Dynamic Systems and Applications 16 (2007) 393-406

OPTIMAL FUSION OF SENSOR DATA FOR DISCRETE KALMAN FILTERING

Z. G. FENG, K. L. TEO, N. U. AHMED, Y. ZHAO, AND W. Y. YAN

College of Mathematics and Computer Science, Chongqing Normal University, Chongqing, P.R.C.
Department of Mathematics and Statistics, Curtin University of Technology, Perth, Australia
School of Information Technology and Department of Mathematics, University of Ottawa, Ottawa, Canada
Department of Mathematics, Zhongshan University, Guangzhou, P.R.C.
Department of Electrical and Computer Engineering, Curtin University of Technology, Perth, Australia

ABSTRACT. In this paper we consider the question of optimal fusion of sensor data in discrete time. The basic problem is to design a linear filter whose output provides an unbiased minimum variance estimate of a signal process whose noisy measurements from multiple sensors are available for input to the filter. The problem is to assign weights to each of the sources (sensor data) dynamically so as to minimize estimation errors. We formulate the problem as an optimal control problem where the weight given to each of the sensor data is considered as one of the control variables satisfying certain constraints. There are as many controls as there are sensors. We develop an efficient method for determining the optimal fusion strategy and gives a numerical result for illustration.

Keywords: *a priori* estimate, discrete Kalman filter, optimal control, sensor

1. INTRODUCTION

In many physical problems, a large amount of data is available for collection. The collection is done at discrete time from different sensors with diverse degrees of reliability. On the basis of these collected data, a natural aim is to estimate the needed but unknown signal as accurately as possible. In the case of single sensor and linear system in Gaussian environment, the best estimator is given by the Kalman filter (recursive estimator). In this paper, we consider the situation in which the data is available at discrete time points from multiple sensors with varying degrees of reliability. Our aim is to find a way to assign an appropriate weight to each of the sensor sources so that unreliable data would not have too much an influence in the estimation. We formulate this problem as a discrete time optimal control problem where the control is the weight vector process dynamically assigning degree

Received November 23, 2005

of importance to sensor sources. We develop an efficient computational method for calculating the optimal weight assignment strategy.

Multisensor problem has received considerable attention in literature. In [11], the measurement adaptive problem involving linear systems with Gaussian perturbations and quadratic cost is formulated and solved. In [2], the sensor scheduling problem is considered in continuous time, where scheduling policies are considered as processes adapted to the observation σ-algebra. It is then shown that the optimal scheduling policy can be obtained by solving a quasi-variational inequality. However, this general formulation is much too complex for an optimal solution to be computed. In [10], the sensor scheduling problem considered is in continuous time involving linear systems. It corresponds to the situation, where the control variables are restricted to take values in a discrete set but with switchings being allowed to take place in continuous time. This formulation leads to an optimal discrete valued control problem, which is a special case of the form considered in [9] and [15]. For the case of discrete time system, the sensor scheduling problem is solvable by the tree search type of algorithms, and greedy algorithms. For example, see those reported in [3], [6], [7], [12] and [13]. In [5], by deriving a precise expression of an effective lower bound, a branch and bound method, which is based on the positive semi-definite property of the covariance matrix introduced in [8], is developed to seek the exact optimal solution.

In [4], we consider the optimal fusion problem in continuous time. The problem is formulated as an optimal control problem involving a matrix Riccati differential equation where the weight given to each of the sensor data is considered as one of the control variables satisfying certain constraints. The existence of an optimal weighting function for each of the sensor sources is established and an efficient method to determine an optimal fusion strategy is developed. In this paper, we shall study the discrete time version of the optimal fusion problem considered in [4]. The modification for this study is that the collection of data is done at discrete time.

The rest of the paper is organized as follows. In Section 2, we present the underlying mathematical problem and formulate the problem for achieving our objectives, which turns out to be a stochastic optimal control problem in discrete time. We then show that this discrete time stochastic optimal control problem is equivalent to a discrete time deterministic optimal control problem. In Section 3, the existence of an optimal control is established. In Section 4, we develop an efficient computational method for the solution of the discrete time deterministic optimal control problem. For illustration, a numerical result is obtained in Section 5.

2. MATHEMATICAL FORMULATION

Let $(\Omega, \mathcal{F}, P)$ be a given probability space. Consider a system governed by the following linear stochastic difference equation

$$x(t+1) = A(t)x(t) + B(t)V(t), \qquad t \in I_1, \tag{2.1a}$$

with initial condition

$$x(0) = x_0, \tag{2.1b}$$

where $I_1 = \{0, 1, \cdots, T-1\}$, and for each $t \in I_1$, $A(t) \in \mathbb{R}^{n\times n}$ and $B(t) \in \mathbb{R}^{n\times d}$ are matrices with real elements. The process $\{V(t), t \in I_1\}$ is a sequence of independent standard Gaussian random vectors with values in $\mathbb{R}^d$ on $(\Omega, \mathcal{F}, P)$, and the mean and covariance are given by

$$E\{V(t)\} = 0 \quad \text{and} \quad E\{(V(t)V(s))\} = \delta_{t,s}, \qquad t, s \in I_1.$$

The initial state x_0 is a $\mathbb{R}^n$-valued Gaussian random vector on $(\Omega, \mathcal{F}, P)$ with mean $E(x_0) = \widehat{x}_0$ and covariance matrix $E\{(x^0 - \widehat{x}^0)(x^0 - \widehat{x}^0)^\intercal\} = P_0$.

Denote $I_2 = \{1, 2, \cdots, T\}$. Our aim is to estimate the process $\{x(t), t \in I_2\}$ based on the measurement data obtained by N sensors, which are governed by the following family of linear stochastic difference equations given by

$$y_i(t) = C_i(t)x(t) + D_i(t)W_i(t), \qquad t \in I_2, \tag{2.2}$$

where, for each $t \in I_1$,

$$C_i(t) \in \mathbb{R}^{m\times n}, \;\; D_i(t) \in \mathbb{R}^{m\times m}, \;\; y_i(t) \in \mathbb{R}^m,$$

and for each i, $1 \le i \le N$, $\{W_i(t), t \in I_2\}$ is a sequence of independent standard $\mathbb{R}^m$-valued Gaussian random vectors.

By virtue of all these available data

$$\{y_i(t), \;\; t \in I_2, \;\; i = 1, \cdots, N\},$$

our objective is to find a dynamical strategy to assign appropriate weights $\{\alpha_i(t), t \in I_2\}$ to all the individual sensor sources. We then use the weighted measurement given by

$$y(t) = \sum_{i=1}^{N} \alpha_i(t) y_i(t), \qquad t \in I_2, \tag{2.3a}$$

where

$$\alpha_i(t) \ge 0, \qquad i = 1, \cdots, N, \tag{2.3b}$$

$$\sum_{i=1}^{N} \alpha_i(t) = 1, \tag{2.3c}$$

to estimate the signal so that the estimation error is minimum. This objective can be formulated as a stochastic optimal control problem. For this, let us define the set

$$\Lambda \equiv \{\lambda \in \mathbb{R}^N : \lambda_i \geq 0,\ i = 1, \dots, N,\ \text{and} \sum_{i=1}^{N} \lambda_i = 1\}. \tag{2.4}$$

Clearly, Λ is a compact convex set.

Any $\alpha(t) = [\alpha_1(t), \cdots, \alpha_N(t)]^\intercal \in \mathbb{R}^N$ such that $\alpha(t) \in \Lambda$ for $t \in I_2$ is called an admissible control. Let $\mathcal{U}$ be the set of all such admissible controls.

For any given $\alpha \in \mathcal{U}$, let

$$\mathcal{F}_\alpha = \sigma\{y(s),\ s \in I_2\} \tag{2.5}$$

denote the smallest σ-algebra relative to which y is measurable. It is well known that the unbiased minimum variance estimate of the process $\{x\}$, given the history $\mathcal{F}_\alpha$, is given by its conditional expectation:

$$\hat{x}(t) = E\{x(t)|\mathcal{F}_\alpha\}. \tag{2.6}$$

We can formulate our basic problem as a stochastic discrete time optimal control problem as follows:

Problem (P1)**.** *Find an $\alpha \in \mathcal{U}$ such that*

$$E\{x(t)\} = E\{\hat{x}(t)\}, \qquad \forall t \in I_2, \tag{2.7}$$

and that it minimizes the cost functional given by

$$J(\alpha) = \sum_{t=0}^{T-1} Tr\{\Sigma(t)P_\alpha(t)\}dt + c_1 Tr\{P_\alpha(T)\} + c_2 \sum_{j=1}^{N}\sum_{t=1}^{T-1} |\alpha_j(t+1) - \alpha_j(t)|, \tag{2.8}$$

where Σ is an $n \times n$-positive definite matrix-valued function satisfying

$$|\Sigma_{ij}(t)| \leq L, \qquad \forall t \in I_1,\ i, j = 1, \cdots, n,$$

while L, c_1, and c_2 are positive constants, and

$$P_\alpha(t) = E\{(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^\intercal |\mathcal{F}_\alpha\} \tag{2.9}$$

is the estimation error covariance matrix.

The first two terms of the cost functional (2.8) aim to minimize estimation errors with a special emphasis on the terminal error, while the third term imposes a penalty on the frequency of switching and the magnitudes of jumps so as to prevent frequent and large changes in control policies.

Given the *a priori* estimate of the state at time t, denoted by $z^-(t)$, we seek an update estimate $z^+(t)$ based on the current measurement data $y(t)$, we impose that the estimator have the following structure

$$z^+(t) = E(t)z^-(t) + \Gamma(t)y(t), \tag{2.10}$$

where the filter inputs are the decision process $\{\alpha\}$, which represents the weight given to each of the noisy measurement channel, and the weighted measurement data $\{y\}$. Our problem is to find the matrices

$$\{E(t), \Gamma(t), F_i(t), \quad i = 1, 2, \cdots, N\}$$

so that for all $t \in I_2$, $z^+(t)$ is the best (unbiased minimum variance = UMV) linear estimate for $x(t)$. Define the estimation error (after and before the measurement) process as follows:

$$e^-(t) = z^-(t) - x(t), \quad t \in I_2, \quad \text{before measurement} \tag{2.11a}$$

$$e^+(t) = z^+(t) - x(t), \quad t \in I_2, \quad \text{after measurement} \tag{2.11b}$$

Substituting (2.11) in (2.10), we obtain

$$\begin{aligned} e^+(t) =& E(t)z^-(t) + \Gamma(t)y(t) - x(t) \\ =& E(t)e^-(t) + \Gamma(t)\sum_{i=1}^{N} \alpha_i(t)y_i(t) + E(t)x(t) - x(t) \\ =& E(t)e^-(t) + \Gamma(t)\sum_{i=1}^{N} \alpha_i(t)(C_i(t)x(t) + D_i(t)W_i(t)) + (E(t) - I)x(t) \\ =& E(t)e^-(t) + \Gamma(t)\sum_{i=1}^{N} \alpha_i(t)D_i(t)W_i(t) + (\Gamma(t)\sum_{i=1}^{N} \alpha_i(t)C_i(t) + E(t) - I)x(t) \end{aligned} \tag{2.12}$$

By assumption

$$E\{W_i(t)\} = 0, \qquad t \in I_2,$$

and hence, if

$$E\{e^-(t)\} = 0, \qquad t \in I_2, \tag{2.13}$$

then we can have an unbiased estimate if and only if

$$E(t) = I - \Gamma(t)\sum_{i=1}^{N} \alpha_i(t)C_i(t), \qquad t \in I_2. \tag{2.14}$$

Substituting (2.14) into (2.12), we have

$$\begin{aligned} e^+(t) &= (I - \Gamma(t)\sum_{i=1}^{N} \alpha_i(t)C_i(t))e^-(t) + \Gamma(t)\sum_{i=1}^{N} \alpha_i(t)D_i(t)W_i(t) \\ &= e^-(t) + \Gamma(t)[y(t) - \sum_{i=1}^{N} \alpha_i(t)C_i(t)z^-(t)] \end{aligned} \tag{2.15}$$

To determine the error covariance, define the error covariance matrices before measurement and after measurement as follows:

$$P^-(t) = E\{e^-(t)(e^-(t))^\intercal\}, \qquad \text{before measurement} \tag{2.16a}$$

$$P^+(t) = E\{e^+(t)(e^+(t))^\intercal\}, \qquad \text{after measurement} \tag{2.16b}$$

Substituting (2.15) into (2.16), we have

$$\begin{aligned}P^+(t) =&E\{[(I-\Gamma(t)\sum_{i=1}^N \alpha_i(t)C_i(t))e^-(t)+\Gamma(t)\sum_{i=1}^N \alpha_i(t)D_i(t)W_i(t)]\cdot\\ &[(I-\Gamma(t)\sum_{i=1}^N \alpha_i(t)C_i(t))e^-(t)+\Gamma(t)\sum_{i=1}^N \alpha_i(t)D_i(t)W_i(t)]^\intercal\}\\ =&(I-\Gamma(t)\sum_{i=1}^N \alpha_i(t)C_i(t))P^-(t)(I-\Gamma(t)\sum_{i=1}^N \alpha_i(t)C_i(t))^\intercal\\ &+\Gamma(t)(\sum_{i=1}^N \alpha_i^2(t)D_i(t)D_i^\intercal(t))\Gamma^\intercal(t)\end{aligned} \tag{2.17}$$

We take the *a priori* estimate of the state as

$$z^-(t+1) = A(t)z^+(t), \tag{2.18}$$

then we compute the error covariance time update

$$P^-(t+1) = A(t)P^+(t)A^\intercal(t) + B(t)B^\intercal(t) \tag{2.19a}$$

with initial condition

$$P^+(0) = P_0. \tag{2.19b}$$

Note that, for fixed but arbitrary $\alpha \in \mathcal{U}$, we may choose Γ that minimizes a given cost functional if an additional assumption is imposed (see [1], Section 5.5, p. 80). Details are given in the following theorem.

Theorem 1. *Consider the system* (2.1), (2.2) *and* (2.3). *Suppose $\alpha \in \mathcal{U}$ is given and the elements of the matrices $\{A(t), B(t), t \in I_1\}$ and $\{C_i(t), D_i(t), t \in I_2\}$ are bounded matrix-valued functions, V, W_i $(i = 1, \cdots, N)$ are a sequence of independent standard Gaussian random vectors on $(\Omega, \mathcal{F}, P)$. Further, suppose that the initial state has finite second moment. Let the cost functional be given by*

$$\mathcal{L}(\alpha) = \sum_{t=0}^{T-1} Tr\{\Sigma(t)P_\alpha^+(t)\} + c_1 Tr\{P_\alpha^+(T)\}, \tag{2.20}$$

where Σ is an $n \times n$-positive definite matrix-valued function satisfying

$$|\Sigma_{ij}(t)| \le L, \qquad \forall t \in I_1,\ i,j = 1,\cdots,n.$$

Then the best linear filter is given by the following set of difference equations

$$z^+(t) = (I - \Gamma_\alpha(t)\Xi_\alpha(t))z^-(t) + \Gamma_\alpha(t)y(t), \tag{2.21a}$$

with initial condition

$$z(0) = \widehat{x}^0. \tag{2.21b}$$

Define

$$R_\alpha(t) = \sum_{i=1}^{N} (\alpha_i(t))^2 D_i(t) D_i^{\intercal}(t), \qquad t \in I_2. \tag{2.22}$$

If $R_\alpha(t)$ is positive definite, then the optimum $\Gamma_\alpha(t)$ is given by

$$\Gamma_\alpha(t) = P_\alpha^-(t)\Xi_\alpha^{\intercal}(t)(\Xi_\alpha(t)P_\alpha^-(t)\Xi_\alpha^{\intercal}(t) + R_\alpha(t))^{-1}, \tag{2.23}$$

where

$$\Xi_\alpha(t) = \sum_{i=1}^{N} \alpha_i(t) C_i(t), \tag{2.24}$$

and the set of error covariance matrices $\{P_\alpha^+(t)\}$ is given by the solution of the matrix difference equation

$$P_\alpha^+(t+1) = (I - \Gamma_\alpha(t+1)\Xi_\alpha(t+1))[A(t)P_\alpha^+(t)A^{\intercal}(t) + B(t)B^{\intercal}(t)], \tag{2.25a}$$

with initial condition

$$P_\alpha^+(0) = P_0. \tag{2.25b}$$

3. **Optimal Fusion**

The main results are presented in the following two theorems.

Theorem 2. *The stochastic optimal control Problem $(P1)$ is equivalent to the following deterministic optimal control problem:*

Problem (P2)**.** *Find an $\alpha \in \mathcal{U}$ such that the cost functional*

$$\mathcal{L}(\alpha) = \sum_{t=0}^{T-1} Tr\{\Sigma(t)P_\alpha^+(t)\} + c_1 Tr\{P_\alpha^+(T)\} + c_2 \sum_{j=1}^{N}\sum_{t=1}^{T-1} |\alpha_j(t+1) - \alpha_j(t)| \tag{3.1}$$

is minimized subject to the dynamic constraint:

$$P_\alpha^+(t+1) = (I - \Gamma_\alpha(t+1)\Xi_\alpha(t+1))[A(t)P_\alpha^+(t)A^{\intercal}(t) + B(t)B^{\intercal}(t)], \tag{3.2a}$$

with initial condition

$$P_\alpha^+(0) = P_0, \tag{3.2b}$$

where c_1, c_2 are positive constants,

$$\Gamma_\alpha(t) = P_\alpha^-(t)\Xi_\alpha^{\intercal}(t)(\Xi_\alpha(t)P_\alpha^-(t)\Xi_\alpha^{\intercal}(t) + R_\alpha(t))^{-1}, \tag{3.3}$$

and R_α, Ξ_α are given by (2.22) and (2.24), respectively.

Proof. From our choice of E, as given by (2.14), it follows that the filter is unbiased and that the error dynamics is given by (2.15). From this equation, we obtained the dynamics for the error covariance matrix given by equation (2.17) and (2.19). The cost functional (2.20) assumes its minimum at Γ_α given by (2.23). Thus, for fixed but arbitrary $\alpha \in \mathcal{U}$, Problem ($P2$) is obtained by substituting the expression for optimum Γ_α given by (2.23) into the error covariance equation (2.17) and (2.19). Hence, Problem ($P1$) is equivalent to Problem ($P2$). □

Note that Problem ($P2$) is a deterministic optimal control problem. The existence of an optimal control is given in the following theorem.

Theorem 3. *Consider the control problem* ($P2$)*. Suppose that* R_λ *given by*

$$R_\lambda(t) = \sum_{i=1}^{N} \lambda_i^2 D_i(t) D_i^\intercal(t) \tag{3.4}$$

is positive definite for all $\lambda \in \Lambda$ *and for all* $t \in I_2$*. Then, Problem* ($P2$) *admits an optimal control* $\alpha^* \in \mathcal{U}$*.*

Proof. Note that the domain $\mathcal{U}$ is a compact and convex subset in

$$\mathcal{A} = \{\alpha : \max_{t \in I_2} \max_{1 \le i \le N} |\alpha_i(t)| < \infty\}. \tag{3.5}$$

Therefore, it suffices to prove the continuity of the cost functional with respect to α.

Clearly, the third term of the cost functional is continuous, as an absolute value function is a continuous function.

Consider the first two terms, we need to show the continuity of $P_\alpha^+(t)$ for all $t \in I_2$. Clearly, $\alpha \to \Xi_\alpha$ and $\alpha \to R_\alpha$ are continuous by their definitions. Continuity of $\alpha \to \Gamma_\alpha(t)$ follows from the expression (3.3) and the facts that $\alpha \to R_\alpha(t)$ is continuous and positive definite and that $\alpha \to P_\alpha^-(t)$ is continuous. Continuity of the later, $\alpha \to P_\alpha^-(t)$, follows from the recursion (2.19a) and the facts that $P_\alpha^+(0) = P_0$ which is independent of α, and $\alpha \to P_\alpha^+(1)$ is continuous. Continuity of $P_\alpha^+(t)$ now follows from the recursion (3.2a) and the continuity of $\alpha \to P_\alpha^+(1)$.

This completes the proof. □

Remark 1. A sufficient condition for $R_\lambda(t)$ to be positive definite for $\lambda \in \Lambda$ is that

$$\bigcup_{i=1}^{N} \{t \in I_2 : Ker(D_i^\intercal(t)) = \{0\}\} = I_2.$$

Physically this means that for each $t \in I_2$ not more than $N-1$ sensors are noiseless or equivalently at least one sensor is noisy. If all the sensors are noiseless over a nonempty subset of $t \in I_2$, we have a singular situation. Using the results reported in ([1], Chapter10), it can be shown that this singular situation will give rise to Riccati type equations subject to algebraic constraints.

4. **Numerical Solution**

In view of Problem $(P2)$, we note that for each $j = 1, \cdots, N$, $|\alpha_j(t+1) - \alpha_j(t)|$ is non-smooth. We choose to use the smoothing technique reported in [14] to smooth it as:

$$S_\rho(y) = \begin{cases} |y|, & \text{if } |y| > \rho \\ [y^2 + \rho^2]/2\rho, & \text{if } |y| \le \rho \end{cases} \tag{4.1}$$

Then, we have the approximate Problem $(P2(\rho))$, which is Problem $(P2)$ with its cost function (3.1) being approximated by

$$\mathcal{L}_\rho(\alpha) = \sum_{t=0}^{T-1} Tr\{\Sigma(t)P_\alpha^+(t)\} + c_1 Tr\{P_\alpha^+(T)\} + c_2 \sum_{j=1}^{N}\sum_{t=1}^{T-1} S_\rho(\alpha_j(t+1) - \alpha_j(t)). \tag{4.2}$$

The cost functional (3.1) contains nonsmooth terms which are now approximated by smooth ones in (4.2). The relationship between the Problem $(P2)$ and its subsequent approximate problems is given in the following theorem.

Theorem 4. *Let $\alpha^{\rho,*}$ and α^* be, respectively, optimal solutions to Problem $(P2(\rho))$ and Problem $(P2)$. Then,*

$$0 \le \mathcal{L}_\rho(\alpha^{\rho,*}) - \mathcal{L}(\alpha^*) \le c_2[N(T-2)]\rho/2. \tag{4.3}$$

Proof. The proof is similar to that given for Theorem 10.4.1 of [14]. □

Problem $(P2(\rho))$ is a standard optimal parameter selection problem. It can be viewed as a mathematical programming problem. For this, we require the gradient formula for the cost functional.

Let $\tilde{\alpha}(t) = \alpha(t+1)$, and

$$\tilde{x}(t) = (P_{11}^+(t), P_{12}^+(t), \cdots, P_{1n}^+(t), P_{22}^+(t), P_{23}^+(t), \cdots, P_{2n}^+(t), \cdots, P_{nn}^+(t))^\intercal.$$

Next, let f be the corresponding vector obtained from the right hand side of (3.2). Then,

$$\tilde{x}(t+1) = f(t, \tilde{x}(t), \tilde{\alpha}(t)) \tag{4.4}$$

and J_ρ denotes the corresponding cost functional obtained from (4.2), i.e.,

$$J_\rho(\tilde{\alpha}) = \tilde{\Phi}_\rho(\tilde{x}(T), \tilde{\alpha}) + \sum_{t=1}^{T-1} \tilde{\mathcal{L}}_\rho(t, \tilde{x}(t), \tilde{\alpha}(t)). \tag{4.5}$$

Thus, the gradient formula of J_ρ with respect to $\tilde{\alpha}$ is given (see [14] for the derivation) by

Theorem 5. *Consider Problem $(P2)(\rho)$. The gradient of the cost functional J_ρ with respect to $\tilde{\alpha}$ is:*

$$\frac{\partial J_\rho(\tilde{\alpha})}{\partial \tilde{\alpha}} = \left[\frac{\partial \tilde{\Phi}_\rho(\tilde{x}(T),\tilde{\alpha})}{\partial \tilde{\alpha}(0)} + \frac{\partial \tilde{H}_\rho(0,\tilde{x}(0),\tilde{\alpha}(0),\lambda(1))}{\partial \tilde{\alpha}(0)}, \cdots, \right.$$

$$\left. \frac{\partial \tilde{\Phi}_\rho(\tilde{x}(T),\tilde{\alpha})}{\partial \tilde{\alpha}(T-1)} + \frac{\partial \tilde{H}_\rho(T-1,\tilde{x}(T-1),\tilde{\alpha}(T-1),\lambda(T))}{\partial \tilde{\alpha}(T-1)}\right], \tag{4.6}$$

where $\tilde{H}_\rho$ is the Hamiltonian function given by

$$\tilde{H}_\rho(t,\tilde{x}(t),\tilde{\alpha}(t),\lambda(t+1)) = \tilde{\mathcal{L}}_\rho(t,\tilde{x}(t),\tilde{\alpha}(t)) + \lambda^\intercal(t+1) f(t,\tilde{x}(t),\tilde{\alpha}(t)), \tag{4.7}$$

and $\lambda(t), t = T-1, T-2, \cdots, 1,$ is the solution of the costate system:

$$\lambda(t) = \frac{\partial \tilde{H}_\rho(t,\tilde{x}(t),\tilde{\alpha}(t),\lambda(t+1))}{\partial \tilde{x}(t)}, \tag{4.8a}$$

with final conditions

$$\lambda(T) = \frac{\partial \tilde{\Phi}_\rho(\tilde{x}(T),\tilde{\alpha})}{\partial \tilde{x}(T)}. \tag{4.8b}$$

To solve the optimal parameter selection Problem (P2(ρ)), we use the optimal control software, MISER 3.2 [14]. For this, we need, at each iteration, the value of the cost functional (4.2) and its gradient (4.6). We propose to solve the optimal fusion Problem (P2) using the following procedure:

Algorithm:

Consider Problem $(P2(\rho))$, we choose constants $\rho_0 > 0, 0 < \gamma < 1$ and set $\rho_{k+1} = \gamma\rho_k$.

Step 1. Solve Problem $(P2(\rho))$ with $\rho = \rho_1$ by using the optimal control software, MISER 3.2 [14]. Let the optimal control obtained be denoted by $\alpha^{\rho_1,*}$.

Step 2. Use $\alpha^{\rho_k,*}$ as initial guess to solve Problem $(P2(\rho))$ with $\rho_{k+1} = \gamma\rho_k$. Let the optimal control obtained be denoted by $\alpha^{\rho_{k+1},*}$.

Step 3. If $|\mathcal{L}_{\rho_{k+1}}(\alpha^{\rho_{k+1},*}) - \mathcal{L}_{\rho_k}(\alpha^{\rho_k,*})| < \varepsilon$, where ε is a pre-specified error constant, go to Step 4. Otherwise, go to Step 2 with $\rho_{k+1} =: \rho_k$.

Step 4. Stop. We have obtained an approximate optimal solution to Problem $(P2)$.

5. Illustrative Example

In this section, we shall apply the method developed in previous sections to the following example.

Example: Consider the following time variant optimal control problem, where the system dynamic is described by:

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ x_3(t+1) \end{bmatrix} = \begin{bmatrix} 0.05t & 0 & 0.1t+0.4 \\ 0.2 & 0.2t & 0 \\ 0 & 1 & 0.1t \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 0.05t+0.5 \\ 0.1t+0.5 \\ 0.2t+0.8 \end{bmatrix} V(t)$$

with initial condition

$$\begin{bmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Assume that there are three sensors given by

$$y_1(t+1) = \begin{bmatrix} 1 & 0.05t & 0.2 \\ 0 & 1 & 0.1t+1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} W_{11}(t) \\ W_{12}(t) \end{bmatrix}$$

$$y_2(t+1) = \begin{bmatrix} 0 & 1 & 2 \\ 0.1t & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} W_{21}(t) \\ W_{22}(t) \end{bmatrix}$$

$$y_3(t+1) = \begin{bmatrix} 0.2t+0.8 & 0 & 1 \\ 1 & 0.1t-0.5 & 0.8 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} W_{31}(t) \\ W_{32}(t) \end{bmatrix}$$

The cost functional is

$$J(u) = \sum_{t=0}^{T-1} Tr\{P^+(t)\} + c_1 Tr\{P^+(T)\} + c_2 \sum_{j=1}^{N} \sum_{t=1}^{T-1} |\alpha_j(t+1) - \alpha_j(t)|.$$

We take $T = 20, c_1 = 1, c_2 = 0.5, \varepsilon = 10^{-2}, \gamma = 0.1$, ρ is reduced gradually from 10^{-3} to 10^{-4}, and the solutions obtained are reported in Table 1.

ρ	10^{-3}	10^{-4}
J_ρ	1.342316e+001	1.341435e+001

TABLE 1. Numerical Results

Figure 1 shows that the optimal weight given to each of the data sources is time varying. Initially, the second data source heavily influences the estimation. Later, the first data becomes more important than the others. Toward the end, the third data source is the most important one. For the purpose of comparing our results with those obtained using the single sensor strategy, we test the error function

$$\sum_{t=0}^{T-1} Tr\{P^+(t)\} + Tr\{P^+(T)\}$$

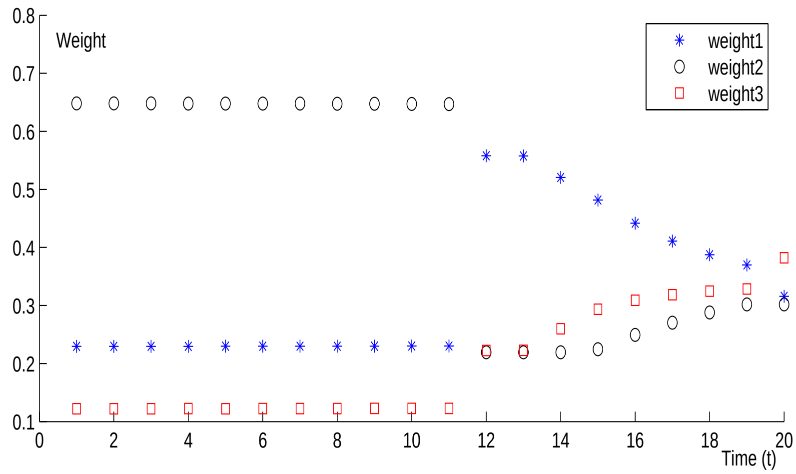

FIGURE 1. Optimal Weights

with different strategies and different values of T. The results are reported in Table 2. In view of the results presented in Table 2, we can see that the fusion strategy is a highly effective way to reduce the accumulative error.

M	1	2	3	4	5
Using only sensor 1	0.2357	0.4593	0.6670	0.8584	1.0391
Using only sensor 2	0.2005	0.4118	0.6315	0.8599	1.0999
Using only sensor 3	0.8485	2.0521	3.4589	4.9834	6.5803
Using optimal strategy	0.1852	0.3790	0.5783	0.7828	0.9943
M	6	7	8	9	10
Using only sensor 1	1.2135	1.3862	1.5629	1.7544	1.9844
Using only sensor 2	1.3549	1.6290	1.9285	2.2649	2.6639
Using only sensor 3	8.2384	9.9739	11.8236	13.8425	16.1156
Using optimal strategy	1.2152	1.4484	1.6982	1.9720	2.2859
M	11	12	13	14	15
Using only sensor 1	2.3153	2.9191	4.2774	7.6960	16.7415
Using only sensor 2	3.1900	4.0181	5.6475	9.5463	19.8723
Using only sensor 3	18.8057	22.1807	26.2752	30.6766	35.1321
Using optimal strategy	2.6778	3.2084	3.9444	4.8665	5.9273
M	16	17	18	19	20
Using only sensor 1	41.5358	100.9794	177.2286	234.6523	275.4797
Using only sensor 2	45.9379	98.5956	179.8121	286.2377	416.9216
Using only sensor 3	39.6206	44.1834	48.8659	53.7075	58.7418
Using optimal strategy	7.1125	8.4030	9.7862	11.2575	12.7449

TABLE 2. Comparative Results

6. Conclusion

In this paper, we construct a model of optimal fusion problem, where the collection of data is collected at discrete time from different sensors with diverse degrees of reliability. A linear filter is designed such that its output provides an unbiased minimum variance estimate of a signal process whose noisy measurements from multiple sensors are available for input to the filter. The problem of assigning appropriate weight to each of the sources dynamically so as to minimize estimation error is formulated as a discrete time deterministic optimal control problem.

Our contributions include showing the existence of an optimal weighting function for each of the sensor sources, and the development of an efficient method for calculating an optimal fusion strategy. From the numerical experience gained, we see that the proposed method is highly efficient.

Acknowledgements

This research is partially supported by a research grant from the Research and Development Office of Curtin University of Technology, and a research grant from the Australian Research Committee.

REFERENCES

[1] N. U. Ahmed, *Linear and nonlinear filtering for scientists and engineers*, World Scientific, Singapore. New Jersey. London. Hong Kong, 1999.

[2] L. S. Baras, and A. Bensoussan, *Optimal sensor scheduling in nonlinear filtering of diffusion process*. SIAM Journal of Control and Optimization, **27** (1989), 786–813.

[3] T. H. Chung, V. Gupta, J. W. Burdick and R. M. Murray, *On a decentralized active sensing strategy using mobile sensor platforms in a network*, Proc. of the 43rd IEEE Conference on Decision and Control, **2** (2004), 1914–1919.

[4] Z. G. Feng, K. L. Teo, N. U. Ahmed, Y. Zhao, W. Y. Yan, *Optimal fusion of sensor data for Kalman filtering*, Discrete and Continuous Dynamical Systems, Series A, **14**(3) 2006, 483–503.

[5] Z. G. Feng, K. L. Teo, Y. Zhao, *Branch and bound method for sensor scheduling in discrete time*, Journal of Industrial and Management Optimization, **1**(4) 2005, 499–512.

[6] V. Gupta, T. H. Chung, B. Hassibi and R. M. Murray, *Sensor scheduling algorithms requiring limited computation*, Proc. of the International Conference on Acoustics, Speech and Signal Processing, **3** (2004), 825–828.

[7] S. Kagami and M. Ishikawa, *A sensor selection method considering communication delays*, Proc. of the 2004 IEEE International Conference on Robotics & Automation, **1** (2004), 206–211.

[8] A. Logothetis, A. Isaksson and R. J. Evans, *An information theoretic approach to observer path design for bearings-only tracking*, Proc. of the 36th Conference on Decision & Control, **4** (1997), 3132–3137.

[9] H. W. J. Lee, K. L. Teo and L. S. Jennings, *Control parametrization enhancing techniques for optimal discrete-valued control problems*, Automatica, **35** (1999), 1401–1407.

[10] H. W. J. Lee, K. L. Teo and A. E. B. Lim, *Sensor scheduling in continuous time*, Automatica, **37** (2001), 2017–2023.

[11] L. Meier, J. Peschon and R. M. Dressler, *Optimal control of measurement subsystems*. IEEE Transactions on Automatic Control, **12** (1967), Issue 5, 528–536.

[12] Y. Oshman, *Optimal sensor selection strategy for discrete-time state estimators*, IEEE Transactions on Aerospace and Electronic Systems, **30** (1994), 307–314.

[13] A. V. Savkin and R. J. Evans and E. Skafidas, *The problem of optimal robust sensor scheduling*, Proc. of the 39th Conference on Decision and Control, **4** (2000), 3791–3796.

[14] K. L. Teo, C. J. Goh and K. H. Wong, A unified computational approach to optimal control problems, Longman Scientific and Technical, New York, 1991.

[15] K. L. Teo, L. S. Jennings, H. W. J. Lee and V. Rehbock, *The control parametrization enhancing transform for constrained optimal control problems*, Journal of Australian Mathematical Society, Series B, **40** (1999), 314–335.