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 filtering 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:

$$x(k+1) = Ax(k) + Bw(k) \tag{1}$$

$$z_j(k) = C_j x(k) + v_j(k) \tag{2}$$

where j =1,2… s denotes sensor index, k denotes measured sequence index. $w(k) \sim N(0,q)$, $v_j(k) \sim N(0,r_j)$,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:

$$\hat{x}(k \mid k) = \overline{a}_1(k)\hat{x}_1(k \mid k) + \ldots + \overline{a}_s(k)\hat{x}_s(k \mid k) \tag{3}$$

Where satisfied:

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

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

The optimal fusion scalar weights $\overline{a}_j(k)$ computation is given as follows:

$$\overline{a}(k) = \frac{\Sigma^{-1} e}{e^T \Sigma^{-1} e} \tag{4}$$

where $\Sigma = (trP_{ij}(k\mid k))$, $i, j = 1, 2, ..., l$,is an $l \times l$ positive definite matrix with $\hat{P}_{jj}(k\mid k) = \hat{P}_j(k\mid k)$. $\overline{a}(k) = [\overline{a}_1(k), ..., \overline{a}_s(k)]^T$, $e = [1, \ldots, 1]^T$ are both l -dimensional vectors. The corresponding variance of fusion estimate $\hat{x}(k\mid k)$ is computed by:

$$\hat{P}(k\mid k) = \sum_{i,j=1}^{s} \overline{a}_i(k)\overline{a}_j(k)\hat{P}_{ij}(k\mid k) \tag{5}$$

and we have $tr\hat{P}(k\mid k) \le tr\hat{P}_{jj}(k\mid k)$, $\hat{P}_{ij}(k\mid k)$ is the cross-covariance matrix of filtering errors between the local estimate $\hat{x}_i(k\mid k)$ and $\hat{x}_j(k\mid k)$, suppose the sensors are independent, we have $\hat{P}_{ij}(k\mid k) = 0$.

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 the finest scale level is generated by reconstruction.

Consider a finite sequence of n-dimensional random vectors at resolution level N with a length of a data block:

$$X(k_N) = [x^T(kM_N+1), x^T(kM_N+2), \cdots, x^T(kM_N+M_N)]^T \tag{6}$$

Where the length $M_N = 2^{N-1}$, k_N denotes for k th data block at resolunion level N .In order to change $X(k_N)$ to the form required by the wavelet transform, we introduce a linear transfromatino. 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 difined such that:

$$[x_{11}, x_{21}, x_{12}, x_{22}]^T = L_N X(k_N) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} X(k_N) \tag{7}$$

Wavelet-based decomposition from level N to level (N -1) can be derived in terms of operators:

$$X_V(k_{N-1}) = L_{N-1}^T diag\{H_{N-1}, \cdots, H_{N-1}\} L_N X(k_N) \tag{8}$$

$$X_D(k_{N-1}) = L_{N-1}^T diag\{G_{N-1}, \cdots, G_{N-1}\} L_N X(k_N) \tag{9}$$

Where subscript V denotes the approximate component at level (N -1) and D denotes the details at level (N -1). H_{N-1} and G_{N-1} are scaling and wavelet operators mapping from level N to 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:

$$\begin{aligned} X(k_N) &= L_N^T diag\{H_{N-1}^T, \cdots, H_{N-1}^T\} L_{N-1} X_V(k_{N-1}) \\ &+ L_N^T diag\{G_{N-1}^T, \cdots, G_{N-1}^T\} L_{N-1} X_D(k_{N-1}) \end{aligned} \tag{10}$$

The multi-scale discrete wavelet transform from level N to level i is illustrated in Fig.1.

The multi-scale decomposition can be obtained by:

$$\begin{bmatrix} X_V(k_i) \\ X_D(k_i) \\ \vdots \\ X_D(k_{N-1}) \end{bmatrix} = L_i^T diag\{T(k_i), \cdots, T(k_i)\} L_N X(k_N) \tag{11}$$

where $T(k_i)$ is an orthogonal matrix:

$$T(k_i) = \begin{bmatrix} \prod_{j=N-1}^{i} H_j \\ G_i \prod_{j=N-1}^{i} H_j \\ \vdots \\ G_{N-1} \end{bmatrix}$$

The inverse transform of Equ.(11):

$$X^i(k_N) = L_N^T diag\{T^T(k_i), \cdots, T^T(k_i)\} L_i \begin{bmatrix} X_V(k_i) \\ X_D(k_i) \\ \vdots \\ X_D(k_{N-1}) \end{bmatrix} \tag{12}$$

The superscript i denotes the reconstruction from level i to the finest level N.

Assume $X(k_N)$ a white noise contaminated signal. Through multi-scale transform, we get an orthogonal projection $X_V(k_i)$ and $X_D(k_{N-1})$, $X_D(k_{N-2})$,…, $X_D(k_i)$. Meanwhile, the white noise is decomposed at different levels. Due to its uncorrelated and normal distribution, white noise affects the signal in the entire $t \sim w$ 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 $X_V(k_i)$ 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 Equ.(1) and Equ.(2) at resolution level N , the block length is $M_N = 2^{N-1}$, we use data block to present the system as:

$$X(k_N+1) = A(k_N)X(k_N) + B(k_N)W(k_N) \tag{13}$$

$$Z_j(k_N) = C_j(k_N)X(k_N) + V_j(k_N) \tag{14}$$

where:

$$A(k_N) = diag\{A, \cdots, A\} \tag{15}$$

$$C_j(k_N) = diag\{C_j, \cdots, C_j\} \tag{16}$$

$$B(k_N) = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1M_N} & 0 & \cdots & 0 \\ 0 & b_{21} & b_{22} & \cdots & b_{2M_N} & \cdots & 0 \\ . & . & . & \cdots & . & . & . \\ 0 & \cdots & 0 & b_{M_N 1} & b_{M_N 2} & \cdots & b_{M_N M_N} \end{bmatrix} \tag{17}$$

and $b_{uv} = A^{M_N - v} B^{M_N - v}$

$$W(k_N) \sim N(0, Q(k_N)) \quad Q(k_N) = diag\{q, \cdots, q\} \tag{18}$$

$$V_j(k_N) \sim N(0, R_j(k_N)) \quad R_j(k_N) = diag\{r_{jN}, \cdots, r_{jN}\} \tag{19}$$

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

Theorem 1 Assume that $X(k_N)$ is interfered by white noise $W(k_N)$, $X_V(k_{N-1})$ and $X_D(k_{N-1})$ derived from Equ.(8,9) are updated by Kalman filter, $\hat{X}_V(k_{N-1})$ and $\hat{X}_D(k_{N-1})$ are the filter estimate, the reconstruction $\hat{X}(k_N)$ from Equ.(10) is optimal in the linear minimum variance(LMV) sense.

Proof:

Assume

$$TH_{N-1} = diag\{H_{N-1}^T, \cdots, H_{N-1}^T\} \quad TG_{N-1} = diag\{G_{N-1}^T, \cdots, G_{N-1}^T\} \tag{20}$$

equation(10)can be rewritten:

$$\hat{X}(k_N) = L_N^T TH_{N-1}^T L_{N-1}\hat{X}_V(k_{N-1}) + L_N^T TG_{N-1}^T L_{N-1}\hat{X}_D(k_{N-1}) \tag{21}$$

$$\begin{aligned} P(k_N) &= E\{[\hat{X}(k_N) - X(k_N)][\hat{X}(k_N) - X(k_N)]^T\} \\ &= L_N^T TH_{N-1}^T L_{N-1} P_V(k_{N-1}) L_{N-1}^T TH_{N-1} L_N \\ &+ L_N^T TG_{N-1}^T L_{N-1} P_D(k_{N-1}) L_{N-1}^T TG_{N-1} L_N \end{aligned} \tag{22}$$

where

$$P_V(k_{N-1}) = E\{[\hat{X}_V(k_{N-1}) - X_V(k_{N-1})][\hat{X}_V(k_{N-1}) - X_V(k_{N-1})]^T\} \tag{23}$$

$$P_D(k_{N-1}) = E\{[\hat{X}_D(k_{N-1}) - X_D(k_{N-1})][\hat{X}_D(k_{N-1}) - X_D(k_{N-1})]^T\} \tag{24}$$

As Kalman filter is optimal based on the linear minimum variance sense, $P_V(k_{N-1})$ and $P_D(k_{N-1})$ is the minimum variance, $P(k_N)$ can be linear represented by $P_V(k_{N-1})$ and $P_D(k_{N-1})$, hence, $P(k_N)$ is optimal in the linear minimum variance sense.

In [7] the details $P_D(k_{N-1})$ is not updated , seen from Therom 1, the result is not optimal in LMV sense. In Part 4 ,we presented a wavelet-based algorithm, which update $P_V(k_{N-1})$ and $P_D(k_{N-1})$ for each local sensor.

IV. WAVELET-BASED OPTIMAL FUSION FILTER WEIGHTED BY SCALARS

Consider a discrete-time system with multiple sensors described by Equ.(13,14). j =1,2… 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.

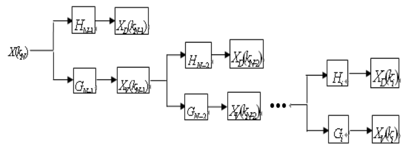

FIGURE Ⅰ.MULTI-SCALE TRANSFORM.

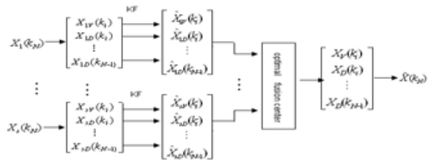

FIGURE Ⅱ.WAVELET-BASED OPTIMAL FUSION ALGORITHM.

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, $X_j(k_N)$ and $Z_j(k_N)$ is decomposed from finer scale to coarser scale using Equ.(11), such as $X_j(k_N)$ is decomposed to approximate $X_{jV}(k_i)$ and detailed $X_{jD}(k_{N-1}), X_{jD}(k_{N-2}), \ldots, X_{jD}(k_i)$,(j=1,2,…s);

(3)Kalman filter update at different scale, local output $\hat{X}_{jD}(k_{N-1})$, $\hat{X}_{jD}(k_{N-2}), \ldots, \hat{X}_{jD}(k_i)$ and $\hat{X}_{jV}(k_i)$,(j=1,2,…s);

(4)Compute the saclar $\overline{a}_D(k_i)$ of $\hat{X}_D(k_i)$ from Equ.(4)

and $\hat{P}_{1D}(k_i),\ldots\hat{P}_{sD}(k_i)$;

(5) Fusion by scalars from Equ.(3) ,(5), $\hat{X}_D(k_i)$ and $\hat{P}_D(k_i)$ can be get;

(6)Repeat step(4)-(6), compute fusion estimate $\hat{X}_D(k_{i+1}),\ldots,\hat{X}_D(k_{N-1})$ and $\hat{X}_V(k_i)$;

(7)Inverse wavelet transform from Equ.(12), the optimal fusion estimate $\hat{X}(k_N)$ 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, initial state $x(0)=[0m,1m/s]$, sample period T=1s, 150 sampling data is taken. Q=1, R1=9,R2=16."New" denotes wavelet-based optimal fusion algorithm proposed in the paper, "Old" denotes the optimal fusion filter weighted by scalars in time domain[5], "Cen" denotes centralized multi-sensor fusion algorithm(also well known as measurement fusion method[12]). We compared each sensor and the fusion result with three different algorithm respectively.Table 1 is the position error std comparison, and table 2 is the velocity error std comparison. 100 Monte-Carlo runs.

TABLE Ⅰ.POSITION ERROR STD (UNIT: M).

	Sensor 1	Sensor 2	fusion
Old	2.1736	2.7653	1.8486
New	1.9878	2.5543	1.5982
Cen			1.8123

TABLEⅡ. VELOCITY ERROR STD (UNIT: M/S).

	Sensor 1	Sensor 2	fusion
Old	1.3808	1.5167	1.3212
New	1.3676	1.5085	1.1197
Cen			1.2947

Analysis: 1)multi-sensor optimal fusion weighted by scalars improves the local estimate precision; 2) In time-domain, the precision of optimal fusion filter weighted by scalars is inferior to that of centralized fusion(weighted by matrix); 3)wavelet-based optimal fusion filter weighted by scalars outperforms the former two algorithms, the accuracy of position and velocity are both enhanced, especially the new algorithm has better performance than centralized fusion one.

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

[1] Carlon N A, Federated Square Root Filter for Decentralized Parallel Processes, IEEE Trans Aerospace and Electronic Systems, 26(2), pp.517-525.1990.

[2] Kim K H, Development of track to track fusion algorithm, Proceeding of the American Control Conference, Mary land,pp.1037-1041,1994.

[3] Deng, Z.L, Qi, R. B, Multi-sensor information fusion suboptimal steady-state Kalman filter, Chinese Science Abstracts, 6(2),pp.183–184,2000.

[4] Sun.Shuli, Multi-sensor information fusion white noise filter weighted by scalars based on Kalman predictor, Automatica, 40(8),pp.1447-1453,2004.

[5] Sun Shuli, Deng,Z.l, Multi-sensor Optimal Information Fusion Criterion in Linear Minimum Variance Sense, Science Technology and Engineering, 4(5),pp.334-337, 2004.

[6] Chou. K. C.,Golden,S, Willsky.A.S, Modeling and Estimation of Multisacle Sochastic Processes, Proceddings of IEEE International Conference on Acoustics, Speech and Signal Processing ,pp.1709-1712,1991.

[7] Hong.L. Mulitiresolutional Distributed Filtering, IEEE Trans. on Automatic Control,39(4), pp853-856,1994.

[8] Hong.L,Chen.Guanrong,Chui,Charles.K. Filter-bank-based Kalman Filtering Technique for Wavelet Estimation and Decomposition of Random Signals, IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing, 25(2), pp.237-241,1998.

[9] Benazza-Benyahia, A. Pesquet, J.-C. et al. A Nonlinear Diffusion-based Three-band Filter Bank. IEEE, Signal Processing Letters, 10(12), pp.360-363,2003.

[10] Lei Zhang,Quan Pan,Paul Bao et al, The Discrete Kalman Filtering of a Class of Dynamic Multiscale Systems, IEEE Trans. on Circuits and Systems II:Analog and Digital Signals Proccessing, 49(10),pp.668-676, 2002.

[11] D.L.Donoho, I.M.Johnstone, Ideal Spatial Adaptation by Wavelet Shrinkage, Biometrika, 81(3),pp.425-455,1994.

[12] Gan, Q. Harris, C.J. Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion. Source. IEEE Transactions on Aerospace and Electronic Systems, 37(1), pp. 273-280,2001

[13] Johnson, A.E.W.,Cholleti, S.R. ,Buchman, T.G.,Clifford, G.D. Improved respiration rate estimation using a Kalman filter and wavelet cross-coherence, Computing in Cardiology Conference (CinC), pp.791 – 794,2013

[14] www.fusion2014.org