Multiresolution Modeling and Estimation of Multisensor Data

Lei Zhang, Xiaolin Wu, *Senior Member, IEEE*, Quan Pan, and Hongcai Zhang

Abstract—This paper presents a multiresolution multisensor data fusion scheme for dynamic systems to be observed by several sensors of different resolutions. A state projection equation is introduced to associate the states of a system at each resolution with others. This projection equation together with the state transition equation and the measurement equations at each of the resolutions construct a continuous-time model of the system. The model meets the requirements of Kalman filtering. In discrete time, the state transition is described at the finest resolution and the sampling frequencies of sensors decrease successively by a factor of two in resolution. We can build a discrete model of the system by using a linear projection operator to approximate the state space projection. This discrete model satisfies the requirements of discrete Kalman filtering, which actually offers an optimal estimation algorithm of the system. In time-invariant case, the stability of the Kalman filter is analyzed and a sufficient condition for the filtering stability is given. A Markov-process-based example is given to illustrate and evaluate the proposed scheme of multiresolution modeling and estimation with multiple sensors.

Index Terms—Kalman filtering, multiresolution analysis, multisensor fusion, optimal estimation.

I. INTRODUCTION

IN many applications, such as system filtering and target tracking, it is often beneficial to employ more than one sensor to acquire sufficient information about an interested object or system [1], [2], [8], [9]. An efficient data fusion algorithm is then necessary to process the obtained measurements aiming at an optimal, or nearly optimal, estimation of the system. In this paper, we consider a class of dynamic systems that are observed by multiple sensors of different resolutions. Suppose the state transition of the system is characterized by

$$\dot{\boldsymbol{x}}(t) = F(t)\boldsymbol{x}(t) + \Gamma(t)\boldsymbol{w}(t) \tag{1.1}$$

where $\boldsymbol{x}(t) \in \Re^{n_x \times 1}$ is the state variable vector to be estimated. Matrices $F(t)$ and $\Gamma(t)$ are system and input matrices, respectively. System noise $\boldsymbol{w}(t)$ is assumed to be Gaussian white process with zero mean and variance $Q(t)$.
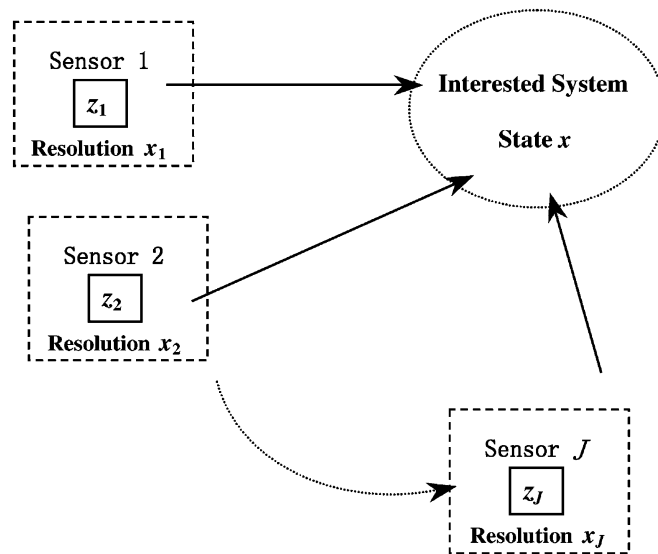
Fig. 1. System is observed by J sensors with different resolutions.

Referring to Fig. 1, the system state vector $\boldsymbol{x}(t)$ is observed by J sensors, and each sensor j has its independent observation $\boldsymbol{z}_j(t)$

$$\boldsymbol{z}_j(t) = H_j(t)\boldsymbol{x}_j(t) + \boldsymbol{v}_j(t), \quad j = 1, 2, \ldots J \tag{1.2}$$

where $H_j(t)$ is the measurement matrix and state $\boldsymbol{x}_j(t)$ belongs to a subspace of $\boldsymbol{x}(t)$. This subspace, in which observation $\boldsymbol{z}_j(t)$ is made, depends on the resolution of sensor j. The higher the resolution is, the more accurately $\boldsymbol{x}_j(t)$ approximates $\boldsymbol{x}(t)$. Observation noises $\boldsymbol{v}_j(t), j = 1, 2, \ldots J$ are independent of each other. They are Gaussian white processes with zero mean and variances $R_j(t)$. The usual principle dealing with the mentioned multisensor dynamic system is to insert the measurements of coarser resolution sensors into the measurements of the finest sensor by time adjustment and then implement Kalman filtering. Such processing does not exploit the multiresolution structure of the multisensor. In this paper, we will introduce the state projection equation to link the multiresolution states and construct a model to estimate the system optimally in sense of minimum mean-square error.

Indexing the sensors by their resolution from 1 to J and supposing that sensor 1 has the highest resolution, we approximate the state transition equation (1.1) by

$$\dot{\boldsymbol{x}}_1(t) = F(t)\boldsymbol{x}_1(t) + \Gamma(t)\boldsymbol{w}(t). \tag{1.3}$$
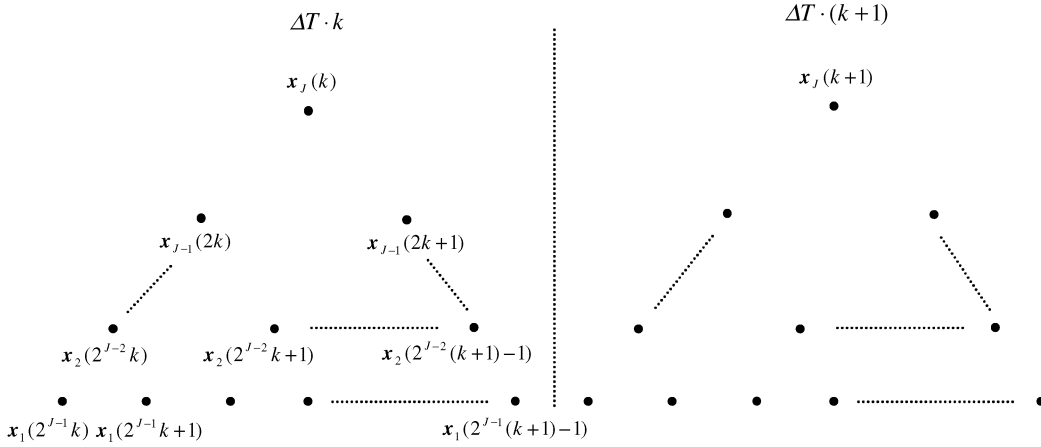
Fig. 2. Structure of the multiresolution multisensor system state nodes in discrete time.

Obviously, when $J = 1$, (1.3) and (1.2) reduce to the classical state-space model of an ordinary dynamic system, for which the Kalman filtering [16], [18]–[20] is the optimal linear minimum mean-square-error estimation (LMMSE) algorithm. In real applications, the system should be discrete. Denote the discrete time versions of (1.3) and (1.2) by

$$\boldsymbol{x}_1(k_1+1) = A(k_1)\boldsymbol{x}_1(k_1) + B(k_1)\boldsymbol{w}(k_1) \tag{1.4}$$

$$\boldsymbol{z}_j(k_j) = C_j(k_j)\boldsymbol{x}_j(k_j) + \boldsymbol{v}_j(k_j), \quad j = 1, 2, \ldots J \tag{1.5}$$

where k_j is the sampling time of sensor j. $A(k_1)$ and $B(k_1)$ are system and input matrices. $C_j(k_j)$ is the measurement matrix of sensor j. $\boldsymbol{w}(k_1)$ and $\boldsymbol{v}_j(k_j)$ are independent Gaussian white processes with zero mean and variances $Q(k_1)$ and $R_j(k_j)$.

Suppose that the sampling frequencies of the sensors decrease successively by a factor of two from 1 to J. Fig. 2 illustrates the structure of the system state nodes. In each time block $\Delta T \cdot k$, there are 2^{J-1} state nodes for sensor 1, 2^{J-2} nodes for sensor $2, \ldots$, and so forth, to only one node for sensor J. The goal is to get the real-time optimal estimation of the state nodes based on the measurements of all sensors. Hong [8] discussed the filtering of such a system in the application of target tracking. He used wavelet transform [11]–[15] to link the state nodes of different sensors. The state nodes of sensor j are derived by filtering the nodes of sensor $j - 1$ with the lowpass wavelet filter $F_0 = \{f_0(n)\}$ and then subsampling by two

$$\boldsymbol{x}_j(n) = \sum_k f_0(2n - k)\boldsymbol{x}_{j-1}(k). \tag{1.6}$$

Some details are lost from $\boldsymbol{x}_{j-1}(n)$ due to lowpass filtering, which could be computed by filtering $\boldsymbol{x}_{j-1}(n)$ with the highpass wavelet filter $F_1 = \{f_1(n)\}$ and then subsampling by two

$$\boldsymbol{y}_j(n) = \sum_k f_1(2n - k)\boldsymbol{x}_{j-1}(k) \tag{1.7}$$

where $\boldsymbol{y}_j(n)$ is called the detail wavelet coefficient.

In each time block $\Delta T \cdot k$, Hong [8] first estimated $\boldsymbol{x}_1(n)$ by the measurements of sensor 1, and then, he wavelet transformed

the estimate to scales $2 \sim J$ as the prediction of $\boldsymbol{x}_j(n)$. The updates of the prediction were conducted on each sensor with the local measurements. At last, the locally updated estimates were inversely transformed to sensor 1 and fused together. In Hong's algorithm, the updating is only performed on the prediction of $\boldsymbol{x}_j(n)$ but not on the detail wavelet coefficient $\boldsymbol{y}_j(n)$. In fact, $\boldsymbol{y}_j(n)$ would contribute to the estimation of $\boldsymbol{x}_1(n)$ because it is correlated with $\boldsymbol{x}_j(n)$ and the observation $\boldsymbol{z}_j(n)$. Therefore, $\boldsymbol{y}_j(n)$ should be updated too. The estimation by Hong's scheme is not optimal. In [9] Zhang *et al.* proposed an optimal estimation scheme of the system by using Haar wavelet transform to link the state nodes of each sensor. The advantage of using Haar wavelet is that the state nodes of sensor j can be represented by the state nodes of the finest sensor within a time block $\Delta T \cdot k$, so a more compact modeling of the system becomes possible. But this restricts the exploitation of the dependencies of node $\boldsymbol{x}_j(n)$ with other nodes outside time block $\Delta T \cdot k$.

In this paper, we present a general modeling scheme of multisensor data. First, we model multisensor data in continuous time by introducing a projection equation that relates multiresolution states to one another. A corresponding discrete model is realized by linking the state nodes of each sensor via a linear projection. It is shown that the model meets the requirements of discrete Kalman filtering. Consequently, the real-time optimal estimation of the system can be carried out by Kalman filtering.

The next section is devoted to the modeling of the multisensor data in continuous time. Section III presents the discrete multiresolution modeling of the system. Since the Kalman filtering is an optimal estimation algorithm of the realized discrete model, the stability of the Kalman filter in time invariant case is discussed in Section IV. A sufficient condition for stable Kalman filtering is presented. In Section V, an example is presented to illustrate the proposed modeling and estimation scheme. Section VI concludes the paper.

II. MODELING OF THE MULTISENSOR DATA IN CONTINUOUS TIME

As shown in Fig. 1, denote by U_0 the space expanded by state variable $\boldsymbol{x}(t)$, and then, state variable $\boldsymbol{x}_1(t)$ of the finest

resolution belongs to space U_1, which is a subspace of U_0. Similarly, $\boldsymbol{x}_j(t) \in U_j$, and we have the following subspace sequence:

$$U_J \subset U_{J-1} \subset \cdots \subset U_1 \subset U_0 \qquad (2.1)$$

in which $\boldsymbol{x}_1(t)$ can be considered as the projection of $\boldsymbol{x}(t)$ from space U_0 to U_1:

$$\boldsymbol{x}_1(t) = P(0,1) \cdot \boldsymbol{x}(t) \qquad (2.2)$$

where $P(0,1)$ is the projection operator. Similarly, $\boldsymbol{x}_2(t)$ can be considered as the projection of $\boldsymbol{x}_1(t)$ from space U_1 to U_2 with projection operator $P(1,2)$: $\boldsymbol{x}_2(t) = P(1,2) \cdot \boldsymbol{x}_1(t)$. Denote $P(j-1, j)$ as the projection operator from space U_{j-1} to U_j; there is

$$\boldsymbol{x}_j(t) = P(j-1, j) \cdot \boldsymbol{x}_{j-1}(t). \qquad (2.3)$$

Then, we have

$$\boldsymbol{x}_j(t) = P_j \cdot \boldsymbol{x}(t) \qquad (2.4)$$

where

$$P_j = P(j-1, j) \cdot P(j-2, j-1) \cdot \ldots \cdot P(0,1). \qquad (2.5)$$

The observation (1.2) can be rewritten as

$$z_j(t) = H_j(t) P_j \cdot \boldsymbol{x}(t) + \boldsymbol{v}_j(t), \quad j = 1, 2, \ldots J. \qquad (2.6)$$

Denote

$$\boldsymbol{z}(t) = \begin{bmatrix} \boldsymbol{z}_1(t) \\ \boldsymbol{z}_2(t) \\ \vdots \\ \boldsymbol{z}_J(t) \end{bmatrix}, \quad H(t) = \begin{bmatrix} H_1(t) P_1 \\ H_2(t) P_2 \\ \vdots \\ H_J(t) P_J \end{bmatrix}$$

$$\boldsymbol{v}(t) = \begin{bmatrix} \boldsymbol{v}_1(t) \\ \boldsymbol{v}_2(t) \\ \vdots \\ \boldsymbol{v}_J(t) \end{bmatrix}. \qquad (2.7)$$

We have

$$\boldsymbol{z}(t) = H(t) \cdot \boldsymbol{x}(t) + \boldsymbol{v}(t). \qquad (2.8)$$

The covariance of $\boldsymbol{v}(t)$ is

$$R(t) = \mathrm{diag}[R_1(t) \quad R_2(t) \quad \cdots \quad R_J(t)]. \qquad (2.9)$$

Together with the state transition equation (1.1)

$$\dot{\boldsymbol{x}}(t) = F(t)\boldsymbol{x}(t) + \Gamma(t)\boldsymbol{w}(t). \qquad (2.10)$$

Equations (2.10) and (2.8) form a new state-space model of the multiresolution multisensor system. Obviously, the model meets the requirements of Kalman filtering, and then, the system state
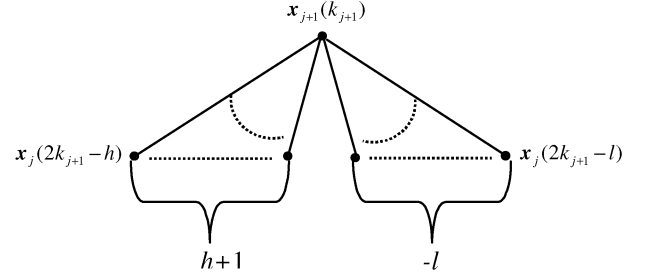

Fig. 3. Link of a coarser resolution state node with the finer resolution nodes by the N-tap lowpass filter, where $h - l + 1 = N$.

$\boldsymbol{x}(t)$ can be estimated by the standard continuous-time Kalman filtering algorithm.

III. DISCRETE MODEL OF THE MULTISENSOR DATA

The continuous-time model of the multiresolution multisensor systems has been presented in the previous section. In practice, however, it is more convenient to use discrete-time models for computation purposes. As mentioned in Section I, the sampling frequency of sensors decreases by a factor of two in each coarser resolution of sensors (refer to Fig. 2). The projection operator P_j relates a state node $\boldsymbol{x}_j(k_j)$ to the nodes $\boldsymbol{x}_1(k_1)$ of the sensor of the finest resolution. In this section, we will use a linear projection operator to represent the relation analytically and then realize the discrete model of the system. In [9], Zhang et al. has employed the discrete Haar wavelet transform to link the state nodes. The filter of Haar wavelet has only two taps. By the Haar wavelet transform, a node $\boldsymbol{x}_j(k_j)$ can be written as the linear combination of the nodes of sensor 1 within the time block $\Delta T \cdot k$. It exactly forms a dyadic tree. Although the dyadic tree is a convenient structure for data processing, the Haar wavelet filter has only one order vanish moment (two taps), and the Haar wavelet transform may not approximate the linear projection P_j sufficiently well. In many situations, more accurate approximation of P_j can be obtained with a longer filter.

Denote by D_0 a lowpass filter with N taps

$$D_0 = \{d_l, d_{l+1}, \ldots, d_0, \ldots, d_h\} \qquad (3.1)$$

where $h - l + 1 = N$, and we have (3.2), shown at the bottom of the page. We use filter D_0 to link the state nodes of adjacent sensors.

A. Formalization of the Measurement Equation

Fig. 3 shows the relation of a node of sensor $j + 1$ to those nodes of sensor j. Note that $\boldsymbol{x}_{j+1}(k_{j+1})$ is the output of $\boldsymbol{x}_j(k_j)$ passing through D_0. It can be represented by the linear combination of N nodes: $\boldsymbol{x}_j(2k_{j+1} - h), \boldsymbol{x}_j(2k_{j+1} - h + 1), \ldots, \boldsymbol{x}_j(2k_{j+1} - l)$. When $N > 2$, the linking structure of

$$l = \begin{cases} -N/2, & N \text{ is even} \\ -(N-1)/2, & N \text{ is odd} \end{cases} \quad h = \begin{cases} N/2 - 1, & N \text{ is even} \\ (N-1)/2, & N \text{ is odd}. \end{cases} \qquad (3.2)$$
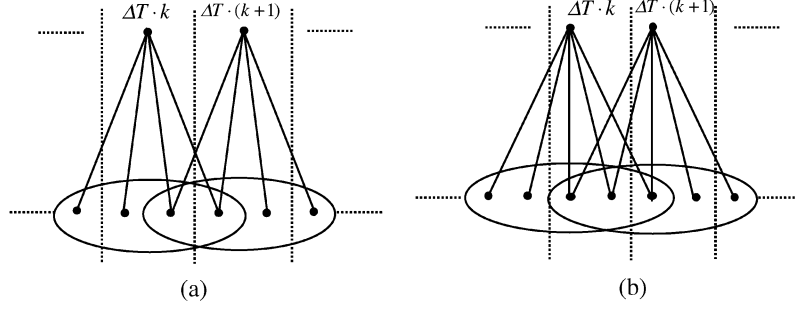
Fig. 4. (a) State nodes linking structure with $J = 2$ and $N = 4$. (b) State nodes linking structure with $J = 2$ and $N = 5$.
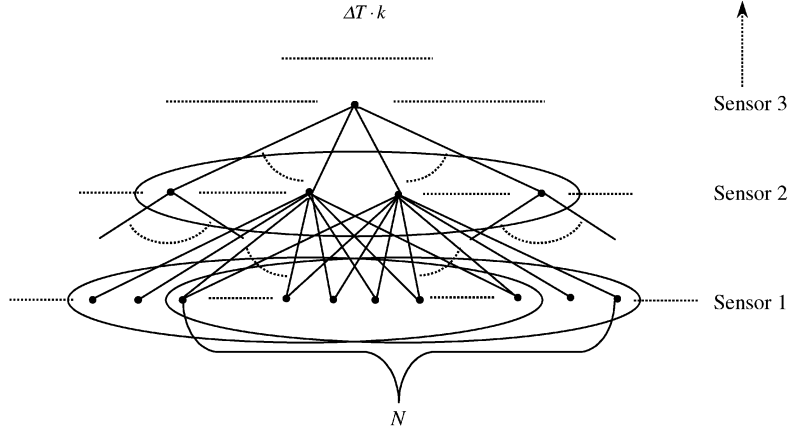
Fig. 5. State node-linking structure of the multiresolution multisensor system by the N-tap lowpass filter.

the state nodes will no longer be a dyadic tree. Supposing that $J = 2$, Fig. 4(a) and (b) show the system nodes structures when $N = 4$ and $N = 5$, respectively. For example, when $N = 4$, to represent the coarser sensor node in time $\Delta T \cdot k$, not only the two nodes in $\Delta T \cdot k$ of the finer sensor but also one node in time $\Delta T \cdot (k - 1)$ and one node in time $\Delta T \cdot (k + 1)$ should be employed. Fig. 5 illustrates the general state node-linking structure of the system by the N-tap lowpass filter D_0.

For compact denotation, we let m_j be an integer variable that takes on values from the set $\{0, 1, \ldots, 2^{J-j} - 1\}, j = 1, 2, \ldots, J$. In time block $\Delta T \cdot k$, a node $\boldsymbol{x}_2(2^{J-2}k + m_2)$ of sensor 2 can be represented as a linear combination of the nodes of sensor 1:

$$\boldsymbol{x}_2(2^{J-2}k + m_2) = (\boldsymbol{x}_1(\bullet) * D_0) \downarrow_2$$
$$= \sum_{i=l}^{h} d_i \cdot \boldsymbol{x}_1(2^{J-1}k + 2m_2 - i) \quad (3.3)$$

where $*$ denotes convolution operation, and $\downarrow_p$ denotes subsampling by factor p. Denote by (3.4), shown at the bottom of the page, the dilation of D_0 by inserting 2^{i-1} zeros between each of the coefficients of D_0. Referring to Fig. 5, for sensor 3, its node

$\boldsymbol{x}_3(2^{J-3}k + m_3)$ can be represented by convoluting $\boldsymbol{x}_1(\bullet)$ with D^3, which is the convolution of D_1 with D_0 and then subsampled by 4. Similarly, for sensor j, its node $\boldsymbol{x}_j(\bullet)$ is represented as follows:

$$\boldsymbol{x}_j(\bullet) = (\boldsymbol{x}_1(\bullet) * D^j) \downarrow_{2^{j-1}} \quad (3.5)$$

where the filter D^j is obtained by

$$D^j = D_0 * D_1 * \ldots * D_{j-2}, \quad j = 2, 3, \ldots, J. \quad (3.6)$$

We let $D^1 = \{1\}$. The total coefficients number N_j of filter D^j is

$$N_j = (2^{j-1} - 1)(N - 1) + 1, \quad j = 1, 2, \ldots, J. \quad (3.7)$$

Denote

$$D^j = \left\{ d_{l_j}^j, d_{l_j+1}^j, \ldots, d_{h_j}^j \right\}, \quad j = 2, 3, \ldots, J. \quad (3.8)$$

where

$$\begin{cases} l_j = (2^{j-1} - 1)l \\ h_j = (2^{j-1} - 1)h \end{cases} . \quad (3.9)$$

$$D_i = \left\{ d_l, \underbrace{0, \ldots, 0}_{2^{i-1}}, d_{l+1}, \underbrace{0, \ldots, 0}_{2^{i-1}}, d_{l+2}, \ldots, d_{h-1}, \underbrace{0, \ldots, 0}_{2^{i-1}}, d_h \right\} \quad (3.4)$$

For the coarsest sensor J, the root state node $\boldsymbol{x}_J(k)$ can be represented as

$$\boldsymbol{x}_J(k) = (\boldsymbol{x}_1(\bullet) * D^J) \downarrow_{2^{J-1}} = \sum_{i=l_J}^{h_J} d_i^J \cdot \boldsymbol{x}_1(2^{J-1}k - i). \quad (3.10)$$

It needs N_J nodes of sensor 1 in total. Obviously, N_J is the maximum number of the finest state nodes that can be utilized to represent the other nodes of coarser sensors in a time block $\Delta T \cdot k$. Define (3.11) and (3.12), shown at the bottom of the page, where col means column vector, $I_{n_{\boldsymbol{x}}}$ is the $n_{\boldsymbol{x}} \times n_{\boldsymbol{x}}$ identity matrix, and

$$\begin{cases} Z_{j,1}(m_j) = 2^{j-1} m_j + h \cdot (2^{J-1} - 2^{j-1}) \\ Z_{j,2}(m_j) = l \cdot (2^{j-1} - 2^{J-1}) - 2^{j-1} m_j \end{cases}. \quad (3.13)$$

Then, the node $\boldsymbol{x}_j(2^{J-j}k + m_j)$ and its measurement $\boldsymbol{z}_j(2^{J-j}k + m_j)$ can be written as

$$\boldsymbol{x}_j(2^{J-j}k + m_j) = M_j(m_j) \cdot \bar{\boldsymbol{x}}(k) \quad (3.14)$$

$$\boldsymbol{z}_j(2^{J-j}k + m_j) = C_j(2^{J-j}k + m_j) M_j(m_j) \bar{\boldsymbol{x}}(k) \\ + \boldsymbol{v}_j(2^{J-j}k + m_j). \quad (3.15)$$

By defining

$$\bar{\boldsymbol{z}}_j(k) = \text{col}\{\boldsymbol{z}_j(2^{J-j}k), \boldsymbol{z}_j(2^{J-j}k + 1), \ldots \\ \boldsymbol{z}_j(2^{J-j}(k+1) - 1)\} \quad (3.16)$$

$$\bar{C}_j(k) = \begin{bmatrix} C_j(2^{J-j}k) M_j(0) \\ C_j(2^{J-j}k + 1) M_j(1) \\ \vdots \\ C_j(2^{J-j}(k+1) - 1) M_j(2^{J-j} - 1) \end{bmatrix} \quad (3.17)$$

$$\bar{\boldsymbol{v}}_j(k) = \text{col}\{\boldsymbol{v}_j(2^{J-j}k), \boldsymbol{v}_j(2^{J-j}k + 1), \ldots \\ \boldsymbol{v}_j(2^{J-j}(k+1) - 1)\} \quad (3.18)$$

we have

$$\bar{\boldsymbol{z}}_j(k) = \bar{C}_j(k) \bar{\boldsymbol{x}}(k) + \bar{\boldsymbol{v}}_j(k) \quad (3.19)$$

and finally

$$\bar{\boldsymbol{z}}(k) = \bar{C}(k) \bar{\boldsymbol{x}}(k) + \bar{\boldsymbol{v}}(k) \quad (3.20)$$

where

$$\bar{\boldsymbol{z}}(k) = \text{col}\{\bar{\boldsymbol{z}}_J(k), \bar{\boldsymbol{z}}_{J-1}(k), \ldots, \bar{\boldsymbol{z}}_1(k)\} \quad (3.21)$$

$$\bar{C}(k) = \text{col}\{\bar{C}_J(k), \bar{C}_{J-1}(k), \ldots, \bar{C}_1(k)\} \quad (3.22)$$

$$\bar{\boldsymbol{v}}(k) = \text{col}\{\bar{\boldsymbol{v}}_J(k), \bar{\boldsymbol{v}}_{J-1}(k), \ldots, \bar{\boldsymbol{v}}_1(k)\}. \quad (3.23)$$

Equation (3.20) is the state augmented measurement equation of the system. In the next subsection, we develop the corresponding state transition equation to complete the state-space model.

B. Formalization of the State Transition Equation

For $N = 2$, the linking structure of state nodes forms a dyadic tree within a time block. The states vector $\bar{\boldsymbol{x}}(k)$ can be transited from one time block to the next block without overlap, but if $N > 2$, the system state nodes structure is no longer a dyadic tree. There are some overlapped nodes between the augmented state vectors $\bar{\boldsymbol{x}}(k)$ and $\bar{\boldsymbol{x}}(k+1)$. This means that when the measurement $\bar{\boldsymbol{z}}(k+1)$ is obtained, the estimation of the overlapped nodes in $\bar{\boldsymbol{x}}(k)$ should be further updated.

The last element of $\bar{\boldsymbol{x}}(k)$ is $\boldsymbol{x}_1(2^{J-1}k - l_J)$, and the first element of $\bar{\boldsymbol{x}}(k+1)$ is $\boldsymbol{x}_1(2^{J-1}(k+1) - h_J)$. The number of their overlapped nodes is $n_p = (2^{J-1} - 1)(N - 2)$, and the number of nonoverlapped nodes is $n_{\bar{p}} = 2^{J-1}$. For convenience, we let

$$n_k = 2^{J-1}k - l_J. \quad (3.24)$$

Then, the last element of $\bar{\boldsymbol{x}}(k)$ is $\boldsymbol{x}_1(n_k)$, or $\boldsymbol{x}_1(n_{k+1} - n_{\bar{p}})$ in another form. Denote the overlapped and nonoverlapped parts in $\bar{\boldsymbol{x}}(k+1)$ by

$$\bar{\boldsymbol{x}}_p(k+1) \\ = \text{col}\{\boldsymbol{x}_1(2^{J-1}(k+1) - h_J), \ldots, \boldsymbol{x}_1(n_{k+1} - n_{\bar{p}})\} \quad (3.25)$$

$$\bar{\boldsymbol{x}}_{\bar{p}}(k+1) \\ = \text{col}\{\boldsymbol{x}_1(n_{k+1} - n_{\bar{p}} + 1), \ldots, \boldsymbol{x}_1(n_{k+1})\}. \quad (3.26)$$

For the nonoverlapped nodes in $\bar{\boldsymbol{x}}(k+1)$, we have

$$\boldsymbol{x}_1(n_{k+1} - n_{\bar{p}} + 1 + m_1) \\ = \prod_{n=0}^{m_1} A(n_{k+1} - n_{\bar{p}} + m_1 - n) \bullet \boldsymbol{x}_1(n_{k+1} - n_{\bar{p}}) \\ + \sum_{i=0}^{m_1} \prod_{n=0}^{m_1 - i - 1} A(n_{k+1} - n_{\bar{p}} + m_1 - n) \\ \bullet B(n_{k+1} - n_{\bar{p}} + i) \boldsymbol{w}(n_{k+1} - n_{\bar{p}} + i). \quad (3.27)$$

Define (3.28)–(3.31), shown at the bottom of the next page, where O is the zero matrix. Then, we have

$$\bar{\boldsymbol{x}}_{\bar{p}}(k+1) = [O \quad \bar{A}_{\bar{p}}(k)] \cdot \bar{\boldsymbol{x}}(k) + \bar{B}_{\bar{p}}(k) \cdot \bar{\boldsymbol{w}}(k). \quad (3.32)$$

$$\bar{\boldsymbol{x}}(k) = \text{col}\{\boldsymbol{x}_1(2^{J-1}k - h_J), \boldsymbol{x}_1(2^{J-1}k - h_J + 1), \ldots, \boldsymbol{x}_1(2^{J-1}k - l_J)\} \quad (3.11)$$

$$M_j(m_j) = \begin{bmatrix} \underbrace{0 \cdot I_{n_{\boldsymbol{x}}}, \ldots, 0 \cdot I_{n_{\boldsymbol{x}}}}_{Z_{j,1}(m_j)} & \underbrace{d_{h_j}^j \cdot I_{n_{\boldsymbol{x}}}, \ldots, d_{l_j}^j \cdot I_{n_{\boldsymbol{x}}}}_{N_j} & \underbrace{0 \cdot I_{n_{\boldsymbol{x}}}, \ldots, 0 \cdot I_{n_{\boldsymbol{x}}}}_{Z_{j,2}(m_j)} \end{bmatrix} \quad (3.12)$$

For the overlapped part $\bar{\boldsymbol{x}}_p(k+1)$, we have

$$\bar{\boldsymbol{x}}_p(k+1) = \begin{bmatrix} O & \bar{I} & O \\ O & O & I_{n_{\boldsymbol{x}}} \end{bmatrix} \cdot \bar{\boldsymbol{x}}(k) + [O] \cdot \bar{\boldsymbol{w}}(k) \qquad (3.33)$$

where $\bar{I} = \text{diag}[\underbrace{I_{n_{\boldsymbol{x}}} \cdots I_{n_{\boldsymbol{x}}}}_{n_p - 1}]$. Let

$$\bar{A}(k) = \begin{bmatrix} O & \bar{I} & O \\ O & O & I_{n_{\boldsymbol{x}}} \\ O & O & \bar{A}_{\bar{p}}(k) \end{bmatrix}, \quad \bar{B}(k) = \begin{bmatrix} O \\ \bar{B}_{\bar{p}}(k) \end{bmatrix}. \quad (3.34)$$

By combining $\bar{\boldsymbol{x}}_p(k+1)$ and $\bar{\boldsymbol{x}}_{\bar{p}}(k+1)$, we have the system state transition equation

$$\bar{\boldsymbol{x}}(k+1) = \begin{bmatrix} \bar{\boldsymbol{x}}_p(k+1) \\ \bar{\boldsymbol{x}}_{\bar{p}}(k+1) \end{bmatrix} = \bar{A}(k) \cdot \bar{\boldsymbol{x}}(k) + \bar{B}(k) \cdot \bar{\boldsymbol{w}}(k). \quad (3.35)$$

Equations (3.35) and (3.20) form the discrete state space model of the multisensor system

$$\begin{cases} \bar{\boldsymbol{x}}(k+1) = \bar{A}(k)\bar{\boldsymbol{x}}(k) + \bar{B}(k)\bar{\boldsymbol{w}}(k) \\ \bar{\boldsymbol{z}}(k) = \bar{C}(k)\bar{\boldsymbol{x}}(k) + \bar{\boldsymbol{v}}(k) \end{cases} \quad (3.36)$$

where $\bar{\boldsymbol{w}}(k)$ and $\bar{\boldsymbol{v}}(k)$ are independent Gaussian white processes. The model meets the requirement of standard discrete Kalman filtering [18]–[20] with which the optimal LMMSE of the system can be computed.

Let the output of Kalman filtering be $\hat{\bar{\boldsymbol{x}}}(k)$; this is the LMMSE of vector $\bar{\boldsymbol{x}}(k)$, which is actually the LMMSE of the finest sensor nodes. The LMMSE of the state nodes of other sensors can be directly obtained from $\hat{\bar{\boldsymbol{x}}}(k)$. In [9], Zhang has proved that 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

$$\hat{\boldsymbol{x}}_j(2^{J-j}k + m_j) = M_j(m_j) \cdot \hat{\bar{\boldsymbol{x}}}(k). \quad (3.37)$$

The dimension of the system model (3.36) is N_J times that of the subsystem model of the finest sensor. Supposing that matrix C_j is of dimension $n_{\boldsymbol{z}_j} \times n_{\boldsymbol{x}}$, then the dimension of matrix $\bar{C}$ is $n_{\bar{\boldsymbol{z}}} \times (N_J \cdot n_{\boldsymbol{x}})$, where $n_{\bar{\boldsymbol{z}}} = \sum_{j=1}^{J} 2^{j-1} n_{\boldsymbol{z}_j}$. In Kalman filtering, one needs to compute the inverse of a $n_{\bar{\boldsymbol{z}}} \times n_{\bar{\boldsymbol{z}}}$ matrix to obtain the gain matrix. The computation burden will increase rapidly in the number of sensors and in the number of taps of filter D_0. To reduce the computational complexity, a sequential Kalman filtering algorithm for the model was presented in [10]. Since the measurements $\boldsymbol{z}_j(k_j)$ are captured independently and the observation noise $\boldsymbol{v}_j(k_j)$ is uncorrelated interscale and intrascale, then the augmented state vector $\bar{\boldsymbol{x}}(k)$ can be updated by $\boldsymbol{z}_j(k_j)$ one by one instead of putting all $\boldsymbol{z}_j(k_j)$ within a time block into one vector $\bar{\boldsymbol{z}}(k)$. The sequential filtering algorithm divides the inverse computation of a huge dimension matrix into the inverses of many small dimension matrices. It reduces the computation greatly. Especially at the finest scale, the sequential filtering is equivalent to the classical Kalman smoothing.

In the discrete model construction described in this section, we supposed that the sensor resolutions decrease by a power of two. In fact, this condition can be relaxed. For any fixed rate of resolution decrease, such as a power of three, etc., the corresponding discrete system model can be built in a similar way. Even for the case where the resolution decrease rates vary along scales, the associated model can be constructed as well but with a much more complex form.

IV. STABILITY OF THE KALMAN FILTER FOR TIME-INVARIANT SYSTEM

Stability [16], [18]–[20] is a critical property to the Kalman filter, and it guarantees the filtering error to be convergent. It is well known that if the state-space model (3.36) is stochastically controllable and observable, then the associated Kalman filter is stable. In the discrete model (3.36), the matrices $\bar{A}(k), \bar{B}(k),$

$$\bar{\boldsymbol{w}}(k) = \text{col}\{\boldsymbol{w}(n_{k+1} - n_{\bar{p}}), \boldsymbol{w}(n_{k+1} - n_{\bar{p}} + 1), \ldots, \boldsymbol{w}(n_{k+1} - 1)\} \qquad (3.28)$$

$$\bar{A}_{\bar{p}}(k, m_1) = \prod_{n=0}^{m_1} A(n_{k+1} - n_{\bar{p}} + m_1 - n) \qquad (3.29)$$

$$\bar{B}_{\bar{p}}(k, m_1) = \begin{bmatrix} \left(\prod_{n=0}^{m_1-1} A(n_{k+1} - n_{\bar{p}} + m_1 - n) \bullet B(n_{k+1} - n_{\bar{p}}) \right)^T \\ \left(\prod_{n=0}^{m_1-2} A(n_{k+1} - n_{\bar{p}} + m_1 - n) \bullet B(n_{k+1} - n_{\bar{p}} + 1) \right)^T \\ \vdots \\ (B(n_{k+1} - n_{\bar{p}} + m_1))^T \\ O \\ \vdots \\ O \end{bmatrix}^T \qquad (3.30)$$

$$\bar{A}_{\bar{p}}(k) = \begin{bmatrix} \bar{A}_{\bar{p}}(k, 0) \\ \bar{A}_{\bar{p}}(k, 1) \\ \vdots \\ \bar{A}_{\bar{p}}(k, 2^{J-1} - 1) \end{bmatrix}, \quad \bar{B}_{\bar{p}}(k) = \begin{bmatrix} \bar{B}_{\bar{p}}(k, 0) \\ \bar{B}_{\bar{p}}(k, 1) \\ \vdots \\ \bar{B}_{\bar{p}}(k, 2^{J-1} - 1) \end{bmatrix} \qquad (3.31)$$

and $\bar{C}(k)$ contain many zero elements, and the system matrix $\bar{A}(k)$ is singular. Except for some special cases, the system model (3.36) will be neither completely controllable nor completely observable, even if the subsystem of the finest sensor is stochastically controllable and observable [10]. Therefore, it is necessary to determine whether the Kalman filter is still stable or not if the model is neither completely controllable nor observable. In this section, we restrict our discussion to time-invariant systems, where matrices $A(\bullet), B(\bullet), C_j(\bullet)$ and the variance matrices of $\boldsymbol{w}(\bullet)$ and $\boldsymbol{v}_j(\bullet)$ in (1.4) and (1.5) are all constants. Then, the model (3.36) becomes

$$\begin{cases} \bar{\boldsymbol{x}}(k+1) = \bar{A}\cdot\bar{\boldsymbol{x}}(k) + \bar{B}\cdot\bar{\boldsymbol{w}}(k) \\ \bar{\boldsymbol{z}}(k) = \bar{C}\cdot\bar{\boldsymbol{x}}(k) + \bar{\boldsymbol{v}}(k) \end{cases} \quad (4.1)$$

where

$$\bar{A} = \begin{bmatrix} O & \bar{I} & O \\ O & O & I_{n_x} \\ O & O & \bar{A}_{\bar{p}} \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} O \\ \bar{B}_{\bar{p}} \end{bmatrix}$$

$$\bar{C} = \text{col}\{\bar{C}_J, \bar{C}_{J-1}, \dots, \bar{C}_1\} \quad (4.2)$$

$$\bar{A}_{\bar{p}} = \text{col}\left\{A, A^2, \dots, A^{2^{J-1}}\right\}$$

$$\bar{B}_{\bar{p}} = \begin{bmatrix} B & O & \cdots & O \\ AB & B & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ A^{2^{J-1}-1}B & A^{2^{J-1}-2}B & \cdots & B \end{bmatrix}. \quad (4.3)$$

Denote by $\bar{Q}$ and $\bar{R}$ the covariance matrices of $\bar{\boldsymbol{w}}(k)$ and $\bar{\boldsymbol{v}}(k)$, respectively. Here, we normalize $\bar{\boldsymbol{w}}(k)$ as a unit white Gaussian process, i.e., $\bar{Q}$ is an identity matrix.

For the time-invariant linear system, the following lemma given in [19] offers a sufficient condition for the stability of its associated Kalman filter.

Lemma 4-1: Given a time-invariant linear system

$$\begin{cases} \boldsymbol{x}(k+1) = \boldsymbol{F}\cdot\boldsymbol{x}(k) + \Gamma\cdot\boldsymbol{w}(k) \\ \boldsymbol{z}(k) = \text{H}\cdot\boldsymbol{x}(k) + \boldsymbol{v}(k) \end{cases} \quad (4.4)$$

where $\boldsymbol{w}(\bullet)$ is unit Gaussian white process with zero mean. If pair $(H, \boldsymbol{F})$ is completely detectable and pair $(\boldsymbol{F}, \Gamma)$ is completely stabilizable, then the system's Kalman filter is asymptotically stable. ∎

The definitions of "completely detectable" and "completely stabilizable" are as follows [19]:

Definition 4-1: Pair $(\boldsymbol{F}, \Gamma)$ is completely stabilizable if there exists a nonsingular matrix T such that

$$T\boldsymbol{F}T^{-1} = \begin{bmatrix} \boldsymbol{F}_{11} & \boldsymbol{F}_{12} \\ O & \boldsymbol{F}_{22} \end{bmatrix}, \quad \Gamma T = \begin{bmatrix} \Gamma_1 \\ O \end{bmatrix} \quad (4.5)$$

where $(\boldsymbol{F}_{11}, \Gamma_1)$ is completely controllable, and $|\lambda_i(\boldsymbol{F}_{22})| < 1$. ∎

Definition 4-2: Pair $(H, \boldsymbol{F})$ is completely detectable if there exists a nonsingular matrix T such that

$$T\boldsymbol{F}T^{-1} = \begin{bmatrix} \boldsymbol{F}_{11} & O \\ \boldsymbol{F}_{21} & \boldsymbol{F}_{22} \end{bmatrix}, \quad HT^{-1} = \begin{bmatrix} H_1 & O \end{bmatrix} \quad (4.6)$$

where $(H_1, \boldsymbol{F}_{11})$ is completely observable, and $|\lambda_i(\boldsymbol{F}_{22})| < 1$. ∎

The above tells us that if the eigenvalues of the subsystem matrix for uncontrollable elements are within the unit circle, then the whole system is stabilizable, and if the eigenvalues of the subsystem matrix for unobservable elements are within the unit circle, then the whole system is detectable. With Lemma 4-1, a theorem can be presented to guarantee the stability of the associated Kalman filter for the multiresolution multisensor system model (4.1).

Theorem 4-1: If at the finest resolution pair (A, B) is completely controllable and pair (C_1, A) is completely observable, then for the multiresolution multisensor system (4.1), pair $(\bar{A}, \bar{B})$ is completely stabilizable, and pair $(\bar{C}, \bar{A})$ is completely detectable; thus, the associated Kalman filter of system (4.1) is asymptotically stable. ∎

Proof 1) $(\bar{A}, \bar{B})$ Is Completely Stabilizable: Notice that for the n-dimensional pair $(\boldsymbol{F}, \Gamma)$, if there exists a non-singular matrix T such that $T\cdot\Omega = \text{col}\{\Omega_c, O\}$, where Ω_c is of full row rank, and Ω is the controllability matrix $\Omega = [\Gamma, \boldsymbol{F}\Gamma, \dots, \boldsymbol{F}^{m-1}\Gamma]$, then $T\cdot\boldsymbol{x} = \text{col}\{\boldsymbol{x}_c, \boldsymbol{x}_{\bar{c}}\}$, where $\boldsymbol{x}_c$ is controllable and $\boldsymbol{x}_{\bar{c}}$ is uncontrollable [17]. This gives us a way to find T.

Denote $z = \lceil n_p/n_{\bar{p}} \rceil$, i.e., z is the maximum integer no greater than $n_p/n_{\bar{p}}$. We have $n_p = z\cdot n_{\bar{p}} + e$. From the structure of matrix $\bar{A}$, when $1 \leq n \leq z + 1$, there is

$$\bar{A}^n = \begin{bmatrix} O & \bar{I}_n & O \\ \hline O & O & I_{n_x} \\ O & O & A_{\bar{p}} \\ \vdots & \vdots & \vdots \\ O & O & A_{\bar{p}}^{(n)} \end{bmatrix} \quad (4.7)$$

where $A_{\bar{p}}^{(n)} = A_{\bar{p}}\cdot A^{(n-1)2^{J-1}}$, and $\bar{I}_n = \text{diag}\{\underbrace{I_{n_x} \cdots I_{n_x}}_{n_p-(n-1)n_{\bar{p}}-1}\}$.

When $n > z + 1$, there is

$$\bar{A}^n = \begin{bmatrix} O & A_{\bar{p},(e)}^{(n-z-1)} \\ \hline O & A_{\bar{p}}^{(n-z)} \\ \vdots & \vdots \\ O & A_{\bar{p}}^{(n)} \end{bmatrix} \quad (4.8)$$

where $A_{\bar{p},(e)}$ is the last $e\cdot n_x$ rows of $A_{\bar{p}}$, and $A_{\bar{p},(e)}^{(n)} = A_{\bar{p},(e)}\cdot A^{(n-1)2^{J-1}}$.

Denote by $\bar{A}_{\bar{p}}^{(n)} = [O \ A_{\bar{p}}^{(n)}]$ and $\bar{A}_{\bar{p},(e)}^{(n)} = [O \ A_{\bar{p},(e)}^{(n)}]$ the associated square matrices of $A_{\bar{p}}^{(n)}$ and $A_{\bar{p},(e)}^{(n)}$, respectively. Letting $n_{\bar{x}} = N_J\cdot n_x$, the controllability matrix of pair $(\bar{A}, \bar{B})$ is as in (4.9), shown at the bottom of the next page, where $\bar{B}_{\bar{p},(e)}$ is the last $e\cdot n_x$ rows of $\bar{B}_{\bar{p}}$. Define

$$P_1 = \begin{bmatrix} I & O & O & \cdots & O & O \\ O & I & O & \cdots & O & O \\ O & -\bar{A}_{\bar{p}} & I & \cdots & O & O \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & -\bar{A}_{\bar{p}} & I \end{bmatrix}. \quad (4.10)$$

Obviously, P_1 is nonsingular. Transforming Ω by P_1, we have (4.11), shown at the bottom of the page, where Ω_1^u and Ω_1^d are independent in row. Then, there is

$$
\bar{A}_1 = P_1 \cdot \bar{A} \cdot P_1^{-1} = \left[\begin{array}{ccc|cc} O & I & O & O & O \\ O & O & I & O & O \\ O & O & \bar{A}_{\bar{p}} & I & O \\ \hline O & O & O & O & I \\ O & O & O & O & O \end{array}\right]
$$

$$
= \left[\begin{array}{c|c} \bar{A}_1^{lu} & \bar{A}_1^{ru} \\ \hline O & \bar{A}_1^{rd} \end{array}\right]. \tag{4.12}
$$

Define

$$
P_2 = \left[\begin{array}{cccccc} I_{n_x} & O & O & \cdots & O & O \\ -A & I_{n_x} & O & \cdots & O & O \\ O & -A & I_{n_x} & \cdots & O & O \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & -A & I_{n_x} \end{array}\right]. \tag{4.13}
$$

Transforming Ω_1 as

$$
\Omega_2 = \left[\begin{array}{c|c} P_2 & O \\ \hline O & I \end{array}\right] \cdot \Omega_1 = \left[\begin{array}{c|c} P_2 & O \\ \hline O & I \end{array}\right] \cdot \left[\begin{array}{c} \Omega_1^u \\ \Omega_1^d \end{array}\right] = \left[\begin{array}{c} \Omega_2^u \\ \Omega_2^d \end{array}\right] \tag{4.14}
$$

we have (4.15), shown at the bottom of the page. Obviously, Ω_2^{uu} and Ω_2^{ud} are independent in row. Now

$$
\bar{A}_2 = \left[\begin{array}{c|c} P_2 & O \\ \hline O & I \end{array}\right]\left[\begin{array}{c|c} A_1^{lu} & A_1^{ru} \\ \hline O & A_1^{rd} \end{array}\right]\left[\begin{array}{c|c} P_2^{-1} & O \\ \hline O & I \end{array}\right]
$$

$$
= \left[\begin{array}{c|c} P_2 \cdot \bar{A}_1^{lu} \cdot P_2^{-1} & P_2 \cdot \bar{A}_1^{ru} \\ \hline O & \bar{A}_1^{rd} \end{array}\right]. \tag{4.16}
$$

It is easy to validate that

$$
P_2 \cdot \bar{A}_1^{lu} \cdot P_2^{-1} = \left[\begin{array}{ccccccc} A^e & \cdots & A & I_{n_x} & O & \cdots & O \\ O & \cdots & O & O & I_{n_x} & \cdots & O \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ O & \cdots & O & O & O & \cdots & I_{n_x} \\ O & \cdots & O & O & O & \cdots & O \end{array}\right]. \tag{4.17}
$$

$$
\Omega = \begin{bmatrix} \bar{B} & \bar{A}\bar{B} & \cdots & \bar{A}^{n_x-1}\bar{B} \end{bmatrix}
$$

$$
= \left[\begin{array}{c|c|c|c|c|c|c|c} O & O & \cdots & O & \bar{B}_{\bar{p},(e)} & \bar{A}_{\bar{p},(e)}\bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p},(e)}^{(n_x-z-2)}\bar{B}_{\bar{p}} \\ O & O & \ddots & \bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}\bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}^{(2)}\bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p}}^{(n_x-z-1)}\bar{B}_{\bar{p}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ O & \bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p}}^{(z-1)}\bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}^{(z)}\bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}^{(z+1)}\bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p}}^{(n_x-2)}\bar{B}_{\bar{p}} \\ \bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}\bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p}}^{(z)}\bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}^{(z+1)}\bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}^{(z+2)}\bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p}}^{(n_x-1)}\bar{B}_{\bar{p}} \end{array}\right] \tag{4.9}
$$

$$
\Omega_1 = P_1 \cdot \Omega
$$

$$
= \left[\begin{array}{ccccccc} O & \cdots & O & O & \bar{B}_{\bar{p},(e)} & \bar{A}_{\bar{p}(e)}\cdot\bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p},(e)}^{n_x-2-z}\cdot\bar{B}_{\bar{p}} \\ O & \cdots & O & \bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}\cdot\bar{B}_{\bar{p}} & \bar{A}_{\bar{p}}^2\cdot\bar{B}_{\bar{p}} & \cdots & \bar{A}_{\bar{p}}^{n_x-1-z}\cdot\bar{B}_{\bar{p}} \\ \hline O & \cdots & \bar{B}_{\bar{p}} & O & O & O & \cdots & O \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \bar{B}_{\bar{p}} & \cdots & O & O & O & O & \cdots & O \end{array}\right] = \left[\begin{array}{c} \Omega_1^u \\ \Omega_1^d \end{array}\right] \tag{4.11}
$$

$$
\Omega_2^u = P_2 \cdot \Omega_1^u
$$

$$
= \left[\begin{array}{cccccccccccccc} \cdots & O & \cdots & O & A^{n_{\bar{p}}-e}B & A^{n_{\bar{p}}-e-1}B & \cdots & B & O & \cdots & O & A^{2\cdot n_{\bar{p}}-e}B & \cdots \\ \cdots & O & \cdots & O & O & O & \cdots & O & B & \cdots & O & O & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \cdots & O & \cdots & O & O & O & \cdots & O & O & \cdots & B & O & \cdots \\ \cdots & B & \cdots & O & O & O & \cdots & O & O & \cdots & O & O & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \cdots & O & \cdots & B & O & O & \cdots & O & O & \cdots & O & O & \cdots \end{array}\right] = \left[\begin{array}{c} \Omega_2^{uu} \\ \Omega_2^{ud} \end{array}\right] \tag{4.15}
$$

Thus, we can denote

$$\bar{A}_2 = \left[\begin{array}{c|c} A^e & \bar{A}_2^{ru} \\ \hline O & \bar{A}_2^{rd} \end{array} \right]. \qquad (4.18)$$

It is obvious that $\bar{A}_2^{rd}$ is a top triangle matrix with all its eigenvalues being zeros.

Because pair (A, B) is completely controllable, its controllability matrix $[B, AB, \ldots, A^{n_x-1}B]$ is of full row rank; therefore, Ω_2^{uu} will be of full row rank. There exists nonsingular matrix P_3 so that

$$\Omega_3 = \left[\begin{array}{c|c} I & O \\ \hline O & P_3 \end{array} \right] \cdot \Omega_2 = \left[\begin{array}{c|c} I & O \\ \hline O & P_3 \end{array} \right] \cdot \left[\begin{array}{c} \Omega_2^{uu} \\ \hline \Omega_2^{ud} \end{array} \right] = \left[\begin{array}{c} \Omega_2^{uu} \\ \hline \Omega_3^d \\ O \end{array} \right] \qquad (4.19)$$

where Ω_3^d is of full row rank and independent of Ω_2^{uu}. Then

$$\begin{aligned} \bar{A}_3 &= \left[\begin{array}{c|c} I & O \\ \hline O & P_3 \end{array} \right] \cdot \bar{A}_2 \cdot \left[\begin{array}{c|c} I & O \\ \hline O & P_3 \end{array} \right]^{-1} \\ &= \left[\begin{array}{c|c} A^e & \bar{A}_2^{ru} \cdot P_3^{-1} \\ \hline O & P_3 \cdot \bar{A}_2^{rd} \cdot P_3^{-1} \end{array} \right]. \end{aligned} \qquad (4.20)$$

$\bar{A}_3$ is the canonical controllable decomposition of $\bar{A}$; therefore, there must be

$$\bar{A}_3^{rd} = P_3 \cdot \bar{A}_2^{rd} \cdot P_3^{-1} = \left[\begin{array}{c|c} \bar{A}_{3,(1)}^{rd} & \bar{A}_{3,(2)}^{rd} \\ \hline O & \bar{A}_{3,(3)}^{rd} \end{array} \right]. \qquad (4.21)$$

Since $\bar{A}_2^{rd}$ and $\bar{A}_3^{rd}$ are similar matrices, they have the same eigenvalues, and then, $\bar{A}_{3,(3)}^{rd}$, which is the subsystem matrix for the uncontrollable elements, has all zero eigenvalues. According to Definition 4-1, $(\bar{A}, \bar{B})$ is completely stabilizable.

Proof 2) $(\bar{C}, \bar{A})$ Is Completely Detectable: Notice that for the n-dimensional pair (H, F), if there exists a nonsingular matrix T such that $\Lambda \cdot T^{-1} = [\Lambda_o \ \ O]$ where Λ_o is of full column rank, and Λ is the observability matrix $\Lambda = \text{col}\{H, HF, \ldots, HF^{m-1}\}$, then $T \cdot x = \text{col}\{x_o, x_{\bar{o}}\}$, where x_o is observable and $x_{\bar{o}}$ is unobservable [17]. This gives a way to find T. From (4.7) and (4.8), we can denote

$$\bar{A}^n = \left[\begin{array}{cc|c} O & \bar{I}_n & O \\ O & O & I_{n_x} \\ O & O & \bar{A}_{\bar{p}} \\ & & \vdots \\ O & O & \bar{A}_{\bar{p}}^{(n)} \end{array} \right] = \left[\begin{array}{c|c} \bar{A}_2^{(n)} & \bar{A}_1^{(n)} \end{array} \right]$$

$$1 \le n \le z+1 \quad (4.22)$$

$$\bar{A}^n = \left[\begin{array}{c|c} O & \bar{A}_{\bar{p},(l)}^{n-z-1} \\ O & \bar{A}_{\bar{p}}^{n-z} \\ \vdots & \vdots \\ O & \bar{A}_{\bar{p}}^n \end{array} \right] = \left[\begin{array}{c|c} O & \bar{A}_1^{(n)} \end{array} \right]$$

$$n > z+1. \quad (4.23)$$

Therefore, the observability matrix of pair $(\bar{C}, \bar{A})$ is

$$\begin{aligned} \Lambda &= \left[\begin{array}{c} \bar{C} \\ \bar{C} \cdot \bar{A} \\ \vdots \\ \bar{C} \cdot \bar{A}^{n_x-1} \end{array} \right] = \left[\begin{array}{c|c} \bar{C}_l & \bar{C}_r \\ \bar{C} \cdot \bar{A}_2 & \bar{C} \cdot \bar{A}_1 \\ \vdots & \vdots \\ \bar{C} \cdot \bar{A}_2^{z+1} & \bar{C} \cdot \bar{A}_1^{z+1} \\ \hline O & \bar{C} \cdot \bar{A}_1^{z+2} \\ \vdots & \vdots \\ O & \bar{C} \cdot \bar{A}_1^{n_x-1} \end{array} \right] \\ &= \left[\begin{array}{c|c} \Lambda^{lu} & \Lambda^{ru} \\ \hline O & \Lambda^{rd} \end{array} \right] = [\Lambda^l \ \ | \ \ \Lambda^r] \end{aligned} \qquad (4.24)$$

where $\bar{C}_r$ is the last n_x columns of $\bar{C}$, and $\bar{C}_l$ is the other columns of $\bar{C}$. First, we illustrate that Λ^l and Λ^r are independent in column. Because of the special structures of $\bar{C}$ and $\bar{A}$, the rows in Λ^{lu} correspond to those rows in Λ^{ru}, where blocks $C_1, C_1A, \ldots, C_1 \cdot A^n, \ldots$ are located, are all zeros. There exists a row transformation matrix P such that

$$\begin{aligned} \left[\begin{array}{c|c} P & O \\ \hline O & I \end{array} \right] \cdot \left[\begin{array}{c|c} \Lambda^{lu} & \Lambda^{ru} \\ \hline O & \Lambda^{rd} \end{array} \right] &= \left[\begin{array}{c|c} \Lambda_p^{lu} & \Lambda_p^{ru} \\ O & \Lambda_p^{rd} \\ \hline O & \Lambda^{rd} \end{array} \right] \\ &= [\Lambda_p^l \ \ | \ \ \Lambda_p^r] \end{aligned} \qquad (4.25)$$

where $\Lambda_p^{rd} = \text{col}\{C_1, C_1A, \ldots, C_1A^n, \ldots\}$. Since pair (C_1, A) is complete observable, its observability matrix $\Lambda_{C_1} = \text{col}\{C_1, C_1A, \ldots, C_1A^{n_x-1}\}$ is of full column rank. It is observed that all the rows of Λ_{C_1} are contained in $\text{col}\{\Lambda_p^{rd}, \Lambda^{rd}\}$; therefore, $\text{col}\{\Lambda_p^{rd}, \Lambda^{rd}\}$ must be of full column rank, and Λ_p^l and Λ_p^r are independent in column. Since P is just a linear row transformation, Λ^l and Λ^r are then independent in column, and Λ^r is of full column rank. There exists a column transformation matrix P_1 such that

$$\Lambda_1 = [\Lambda^l \ \ | \ \ \Lambda^r] \cdot \left[\begin{array}{c|c} P_1 & O \\ \hline O & I \end{array} \right] = [O \ \ \Lambda_{p_1}^l \ \ | \ \ \Lambda^r] \qquad (4.26)$$

and it can be written as

$$\Lambda_1 = [O \ \ | \ \ \Lambda_{p_1}^l \ \ \Lambda^r] = [O \ \ | \ \ \Lambda_1^r] \qquad (4.27)$$

where Λ_1^r is of full column rank. Denote $\bar{A}$ as

$$\bar{A} = \left[\begin{array}{cc} \bar{A}^{lu} & \bar{A}^{ru} \\ O & A^{2^{J-1}} \end{array} \right]. \qquad (4.28)$$

$\bar{A}^{lu}$ is a top triangle matrix with all zeros on the diagonal. We have

$$\begin{aligned} \bar{A}_1 &= \left[\begin{array}{cc} P_1^{-1} & O \\ O & I \end{array} \right] \left[\begin{array}{cc} \bar{A}^{lu} & \bar{A}^{ru} \\ O & A^{2^{J-1}} \end{array} \right] \left[\begin{array}{cc} P_1 & O \\ O & I \end{array} \right] \\ &= \left[\begin{array}{cc} P_1^{-1} \bar{A}^{lu} P_1 & P_1^{-1} \bar{A}^{ru} \\ O & A^{2^{J-1}} \end{array} \right] \end{aligned} \qquad (4.29)$$

Divide $P_1^{-1}\bar{A}^{lu}P_1$ as

$$P_1^{-1}\bar{A}^{lu}P_1 = \begin{bmatrix} \bar{A}_1^{lu} & \bar{A}_{p_1}^{ru} \\ \bar{A}_{p_1}^{ld} & \bar{A}_{p_1}^{rd} \end{bmatrix} \tag{4.30}$$

where the square matrix $\bar{A}_{p_1}^{rd}$ has the same number of columns as $\Lambda_{p_1}^l$. Then, $\bar{A}_1$ can be rewritten as

$$\bar{A}_1 = \left[\begin{array}{c|c} \bar{A}_1^{lu} & \bar{A}_1^{ru} \\ \hline \bar{A}_1^{ld} & \bar{A}_1^{rd} \end{array} \right] \tag{4.31}$$

Apply column transformation to Λ_1 such that

$$\Lambda_2 = [O \mid \Lambda_1^r] \cdot \left[\begin{array}{c|c} O & I \\ \hline I & O \end{array} \right] = [\Lambda_1^r \mid O]. \tag{4.32}$$

Now

$$\begin{aligned} \bar{A}_2 &= \left[\begin{array}{c|c} O & I \\ \hline I & O \end{array} \right] \cdot \left[\begin{array}{c|c} \bar{A}_1^{lu} & \bar{A}_1^{ru} \\ \hline \bar{A}_1^{ld} & \bar{A}_1^{rd} \end{array} \right] \cdot \left[\begin{array}{c|c} O & I \\ \hline I & O \end{array} \right]^{-1} \\ &= \left[\begin{array}{c|c} \bar{A}_1^{rd} & \bar{A}_1^{ld} \\ \hline \bar{A}_1^{ru} & \bar{A}_1^{lu} \end{array} \right]. \end{aligned} \tag{4.33}$$

Since $\bar{A}_2$ is the canonical observable decomposition of $\bar{A}$, $\bar{A}_1^{ld}$ must be zero, and because $\bar{A}_{p_1}^{ld}$ is the submatrix of $\bar{A}_1^{ld}$, it is also a zero matrix, i.e.,

$$P_1^{-1}\bar{A}^{lu}P_1 = \begin{bmatrix} \bar{A}_1^{lu} & \bar{A}_{p_1}^{ru} \\ O & \bar{A}_{p_1}^{rd} \end{bmatrix}.$$

Then, we can see that the eigenvalues of $\bar{A}_1^{lu}$ are part of those of $P_1^{-1}\bar{A}^{lu}P_1$, which is the similar matrix of $\bar{A}^{lu}$. Because $\bar{A}^{lu}$ is a top triangle matrix with all zeros on the diagonal, the eigenvalues of $\bar{A}_1^{lu}$ are all zeros. According to Definition 4-2, $(\bar{C}, \bar{A})$ is completely detectable.

Since pair $(\bar{A}, \bar{B})$ is completely stabilizable and pair $(\bar{C}, \bar{A})$ is completely detectable, according to Lemma 4-1, the corresponding Kalman filter of model (4.1) is asymptotically stable. ∎

V. EXAMPLE

In the implementation of the proposed multiresolution multisensor data fusion scheme, the coefficients of the lowpass filter D_0 should be first determined. In some systems, D_0 is known priorly. Otherwise, D_0 needs to be estimated. Actually, the correlation information of $\boldsymbol{x}_j$ across sensors is also hidden in $\boldsymbol{z}_j$ to some extent. In applications, D_0 can be first estimated by the prior knowledge of the system and then updated by the measurements $\boldsymbol{z}_j$ via an appropriate estimation approach. In recent years, the wavelet [11]–[15] has been successfully used to represent and model a variety of random processes [3]–[7], such as the multiscale autoregressive (MAR) framework [3]–[5] of statistical signals. The wavelet transform has a natural multiresolution and multiscale structure. In the following example, we set the original state projection filter D_0 to be $\{(1/3),(1/3),(1/3)\}$ and use the lowpass CDF(2, 2)

wavelet filter $\{(1/4),(1/2),(1/4)\}$ of [11] to approximate D_0. Interestingly, the experimental results show that the approximation error of D_0 only slightly affects the performance of the system estimation. In contrast, if we use the Haar wavelet filter $\{(1/2),(1/2)\}$ to approximate D_0, as what was done in [9], the system estimation error increases significantly.

The first-order scalar Markov processes are used in our simulation. The state transition equation is

$$x_1(k_1 + 1) = ax_1(k_1) + bw(k_1) \tag{5.1}$$

where $w(\bullet)$ is Gaussian white noise with zero mean and variance $q = 1$. Suppose there are two sensors available to measure the system, i.e.,

$$\begin{cases} z_1(k_1) = c_1 x_1(k_1) + v_1(k_1) \\ z_2(k_2) = c_2 x_2(k_2) + v_2(k_2) \end{cases} \tag{5.2}$$

where Gaussian white noises $v_1(\bullet)$ and $v_2(\bullet)$ are of zero-mean, their variances are r_1 and r_2, respectively, and $v_1(\bullet), v_2(\bullet)$ and $w(\bullet)$ are independent of each other. Letting $D_0 = \{(1/3),(1/3),(1/3)\}$ and referring to Fig. 6, the states at the second scale $x_2(k)$ are generated from the states at the first scale by

$$x_2(k) = x_1(2k-1)/3 + x_1(2k)/3 + x_1(2k+1)/3. \tag{5.3}$$

According to the modeling scheme of Section III, we have

$$\begin{aligned} \bar{\boldsymbol{x}}(k) &= \text{col}\{x_1(2k-1), x_1(2k), x_1(2k+1)\} \\ \bar{\boldsymbol{z}}(k) &= \text{col}\{z_2(k), z_1(2k), z_1(2k+1)\} \\ \bar{\boldsymbol{w}}(k) &= \text{col}\{w(2k+1), w(2k+2)\} \\ \bar{\boldsymbol{v}}(k) &= \text{col}\{v_2(k), v_1(2k), v_1(2k+1)\} \end{aligned}$$

and the matrices in time invariant model (4.1) are

$$\bar{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & a \\ 0 & 0 & a^2 \end{bmatrix} \quad \bar{B} = \begin{bmatrix} 0 & 0 \\ b & 0 \\ ab & b \end{bmatrix}$$

$$\bar{C} = \begin{bmatrix} c_2/D_0(1) & c_2/D_0(2) & c_2/D_0(3) \\ 0 & c_1 & 0 \\ 0 & 0 & c_1 \end{bmatrix}$$

$$\bar{Q} = \text{Cov}(\bar{w}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \bar{R} = \text{Cov}(\bar{v}) = \begin{bmatrix} r_2 & 0 & 0 \\ 0 & r_1 & 0 \\ 0 & 0 & r_1 \end{bmatrix}.$$

To evaluate the proposed scheme, we perform experiments by assuming that the true filter D_0 is known and that D_0 is unknown and approximated by the CDF(2, 2) wavelet filter $\{(1/4),(1/2),(1/4)\}$. The experimental results of the method in [9], in which the Haar wavelet was used to approximate D_0, are also given for comparison.

The parameters in the simulations are set as $a = 0.95, b = 1, c_1 = 1, c_2 = 1, r_1 = 2.25$, and $r_2 = 1.69$. By implementing the Kalman filtering of model (4.1), we obtained $\hat{\bar{\boldsymbol{x}}}(k)$, which is the LMMSE of $\bar{\boldsymbol{x}}(k)$. From $\hat{\bar{\boldsymbol{x}}}(k), \hat{x}_1(k_1)$, the LMMSE
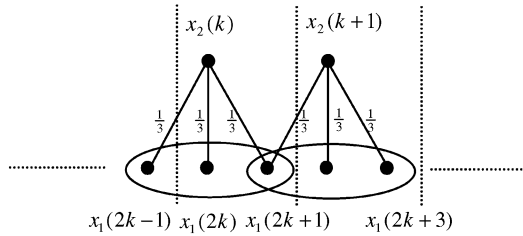
Fig. 6.   State nodes linking by filter $D_0 = \{(1/3), (1/3), (1/3)\}$ of a two sensors system.

TABLE I

AVERAGE NOISE COMPRESSION RATIOS OF THE PROPOSED SCHEME BY TRUE STATE PROJECTION FILTER $D_0 = \{(1/3), (1/3), (1/3)\}$ AND CDF(2, 2) WAVELET FILTER $D_0 = \{(1/4), (1/2), (1/4)\}$, AND THAT OF THE METHOD IN [9], WHERE THE HAAR WAVELET FILTER IS USED TO APPROXIMATE THE STATE PROJECTION FILTER. WE RAN MONTE CARLO SIMULATIONS 200 TIMES TO CALCULATE THE AVERAGE RESULTS

Noise compression ratio	Proposed scheme By true filter D_0	Proposed scheme By CDF(2,2) wavelet filter	Method in [9]
c_1	1.93	1.92	1.72
c_2	2.28	2.24	1.99

of $x_1(k_1)$, can be directly derived, and $\hat{x}_2(k_2)$, which is the LMMSE of $x_2(k_2)$, can be computed by (3.37). The estimation error is $\tilde{x}_j(k_j) = x_j(k_j) - \hat{x}_j(k_j)$, and the noise-suppression ratio c_j is defined as the ratio of the norm of observation noise v_j to that of $\tilde{x}_j$:

$$c_j = \|v_j\|/\|\tilde{x}_j\|. \qquad (5.4)$$

After running Monte Carlo simulations for 200 times, we listed the average values of noise suppression ratio c_j by setting $D_0 = \{(1/3), (1/3), (1/3)\}$ and $D_0 = \{(1/4), (1/2), (1/4)\}$ in Table I. The result of the scheme in [9] was also listed for comparison. It is interesting that the result of the approximation filter $\{(1/4), (1/2), (1/4)\}$ is only slightly worse than that of the true filter $\{(1/3), (1/3), (1/3)\}$. This implies that the proposed scheme is robust to the approximation errors of state projection filters to some extent, whereas the result of [9] is much worse than the new scheme. The reason for this is that the short Haar wavelet filter is incapable of sufficiently representing the state node correlation information cross scales.

Fig. 7(a) shows a sequence of true states $x_1(k_1)$ and its measurements $z_1(k_1)$, and Fig. 7(b) shows the associated state sequence $x_2(k_2)$, which is calculated by (5.3) and its measurement $z_2(k_2)$. Fig. 8(a) illustrates the estimation errors $\tilde{x}_1(k_1)$ by true filter $D_0 = \{(1/3), (1/3), (1/3)\}$ (solid line) and by approximation filter $D_0 = \{(1/4), (1/2), (1/4)\}$ (dotted line) for visual comparison. The noise suppression ratios c_1 are 1.978 and 1.969, respectively. Obviously, the two error sequences are almost the same. In Fig. 8(b), we illustrated the errors $\tilde{x}_1(k_1)$ by true filter $D_0 = \{(1/3), (1/3), (1/3)\}$ (solid line) and by Haar wavelet filter $D_0 = \{(1/2), (1/2)\}$ (dotted line), which is used in [9]. The noise suppression ratio c_1 by the Haar wavelet filter is 1.721. It is seen that the error sequence by the Haar wavelet has a much higher magnitude. Similarly, in Fig. 9(a) and (b), we showed the estimation error sequences
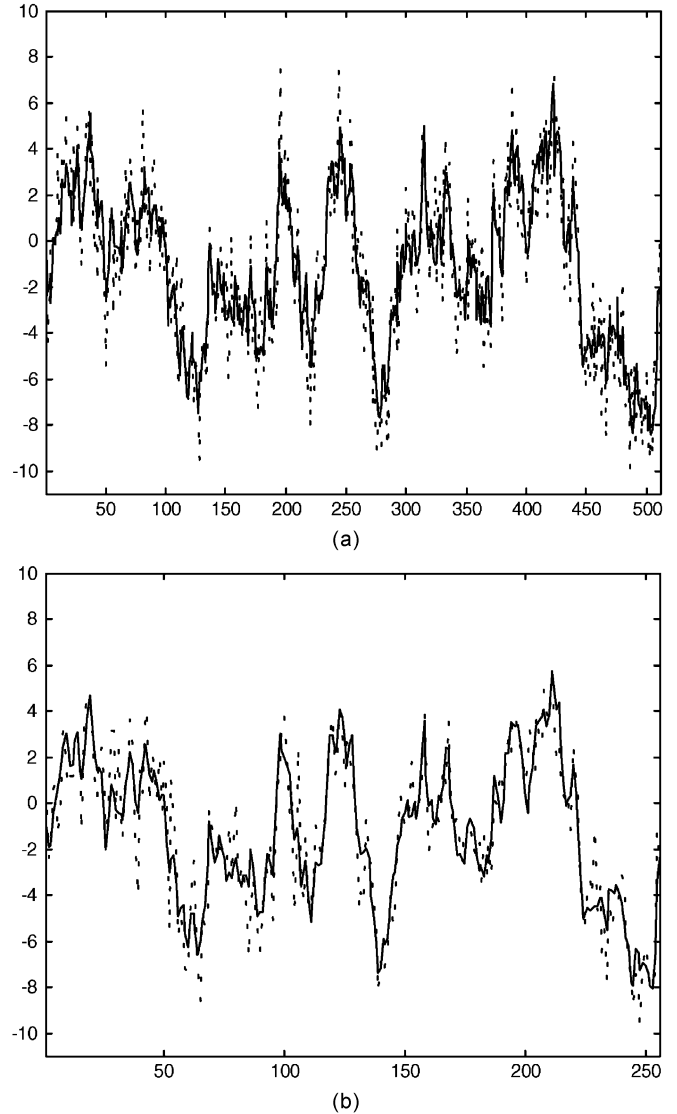

Fig. 7.   Data of a one-order scalar Markov process. (a) True state $x_1(k_1)$ (solid) and observation $z_1(k_1)$ (dotted). (b) True state $x_2(k_2)$ (solid) and observation $z_2(k_2)$, where $x_2(k_2)$ is obtained by filtering $x_1(k_1)$ with the lowpass filter $D_0 = \{(1/3), (1/3), (1/3)\}$.

$\tilde{x}_2(k_2)$ by the three filters $D_0 = \{(1/3), (1/3), (1/3)\}, D_0 = \{(1/4), (1/2), (1/4)\}$ and $D_0 = \{(1/2), (1/2)\}$. The noise suppression ratios c_2 are 2.329, 2.302, and 2.042, respectively.

VI. CONCLUSION

In this paper, we developed a modeling and estimation approach for a class of multiresolution multisensor dynamic systems, whose states are observed by several sensors of different resolutions. By introducing the state space projection equation to relate the states in each resolution space, we constructed the continuous-time model of the system. In discrete time applications, the sampling frequencies of sensors are made to decrease by a factor of two for each coarser resolution. We employed a linear projection to be associated with the state nodes of each of the sensors and constructed a discrete model of the system. It is shown that the Kalman filtering is the
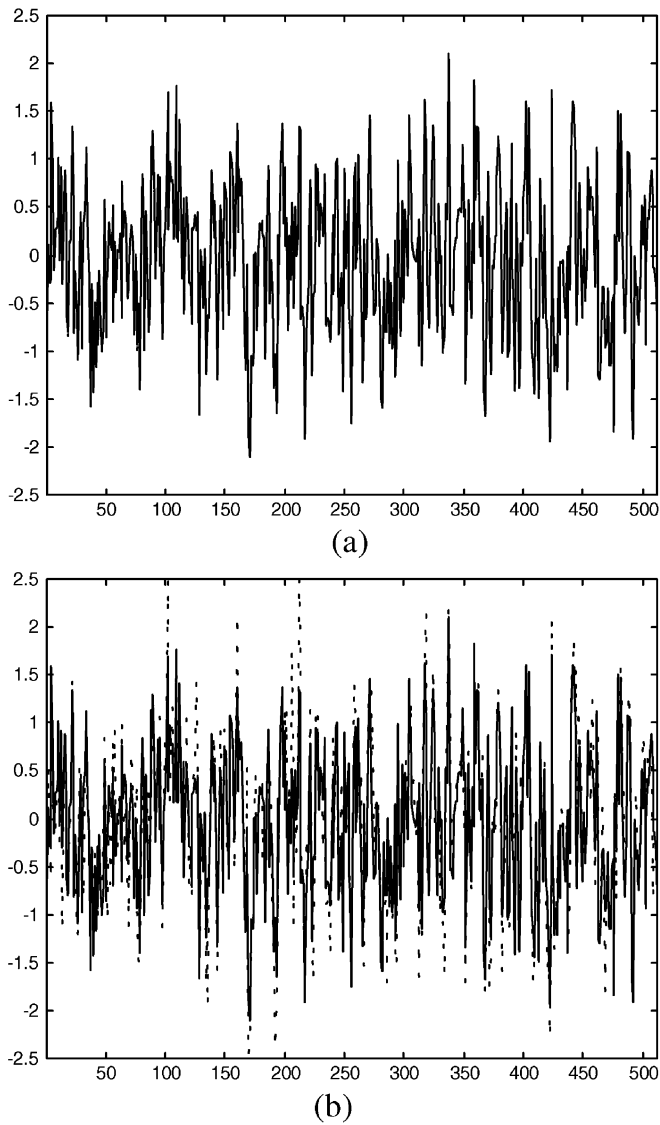
Fig. 8. Kalman filtering errors at the first scale. (a) Estimation error sequence $\bar{x}_1(k_1)$ by true filter $D_0 = \{(1/3),(1/3),(1/3)\}$ (solid) and that by approximation filter $D_0 = \{(1/4),(1/2),(1/4)\}$ (dotted). The two curves are almost the same, and the noise compression ratios c_1 are 1.978 and 1.969, respectively. (b) Estimation error sequence $\bar{x}_1(k_1)$ by true filter $D_0 = \{(1/3),(1/3),(1/3)\}$ (solid) and that by Haar wavelet filter $D_0 = \{(1/2),(1/2)\}$ (dotted). Obviously, the error sequence by the Haar wavelet has higher magnitude, and the associated noise compression ratio c_1 is 1.721.

Fig. 9. Kalman filtering errors at the second scale. (a) Estimation error sequence $\bar{x}_2(k_2)$ by true filter $D_0 = \{(1/3),(1/3),(1/3)\}$ (solid) and that by approximation filter $D_0 = \{(1/4),(1/2),(1/4)\}$ (dotted). The noise compression ratios c_2 are 2.329 and 2.302, respectively. (b) Estimation error sequence $\bar{x}_2(k_2)$ by true filter $D_0 = \{(1/3),(1/3),(1/3)\}$ (solid) and that by the Haar wavelet filter $D_0 = \{(1/2),(1/2)\}$ (dotted). The noise compression ratio c_2 by the Haar wavelet is 2.042.

optimal LMMSE algorithm for the developed system model. We proved that as long as the subsystem at the finest resolution is completely controllable and observable, the associated Kalman filtering of our system model is asymptotically stable. The first-order Markov processes were used to evaluate the new scheme in simulations. Our empirical results showed that the proposed scheme is robust to the approximation error of the state projection filter. Future work will develop an adaptive estimation algorithm of the state projection filter. The goal is to make filter estimation stable and optimal in some sense. One idea is to update iteratively the filter, which is initialized at the beginning, once a new measurement of the system is made.
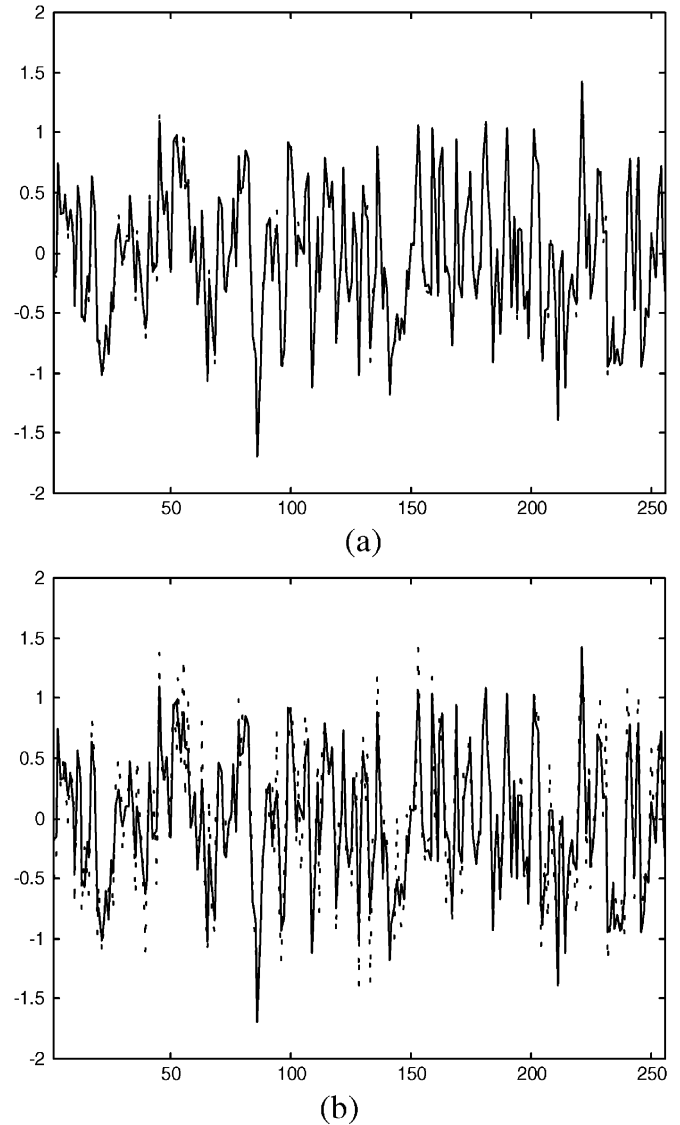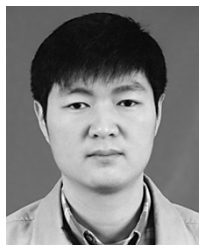
REFERENCES

[1] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, pp. 6–23, Jan. 1997.

[2] E. Weinstein *et al.*, "Iterative and sequential algorithms for multisensor signal enhancement," *IEEE Trans. Signal Processing*, vol. 42, pp. 846–859, Apr. 1994.

[3] M. Basseville, A. Benveniste, K. Chou, S. Golden, R. Nikoukhah, and A. S. Willsky, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory*, vol. 38, pp. 766–784, Mar. 1992.

[4] K. Chou, A. S. Willsky, and R. Nikoukhah, "Multiscale systems, Kalman filters, and Riccati equations," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 479–492, Mar. 1994.

[5] K. Daoudi, A. Frakt, and A. S. Willsky, "Multiscale autoregressive models and wavelets," *IEEE Trans. Inform. Theory*, vol. 45, pp. 828–845, Apr. 1999.

[6] G. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proc. IEEE*, vol. 81, pp. 1428–1450, Oct. 1993.

[7] M. Luettgen, W. Karl, A. S. Willsky, and R. Tenney, "Multiscale representations of Markov random fields," *IEEE Trans. Signal Processing*, vol. 41, pp. 3377–3396, Dec. 1993.

[8] L. Hong, "Multiresolutional distributed filtering," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 853–856, Apr. 1994.

[9] L. Zhang *et al.*, "The discrete Kalman filtering of a class of dynamic multiscale systems," *IEEE Trans. Circuits Syst. II*, vol. 49, pp. 668–676, Oct. 2002.

[10] L. Zhang, "The optimal estimation of a class of dynamic systems," Ph.D. dissertation, Northwestern Polytech. Univ., Xi'an, China, Oct. 2001.

[11] I. Daubechies, *Ten Lectures on Wavelets*, ser. CBMS-NSF Series in Appl. Math.. Philadelphia, PA: SIAM, 1992.

[12] S. Mallat, "A theory for multiresolution signal decmopositon: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, Nov. 1989.

[13] M. Vetterli and C. Herley, "Wavelet and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, pp. 2207–2232, Sept. 1992.

[14] B. Jawerth and W. Sweldens, "An overview of wavelet based multiresolution analyzes," in *SIAM Rev.*, vol. 36, 1994, pp. 377–412.

[15] G. Strang and T. Nguyen, *Wavelet and Filter Banks*. Wellesley, MA: Wellesley-Cambridge, 1996.

[16] F. L. Lewis, *Optimal Estimation*. New York: Wiley, 1986.

[17] C.-T. Chen, *Linear System Theory and Design*. New York: Holt, Rinehart, and Winston, 1970.

[18] J. M. Mendal, *Lessons in Digital Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

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

[20] S. S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[21] P. Lancaster *et al.*, *The Theory of Matrices*, Second ed. New York: Academic, 1985.

Xiaolin Wu (SM'96) received the B.Sc. degree from Wuhan University, Wuhan, China in 1982 and the Ph.D. degree from the University of Calgary, AB, Canada, in 1988, both in computer science.

He is currently a Professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, and a Research Professor of computer science with Polytechnic University, Brooklyn, NY, where he holds the NSERC-DALSA research chair in Digital Cinema. His research interests include image processing, multimedia coding and communications, data compression, and signal quantization. He has published over 100 research papers in these fields.

Quan Pan received the B.S. degree in automatic control from Huazhong Institute of Technology, Wuhan, China, in 1982 and the M.S. and Ph.D. degrees, both in control theory and control engineering, from Northwestern Polytechnical University (NPU), Xi'an, China, in 1991 and 1997, respectively.

He has been a Professor with NP since 1991, where he is the Director of the Information Fusion Institute. His research interests include information fusion, target tracking, automatic target recognition, multiscale system theory, and signal processing.

Dr. Pan received the National Youth Award of Outstanding Contribution of Science and Technology in China in 1998. He is a member of International Society of Information Fusion.

Lei Zhang was born in 1974 in China. He received the B.S. degree in 1995 from Shenyang Institute of Aeronautic Engineering, Shenyang, China, and the M.S. and Ph.D. degrees in electrical and engineering from Northwestern Polytechnical University, Xi'an, China, in 1998 and 2001, respectively.

From 2001 to 2002, he was a research associate with the Department of Computing, The Hong Kong Polytechnic University. Currently, he is a postdoctoral researcher with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. His research interests include digital signal and image processing, optimal estimation theory, information fusion, and wavelet transform.

Hongcai Zhang received the degree in gyroscope engineering in 1961 and the postgraduate degree in nonlinear control theory and it's applications in 1964, both from Northwestern Polytechnical University (NPU), Xi'an, China.

He joined the Department of Automatic Control, NPU, in 1964, where, since 1988, he has been a Professor. He is the Director of the Control Engineering Institute at NPU. He was a vice dean of the Department of Automatic Control from 1987 to 1992 and a curator of the Library at NPU from 1996 to 1998. His research interests include information fusion, target tracking, multiscale system theory, and signal processing.