Wavelet-Based Multi-Sensor Optimal Information Fusion

M. Cai
Luoyang Institute of Electro-optical Equipment, AVIC,
Science and Technologyon Electron-optic Control Laboratory,
Luoyang, China

J.X. Li
Shanghai Jiaotong University,
Shanghai, China

Abstract—The problem of multi-sensors information fusion is studied in time-frequency domain, and a new optimal criteria weighted by scalars for unify multi-sensor systems is presented. Wavelet transform is introduced in multi-scale signal filtering, the approximate component and details are both updated. The local sensor estimate is fused via an optimal algorithm weighted by scalars, then reconstruct at the finest scale. The method proposed: (1) “complete” multi-scale filtering for unify multi-sensor system, estimation performance is greatly enhanced; (2) distributed fusion weighted by scalars, only requires the computation of scalar weights, avoids the computation of matrix weights, the computational burden can obviously be reduced; (3) the simulation also shows it outperforms optimal fusion filter weighted by scalars and centralized multi-sensor fusion.

Keywords—wavelet transform; multi-scale filtering; optimal information fusion; multi-sensor

I. INTRODUCTION

In the recent 20 years, multi-sensor information fusion has becoming an important research area[13,14]. There are many effective method: weighted average estimate, Bayes estimate, Kalman filter(centralized or distributed), maximum likelihood estimate, neural network fusion. Weighted average estimate is the most simple and intuitive fusion method. Carlson[1] presents the famous federated square root filter weighted by matrices. Kim[2] give the multi-sensor optimal information fusion estimate in the maximum likelihood sense, which also needs computation of matrix weights. In[3][4] the optimal information fusion criterion weighted by scalars is derived. In[5] the information fusion criterions weighted by matrices and weighted by scalar are discussed, their precision and computational burdens are compared. The fusion criterion weighted by scalar is a better choice in engineering application for its computational advantages.

The former information fusion are mostly derived from time-domain, the abundant frequency information contained in signal are not utilized during the processing. Recently, the introduction wavelet transform started multi-scale signal processing. Chou and Willsyk[6] started pioneer research on modeling and estimation of multi-scale stochastic processes using wavelet transforms, a dynamic multi-scale model and algorithm was built. The correlative research involves[7,8,9,10]Hong proposed an optimal and dynamic multi-scale distributed algorithm for multiresolutional sensory information. For uniform resolution multi-sensor systems, it is invalid. And during the multi-scale filtering, the details which also contains noise ,are not updated.

In the paper, multi-sensor information fusion is studied in the view of application needs. A wavelet-based optimal fusion filter weighted by scalars is proposed. In section 2 the necessity of details filtering in multi-scale signal filtering is discussed. The optimal fusion filter weighted by scalars in time domain and wavelet domain are presented respectively in section 3 and section 4. An simulation is given in section 5 to illustrate the algorithm. Finally, section 6 concludes the paper.

II. OPTIMAL FUSION FILTER WEIGHTED BY SCALARS IN TIME DOMAIN

Consider the discrete-time stochastic system with multiple sensors:

\[ x(k+1) = Ax(k) + Bw(k) \]

\[ z_j(k) = C_j x(k) + v_j(k) \]

where \( j = 1,2...s \) denotes sensor index, \( k \) denotes measured sequence index. \( w(k) \sim N(0,q) \), \( v_j(k) \sim N(0,r) \), The initial value of \( x(0) \) is a random vector with a mean and a variance matrix given by: \( E\{x(0)\} = x_0 \), \( E\{(x(0) - x_0)(x(0) - x_0)^T\} = P_0 \), it is assumed that \( x(0) \), \( w(k) \), \( v_j(k) \) are independent of each other.

We assume all local sensors are faultless. Based on the local optimal Kalman estimation \( \hat{x}_j(k \mid k) \), \( j = 1,2...s \), multi-sensor information optimal(i.e. linear minimum variance) fusion with scalar weights:

\[ \hat{x}(k \mid k) = \sum_{j=1}^{s} \alpha_j \hat{x}_j(k \mid k) \]

Where satisfied:

(a) Unbiasedness, namely \( E\hat{x}(k \mid k) = E\hat{x}(k) \)

(b) Optimality, namely, to find the optimal scalar...
weights \( \overline{\alpha}_j(k) \) to minimize the performance index
\( J = \text{tr}P(k | k) \), i.e. \( \text{tr} \hat{P}(k | k) = \min \{ \text{tr}P(k | k) \} \),
where \( P(k | k) \) denotes the variance of arbitrary fusion filter
with scalar weights, \( \hat{P}(k | k) \) denotes the variance of the
optimal fusion filter with scalar weights, symbol \( \text{tr} \)
refers to the trace of a matrix.

The optimal fusion scalar weights \( \overline{\alpha}_j(k) \) computation is
given as follows:
\[
\overline{\alpha}_j(k) = \frac{\sum_{i=1}^e \omega \pi_j(k) \overline{\pi}_j(k) \hat{P}_j(k | k)}{e \sum_{i=1}^e \omega_j(k) \overline{\pi}_j(k) \hat{P}_j(k | k)}
\tag{4}
\]
where \( \sum = \{ \text{tr}P_i(k | k) \}, i,j = 1,2,...,l \) is an \( l \times l \) positive
definite matrix with \( \hat{P}_j(k | k) = \hat{P}_j(k | k) \),
\( \overline{\alpha}(k) = [\overline{\alpha}(k),...\overline{\alpha}(k)]^T \),
e = \{1,...,1\} \are both \( l \)-dimensional vectors. The
corresponding variance of fusion estimate \( \hat{x}(k | k) \) is
computed by:
\[
\hat{P}(k | k) = \sum_{j=1}^l \pi_j(k) \overline{\pi}_j(k) \hat{P}_j(k | k)
\tag{5}
\]
and we have \( \text{tr} \hat{P}(k | k) \leq \text{tr} \hat{P}_j(k | k) \),
\( \hat{P}_j(k | k) \) is the cross-covariance matrix of filtering errors
between the local estimate \( \hat{x}_j(k | k) \) and \( \hat{x}_j(k | k) \), suppose
the sensors are independent, we have \( \hat{P}_j(k | k) = 0 \).

III. WAVELET-BASED MULTI-SCALE SIGNAL FILTERING

The idea of multi-scale signal filtering[6,7]: decomposing the
signal measured at the finest scale onto an orthonormal
basis via wavelet transform, multi-scale signal can be obtained.
Updating the multi-scale signal , then an optimal estimate at
basis via wavelet transform, multi-scale signal can be obtained.

Consider a finite sequence of n-dimensional random
vectors at resolution level \( N \) with a length of a data block:
\[
X(k_N)=[x^T \big( kM_N+1 \big),x^T \big( kM_N+2 \big),...x^T \big( kM_N+M_N \big)]^T
\tag{6}
\]
Where the length \( M_N = 2^{N-1} \), \( k_N \) denotes for \( k \) th data
block at resolution level \( N \). In order to change \( X(k_N) \) to
the form required by the wavelet transform, we introduce a
linear transformatino. For instance, for a sequence with two
two-dimensional vectors \( X(k_N)=[(x_{11},x_{12}), (x_{21},x_{22})]^T \), the
transformation \( L_N \) can be defined such that:
\[
[x_{11},x_{12},x_{21},x_{22}]^T = L_N X(k_N) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix} X(k_N)
\tag{7}
\]
Wavelet-based decomposition from level \( N \) to level \((N - 1)\) can be derived in terms of operators:
\[
X_r(k_{N-1}) = L_N^T \text{diag}[H_N,\ldots,H_N]L_N X(k_N)
\tag{8}
\]
\[
X_s(k_{N-1}) = L_N^T \text{diag}[G_N,\ldots,G_N]L_N X(k_N)
\tag{9}
\]
Where subscript \( V \) denotes the approximate component at
level \((N - 1) \), and \( D \) denotes the details at level \((N - 1) \). Linear
transformation \( L_N \) and \( L_{N-1} \) are introduced to change
\( X(k_N) \) and \( X(k_{N-1}) \) to the form required by the wavelet
transform[2]. The inverse transformation form:
\[
X(k_N) = L_N^T \text{diag}[H_N^T,\ldots,H_N^T]L_N X_r(k_{N-1})
+ L_N^T \text{diag}[G_N^T,\ldots,G_N^T]L_N X_s(k_{N-1})
\tag{10}
\]
The multi-scale discrete wavelet transform from level \( N \)
to level \( l \) is illustrated in Fig.1.

The multi-scale decomposition can be obtained by:
\[
\begin{bmatrix}
X_r(k_l) \\
X_s(k_{l+1})
\end{bmatrix} = L_{l-1}^T \text{diag}[H_{l-1},H_{l-1}]L_{l-1} X_r(k_{l+1})
\tag{11}
\]
where \( T(k_l) \) is an orthogonal matrix:
\[
T(k_l) = \begin{bmatrix}
\prod_{j=1}^{N-l+1} H_j \\
\prod_{j=1}^{N-l+1} H_j \\
\end{bmatrix} \\
\begin{bmatrix}
G_{l+1} \prod_{j=1}^{N-l+1} H_j \\
G_{l+1} \prod_{j=1}^{N-l+1} H_j \\
\end{bmatrix}
\]
The inverse transform of Eqn.(11):
\[
X'(k_l) = L_{l+1}^T \text{diag}[H_{l+1},H_{l+1}]L_{l+1} X_r(k_l)
+ L_{l+1}^T \text{diag}[G_{l+1},G_{l+1}]L_{l+1} X_s(k_l)
\tag{12}
\]
The superscript \( l \) denotes the reconstruction from level \( l \)
to the finest level \( N \).

Assume \( X(k_N) \) a white noise contaminated signal. Through multi-scale transform, we get an orthogonal projection \( X_r(k_N) \) and \( X_s(k_{N-1}) \), \( X_s(k_{N-2}) \) ... \( X_s(k_1) \). Meanwhile, the white noise is decomposed at different levels.
Due to its uncorrelated and normal distribution, white noise
affects the signal in the entire \( l \sim w \) domain, the noise at
arbitrary level has the same statistical characteristic of the
original signal[11].

The decomposed multi-scale signal contains both useful
signal and noise component. In [7,10] only approximate
component $X_j(j = 1, \ldots, s)$ is updated, but the details are not. Hence, the multi-scale filtering is incomplete. Especially in the non-stationary environment or the noise has large amplitude, the noise is the main component of details, estimation accuracy is greatly influenced by the filtering of details. In the following part, we give proof that details filtering can improve the whole performance of estimation.

Firstly, for a given system described by Eq.(1) and Eq.(2) at resolution level $N$, the block length is $M_N = 2^{N-1}$, we use data block to present the system as:

$$X(k+1) = A(k)x(k) + B(k)w(k)$$

(13)

$$Z_j(k) = C_j(k)x(k) + V_j(k)$$

(14)

where:

$$A(k) = diag\{A, \ldots, A\}$$

(15)

$$C_j(k) = diag\{C_j, \ldots, C_j\}$$

(16)

$$b(k) = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ 0 & b_{22} & \cdots & b_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{ss} \end{bmatrix}$$

(17)

and $b_{vv} = A^{M_N-r}B^{M_N-r}$

$$W(k) \sim N(0,Q(k)) \quad Q(k) = diag\{q_1, \ldots, q_s\}$$

(18)

$$V_j(k) \sim N(0,R_j(k)) \quad R_j(k) = diag\{r_{j1}, \ldots, r_{js}\}$$

(19)

Suppose the system Eq.(13) and Eq.(14) is mapped from level $N$ to level $N-1$.

Theorem 1 Assume that $X(k)$ is interfered by white noise $W(k)$, $X_i(k|N)$ and $X_i(k|N-1)$ derived from Eq.(8,9) are updated by Kalman filter, $\hat{X}_i(k|N)$ and $\hat{X}_i(k|N-1)$ are the filter estimate, the reconstruction $\hat{X}_i(k)$ from Eq.(10) is optimal in the linear minimum variance(LMV) sense.

Proof:

Assume

$$TH_{v1} = diag\{H_{v1}^{(1)}, \ldots, H_{v1}^{(s)}\} \quad TG_{v1} = diag\{G_{v1}^{(1)}, \ldots, G_{v1}^{(s)}\}$$

(20)

equation(10) can be rewritten:

$$\hat{X}(k) = L_k^{T}\hat{H}_k^{T}L_k^{Q_k}\hat{X}_k(k) + L_k^{T}G_k^{T}L_k^{Q_k}\hat{X}_k(k)$$

(21)

$$P(k) = E[(\hat{X}(k) - X(k))(\hat{X}(k) - X(k))^T]$$

$$= L_k^{T}H_k^{T}J_kP_kL_k^{Q_k}H_k + L_k^{T}G_k^{T}L_k^{Q_k}G_k$$

(22)

where

$$P_k(k) = E[\hat{X}_k(k)|X(k)]E[\hat{X}_k(k)|X(k)]^T$$

(23)

$$P_k(k) = E[\hat{X}_k(k+1)|X(k+1)]E[\hat{X}_k(k+1)|X(k+1)]^T$$

(24)

As Kalman filter is optimal based on the linear minimum variance sense, $P(k|N)$ and $P(k|N-1)$ is the minimum variance, $P(k|N)$ can be linear represented by $P(k|N)$ and $P(k|N-1)$, hence, $P(k|N)$ is optimal in the linear minimum variance sense.

In [7] the details $P(k|N)$ is not updated, seen from Threm 1, the result is not optimal in LMV sense. In Part 4 , we presented a wavelet-based algorithm, which update $P(k|N)$ and $P(k|N)$ for each local sensor.

IV. WAVELET-BASED OPTIMAL FUSION FILTER WEIGHTED BY SCALARS

Consider a discrete-time system with multiple sensors described by Eq.(13,14). $j = 1,2 \ldots s$ denotes the sensor index, $k$ denotes data block index. Fig.2 is the flow chart for wavelet-based optimal fusion algorithm, $N$ and $i$ denotes different resolution level.

**FIGURE I. MULTI-SCALE TRANSFORM.**

**FIGURE II. WAVELET-BASED OPTIMAL FUSION ALGORITHM.**

The multi-scale signal filtering includes approximate component and detailed component, the distributed filtering estimate are fused by scalars in optimal fusion center, the scalars are computed from the variance of arbitrary local estimate. The output of fusion center is an optimal multi-scale state estimate, then, the optimal estimate at the finest scale can be get via inverse wavelet transform.

In summary, the algorithm can be implemented by going through the following procedure:

1. Wavelet &decomposition layer selection, the wavelet with a good time domain localization property, such as Haar wavelet, is preferred; $i = 1, N = 3 \sim 5$ is enough;

2. Multi-scale decomposition, $X_j(k)$ and $Z_j(k)$ is decomposed from finer scale to coarser scale using Eq.(11), such as $X_j(k)$ is decomposed to approximate $X_{j0}(k)$ and detailed $X_{j0}(k), X_{j0}(k), \ldots, X_{j0}(k) (j = 1,2 \ldots s);$

3. Kalman filter update at different scale, local output $\hat{X}_{j0}(k), \hat{X}_{j0}(k), \ldots, \hat{X}_{j0}(k)$ and $\hat{X}_{j0}(k) (j = 1,2 \ldots s);$

4. Compute the scalar $\bar{a}_{j0}(k)$ of $\hat{X}_{j0}(k)$ from Eq.(4)
and \( \hat{P}_D(k) \) can be get;

(6) Repeat step(4)-(6), compute fusion estimate \( \hat{X}_D(k_{i-1}) \), \( \hat{X}_D(k_{i}) \) and \( \hat{X}_\phi(k_{i}) \);

(7) Inverse wavelet transform from Equ.(12), the optimal fusion estimate \( \hat{X}(k_{\phi}) \) is obtained.

Remark: As the filter proposed is a information fusion algorithm in wavelet domain, the abundant frequency information is complete utilized. Especially, we testified the necessity of details updating. Hence, it should have better performance than the fusion filter in time domain. Meanwhile, the algorithm is weighted by scalars, it avoids the computation of matrix weight, and avoids the inverses of higher dimensional matrices, so that the computation burden may be reduced.

V. SIMULATION

As an example, the simulation of two-sensor information fusion is presented. The sensors are assumed to be independent. The target is take a constant velocity movement, fusion is presented. The sensors are assumed to be reduced. Dimensional matrices, so that the computation burden may be reduced.


VI. CONCLUSION

A wavelet-based optimal information fusion criterion weighted by scalar is proposed. The algorithm improves the performance of multisensor state estimation with less computational burden. The simulation also shows its effectiveness. It can be applied in estimate application.

REFERENCES


TABLE 1. POSITION ERROR STD (UNIT: M).

<table>
<thead>
<tr>
<th>Sensor 1</th>
<th>Sensor 2</th>
<th>Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old</td>
<td>2.1736</td>
<td>2.7653</td>
</tr>
<tr>
<td>New</td>
<td>1.9878</td>
<td>2.5543</td>
</tr>
<tr>
<td>Cen</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2. VELOCITY ERROR STD (UNIT: M/S).

<table>
<thead>
<tr>
<th>Sensor 1</th>
<th>Sensor 2</th>
<th>Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old</td>
<td>1.3808</td>
<td>1.5167</td>
</tr>
<tr>
<td>New</td>
<td>1.3676</td>
<td>1.5085</td>
</tr>
<tr>
<td>Cen</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>