ is a random vector, $Ex = \bar{x}$ and C_x are known. Suppose $\hat{x}_i \in R^n, i = 1, 2, ..., L$, are unbiased estimators of x, and $\tilde{x}_i \perp \hat{x}_i$. Assume that $\tilde{x}_i$ and $\tilde{x}_j$, $(i \neq j)$ are correlated, and the error variance and cross covariance matrices are denoted by P_i and P_{ij} , respectively. Define $Y = [\hat{x}_1^T, \hat{x}_2^T, ..., \hat{x}_L^T]^T$,$e = [I, I, ..., I]^T \in R^{nL \times n}$, then $E(Y) = e\bar{x}$, the optimal (i.e. linear minimum variance) fusion estimator is given as

$$\hat{x} = \bar{x} + C_{xY}C_Y^{-1}(Y - e\bar{x}), \tag{12}$$

where

$$C_{xY} = [C_{x\hat{x}_1}, ..., C_{x\hat{x}_L}], \tag{13}$$

$$C_{x\hat{x}_i} = C_x - P_i, \tag{14}$$

and

$$[C_Y]_{ij} = C_x - P_i - P_j + P_{ij}, \tag{15}$$

Proof: From Lemma 1, we have (12). From the definition of Y and Lemma 2, we have (13) and (14). Now, from (6), we have

$$\begin{aligned} [C_Y]_{ij} &= E[(\hat{x}_i - \bar{x})(\hat{x}_j - \bar{x})^T] \\ &= E[(x - \bar{x} - \tilde{x}_i)(x - \bar{x} - \tilde{x}_j)^T] \\ &= C_x - P_i - P_j + P_{ij}. \end{aligned} \tag{16}$$

The proof is completed.

Remark 2: In Lemma 3, the orthogonality of $\tilde{x}_i$ and $\hat{x}_i$ is necessary. Without this condition, this lemma do not hold. Because, in the proof above, we use Lemma 2, where the orthogonality of $\tilde{x}_i$ and $\hat{x}_i$ is indispensable.

For ease of comparison, we present the fusion rule weighted by matrices from [9] in the following.

Lemma 4 [9]: Let $\hat{x}_i, i = 1, 2, \ldots, L$, be unbiased estimators of the stochastic column vector $x \in R^n$, i.e. $E(\hat{x}_i) = E(x)$. Let estimation errors be $\tilde{x}_i = x - \hat{x}_i$. Assume that $\tilde{x}_i$ and $\tilde{x}_j$ $(i, j = 1, 2, \ldots, L, i \neq j)$ are correlated, and the error variance and cross-covariance are $P_i (= P_{ii})$ and P_{ij}, respectively. Then the optimal fusion estimator weighted by matrices can be computed as follows:

$$\hat{x}_0 = \sum_{i=1}^{L} A_i \hat{x}_i, \tag{17}$$

where, the optimal matrix weights $A_i \in R^{n\times n}, i = 1,2,\ldots,L$, are computed by

$$[A_1,\ldots,A_L] = (e^T P^{-1} e)^{-1} e^T P^{-1}, \tag{18}$$

where, $[A_1,\ldots,A_L] \in R^{n\times nL}$, $P = (P_{ij}) \in R^{nL\times nL}$, $e = [I_n,\ldots,I_n]^T \in R^{nL\times n}$. The error variance of the optimal information fusion estimator with matrix weights is computed by

$$P_0 = (e^T P^{-1} e)^{-1}, \tag{19}$$

and we have the relations $trP_0 \le trP_i,\ i = 1,2,\ldots,L$.

Remark 3: In this lemma, the mean and variance information of x is not used or unknown indeed. From the proof of this lemma in [9], we can see that $\sum_{i=1}^{L} A_i = I$. This condition is sufficient for the unbiasedness of the fused filter, when Ex is unknown. But it is unnecessary in Lemma 3, where Ex is known. What's more, if $C_{xY}C_Y^{-1}e = I$, the fusion filter in Lemma 3 is identical to the fusion filter in Lemma 4. When $C_{xY}C_Y^{-1}e \neq I$, from the optimality of the projection theory, the fusion filter in Lemma 3 is better than the fusion filter in Lemma 4.

IV. Main Results

First, for system (1),we give the local Kalman filters as follows.

Lemma 5 [9]: Under Assumptions 1 and 2, the ith local sensor subsystem of system (1) has the local Kalman filter

$$\hat{x}_i(t+1|t+1) = \hat{x}_i(t+1|t) + K_i(t+1)\varepsilon_i(t+1), \tag{20}$$

$$\hat{x}_i(t+1|t) = \bar{\Phi}_i \hat{x}_i(t|t) + J_i y_i(t), \tag{21}$$

$$\varepsilon_i(t+1) = y_i(t+1) - H_i \hat{x}_i(t+1|t), \tag{22}$$

$$K_i(t+1) = P_i(t+1|t)H_i^T[H_iP_i(t+1|t)H_i^T + R_i]^{-1}, \tag{23}$$

$$P_i(t+1|t) = \bar{\Phi}_iP_i(t|t)\bar{\Phi}_i^T + \Gamma(Q - S_iR_i^{-1}S_i^T)\Gamma^T, \tag{24}$$

$$P_i(t+1|t+1) = [I_n - K_i(t+1)H_i]P_i(t+1|t), \tag{25}$$

where $\bar{\Phi}_i = \Phi - J_iH_i$, $J_i = S_iR_i^{-1}$. The initial values are $\hat{x}_i(0|0) = \mu_0, P_i(0|0) = P_0$.

The local filtering error cross covariances are given by the following lemma.

Lemma 6 [9]: Under Assumptions 1 and 2, the local Kalman filtering error cross covariance between the ith and the jth sensor subsystems has the following recursive form:

$$\begin{aligned} P_{ij}(t+1|t+1) &= [I_n - K_i(t+1)H_i]\{\bar{\Phi}_iP_{ij}(t|t)\bar{\Phi}_j^T \\ &+\Gamma Q\Gamma^T - J_jR_jJ_j^T - J_iR_iJ_i^T + J_iR_{ij}J_j^T \\ &+\bar{\Phi}_iK_i(t)(R_{ij}J_j^T - S_i^T\Gamma^T) \\ &+(J_iR_{ij} - \Gamma S_j)K_j^T(t)\bar{\Phi}_j^T\} \times [I_n - K_j(t+1)H_j]^T \\ &+K_i(t+1)R_{ij}K_j^T(t+1), \end{aligned} \tag{26}$$

where the initial value is $P_{ij}(0|0) = P_0$.

For system (1), the mean and variance information of $x(t)$ is as follows.

Theorem 1: The mean and variance of $x(t)$ can be computed by

$$E(x(t)) = \Phi E(x(t-1)), \tag{27}$$

$$C_x(t) = \Phi C_x(t-1)\Phi^T + \Gamma Q \Gamma^T. \tag{28}$$

with initial values $E(x(0)) = \mu_0$, $C_x(0) = P_0$. And we also have

$$E(x(t)) = \Phi^t \mu_0, \tag{29}$$

$$C_x(t) = \Phi^t P_0 (\Phi^t)^T + \sum_{k=0}^{t-1} \Phi^{t-1-k}\Gamma Q\Gamma^T(\Phi^{t-1-k})^T. \tag{30}$$

Proof: From the system model (1), and assumptions 1 and 2, we have (27) and (28). Also from the system model (1), we have

$$x(t) = \Phi^t x(0) + \sum_{k=0}^{t-1} \Phi^{t-1-k}\Gamma w(k). \tag{31}$$

So we obtain (29) and (30). This completes the proof.

Now, given $\hat{x}_i(t|t), i = 1,2,...,L$, define

$$Y(t) = [\hat{x}_1^T(t|t),...,\hat{x}_L^T(t|t)]^T, \tag{32}$$

then

$$E(Y(t)) = eE(x(t)) = e\Phi^t\mu_0. \tag{33}$$

Using the fusion algorithm in Lemma 3, we give following optimal fusion filter.

Remark 4: From Theorem 1, we can see that, the mean and variance information of $x(t)$ can be computed off-line, given the initial values values $E(x(0))$ and $C_x(0)$.

Theorem 2: Given $\hat{x}_i(t|t), i = 1,2,...,L$, the optimal distributed fusion filter of $x(t)$ is

$$\hat{x}_0(t|t) = \Phi^t\mu_0 + C_{x(t),Y(t)}C_{Y(t)}^{-1}[Y(t) - e\Phi^t\mu_0], \tag{34}$$

where

$$C_{x(t),Y(t)} = [C_{x(t),\hat{x}_1(t|t)},...,C_{x(t),\hat{x}_L(t|t)}], \tag{35}$$

$$C_{x(t),\hat{x}_i(t|t)} = C_x(t) - P_i(t|t), \tag{36}$$

and

$$[C_{Y(t)}]_{ij} = C_x(t) - P_i(t|t) - P_j(t|t) + P_{ij}(t|t), \tag{37}$$

where $P_i(t|t)$, $P_j(t|t)$ and $P_{ij}(t|t)$ are computed by Lemma 5 and Lemma 6, $C_x(t)$ is given by Theorem 1.

Proof:From the fusion algorithm in Lemma 3, the optimal distributed fusion filter of $x(t)$ is

$$\hat{x}_0(t|t) = E(x(t)) + C_{x(t),Y(t)}C_{Y(t)}^{-1}[Y(t) - E(Y(t))], \tag{38}$$

Substituting (29) and (33) into (38), we have (34). From (13),(14) and (15), we have (35), (36) and (37). This completes the proof.

Remark 5: In Theorem 2, the local filters are Kalman filters, $\tilde{x}_i \perp \hat{x}_i$, so, the conditions in Lemma 3 are satisfied. If the local filters are not Kalman filters, and the orthogonality of $\tilde{x}_i$ and $\hat{x}_i$ does not exist, Lemma 3 and Theorem 2 do not hold.

V. Simulation Example

For ease of performance comparison, we consider the radar tracking system with three sensors presented in [9]

$$x(t+1) = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & T \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} w(t), \quad (39)$$

$$y_i(t+1) = H_i x(t) + v_i(t), \quad (40)$$

$$v_i(t) = \alpha_i w(t) + \xi_i(t), i = 1, 2, 3. \quad (41)$$

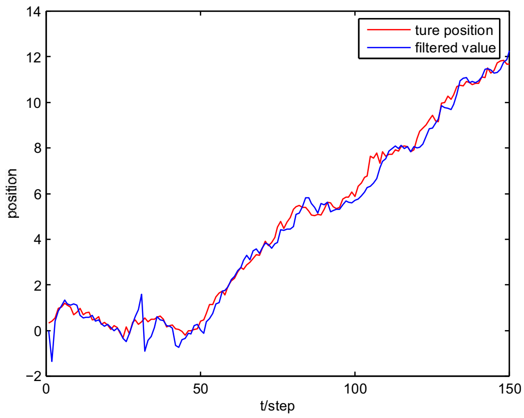

Fig. 1. The fusion filtering result for position.

In this system, the sampling period is T . The state $x(t) = [s(t)\ \dot{s}(t)\ \ddot{s}(t)]^T$, where $s(t)$, $\dot{s}(t)$ and $\ddot{s}(t)$ are the position, velocity and acceleration, respectively, of the target at time t. $y_i(t), i = 1, 2, 3$ are the measurement signals. The process noise $w(t)$ is zero-mean Gaussian white noise and has variance σ_w^2.

The coefficients α_i are constant scalars. And $\xi_i(t), i = 1, 2, 3$ are zero-mean Gaussian white noises with variance matrices $\sigma_{\xi_i}^2$, and are independent of $w(t)$. Our aim is to find the optimal distributed fusion filter $\hat{x}_0(t|t)$.

In this example, we set $T = 0.01$, $\sigma_w^2 = 1$; $\sigma_{\xi_1}^2 = 5$, $\sigma_{\xi_2}^2 = 8$, $\sigma_{\xi_3}^2 = 10$; $\alpha_1 = 0.5$, $\alpha_2 = 0.8$, $\alpha_3 = 0.4$; $H_1 = [1, 0, 0]$, $H_2 = [0, 1, 0]$, $H_3 = [0, 0, 1]$, the initial value $x(0) = 0$, $P_0 = 0.1 I_3$. We take 150 samples.

The three local Kalman filters are given by Lemma 5. The fusion filter is given by Theorem 2. The filtering effects of our fusion filter are shown in Fig.1-Fig.3. The filtering performance comparison of the three local filters, the matrix-weighting fusion filter in [9], and our fusion filter, is shown in Fig.4. From Fig.4, we can see that our fusion filter has better filtering performance than the matrix-weighting fusion filter.

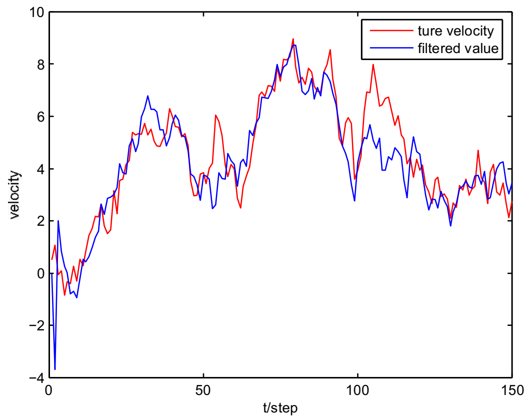

Fig. 2. The fusion filtering result for velocity.

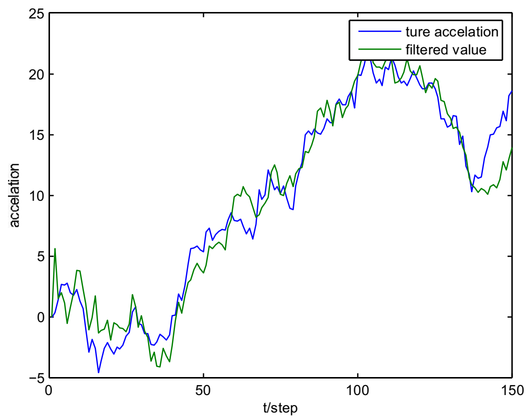

Fig. 3. The fusion filtering result for acceleration.

VI. Conclusions

This paper considers the distributed optimal fusion filtering problem for multi-sensor systems. First, a new fusion algorithm with the mean and variance information of the state is presented, based on the projection theory. Then the mean and variance information of the state is computed. At last, after obtaining the local filtering results and the error variance and cross covariance matrices, the distributed optimal fusion filter is derived by the new fusion algorithm. The performance comparison between this fusion filter and the matrix-weighting fusion filter is made.

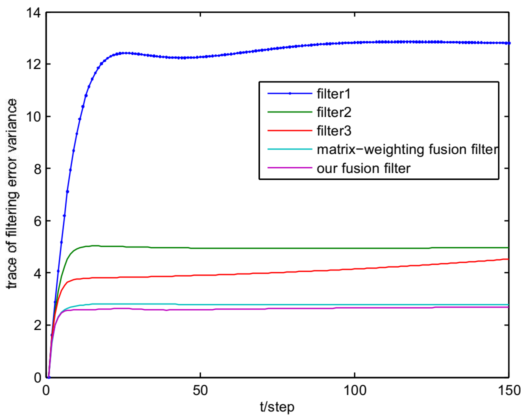

Fig. 4. The filtering performance comparison.

ACKNOWLEDGMENT

The authors are grateful for the reviewers for their constructive comments and suggestion.

REFERENCES

[1] Y. Bar-Shalom, "On the track-to-track correlation problem," IEEE Trans. Automat. Contr., vol. 26, no. 2, pp. 571-572, Apr. 1981.

[2] N. A. Carlson, "Federated square root filter for decentralized parallel processes," IEEE Trans. Aeros. Electron. Syst., vol. 26, no. 3, pp. 517-525, May 1990.

[3] Z.-Q. Luo and J. N. Tsitsiklis, Data fusion with minimal communication, IEEE Trans. Inf. Theory, vol. 40, no. 5, pp. 1551-1563, Sep. 1994.

[4] E. Song, Y. Zhu, and J. Zhou, "Sensors optimal dimensionality compression matrix in estimation fusion," Automatica, vol. 41, no. 12, pp. 2131-2139, Dec. 2005.

[5] K. H. Kim, "Development of track to track fusion algorithm," Proc. Amer. Contr. Conf., pp. 1037-1041, 1994.

[6] Z. Liu, Y. Liu, and C. Li, Distributed sparse recursive least-squares over networks, IEEE Trans. Signal Process., vol. 62, no. 6, pp. 1386-1395, Mar. 2014.

[7] X. Li, Y. Zhu, J. Wang, C. Han, "Optimal linear estimation fusion-Part I: Unified fusion rules," IEEE Trans. Inf. Theory, vol. 49, no. 9, pp. 2192-2208, Sept. 2003.

[8] J. Fang and H. Li, Joint dimension assignment and compression for distributed multisensor estimation, IEEE Signal Process. Lett., vol. 15, pp. 174-177, Jan. 2008.

[9] S. Sun, Z. Deng, "Multi-sensor optimal information fusion Kalman filter," Automatica, vol. 40, no. 6, pp. 1017-1023, Jun. 2004.

[10] J. Sun, C. Zhang, B. Guo, "Distributed full-order optimal fusion filters and smoothers for discrete-time stochastic singular systems," Int. J. Syst. Sci., vol. 42, no. 3, pp. 507-516, Mar 2011.

[11] S. Sun, L. Xie, W. Xiao, Y. C. Soh, "Optimal linear estimation for systems with multiple packet dropouts," Automatica, vol. 44, no. 5, pp. 1333C1342, May 2008.

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