Modeling and Estimation of a Class of Dynamic Multiscale Systems by General Compactly Supported Wavelet[1)]

PAN Quan[1] CUI Pei-Ling[1,2] ZHANG Lei[1] ZHANG Hong-Cai[1]
[1](*Department of Automatic Control, Northwestern Polytechnical University, Xi′an* 710072)
[2](*Department of Automation, Tsinghua University, Beijing* 100084)
(E-mail: quanpan@nwpu.edu.cn)

Abstract This paper addresses the problem of dynamic multiscale system (DMS) estimation. Research achievements in the related area have been reported in the literature, but they either rely on the assumption of stationarity of the observed process or are difficult to be implemented. In this paper, a model of DMS that meets the requirements of the standard discrete time Kalman filtering is built and is realized by general compactly supported wavelet. The introduction of the state space projection equation and the augmentation of measurement equation are a major part of the novelty in our work. A theorem on the optimal filtering output at each scale is put forward. Experimental results are given to verify our methods′ superior performances.

Key words Dynamic multiscale system, compactly supported wavelet, Kalman filtering

1 Introduction

The problem of estimating the state of a dynamic multiscale system (DMS) on the basis of available noisy measurements arises in a variety of contexts, including remote sensing and geophysics. It is also one of the well-known key problems in modern control theory.

A well-known achievement is the multiscale stochastic model that stems from the work of Willsky *et al* and that allows the modeling of multiresolution data at different levels in the bintree[1,2]. This model has made success in a number of applications[3~13]. Motivated by the success of this methodology in solving static estimation problems, a dynamic algorithm was proposed by propagating the static estimator over time with alternating update and prediction steps in a manner analogous to Kalman filtering[14,15]. In [16~20], measurements available at multiple resolution levels were integrated by the wavelet transform to deal with the target tracking by L. Hong.

Methods mentioned above are unrealistic for many real-world signals due to the time delay, complex realization process and so on. The application of dynamic estimation methods are still a challenging goal. In this paper, a novel model of DMS is built and is realized by general compactly supported wavelets except Haar wavelet. In this paper, this kind of wavelets is denoted as general compactly supported wavelet. It is our strong belief that the idea developed here has a far broader range of applicability.

2 State-space models

2.1 Problem formulation

The object considered in our DMS is information source S with real state $\boldsymbol{x} \in U_0 \subset R^n$, and we want to estimate its state based on all of the observations optimally in real-time. The states are characterized by a differential equation and the observations are obtained in a sequence of subspaces of the state-spaces. That is to say, the sensors are distributed in different resolution spaces, and the observations are projections of real state on every subspace. Ranking the sensors from high resolution to low, namely, from 1 to J, the resolution state of the first sensor is $\boldsymbol{x}_1 \in U_1$. Similarly, $\boldsymbol{x}_2 \in U_2, \ldots, \boldsymbol{x}_J \in U_J$, a set of linear subspaces $U_J \subset U_{J-1} \subset \ldots \subset U_1 \subset U_0$ in R^n space can be got.

2.2 State-space model of the continuous time DMS

1) the state space projection equation is

$$\boldsymbol{x}_j = P_j^0 \boldsymbol{x}, \quad j = 1, 2, \ldots, J \tag{1}$$

1) Supported by National Natural Science Foundation of P. R. China (60172037)
Received March 24, 2003; in revised form June 12, 2004

Assuming the projection operator $U_0 \to U_j$ is P_j^0, the projection of real state $\boldsymbol{x}$ on subspace U_j can be expressed as (1).

The resolution state $\boldsymbol{x}_2$ can be seen as linear projection of $\boldsymbol{x}_1$ from spaces U_1 to U_2 with the projection operator $P(1,2)$. Similarly, $\boldsymbol{x}_3$ can be viewed as the projection of $\boldsymbol{x}_2$ from U_2 to U_3 with the projection operator $P(2,3)$. (1) can in turn be written as

$$\boldsymbol{x}_j = P(j-1,j)x_{j-1}, \quad j = 2,3,\dots,J \tag{2}$$

We have the following simplified form of (2)

$$\boldsymbol{x}_j = P_j x_1, \quad j = 1,2,\dots,,J \tag{3}$$

where $P_j = P(j-1,j)P(j-2,j-1)\cdots P(1,2)$, $j = 1,2,\dots,J$, P_1 is I.

2) state differential equation

$\boldsymbol{x}_1$ is the most precise description of real state $\boldsymbol{x}$ that can be obtained. In other words, the estimation of $\boldsymbol{x}$ can only be replaced with that of $\boldsymbol{x}_1$. Mapping the state differential equation to space U_1, we have

$$\dot{\boldsymbol{x}}_1(t) = A(t)\boldsymbol{x}_1(t) + B(t)\boldsymbol{w}(t) \tag{4}$$

where $A(t)$ is the state transition matrix, $B(t)$ is the noise stimulus matrix, and $\boldsymbol{w}(t)$ is a Gaussian white noise process with covariance I.

3) augmented measurement equation

The measurement equation of j-th sensor has the form

$$\boldsymbol{z}_j(t) = C_j(t)\boldsymbol{x}_j(t) + \boldsymbol{v}_j(t), \quad j = 1,2,\dots,J \tag{5}$$

where $C_j(t)$ is the measurement matrix. Measurement noise $\boldsymbol{v}_j(t)$ is Gaussian white noise with variance $R_j(t)$. They are independent of each other and uncorrelated with system noise $\boldsymbol{w}(t)$.

Substituting (3) into (5), we derive

$$\boldsymbol{z}_j(t) = C_j(t)P_j\boldsymbol{x}_1(t) + \boldsymbol{v}_j(t), \quad j = 1,2,\dots,J$$

Then we can get the augmented measurement equation

$$\boldsymbol{z}_j(t) = C_j(t)\boldsymbol{x}_1(t) + \boldsymbol{v}(t) \tag{6}$$

where

$$\boldsymbol{z}(t) = \begin{bmatrix} \boldsymbol{z}_1(t) \\ \boldsymbol{z}_2(t) \\ \vdots \\ \boldsymbol{z}_J(t) \end{bmatrix}, \ C(t) = \begin{bmatrix} C_1(t)P_1 \\ C_2(t)P_2 \\ \vdots \\ C_J(t)P_J \end{bmatrix}, \ \boldsymbol{v}(t) = \begin{bmatrix} \boldsymbol{v}_1(t) \\ \boldsymbol{v}_2(t) \\ \vdots \\ \boldsymbol{v}_J(t) \end{bmatrix}$$

$\boldsymbol{v}(t)$ and $\boldsymbol{w}(t)$ are independent, and

$$Cov(\boldsymbol{v}(t)) = diag[R_1(t), R_2(t), \dots, R_J(t)]$$

(4) and (6) constitute the continuous time DMS model. The key to build this model is to map the estimated state to the finest scale and to introduce the state space projection equation.

3 Realization by general compactly supported wavelet

3.1 Typical discrete DMS

Discrete system can answer for many practical issues. Here, we discuss a class of typical discrete DMS. In frequency domain, resolution stands for the system bandwidth. According to the Shannon theorem, narrow bandwidth corresponds to low sampling rate. For convenience, we let the sampling rate decrease from sensor 1 to sensor J by a factor of two. Obviously, sensor 1 corresponds to the finest scale. The state at all scales in time interval ΔT is called a state block, and the measurement a data block. Every other ΔT, the state estimation must be updated when a new data block is available.

The state $\boldsymbol{x}_j(\bullet)$ is an approximation of resolution state $\boldsymbol{x}_1$ at the finest scale. It can be expressed in the form of the linear combination of the finest scale state $\boldsymbol{x}_1(\bullet)$, and the linear operator P_j in (3) is actually a low-pass filter H_j.

There is crude similarity between the multiscale analysis based on wavelet transform and the multiscale system. In wavelet transform, filters must meet some constrains to have perfect reconstruction, but here the filter H_j is determined by the physical characteristics of the sensors. In fact, determining the filter H_j from their physical characteristics is not easy, and it is not necessary to do so because we have rich classes of wavelet basis.

3.2 Realization of discrete time variant DMS models by general compactly supported wavelet

Since the sampling rate of sensors decreases by a factor of two, their frequency bandwidths are then presumed to decrease in the same ratio. For convenience, the filters are supposed to be finite impulse response (FIR) digital filter, a general choice is the Dau(N) compactly supported orthogonal wavelets by Daubechies[21] and compactly supported bi-orthogonal wavelets by Cohen[22]. Haar wavelet has only one vanishing moment. That is to say, its approximation ability is not good. If it cannot meet the requirement of a practical system, the state node at coarse scale must be approximated by wavelet with higher vanishing moment. In this case, the approximation precision might be much higher, but we must pay more cost, because the node approximation at time $k\Delta T$ must use the nodes outside it, thus the model becomes complex.

Suppose that the low-pass smoothing filter corresponding to the wavelet is N-tap D_0, and d_0 denotes the impulse response at moment zero. Let us define

$$D_0 = [d_m, d_{m+1}, \ldots, d_0, \ldots, d_n]$$

Then we have

$$n - m + 1 = N \tag{7}$$

There are N finer nodes $\boldsymbol{x}_j(2k-n), \boldsymbol{x}_j(2k-n+1), \ldots, \boldsymbol{x}_j(2k), \ldots, \boldsymbol{x}_j(2k-m)$ that relate to one coarse scale node $\boldsymbol{x}_{j+1}(k)$. Generally, the scale function is preferably to be symmetric. In this case, let

$$m = \begin{cases} -N/2, & \text{if } N \text{ is an even number} \\ -(N-1)/2, & \text{if } N \text{ is an odd number} \end{cases}, \quad n = \begin{cases} N/2-1, & \text{if } N \text{ is an even number} \\ (N-1)/2, & \text{if } N \text{ is an odd number} \end{cases}$$

For $N = 4$, $m = -2$, and $n = 1$, one node at coarse scale is related to four nodes at a finer scale. Here, the state structure is not standard bintree, the build of multiscale structure cannot be finished in one data block. As shown in Fig. 1, the approximation of coarse scale nodes not only need two nodes in the same time block, but also one node in the former time block and one in the latter.

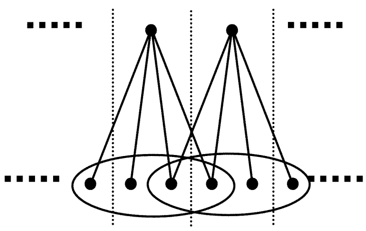


Fig. 1 The state nodes′ relation between adjacent scales ($N = 4, m = -2$, and $n = 1$)

For $m_j = 0, 1, \ldots, 2^{J-j} - 1$, the node $\boldsymbol{x}_2(2^{J-2}k + m_2)$ at scale 2 can be denoted as

$$\boldsymbol{x}_2(2^{J-2}k + m_2) = (D_0 * \boldsymbol{x}_1) \downarrow_2 = \sum_{h=m}^{n} d_h \boldsymbol{x}_1(2^{J-1}k + 2m_2 - h)$$

where $*$ represents convolution, $\downarrow_2$ denotes the downsampling by a factor of two. In the following, $\downarrow_p$ denotes the downsampling by a factor of p.

Letting

$$D_j = \Big[d_m, \underbrace{0, 0, \ldots, 0}_{2^j-1}, d_{m+1}, \underbrace{0, 0, \ldots, 0}_{2^j-1}, d_{m+2}, \ldots, d_{n-1}, \underbrace{0, 0, \ldots, 0}_{2^j-1}, d_n\Big] \tag{8}$$

which is obtained by inserting $2^j - 1$ zeros between every two coefficients of the filter, its zero impulse response is still d_0, for example

$$D_1 = [d_m, 0, d_{m+1}, 0, d_{m+2}, 0, \ldots, d_{n-1}, 0, d_n]$$

Similarly, the node at scale j is obtained as follows.

$$D^j = D_0 * D_1 * \cdots * D_{j-2}, \quad j = 2, 3, \ldots, J \tag{9}$$

Let $D^1 = [1]$

$$\boldsymbol{x}_j(\bullet) = (D^j * \boldsymbol{x}_1) \downarrow_{2^{j-1}}, \quad j = 1, 2, \ldots, J \tag{10}$$

Then the tap number of filter D^j can be derived as

$$N^j = (2^{j-1} - 1)(N - 1) + 1$$

Obviously, N^j and N have the same parity. Letting $D^j = [d^j_{m^j}, d^j_{m^j+1}, \ldots, d^j_{n^j}]$, the relations between m^j, n_j and m, n are

$$\begin{cases} m^j = (2^{j-1} - 1)m \\ n^j = (2^{j-1} - 1)n \end{cases}$$

From (10), we can get

$$\boldsymbol{x}_j(2^{J-j}k + m_j) = \sum_{h=m^j}^{n^j} d^j_h \boldsymbol{x}_1(2^{J-1}k + 2^{j-1}m_j - h)$$

At time $k\Delta T$, the root node of $x_J(k)$ is of the form

$$\boldsymbol{x}_J(k) = (D^J * \boldsymbol{x}_1) \downarrow_{2^{J-1}} = \sum_{h=m^J}^{n^J} d^J_h \boldsymbol{x}_1(2^{J-1}k - h)$$

It needs $N^J = (2^{J-1} - 1)(N - 1) + 1$ nodes at scale 1. Obviously, this is just the maximum node number to approximate all of the nodes at time $k\Delta T$. Defining

$$\bar{\boldsymbol{x}}(k) = \begin{bmatrix} \boldsymbol{x}_1(2^{J-1}k - (2^{J-1} - 1)n) \\ \boldsymbol{x}_1(2^{J-1}k - (2^{J-1} - 1)n + 1) \\ \vdots \\ \boldsymbol{x}_1(2^{J-1}k - (2^{J-1} - 1)m \end{bmatrix}$$

$\bar{\boldsymbol{x}}(k)$ has N^J fine-scale state components. Denoting

$$M_j(m_j) = \Big[\underbrace{O, \ldots, O}_{Z^j_1(m_j)}\ \underbrace{d^j_{n^j} \cdot I, d^j_{n^j-1} \cdot I, \ldots, d^j_{m^j} \cdot I,}_{N^j}\ \underbrace{O, \ldots, O}_{Z^j_2(m_j)}\Big]$$

where I is $N_x \times N_x$ identity matrix, O is $N_x \times N_x$ zero matrix, $Z^j_1(m_j)$ and $Z^j_2(m_j)$ are

$$Z^j_1(m_i) = 2^{j-1}m_j + n(2^{J-1} - 2^{j-1})$$
$$Z^j_2(m_j) = m(2^{j-1} - 2^{J-1}) - 2^{j-1}m_j$$

the node at scale j can be written as

$$\boldsymbol{x}_j(2^{J-j}k + m_j) = M_j(m_j)\bar{\boldsymbol{x}}(k) \tag{11}$$

From (11), the measurement of $\boldsymbol{x}_j(2^{J-j}k + m_j)$ can be obtained

$$\boldsymbol{z}_j(2^{J-j}k + m_j) = C_j(2^{J-j}k + m_j)M_j(m_j)\bar{\boldsymbol{x}}(k) + \boldsymbol{v}_j(2^{J-j}k = m_j)$$

Defining

$$\bar{\boldsymbol{z}}_j = col(\boldsymbol{z}_j(2^{J-j}k), \ldots, \boldsymbol{z}_j(2^{J-j}k + m_j), \ldots, \boldsymbol{z}_j(2^{J-j}k + 2^{J-j} - 1))$$

$$\bar{C}_j(k) = col(C_j(2^{J-j}k)M_j(0), C_i(2^{J-j}k+1)M_j(1), \ldots, C_j(2^{J-j}(k+1)-1)M_j(2^{J-j}-1))$$
$$\bar{\boldsymbol{v}}_j(k) = col(\boldsymbol{v}_j(2^{J-j}k), \boldsymbol{v}_j(2^{j-j}k+1), \ldots, \boldsymbol{v}_j(2^{J-j}(k+1)-1))$$

the covariance matrix of $\bar{\boldsymbol{v}}_j(k)$ is

$$\bar{R}_j(k) = diag[R_j(\bullet), R_j(\bullet), \ldots, R_j(\bullet)]$$

The measurement equation can be written in matrix form

$$\bar{\boldsymbol{z}}_j(k) = \bar{C}_j(k)\bar{\boldsymbol{x}}(k) + \bar{\boldsymbol{v}}_j(k) \tag{12}$$

Letting

$$\bar{\boldsymbol{z}}(k) = col(\bar{\boldsymbol{z}}_J(k), \bar{\boldsymbol{z}}_{J-1}(k), \ldots, \bar{\boldsymbol{z}}_1(k))$$
$$\bar{C}(k) = col(\bar{C}_J(k), \bar{C}_{J-1}(k), \ldots, \bar{C}_1(k))$$
$$\bar{\boldsymbol{v}}(k) = col(\bar{\boldsymbol{v}}_J(k), \bar{\boldsymbol{v}}_{J-1}(k), \ldots, \bar{\boldsymbol{v}}_1(k))$$

the variance of Gaussian white noise $\bar{\boldsymbol{v}}(k)$ is

$$\bar{R}(k) = diag[\bar{R}_J(k), \bar{R}_{J-1}(k), \ldots, \bar{R}_1(k)]$$

Then

$$\bar{\boldsymbol{z}}(k) = \bar{C}(k)\bar{\boldsymbol{x}}(k) + \bar{\boldsymbol{v}}(k) \tag{13}$$

For the general compactly supported wavelet, the multiscale system does not constitute bintree, a portion of the state vector $\bar{\boldsymbol{x}}(k)$ overlap with that of $\bar{\boldsymbol{x}}(k+1)$. When the measurements of time $(k+1)\Delta T$ are available, the overlapped part is performed by a smoothing process. The last element of $\bar{\boldsymbol{x}}(k)$ is $\boldsymbol{x}_1(2^{J-1}k - (2^{J-1}-1)m)$, the first element of $\bar{\boldsymbol{x}}(k+1)$ is $\boldsymbol{x}_1(2^{J-1}(k+1) - (2^{J-1}-1)n)$, the overlapping length is $l1 = (2^{J-1}-1)(N-2)$, and the non-overlapping length $l2 = 2^{J-1}$. In order to simplify the expression, we define $l(k) = 2^{J-1}k - (2^{J-1}-1)m$. Then the last element of $\bar{\boldsymbol{x}}(k+1)$ is $\boldsymbol{x}_1(l(k+1))$, and that of $\bar{\boldsymbol{x}}(k)$ is $\boldsymbol{x}_1(l(k))$, namely, $\boldsymbol{x}_1(l(k+1)-l2)$.

For the non-overlapping parts of $\bar{\boldsymbol{x}}(k+1)$ and $\bar{\boldsymbol{x}}(k)$, the following can be obtained according to the state equation

$$\boldsymbol{x}_1(l(k+1)-l2+1+m_1) = \prod_{n=0}^{m_1} A(l(k+1)-l2+1+m_1-n) \cdot \boldsymbol{x}_1(l(k+1)-l2) + \sum_{h=0}^{m_1} \prod_{n=0}^{m_1-h-1} A(l(k+1)-l2+1+m_1-n) \cdot B(l(k+1)-l2+h)\boldsymbol{w}(l(k+1)-l2+h)$$

where $m_1 = 0, 1, \ldots, 2^{J-1}$. Define

$$\bar{\boldsymbol{w}}(k) = col(\boldsymbol{w}(l(k+1)-l2), \boldsymbol{w}(l(k+1)-l2+1), \ldots, \boldsymbol{w}(l(k+1)-1))$$
$$A(k, m_1) = \prod_{n=0}^{m_1} A(l(k+1)-l2+m_1-n)$$
$$B(k, m_1) = \begin{bmatrix} \left(\prod_{n=0}^{m_1-1} A(l(k+1)-l2+m_1-n) \cdot B(l(k+1)-l2)\right)^{\mathrm{T}} \\ \left(\prod_{n=0}^{m_1-2} A(l(k+1)-l2+m_1-n) \cdot B(l(k+1)-l2+1)\right)^{\mathrm{T}} \\ \vdots \\ (B(l(k+1)-l2+m_1))^{\mathrm{T}} \\ O \\ \vdots \\ O \end{bmatrix}^{\mathrm{T}}$$

where $\bar{\boldsymbol{w}}(k)$ is $2^{J-1}u\times 1$ Gaussian white noise with covariance I, $A(k,m_1)$ is $N_x\times N_x$ matrix, $B(k,m_1)$ is $N_x\times 2^{J-1}u$ matrix, the elements in the last $(2^{J-1}-m_1-1)u$ columns are all zeros. Let

$$A(k)=\begin{bmatrix}A(k,0)\\A(k,1)\\\vdots\\A(k,2^{J-1}-1)\end{bmatrix},\ B(k)=\begin{bmatrix}B(k,0)\\B(k,1)\\\vdots\\B(k,2^{J-1}-1)\end{bmatrix},\ \bar{\boldsymbol{x}}(2,k+1)=\begin{bmatrix}\boldsymbol{x}_1(l(k+1)-l2+1)\\\boldsymbol{x}_1(l(k+1)-l2+2)\\\vdots\\\boldsymbol{x}_1(l(k+1))\end{bmatrix}$$

$A(k)$ is $l2\cdot N_x\times N_x$ matrix, $B(k)$ is $l2\cdot N_x\times 2^{J-1}u$ matrix, and $\bar{\boldsymbol{x}}(2,k+1)$ is $l2\cdot N_x\times 1$ matrix. Then $\bar{\boldsymbol{x}}(2,k+1)$ is non-overlapping parts of $\bar{\boldsymbol{x}}(k+1)$ and $\bar{\boldsymbol{x}}(k)$, which can be written as

$$\bar{\boldsymbol{x}}(2,k+1)=[O\quad A(k)]\bar{\boldsymbol{x}}(k)+B(k)\bar{\boldsymbol{w}}(k) \tag{14}$$

where O is a matrix with the indicated dimension whose elements are all zeros.

For simple expression, letting $f(k+1)=2^{J-1}(k+1)-(2^{J-1}-1)n$, the overlapping part of $\bar{\boldsymbol{x}}(k+1)$ and $\bar{\boldsymbol{x}}(k)$ is

$$\bar{\boldsymbol{x}}(1,k+1)=\begin{bmatrix}\boldsymbol{x}_1(f(k+1))\\\boldsymbol{x}_1(f(k+1)+1)\\\vdots\\\boldsymbol{x}_1(f(k+1)+l1-1)\end{bmatrix}$$

If $\bar{I}=diag\Big(\underbrace{I,I,\ldots,I}_{l1-1}\Big)$, then

$$\bar{\boldsymbol{x}}(1,k+1)=\begin{bmatrix}O&\bar{I}&O\\O&O&I\end{bmatrix}\bar{\boldsymbol{x}}(k)+[O_1]\bar{\boldsymbol{w}}(k)$$

where O_1 is $l1\cdot N_x\times 2^{J-1}u$ zero matrix. The state transition equation has the form

$$\bar{\boldsymbol{x}}(k+1)=\begin{bmatrix}\bar{\boldsymbol{x}}(1,k+1)\\\bar{\boldsymbol{x}}(2,k+1)\end{bmatrix}=\begin{bmatrix}O&\bar{I}&O\\O&O&I\\O&O&A(k)\end{bmatrix}\bar{\boldsymbol{x}}(k)+\begin{bmatrix}O_1\\B(k)\end{bmatrix}\bar{\boldsymbol{w}}(k) \tag{15}$$

Letting

$$\bar{A}(k)=\begin{bmatrix}O&\bar{I}&O\\O&O&I\\O&O&A(k)\end{bmatrix},\ \bar{B}(k)=\begin{bmatrix}O_1\\B(k)\end{bmatrix}$$

(15) can be written as

$$\bar{\boldsymbol{x}}(k+1)=\bar{A}(k)\bar{\boldsymbol{x}}(k)+\bar{B}(k)\bar{\boldsymbol{w}}(k) \tag{16}$$

For convenience, (13) and (16) are written together as follows.

$$\begin{cases}\bar{\boldsymbol{x}}(k+1)=\bar{A}(k)\bar{\boldsymbol{x}}(k)+\bar{B}(k)\bar{\boldsymbol{w}}(k)\\\bar{\boldsymbol{z}}(k)=\bar{C}(k)\bar{\boldsymbol{x}}(k)+\bar{\boldsymbol{v}}(k)\end{cases} \tag{17}$$

where $\bar{\boldsymbol{w}}(k)$ and $\bar{\boldsymbol{v}}(k)$ are uncorrelated Gaussian white noise. For the optimal filtering output at each scale, we have Theorem 1[23].

Theorem 1. If $\hat{\bar{\boldsymbol{x}}}(k)$ is the LMMSE of $\bar{\boldsymbol{x}}(k)$, then the LMMSE of node $\boldsymbol{x}_j(2^{J-j}k+m_j)$ is $M_j(m_j)\cdot\hat{\bar{\boldsymbol{x}}}(k)$.

4 Simulation

We take constant velocity motion as illustrative examples, whose system equation at the finest scale is

$$\boldsymbol{x}_1(k_1)=A\boldsymbol{x}_1(k_1)+B\boldsymbol{w}(k_1)$$

where

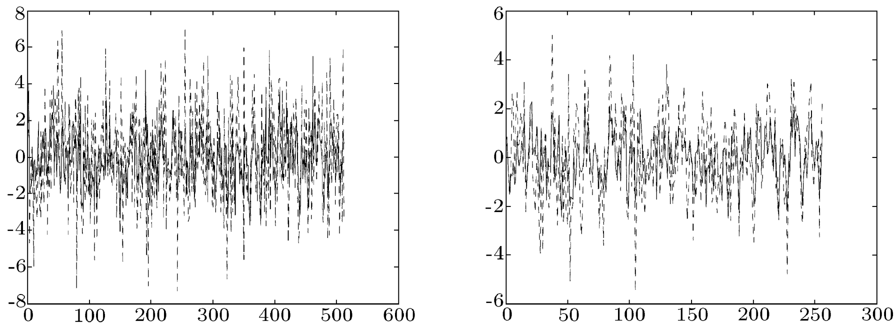
$$A=\begin{bmatrix}1&T\\0&1\end{bmatrix},\ B=\begin{bmatrix}\frac{1}{2}T^2&T\end{bmatrix}^{\mathrm{T}}$$

T is the sampling rate, $\boldsymbol{w}(\bullet)$ is Gaussian white noise with zero mean and its variance is $q(\bullet)$. Suppose the DMS has two scales, and the measurements are available at both scales, *i.e.*,

$$\begin{cases} \boldsymbol{z}_1(k_1) = \boldsymbol{C}_1\boldsymbol{x}_1(k_1) + \boldsymbol{v}_1(k_1) \\ \boldsymbol{z}_2(k_2) = \boldsymbol{C}_2\boldsymbol{x}_2(k_2) + \boldsymbol{v}_2(k_2) \end{cases}$$

where k_j denotes the sampling time at scale j, $\boldsymbol{C}_1 = \boldsymbol{C}_2 = [1 \quad 0]$, Gaussian white noise $\boldsymbol{v}_1(\bullet)$ and $\boldsymbol{v}_2(\bullet)$ are with zero-mean and variances $r_1(\bullet)$ and $r_2(\bullet)$, respectively. They are uncorrelated with $\boldsymbol{w}(\bullet)$. The first component of $\boldsymbol{x}_1$ is displacement, and the second is velocity.

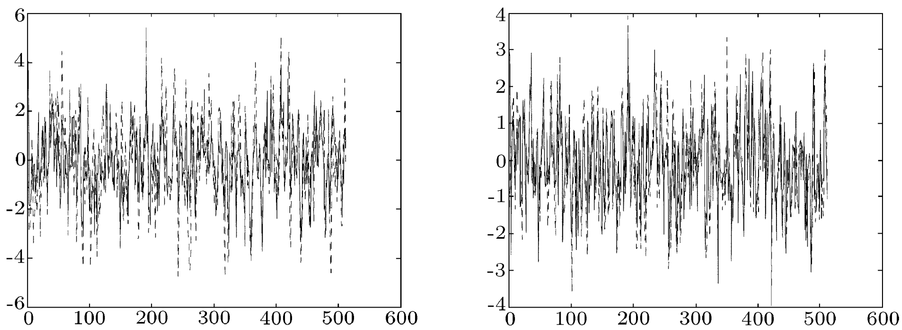
Here CDF(2,2) wavelet is used, and let $T = 1$, $q = 1$, $r_1 = 6.25$, $r_2 = 3.24$. Figs. 2(a) and 2(b) compare the measurement noise with the estimation error of displacement at scale 1 and scale 2. Figs. 3(a) and 3(b) compare the estimation error of displacement and velocity at scale 1 obtained by performing Kalman filter directly and by the algorithm in this paper. The noise compression ratio of the algorithm in this paper is 2.4102 db higher than that by performing Kalman filter directly.



(a) Scale 1, the noise compression ratio is 5.4666 db

(b) Scale 2, the noise compression ratio is 5.0171 db

Fig. 2 Measurement noise (dotted) and estimation error (solid) of displacement



(a) Displacement

(b) Velocity

Fig. 3 Estimation error at scale 1 obtained by performing Kalman filter directly (dotted) and by the algorithm in this paper (solid)

Using our model and several groups of different parameters, the algorithm in this paper is compared with that of L. Hong[16] (Table 1). It can be seen that the estimation accuracy of our algorithm is better than that of L. Hong at each scale.

Table 1 Noise compression ratios of L. Hong and this paper (db)

parameters				scale 1		scale 2	
T	q	r_1	r_2	L. Hong	this paper	L. Hong	this paper
1	1	6.25	3.24	2.6874	5.8005	0.4620	5.1124
1	2	6.25	3.24	2.0948	5.3227	0.3359	4.9952
1	4	6.25	3.24	1.5579	4.7975	0.2250	4.8729
1	1	4	2.25	2.2946	4.3560	0.4088	5.0048
1	1	9	6.25	2.9739	4.7624	0.6964	5.7632
1	1	16	9	3.5452	5.4000	0.7118	5.2610

5 Conclusion

The modeling and optimal estimation of a class of DMS that is observed independently by several sensors at different scales are proposed. Using general compactly supported wavelet transform to approximate the state space projection between scales, we generalize the DMS into the standard state space model. Then the Kalman filtering is employed as the LMMSE algorithm. An example is presented to illustrate the proposed scheme and its relationship with the traditional Kalman filtering and the multiresolutional filtering algorithm of L. Hong.

References

1 Basseville M, Benveniste A, Chou K, Golden S, Nikoukhah R, Willsky A S. Modeling and estimation of multiresolution stochastic processes. *IEEE Transactions on Information Theory*, 1992, **38**(2): 766∼784

2 Chou K, Willsky A S, Benveniste A. Multiscale recursive estimation, data fusion, and regularization. *IEEE Transactions on Automatic Control*, 1994, **39**(3): 464∼478

3 Tsai A, Zhang J, Willsky A S. Expectation-maximization algorithms for image processing using multiscale models and mean-field theory, with applications to laser radar range profiling and segmentation. *Optical engineering*, 2001, **40**(7): 1287∼1301

4 Schneider M K, Fieguth P W, Karl W C, Willsky A S. Multiscale Methods for the Segmentation and Reconstruction of Signals and Images. *IEEE Transactions on Image Processing*, 2000, **9**(3): 456∼468

5 Daniel M, Willsky A S. The modeling and estimation of statistically self-similar processes in a multiresolution framework. *IEEE Transactions on Information Theory*, 1999, **45**(3): 955∼970

6 Frakt A B, Willsky A S. Efficient multiscale stochastic realization. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Seattle, Washington, 1998. **4**: 2249∼2252

7 Fieguth P, Karl W, Willsky A S. Efficient multiresolution counterparts to variational methods for surface reconstruction. *Computer Vision and Image Understanding*, 1998, **70**(2): 157∼176

8 Fosgate C, Krim H, Irving W W, Willsky A S. Multiscale segmentation and anomaly enhancement of SAR imagery. *IEEE Transactions on Image Processing*, 1997, **6**(1): 7∼20

9 Irving W W, Fieguth P, Willsky A S. An overlapping tree approach to multiscale stochastic modeling and estimation. *IEEE Transactions on Image Processing*, 1997: **6**(11): 1517∼1529

10 Fieguth P, Willsky A S. Fractal estimation using models on multiscale trees. *IEEE Transactions on Signal Processing*, 1996, **44**(5): 1297∼1300

11 Luettgen M, Willsky A S. Likelihood calculation for a class of multiscale stochastic models with application to texture discrimination. *IEEE Transactions on Image Processing*, 1995, **4**(2): 194∼207

12 Luettgen M, Karl W, Willsky A S. Efficient multiscale regularization with applications to the computation of optical flow. *IEEE Transactions on Image Processing*, 1994, **3**(1): 41∼64

13 Luettgen M, Karl W, Willsky A S, Tenney R. Multiscale representations of Markov random fields. *IEEE Transactions on Signal Processing*, 1993, **41**(12): 3377∼3396

14 Ho T. Large-scale multiscale estimation of dynamic systems. [Ph. D thesis]. Massachusetts: Massachusetts Institute of Technology, 1998

15 Luettgen M, Willsky A S. Multiscale smoothing error models. *IEEE Transactions on Automatic Control*, 1995, **40**(1): 173∼175

16 Hong L. Multiresolution distributed filtering. *IEEE Transactions on Automatic Control*, 1994, **39**(4): 853∼856

17 Hong L, Cheng G, Chui C K. A filter-bank-based Kalman filtering technique for wavelet estimation and decomposition of random signals. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 1998, **45**(2): 237∼241

18 Hong L. Multiresolution target tracking using the wavelet transform. In: Proceedings of 32nd IEEE Conference on Decision and Control, San Antonio, Texas, 1993. **1**: 924∼929

19 Hong L, Scaggs T. Real-time optimal filtering for stochastic systems with multiresolutional measurements. *Systems & Control Letters*, 1993, **20**: 381∼387

20 Hong L. Multiresolution filtering using wavelet transform. *IEEE Transactions on Aerospace and Electronic System*, 1993, **29**(4): 1244~1251
21 Daubechies I. Ten Lectures on Wavelets. Philadephia, Pennsylvania: Society for Industrial and Applied Mathematics, 1992
22 Cohen, Daubechies I. Bi-orthogonal bases of compactly supported wavelets. *Communications Pure and Applied Mathematics*, 1992, **45**: 485~560
23 Zhang Lei. The optimal estimation of a class of dynamic multiscale systems. [Ph. D. thesis] Xi′an: Northwestern Polytechnical University, 2001

PAN Quan Ph. D. He is now a professor of Northwestern Polytechnical University, His research interest include information fusion theories and applications, modeling and simulation of C^3I systems, modeling of dynamic system, estimation and control, modeling and estimation of hybrid system and dynamic multiscale system, intelligent information process, multitarget tracking, and wavelet theories and application in image processing.

CUI Pei-Ling Postdoctor of Tsinghua University. Her research interests include dynamic multiscale system estimation, multisensor information fusion.

ZHANG Lei Received his bachelor degree from Shenyang Institute of Aeronautic Engineering, Shenyang, China, in 1995 and the master and Ph. D. degrees from Northwestern Polytechnical University, Xi′an, China, in 1998 and 2001, respectively. His research interests include optimal estimation theory, information fusion, and wavelet transform.

ZHANG Hong-Cai Received his bachelor and master degrees from Northwestern Polytechnical University, Xi′an, China, in 1961 and 1964, respectively. His research interests include information fusion, target tracking, multiscale system theory, and signal processing.