1

Real-time Recursive State Estimation for Nonlinear Discrete Dynamic Systems with Gaussian or non-Gaussian Noise

Kerim Demirbaş
Department of Electrical and Electronics Engineering
Middle East Technical University
Inonu Bulvari, 06531 Ankara
Turkey

1. Introduction

Many systems in the real world are more accurately described by nonlinear models. Since the original work of Kalman (Kalman, 1960; Kalman & Busy, 1961), which introduces the Kalman filter for linear models, extensive research has been going on state estimation of nonlinear models; but there do not yet exist any optimum estimation approaches for all nonlinear models, except for certain classes of nonlinear models; on the other hand, different suboptimum nonlinear estimation approaches have been proposed in the literature (Daum, 2005). These suboptimum approaches produce estimates by using some sorts of approximations for nonlinear models. The performances and implementation complexities of these suboptimum approaches surely depend upon the types of approximations which are used for nonlinear models. Model approximation errors are an important parameter which affects the performances of suboptimum estimation approaches. The performance of a nonlinear suboptimum estimation approach is better than the other estimation approaches for specific models considered, that is, the performance of a suboptimum estimation approach is model-dependent.

The most commonly used recursive nonlinear estimation approaches are the extended Kalman filter (EKF) and particle filters. The EKF linearizes nonlinear models by Taylor series expansion (Sage & Melsa, 1971) and the unscented Kalman filter (UKF) approximates *a posteriori* densities by a set of weighted and deterministically chosen points (Julier, 2004). Particle filters approximates *a posterior* densities by a large set of weighted and randomly selected points (called particles) in the state space (Arulampalam et al., 2002; Doucet et al., 2001; Ristic et al., 2004). In the nonlinear estimation approaches proposed in (Demirbaş, 1982; 1984; Demirbaş & Leondes, 1985; 1986; Demirbaş, 1988; 1989; 1990; 2007; 2010): the disturbance noise and initial state are first approximated by a discrete noise and a discrete initial state whose distribution functions the best approximate the distribution functions of the disturbance noise and initial state, states are quantized, and then multiple hypothesis testing is used for state estimation; whereas Grid-based approaches approximate *a posteriori* densities by discrete densities, which are determined by predefined gates (cells) in the predefined state space; if the state space is not finite in extent, then the state space necessitates some truncation of the state space; and grid-based estimation approaches assume the availability of the state

transition density $p(x(k)|x(k-1))$, which may not easily be calculated for state models with nonlinear disturbance noise (Arulampalam et al., 2002; Ristic et al., 2004). The Demirbaş estimation approaches are more general than grid-based approaches since 1) the state space need not to be truncated, 2) the state transition density is not needed, 3) state models can be any nonlinear functions of the disturbance noise.
This chapter presents an online recursive nonlinear state filtering and prediction scheme for nonlinear dynamic systems. This scheme is recently proposed in (Demirbaş, 2010) and is referred to as the DF throughout this chapter. The DF is very suitable for state estimation of nonlinear dynamic systems under either missing observations or constraints imposed on state estimates. There exist many nonlinear dynamic systems for which the DF outperforms the extended Kalman filter (EKF), sampling importance resampling (SIR) particle filter (which is sometimes called the bootstrap filter), and auxiliary sampling importance resampling (ASIR) particle filter. Section 2 states the estimation problem. Section 3 first discusses discrete noises which approximate the disturbance noise and initial state, and then presents approximate state and observation models. Section 4 discusses optimum state estimation of approximate dynamic models. Section 5 presents the DF. Section 6 yields simulation results of two examples for which the DF outperforms the EKF, SIR, and ASIR particle filters. Section 7 concludes the chapter.

2. Problem statement

This section defines state estimation problem for nonlinear discrete dynamic systems. These dynamic systems are described by

State Model

$$x(k+1) = f(k, x(k), w(k)) \tag{1}$$

Observation Model

$$z(k) = g(k, x(k), v(k)), \tag{2}$$

where k stands for the discrete time index; $f : \mathbb{R}x\mathbb{R}^m x\mathbb{R}^n \rightarrow \mathbb{R}^m$ is the state transition function; $\mathbb{R}^m$ is the m-dimensional Euclidean space; $w(k) \in \mathbb{R}^n$ is the disturbance noise vector at time k; $x(k) \in \mathbb{R}^m$ is the state vector at time k; $g : \mathbb{R}x\mathbb{R}^m x\mathbb{R}^p \rightarrow \mathbb{R}^r$ is the observation function; $v(k) \in \mathbb{R}^p$ is the observation noise vector at time k; $z(k) \in \mathbb{R}^r$ is the observation vector at time k; $x(0)$, $w(k)$, and $v(k)$ are all assumed to be independent with known distribution functions. Moreover, it is assumed that there exist some constraints imposed on state estimates. The DF recursively yields a predicted value $\hat{x}(k|k-1)$ of the state $x(k)$ given the observation sequence from time one to time $k-1$, that is, $Z^{k-1} \triangleq \{z(1),\ z(2), \ldots, z(k-1)\}$; and a filtered value $\hat{x}(k|k)$ of the state $x(k)$ given the observation sequence from time one to time k, that is, Z^k. Estimation is accomplished by first approximating the disturbance noise and initial state with discrete random noises, quantizing the state, that is, representing the state model with a time varying state machine, and an online suboptimum implementation of multiple hypothesis testing.

3. Approximation

This section first discusses an approximate discrete random vector which approximates a random vector, and then presents approximate models of nonlinear dynamic systems.

3.1 Approximate discrete random noise

In this subsection: an approximate discrete random vector with n possible values of a random vector is defined; approximate discrete random vectors are used to approximate the disturbance noise and initial state throughout the chapter; moreover, a set of equations which must be satisfied by an approximate discrete random variable with n possible values of an absolutely continuous random variable is given (Demirbaş, 1982; 1984; 2010); finally, the approximate discrete random variables of a Gaussian random variable are tabulated.

Let w be an m-dimensional random vector. **An approximate discrete random vector with n possible values** of w, denoted by w_d, is defined as an m-dimensional discrete random vector with n possible values whose distribution function the best approximates the distribution function of w over the distribution functions of all m-dimensional discrete random vectors with n possible values, that is

$$w_d = \min_{y \epsilon D}{}^{-1} \{ \int_{\mathbb{R}^n} [F_y(a) - F_w(a)]^2 da \} \tag{3}$$

where D is the set of all m-dimensional discrete random vectors with n possible values, $F_y(a)$ is the distribution function of the discrete random vector y, $F_w(a)$ is the distribution function of the random vector w, and $\mathbb{R}^m$ is the m-dimensional Euclidean space. An approximate discrete random vector w_d is, in general, numerically, offline-calculated, stored and then used for estimation. The possible values of w_d are denoted by w_{d1}, w_{d2}, ..., and w_{dn} ; and the occurrence probability of the possible value w_{di} is denoted by $P_{w_{di}}$, that is

$$P_{w_{di}} \triangleq Prob\{w_d = w_{di}\}. \tag{4}$$

where $Prob\{w_d(0) = w_{di}\}$ is the occurrence probability of w_{di}.

Let us now consider the case that w is an absolutely continuous random variable. Then, w_d is an approximate discrete random variable with n possible values whose distribution function the best approximates the distribution function $F_w(a)$ of w over the distribution functions of all discrete random variables with n possible values, that is

$$w_d = \min_{y \epsilon D}{}^{-1} \{ J(F_y(a)) \}$$

in which the distribution error function (the objective function) $J(F_y(a))$ is defined by

$$J(F_y(a)) \triangleq \int_{\mathbb{R}} [F_y(a) - F_w(a)]^2 da$$

where D is the set of all discrete random variables with n possible values, $F_y(a)$ is the distribution function of the discrete random variable y, $F_w(a)$ is the distribution function of the absolutely continuous random variable w, and $\mathbb{R}$ is the real line. Let the distribution function $F_y(a)$ of a discrete random variable y be given by

$$F_y(a) \triangleq \begin{cases} 0 & \text{if } a < y_1 \\ F_{y_i} & \text{if } y_i \leq a < y_{i+1},\ i = 1,\ 2,\ \ldots, n-1 \\ 1 & \text{if } a \geq y_n. \end{cases}$$

Then the distribution error function $J(F_y(a))$ can be written as

$$J(F_y(a)) = \int_{-\infty}^{y_1} F_w^2(a) da + \sum_{i=1}^{n-1} \int_{y_i}^{y_{i+1}} [F_{y_i} - F_w(a)]^2 da + \int_{y_n}^{\infty} [1 - F_w(a)]^2 da.$$

Let the distribution function $F_{w_d}(a)$ of an approximate discrete random variable w_d be

$$F_{w_d}(a) \triangleq \begin{cases} 0 & \text{if } a < w_{d1} \\ F_{w_{di}} & \text{if } w_{di} \le a < w_{di+1},\ i = 1,\ 2,\ \ldots, n-1 \\ 1 & \text{if } a \ge w_{dn}. \end{cases}$$

It can readily be shown that the distribution function $F_{w_d}(a)$ of the approximate discrete random variable w_d must satisfy the set of equations given by

$$\begin{aligned} F_{w_{d1}} &= 2F_w(w_{d1}); \\ F_{w_{di}} + F_{w_{di+1}} &= 2F_w(w_{di+1}),\ \ i = 1,\ 2,\ \ldots,\ n-2; \\ 1 + F_{w_{dn-1}} &= 2F_w(w_{dn}); \\ F_{w_{di}}[w_{di+1} - w_{di}] &= \int_{w_{di}}^{w_{di+1}} F_w(a)da,\ \ i = 1,\ 2,\ \ldots,\ n-1. \end{aligned} \tag{5}$$

The values w_{d1}, w_{d2}, ..., w_{dn}, $F_{w_{d1}}$, $F_{w_{d2}}$, ...,$F_{w_{dn}}$ satisfying the set of Eqs. (5) determine the distribution function of w_d. These values can be, in general, obtained by numerically solving Eqs. (5). Then the possible values of the approximate discrete random variable w_d become w_{d1}, w_{d2}, ..., and w_{dn} ; and the occurrence probabilities of these possible values are obtained by

$$P_{w_{di}} = \begin{cases} F_{w_{d1}} & \text{if } i = 1 \\ F_{w_{di}} - F_{w_{di-1}} & \text{if } i = 2,\ 3,\ \ldots, n-1 \\ 1 - F_{w_{dn}} & \text{if } i = n. \end{cases}$$

where $P_{w_{di}} = Prob\{w_d = w_{di}\}$, which is the occurrence probability of w_{di}.
Let y be a Gaussian random variable with zero mean and unit variance. An approximate discrete random variable y_d with n possible values was numerically calculated for different n's by using the set of Eqs. (5). The possible values y_{d1}, y_{d2}, ..., y_{dn} of y_d and the occurrence probabilities $P_{y_{d1}}$, $P_{y_{d2}}$, ..., $P_{y_{dn}}$ of these possible values are given in Table 1, where $P_{y_{di}} \triangleq Prob\{y_d = y_{di}\}$. As an example, the possible values of an approximate discrete random variable with 3 possible values of a Gaussian random variable with zero mean and unit variance are -1.005, 0.0, and 1.005; and the occurrence probabilities of these possible values are 0.315, 0.370, and 0.315, respectively. Let w be a Gaussian random variable with mean $E\{w\}$ and variance σ^2. This random variable can be expressed as $w = y\sigma + E\{w\}$. Hence, the possible values of an approximate discrete random variable of w are given by $w_{di} = y_{di}\sigma + E\{w\}$, *where* $i = 1, 2, 3, ..., n$; and the occurrence probability of the possible value w_{di} is the same as the occurrence probability of y_{di}, which is given in Table 1.

3.2 Approximate models

For state estimation, the state and observation models of Eqs. (1)and (2) are approximated by the time varying finite state model and approximate observation model which are given by

Finite State Model

$$x_q(k+1) = Q(f(k, x_q(k), w_d(k))) \tag{6}$$

Approximate Observation Model

$$z(k) = g(k, x_q(k), v(k)), \tag{7}$$

n	y_{d1} $P_{y_{d1}}$	y_{d2} $P_{y_{d2}}$	y_{d3} $P_{y_{d3}}$	y_{d4} $P_{y_{d4}}$	y_{d5} $P_{y_{d5}}$	y_{d6} $P_{y_{d6}}$	y_{d7} $P_{y_{d7}}$	y_{d8} $P_{y_{d8}}$	y_{d9} $P_{y_{d9}}$	y_{d10} $P_{y_{d10}}$
1	0.000 1.000									
2	-0.675 0.500	0.675 0.500								
3	-1.005 0.315	0.0 0.370	1.005 0.315							
4	-1.218 0.223	-0.355 0.277	0.355 0.277	1.218 0.223						
5	-1.377 0.169	-0.592 0.216	0.0 0.230	0.592 0.216	1.377 0.169					
6	-1.499 0.134	-0.768 0.175	-0.242 0.191	0.242 0.191	0.768 0.175	1.499 0.134				
7	-1.603 0.110	-0.908 0.145	-0.424 0.162	0.0 0.166	0.424 0.162	0.908 0.145	1.603 0.110			
8	-1.690 0.092	-1.023 0.124	-0.569 0.139	-0.184 0.145	0.184 0.145	0.569 0.139	1.023 0.124	1.690 0.092		
9	-1.764 0.079	-1.120 0.106	-0.690 0.121	-0.332 0.129	0 0.130	0.332 0.129	0.690 0.121	1.120 0.106	1.764 0.079	
10	-1.818 0.069	-1.199 0.093	-0.789 0.106	-0.453 0.114	-0.148 0.118	0.148 0.118	0.453 0.114	0.789 0.106	1.199 0.093	1.818 0.069

Table 1. Approximate Discrete Random Variables the best Approximating the Gaussian Random Variable with Zero Mean and Unit Variance

where $w_d(k)$ is an approximate discrete random vector with, say, n possible values of the disturbance noise vector $w(k)$; this approximate vector is pre(offline)-calculated, stored and then used for estimation to calculate quantization levels at time $k+1$; the possible values of $w_d(k)$ are denoted by $w_{d1}(k)$, $w_{d2}(k)$, ..., and $w_{dn}(k)$; $Q : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a quantizer which first divides the m-dimensional Euclidean space into nonoverlapping generalized rectangles (called gates) such that the union of all rectangles is the m-dimensional Euclidean space, and then assigns to each rectangle the center point of the rectangle, Fig. 1 (Demirbaş, 1982; 1984; 2010); $x_q(k)$, $k > 0$, is the quantized state vector at time k and its quantization levels, whose number is (say) m_k, are denoted by $x_{q1}(k)$, $x_{q2}(k)$, ..., and $x_{qm_k}(k)$. The quantization levels of $x_q(k+1)$ are calculated by substituting $x_q(k) = x_{qi}(k)$ $(i = 1, 2, \ldots, m_k)$ for $x_q(k)$ and $w_d(k) = w_{dj}(k)$ $(j = 1, 2, \ldots, n)$ for $w_d(k)$ in the finite state model of Eq. (6). As an example, let the quantization level $x_{qi}(k)$ in the gate G_i be mapped into the gate G_j by the l^{th}-possible value $w_{dl}(k)$ of $w_d(k)$, then, $x(k+1)$ is quantized to $x_{qj}(k+1)$, Fig. 1. One should note that the approximate models of Eqs. (6) and (7) approach the models of Eqs. (1) and (2) as the gate sizes $(GS) \rightarrow 0$ and $n \rightarrow \infty$. An optimum state estimation of the models of Eqs. (6) and (7) is discussed in the next section.

4. Optimum state estimation

This section discuses an optimum estimation of the models of Eqs. (6) and (7) by using multiple hypothesis testing. On the average overall error probability sense, optimum estimation of states of the models of Eqs. (6) and (7) is done as follows: Finite state model

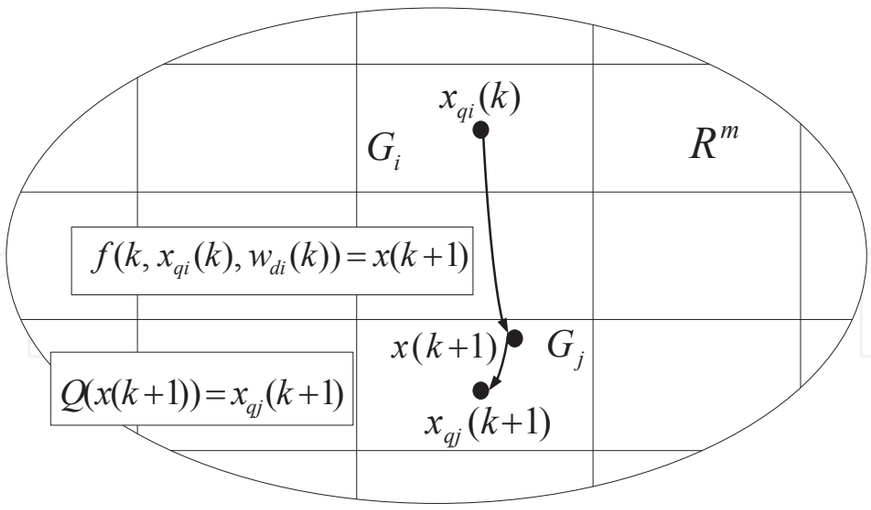


Fig. 1. Quantization of States

of Eq. (6) is represented by a trellis diagram from time 0 to time k (Demirbaş, 1982; 1984; Demirbaş & Leondes, 1985; Demirbaş, 2007). The nodes at time j of this trellis diagram represent the quantization levels of the state $x(j)$. The branches of the trellis diagram represent the transitions between quantization levels. There exist, in general, many paths through this trellis diagram. Let H^i denote the i^{th} path (sometimes called the i^{th} hypothesis) through the trellis diagram. Let $x_q^i(j)$ be the node (quantization level) through which the path H^i passes at time j. The estimation problem is to select a path (sometimes called the estimator path) through the trellis diagram such that the average overall error probability is minimized for decision (selection). The node at time k along this estimator path will be the desired estimate of the state $x(k)$. In Detection Theory (Van Trees, 2001; Weber, 1968): it is well-known that the optimum decision rule which minimizes the average overall error probability is given by

$$Select \;\; H^n \;\; as \; the \; estimator \; path \; if \; M(H^n) \geq M(H^l) \;\; for \;\; all \;\; l \neq n, \tag{8}$$

where $M(H^n)$ is called the metric of the path $M(H^n)$ and is defined by

$$M(H^n) \triangleq \ln\{p(H^n) Prob\{observation \; sequence \mid H^n\}\}, \tag{9}$$

where ln stands for the natural logarithm, $p(H^n)$ is the occurrence probability (or the *a priori probability*) of the path H^n, and $Prob\{observation \;\; sequence \mid H^n\}$ is the conditional probability of the observation sequence given that the actual values of the states are equal to the quantization levels along the path H^n. If the inequality in the optimum decision rule becomes an equality for an observation sequence, anyone of the paths satisfying the equality can be chosen as the estimator path, which is a path having the biggest metric.
It follows, from the assumption that samples of the observation noise are independent, that $Prob\{observation \; sequence \mid H^n\}$ can be expressed as

$$Prob\{observation \; sequence \mid H^n\} = \prod_{j=1}^{k} \lambda(z(j) \mid x_q^n(j)) \tag{10}$$

where

$$\lambda(z(j)|x_q^n)(j)) \triangleq \begin{cases} 1 & \text{if z(j) is neither available nor used for estimation} \\ p(z(j)|x_q^n(j)) & \text{if z(j) is available and used for estimation,} \end{cases} \quad (11)$$

in which, $p(z(j)|x_q^n(j))$ is the conditional density function of $z(j)$ given that the actual value of state is equal to $x_q^n(j)$, that is, $x(j) = x_q^n(j)$; and this density function is calculated by using the observation model of Eq. (2).
It also follows, from the assumption that all samples of the disturbance noise and the initial state are independent, that the *a priori probability* of H^n can be expressed as

$$p(H^n) = Prob\{x_q(0) = x_q^n(0)\} \prod_{j=1}^{k} T(x_q^n(j-1) \rightarrow x_q^n(j)), \quad (12)$$

where $Prob\{x_q(0) = x_q^n(0)\}$ is the occurrence probability of the initial node (or quantization level) $x_q^n(0)$, and $T(x_q^n(j-1) \rightarrow x_q^n(j))$ is the transition probability from the quantization level $x_q^n(j-1)$ to the quantization level $x_q^n(j))$, that is, $T(x_q^j(i-1) \rightarrow x_q^n(j)) \triangleq Prob\{x_q(j) = x_q^n(j)|x_q(j-1) = x_q^n(j-1)\}$, which is the probability that $x_q^n(j-1)$ is mapped to $x_q^n(j)$ by the finite state model of Eq. (6) with possible values of $w_d(j-1)$. Since the transition from $x_q^n(j-1)$ to $x_q^n(j)$ is determined by possible values of $w_d(j-1)$, this transition probability is the sum of occurrence probabilities of all possible values of $w_d(j-1)$ which map $x_q^n(j-1)$ to $x_q^n(j)$.
The estimation problem is to find the estimator path, which is the path having the biggest metric through the trellis diagram. This is accomplished by the Viterbi Algorithm (Demirbaş, 1982; 1984; 1989; Forney, 1973); which systematically searches all paths through the trellis diagram. The number of quantization levels of the finite state model, in general, increases exponentially with time k. As a result, the implementation complexity of this approach increases exponentially with time k (Demirbaş, 1982; 1984; Demirbaş & Leondes, 1985; Demirbaş, 2007). In order to overcome this obstacle, a block-by-block suboptimum estimation scheme was proposed in (Demirbaş, 1982; 1984; Demirbaş & Leondes, 1986; Demirbaş, 1988; 1989; 1990). In this estimation scheme: observation sequence was divided into blocks of constant length. Each block was initialized by the final state estimate from the last block. The initialization of each block with only a single quantization level (node), that is, the reduction of the trellis diagram to one node at the end of each block, results in state estimate divergence for long observation sequences, i.e., large time k, even though the implementation complexity of the proposed scheme does not increase with time (Kee & Irwin, 1994). The online and recursive state estimation scheme which is recently proposed in (Demirbaş, 2010) prevents state estimate divergence caused by one state initialization of each block for the block-by-block estimation. This recently proposed estimation scheme, referred to as the DF throughout this chapter, first prunes all paths going through the nodes which do not satisfy constraints imposed on estimates and then assigns a metric to each node (or quantization level) in the trellis diagram. Furthermore, at each time (step, or iteration), the number of considered state quantization levels (nodes) is limited by a selected positive integer MN, which stands for the maximum number of quantization levels considered through the trellis diagram; in other words , MN nodes having the biggest metrics are kept through the trellis diagram and all the paths going through the other nodes are pruned. Hence, the implementation complexity of the DF does not increase with time. The number MN is one of the parameters determining the implementation complexity and the performance of the DF.

5. Online state estimation

This section first yields some definitions, and then presents the DF.

5.1 Definitions

Admissible initial state quantization level : a possible value $x_{qi}(0) \triangleq x_{di}(0)$ of an approximate discrete random vector $x_q(0) \triangleq x_d(0)$ of the initial state vector $x(0)$ is said to be an admissible quantization level of the initial state vector (or an admissible initial state quantization level) if this possible value satisfies the constraints imposed on the state estimates. Obviously, if there do not exist any constraints imposed on the state estimates, then all possible values of the approximate discrete random vector $x_q(0)$ are admissible.

Metric of an admissible initial state quantization level: the natural logarithm of the occurrence probability of an admissible initial quantization level $x_{qi}(0)$ is referred to as the metric of this admissible initial quantization level. This metric is denoted by $M(x_{qi}(0))$, that is

$$M(x_{qi}(0)) \triangleq \ln\{Prob\{x_q(0) = x_{qi}(0)\}\}. \tag{13}$$

where $Prob\{x_q(0) = x_{qi}(0)\}$ is the occurrence probability of $x_{qi}(0)$.

Admissible state quantization level at time k: a quantization level $x_{qi}(k)$ of a state vector $x(k)$, where $k \geq 1$, is called an admissible quantization level of the state (or an admissible state quantization level) at time k if this quantization level satisfies the constraints imposed on the state estimates. Surely, if there do not exist any constraints imposed on the state estimates, then all the quantization levels of the state vector $x(k)$, which are calculated by Eq. (6), are admissible.

Maximum number of considered state quantization levels at each time: MN stands for the maximum number of admissible state quantization levels which are considered at each time (step or iteration) of the DF. MN is a preselected positive integer. A bigger value of MN yields better performance, but increases implementation complexity of the DF.

Metric of an admissible quantization level (or node) at time k**, where** $k \geq 1$: the metric of an admissible quantization level $x_{qj}(k)$, denoted by $M(x_{qj}(k))$, is defined by

$$\begin{aligned} M(x_{qj}(k)) \triangleq & \max_n \{M(x_{qn}(k-1)) + ln[T(x_{qn}(k-1) \to x_{qj}(k))]\} \\ & + ln[\lambda(z(k)|x_{qj}(k))], \end{aligned} \tag{14}$$

where the maximization is taken over all considered state quantization levels at time $k-1$ which are mapped to the quantization level $x_{qj}(k)$ by the possible values of $w_d(k-1)$; ln stands for the natural logarithm; $T(x_{qn}(k-1) \to x_{qj}(k))$ is the transition probability from $x_{qi}(k-1)$ to $x_{qj}(k)$ is given by

$$T(x_{qi}(k-1) \to x_{qj}(k)) = \sum_n Prob\{w_d(k-1) = w_{dn}(k-1)\}, \tag{15}$$

where $Prob\{w_d(k-1) = w_{dn}(k-1)\}$ is the occurrence probability of $w_{dn}(k-1)$ and the summation is taken over all possible values of $w_d(k-1)$ which maps $x_{qi}(k-1)$ to $x_{qj}(k)$; in

other words, the summation is taken over all possible values of $w_d(k-1)$ such that

$$Q(f(k-1, x_{qi}(k-1), w_{dn}(k-1))) = x_{qj}(k); \tag{16}$$

and

$$\lambda(z(k)|x_{qj}(k)) \triangleq \begin{cases} 1 & \text{if z(j) is neither available nor used for estimation} \\ p(z(k)|x_{qj}(k)) & \text{if z(j) is available and used for estimation,} \end{cases} \tag{17}$$

in which, $p(z(k)|x_{qj}(k))$ is the conditional density function of $z(k)$ given that the actual value of state $x(k) = x_{qj}(k)$, and this density function is calculated by using the observation model of Eq. (2).

5.2 Estimation scheme (DF)

A flowchart of the DF is given in Fig. 3 for given $F_{w(k)}(a)$, $F_{x(0)}(a)$, MN, n, m , and GS; where $F_{w(k)}(a)$ and $F_{x(0)}(a)$ are the distribution functions of $w(k)$ and $x(0)$ respectively, n and m are the numbers of possible values of approximate random vectors of $w(k)$ and $x(0)$ respectively; GS is the gate size; and $z(k)$ is the observation at time k. The parameters MN, n, m , and GS determine the implementation complexity and performance of the DF. The number of possible values of the approximate disturbance noise $w_d(k)$ is assumed to be the same, n , for all iterations, i.e., for all k. The filtered value $\hat{x}(k|k)$ and predicted value $\hat{x}(k|k-1)$ of the state $x(k)$ are recursively determined by considering only MN admissible state quantization levels with the biggest metrics and discarding other quantization levels at each recursive step (each iteration or time) of the DF. Recursive steps of the DF is described below.

Initial Step (Step 0): an approximate discrete random vector $x_d(0)$ with m possible values of the initial state $x(0)$ is offline calculated by Eq. (3). The possible values of this approximate random vector are defined as the initial state quantization levels (nodes). These initial state quantization levels are denoted by $x_{q1}(0)$, $x_{q2}(0)$, ..., and $x_{qm}(0)$, where $x_{qi(0)} \triangleq x_{di(0)}$ $(i = 1\ 2\ ...m)$. Admissible initial state quantization levels, which satisfy the constraints imposed on state estimates, are determined and the other initial quantization levels are discarded. If the number of admissible initial quantization levels is zero, then the number, m, of possible values of the approximate initial random vector $x_d(0)$ is increased and *the initial step* of the DF is repeated from the beginning; otherwise, the metrics of admissible initial quantization levels are calculated by Eq. (13). The admissible initial state quantization levels (represented by $x_{q1}(0)$, $x_{q2}(0)$, ..., and $x_{qN_0}(0)$) and their metrics are considered in order to calculate state quantization levels and their metrics at time $k = 1$. These considered quantization levels are denoted by nodes (at time 0) on the first row (or column) of two rows (or columns) trellis diagram at the first step $k = 1$ of the DF, Fig. 2.

State estimate at time 0*:* if the mean value of $x(0)$ satisfies constraints imposed on state estimates such as the case that there do not exist any estimate constraints , then this mean value is taken as both $\hat{x}(0|0)$ and $\hat{x}(0|0-1)$; otherwise, the admissible initial state quantization level (node) with the biggest metric is taken as both the filtered value $\hat{x}(0|0)$ and predicted value $\hat{x}(0|0-1)$ of the state $x(0)$, given no observation.

Recursive Step (Step k): An approximate discrete disturbance noise vector $w_d(k-1)$ with n possible values of the disturbance noise $w(k-1)$ is offline obtained by Eq. (3). The quantization levels of the state vector at time k are calculated by using the finite state model of Eq. (6) with all the considered quantization levels (or nodes) $x_{q1}(k-1)$, $x_{q2}(k-1)$ ... $x_{qN_{k-1}}(k-1)$ at time $k-1$; and all possible values $w_{d1}(k-1)$, $w_{d2}(k-1)$, ..., $w_{dn}(k-1)$ of the approximate discrete disturbance noise vector $w_d(k-1)$. That is, substituting the

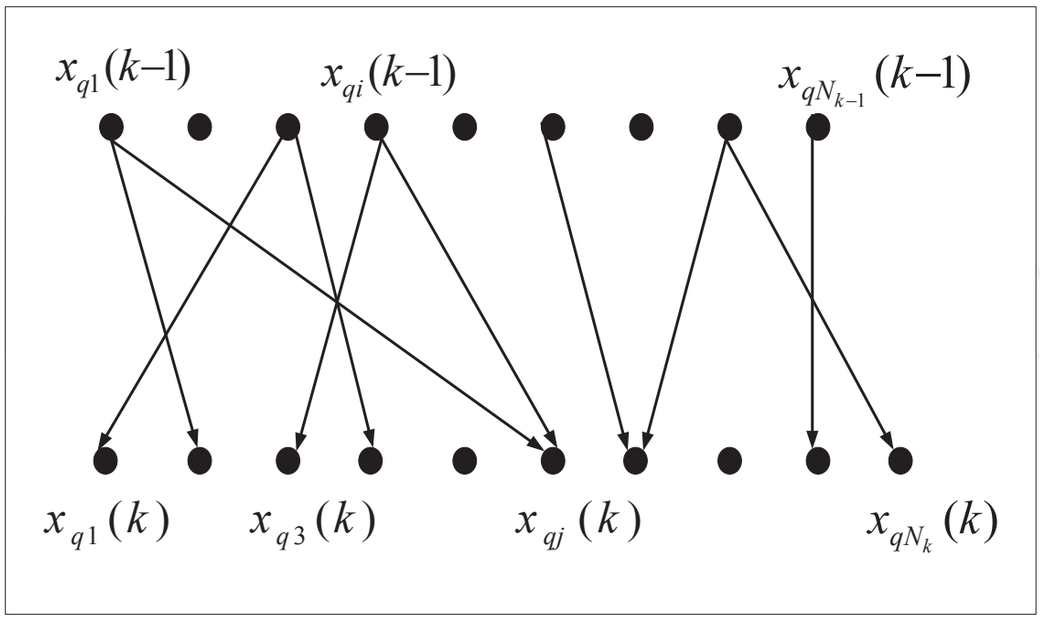


Fig. 2. Two Row Trellis Diagram of Admissible State Quantization Levels

considered state quantization levels $x_{qi}(k-1)$ $(i = 1, 2, \ldots, N_{k-1})$ for $x_q(k-1)$ and the possible values $w_d(k-1) = w_{dj}(k-1)$ $(j = 1, 2, \ldots, n)$ for $w_d(k-1)$ in the finite state model of Eq. (6), the quantization levels of the state at time k are calculated (generated). The admissible quantization levels at time k, which satisfy constraints imposed on state estimates, are determined and non-admissible state quantization levels are discarded. If the number of admissible state quantization levels at time k is zero, then a larger n, MN or smaller GS is taken and *the recursive step at time k* of the DF is repeated from the beginning; otherwise, the metrics of all admissible state quantization levels at time k are calculated by using Eq. (14). If the number of admissible state quantization levels at time k is greater than MN, then only MN admissible state quantization levels with biggest metrics, otherwise, all admissible state quantization levels with their metrics are considered for the next step of the DF. The considered admissible quantization levels (denoted by $x_{q1}(k)$, $x_{q2}(k)$, ...,$x_{qN_k}(k)$) and their metrics are used to calculate the state quantization levels and their metrics at time $k+1$. The considered state quantization levels at time k are represented by the nodes on the second row (or column) of two rows (or columns) trellis diagram at *the recursive step k* and on the first row (or column) of two rows (or columns) trellis diagram at *the recursive step* $k+1$, Fig. 2; where the subscript N_k, which is the number of considered nodes at the end of *Recursive step k*, is less than or equal to MN; and the transition from a node at time $k-1$, say $x_{qi}(k-1)$, to a node at time k , say $x_{qj}(k)$, is represented by a directed line which is called a branch. *Estimate at time k:* the admissible quantization level (node) with the biggest metric at time k is taken as the desired estimate of the state at time k, that is, the node with the biggest metric at time k is the desired predicted value of $x(k)$ if $z(k)$ is neither available nor used for estimation; otherwise, the node at time k with the biggest metric is the filtered value of $x(k)$. If there exist more than one nodes having the same biggest metric, anyone of these nodes can be taken as the desired estimate.

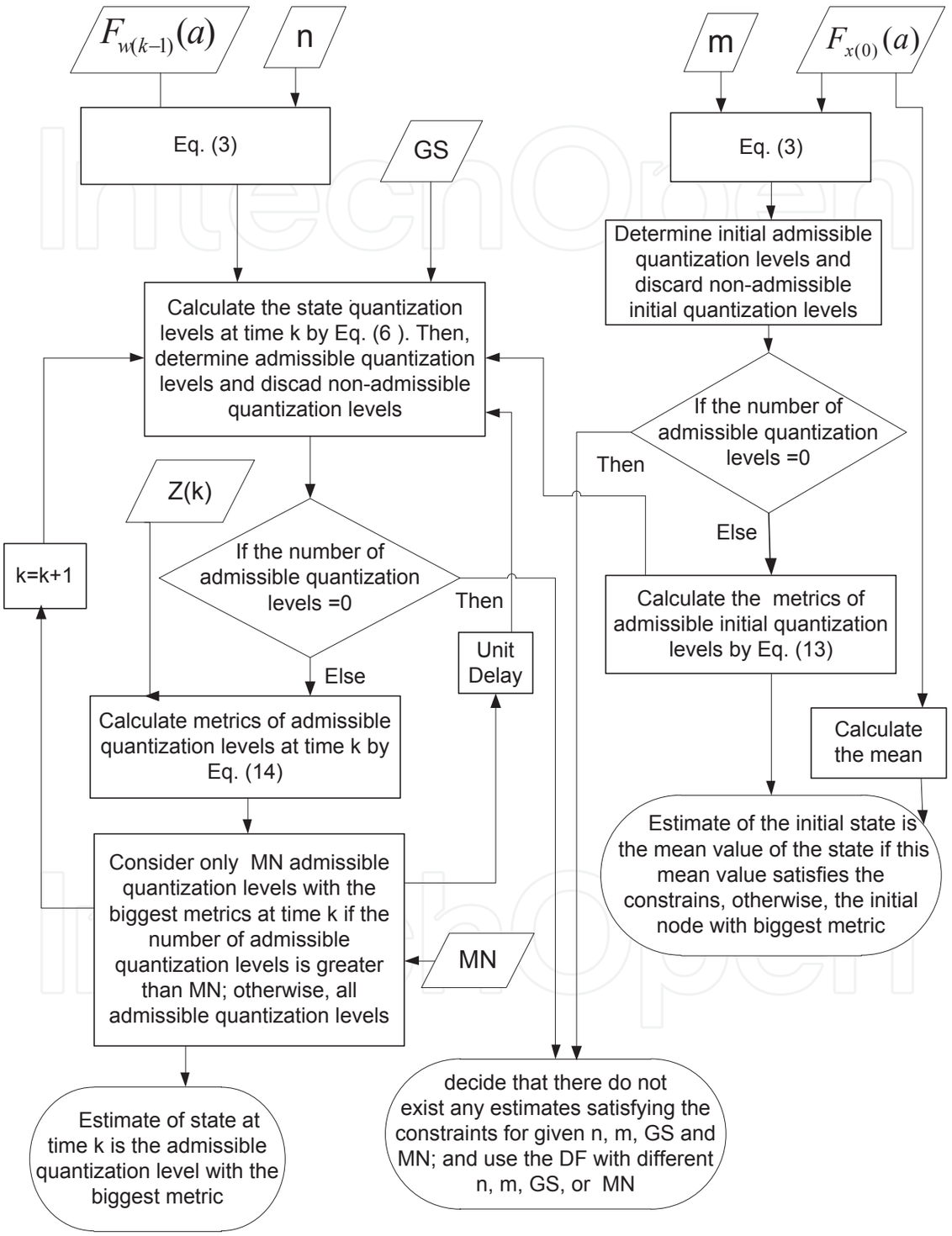


Fig. 3. Flowchart of the DF

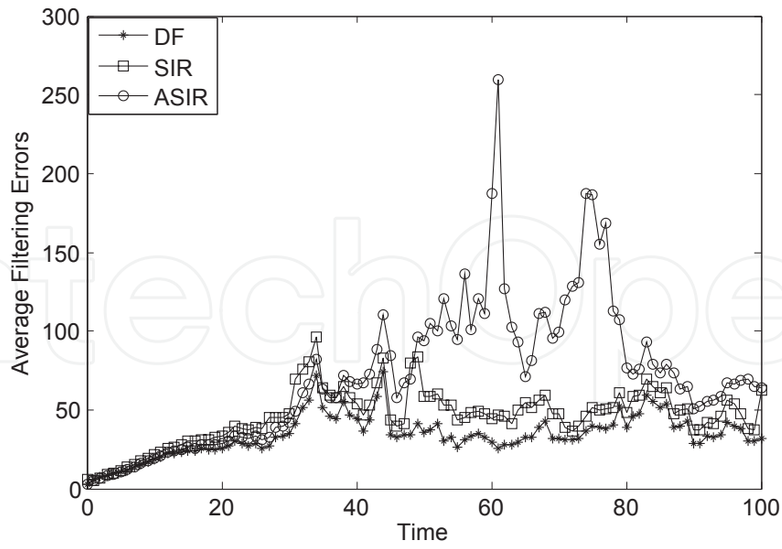


Fig. 4. Average Filtering Errors for Eqs. (18) and (19)

6. Simulations

In this section, Monte Carlo simulation results of two examples are given. More examples are presented in (Demirbaş, 2010). The first example is given by

State Model

$$x(k+1) = x(k)[1 + \frac{k}{k+1}\cos(0.8x(k) + 2w(k))] + w(k) \tag{18}$$

Observation Model

$$z(k) = \frac{6x(k)}{1 + x^2(k)} + v(k), \tag{19}$$

where the random variables $x(0)$, $w(k)$, and $v(k)$ are independent Gaussian random variables with means 6, 0, 0 and variances 13, 20, 15 respectively. It was assumed that there did not exist any constraints imposed on state estimates. The state model of Eq. (18) is an highly nonlinear function of the disturbance noise $w(k)$. The extended Kalman filter (EKF) and the grid-based approaches may not be used for the state estimation of this example, since the EKF assumes a linear disturbance noise in the state model and the grid based approaches assumes the availability of the state transition density $p(x(k)|x(k-1))$ which may not readily calculated (Arulampalam et al., 2002; Ristic et al., 2004). States of this example were estimated by using the DF, the sampling importance resampling (SIR) particle filter (which is sometimes called the bootstrap filter, and auxiliary sampling importance resampling (ASIR) particle filter (Arulampalam et al., 2002; Gordon et al., 1993). Average absolute filtering and prediction errors are sketched in Figs. 4 and 5 for 2000 runs each of which consists of 100 iterations. These estimation errors were obtained by using the SIR and ASIR particle filters with 1000 particles and the DF for which the random variables $x(0)$ and $w(k)$ were approximated by the approximate random variables with 3 possible values (which are given in Section 3); the gate size (GS) and MN were taken as 0.1 and 8 respectively. The average filtering and prediction errors per one estimation (one iteration) were 33.8445, 45.6377, 71.5145 and 34.0660, 45.4395,

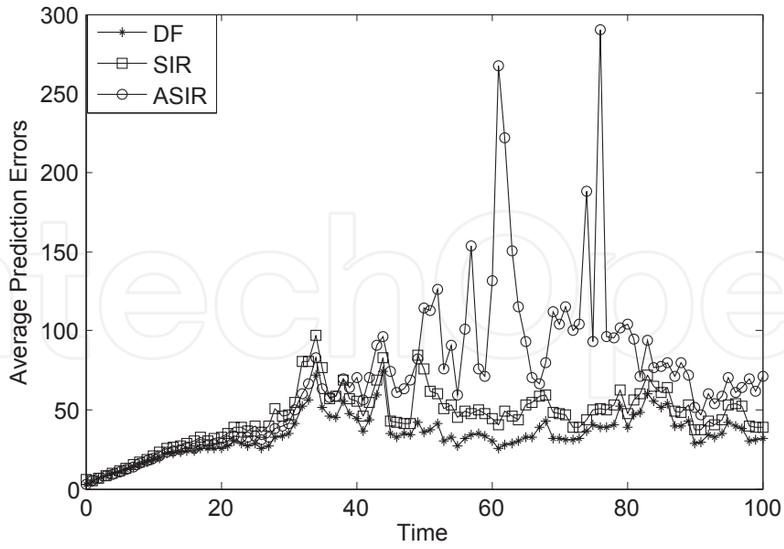


Fig. 5. Average Prediction Errors for Eqs. (18) and (19)

70.2305 respectively. A typical run with 100 iteration took 0.0818, 0.2753, 0.3936 seconds for the DF, SIR and ASIR particle filters, respectively. The DF clearly performs better than both the SIR and ASIR particle filter. Moreover, the DF is much faster than both the SIR and ASIR particle filters with 1000 particles.
The second example is described by

State Model

$$x(k+1) = x(k)[1 + \frac{k}{k+1}\cos(0.8x(k))] + w(k) \tag{20}$$

Observation Model

$$z(k) = \frac{6x(k)}{1+x^2(k)} + v(k), \tag{21}$$

where the random variables $x(0)$, $w(k)$,and $v(k)$ are independent Gaussian random variables with means 3, 0, 0 and variances 8, 9, 9 respectively. It was assumed that there did not exist any constraints imposed on state estimates. Average absolute filtering and prediction errors are sketched in Figs. 6 and 7 for 2000 runs each of which consists of 200 iterations. These estimation errors were obtained by using the SIR and ASIR particle filters with 1000 particles and the DF for which the random variables $x(0)$ and $w(k)$ were approximated by the approximate random variables with 3 possible values (which are given in Section 3); the gate size (GS) and MN were taken as 0.1 and 4 respectively. The average filtering and prediction errors per one estimation (one iteration) were 38.4913, 61.5432, 48.4791 and 38.5817, 61.4818, 48.5088 respectively. A typical run with 200 iteration took 0.0939, 0.5562, 0.8317 seconds for the DF, SIR and ASIR particle filters, respectively. The state model of the second example is a linear function of the disturbance noise. Hence, the extended Kalman filter (EKF) was also used for state estimation, but the EKF estimation errors quickly diverged, hence, the EKF state estimation errors are not sketched. The DF clearly performs better than the EKF, SIR and ASIR particle filters and also the DF is much faster than both the SIR and ASIR particle filters with 1000 particles for the second example.

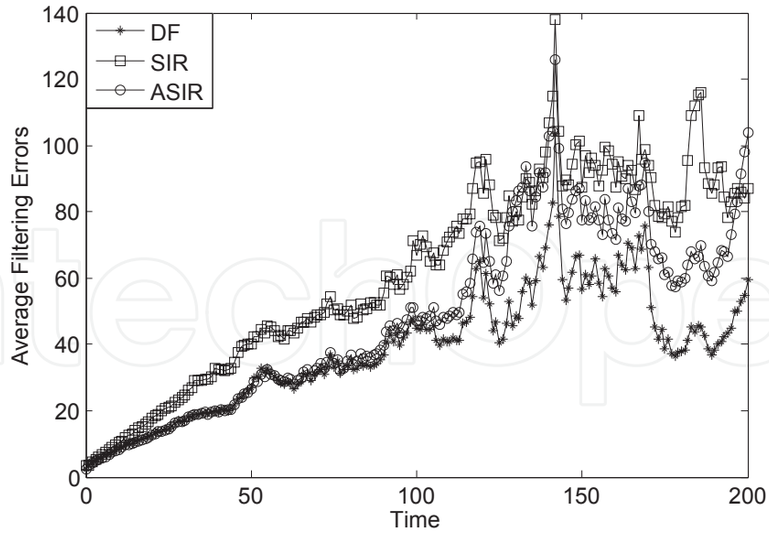


Fig. 6. Average Filtering Errors for Eqs. (20) and (21)

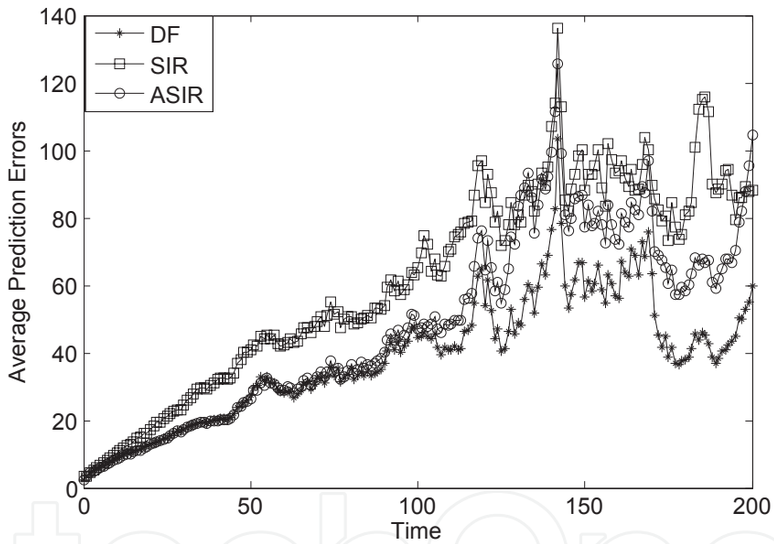


Fig. 7. Average Prediction Errors for Eqs. (20) and (21)

The performance of the DF is determined by the possible values (n and m) of the approximate discrete random disturbance noise and approximate discrete initial state, gate size (GS), maximum number (MN) of considered state quantization levels at each iteration. As GS goes to zero and the parameters n, m, and MN approach infinity, the approximate models of Eq. (6) and (7) approach the models of Eqs. (1) and (2), hence, the DF approaches an optimum estimation scheme, but the implementation complexity of the DF exponentially increases with time k. The parameters n, m, GS, MN which yield the best performance for given models are determined by Monte Carlo simulations for available hardware speed and storage. For given nonlinear models: the performances of the DF, EKF, particle filters, and others must be compared by Monte Carlo simulations with available hardware speed and storage. The estimation scheme yielding the best performance should be used. The EKF is surely much

faster than both the DF and particle filters. The speed of the DF is based upon the parameters n, m, GS, MN; whereas the speeds of particle filters depend upon the number of particles used.

7. Conclutions

Presented is a real-time (online) recusive state filtering and prediction scheme for nonlinear discrete dynamic systems with Gaussian or non-Gaussian disturbance and observation noises. This scheme, referred to as the DF, is recently proposed in (Demirbaş, 2010). The DF is very suitable for state estimation of nonlinear dynamic systems under either missing observations or constraints imposed on state estimates. The DF is much more general than grid based estimation approaches. This is based upon discrete noise approximation, state quantization, and a suboptimum implementation of multiple hypothesis testing , whereas particle filters are based upon sequential Monte Carlo Methods. The models of the DF is as general as the models of particle filters, whereas the models of the extended Kalman filter (EKF) are linear functions of the disturbance and observation noises. The DF uses state models only to calculate transition probabilities from gates to gates. Hence, if these transition probabilities are known or can be estimated, state models are not needed for estimation with the DF, whereas state models are needed for both the EKF and particle filters. The performance and implementation complexity of the DF depend upon the gate size (GS), numbers n and m of possible values of approximate discrete disturbance noise and approximate discrete initial state, and maximum number (MN) of considered quantization levels at each iteration of the DF; whereas the performances and implementation complexities of particle filters depend upon numbers of particles used. The implementation complexity of the DF increases with a smaller value of GS, bigger values of n, m, and MN. These yield more accurate approximations of state and observation models; whereas the implementation complexities of particle filters increase with larger numbers of particles, which yield better approximations of conditional densities. Surely, the EKF is the simplest one to implement. The parameters (GS, n,m, MN) for which the DF yields the best performance for a real-time problem should be determined by Monte Carlo simulations. As $GS \to 0$, $n \to \infty$, $m \to \infty$,and $MN \to \infty$; the DF approaches the optimum one in the average overall error probability sense, but its implementation complexity exponentially increases with time. The performances of the DF, particle filters, EKF are all model-dependent. Hence, for a real-time problem with available hardware speed and storage; the DF, particle filters, and EKF (if applicable) should all be tested by Monte Carlo simulations, and the one which yields the best results should be used. The implementation complexity of the DF increases with the dimensions of multidimensional systems, as in the particle filters.

8. References

Arulampalam, M.S.; Maskell, S.; Gordon, N.; and Clapp, T. (2002), A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, Vol.50, pp. 174-188.

Daum, F.E. (2005). Nonlinear Filters: Beyond the Kalman Filter, *IEEE A&E Systems Magazine*, Vol. 20, No. 8, Part 2, pp. 57-69

Demirbaş K. (1982), New Smoothing Algorithms for Dynamic Systems with or without Interference, *The NATO AGARDograph Advances in the Techniques and Technology of Applications of Nonlinear Filters and Kalman Filters*, C.T. Leonde, (Ed.), AGARD, No. 256, pp. 19-1/66

Demirbaş, K. (1984). Information Theoretic Smoothing Algorithms for Dynamic Systems with or without Interference, *Advances in Control and Dynamic Systems,* C.T. Leonde, (Ed.), Volume XXI, pp. 175-295, Academic Press, New York

Demirbaş, K.; Leondes, C.T. (1985), Optimum decoding based smoothing algorithm for dynamic systems, *The International Journal of Systems Science,* Vol.16, No. 8, pp. 951-966

Demirbaş, K.; Leondes, C.T. (1986). A Suboptimum decoding based smoothing algorithm for dynamic Systems with or without interference, *The International Journal of Systems Science,* Vol.17, No.3, pp. 479-497.

Demirbaş, K. (1988). Maneuvering target tracking with hypothesis testing, *I.E.E.E. Transactions on Aerospace and Electronic Systems,* Vol.23, No.6, pp. 757-766.

Demirbaş, K. (1989). Manoeuvring-target Tracking with the Viterbi Algorithm in the Presence of Interference, *the IEE Proceedings-PartF, Communication, Radar and Signal Processing*, Vol. 136, No. 6, pp. 262-268

Demirbaş, K. (1990), Nonlinear State Smoothing and Filtering in Blocks for Dynamic Systems with Missing Observation, *The International Journal of Systems Science,* Vol. 21, No. 6, pp. 1135-1144

Demirbaş, K. (2007), A state prediction scheme for discrete time nonlinear dynamic systems, *the international journal of general systems,* Vol. 36, No. 5, pp. 501-511

Demirbaş, K. (2010). A new real-time suboptimum filtering and prediction scheme for general nonlinear discrete dynamic systems with Gaussian or non-Gaussian noise, to appear in *the international journal of Systems Science,* DOI: 10.1080/00207721003653682, first published in *www.informaworld.com* on 08 September 2010.

Doucet, A.; De Freitas, J.F.G.; and Gordon, N.J. (2001), An introduction to sequential Monte Carlo methods, *in Sequential Monte Carlo Methods in Practice,* A. Doucet, J.F.G de Freitas, and N.J.Gordon (Eds.), Springer-Verlag, New York

Forney, G.D. (1973). The Viterbi algorihm, *Proceedings of the IEEE*, Vol. 61, pp. 268-278

Gordon, N.; Salmond, D.; Smith, A.F.M. (1993). Novel approach to nonlinear and non-Gaussian Bayesian state estimation. *Proceedings of the Institute of Electrical Engineering F,* Vol.140, pp. 107-113

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems, *Trans. ASME, Journal of Basic Engineering,* Vol. 82, pp. 35-45.

Kalman, R.E.; Busy R.S. (1960). New Results in Linear Filtering and Prediction Theory, *Trans. ASME, Journal of Basic Engineering,* Series D, Vol. 83, pp. 95-108

Kee, R.J.; Irwin, G.W. (1994). Investigation of trellis based filters for tracking, *IEE Proceedings Radar, Sonar and Navigation,* Vol. 141, No.1 pp. 9-18.

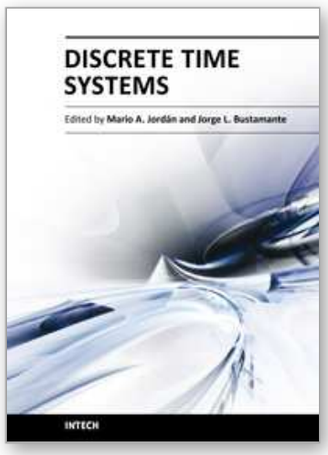
Julier S. J.; Uhlmann J. K. (2004). Unscented Filtering and Nonlinear Estimation *Proceedings of the IEEE,* **92**, pp. 401-422

Ristic B., Arulampalam S.; Gordon N. (2004). *Beyond the Kalman Filter: Particle Filters for Tracking Applications* , Artech House, London

Sage, A.P.; Melsa, J.L. (1971). *Estimation Theory with Applications to Communications and Control,* McGraw-Hill, New York

Van Trees, H. L. (2001). *Detection, Estimation and Modulation: Part I. Detection, Estimation, and Linesr Modulation Theory,* Jhon Wiley and Sons, Inc., New York, ISBN 0-471-09517-6

Weber, C. L. (1968), *Elements of Detection and Signal Design,* McGraw-Hill, New York



Discrete Time Systems

Edited by Dr. Mario Alberto JordÃ¡n

ISBN 978-953-307-200-5
Hard cover, 526 pages
Publisher InTech
Published online 26, April, 2011
Published in print edition April, 2011

Discrete-Time Systems comprehend an important and broad research field. The consolidation of digital-based computational means in the present, pushes a technological tool into the field with a tremendous impact in areas like Control, Signal Processing, Communications, System Modelling and related Applications. This book attempts to give a scope in the wide area of Discrete-Time Systems. Their contents are grouped conveniently in sections according to significant areas, namely Filtering, Fixed and Adaptive Control Systems, Stability Problems and Miscellaneous Applications. We think that the contribution of the book enlarges the field of the Discrete-Time Systems with signification in the present state-of-the-art. Despite the vertiginous advance in the field, we also believe that the topics described here allow us also to look through some main tendencies in the next years in the research area.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen