Target Tracking With Glint Noise

WENG-RONG WU
MIST Research Center

In tracking targets, there can be an uncertainty associated with the measurements in addition to their inaccuracy, which is usually modeled by additive Gaussian noise. However, the Gaussian modeling of the noise may not be true. Noise can be non-Gaussian. The non-Gaussian noise arising in a radar system is known as glint noise. The distribution of glint noise is long tailed and will seriously affect the tracking performance. A new algorithm is developed here which can significantly improve the tracking performance when the glint noise is present.

Manuscript received June 21, 1991; revised February 6, 1992.

IEEE Log No. T-AES/29/1/03777.

Author's address: MIST Research Center, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

I. INTRODUCTION

The Kalman filter is widely used in the tracking problem. It can optimally estimate the target motion from noisy radar data. The optimality of the Kalman filter is based on the assumption of the Gaussian noise. If the assumption is violated, the Kalman filter is no longer the optimal filter.

In a radar system, due to the target glint, the measurement noise may present non-Gaussian behavior. This is referred to as the glint noise. A typical glint noise record is shown in Fig. 1 [6]. From the figure, we can clearly see the non-Gaussian nature of the noise. The spiky character manifests itself in the long-tailed distribution. It is well known that the least square estimate can be seriously degraded when the observation noise is non-Gaussian [26]. Unfortunately, very few results has been reported regarding this problem and the standard Kalman filter is continuously used in tracking applications. Hewer, Martin, and Zeh [6] treat the glint noise as the mixture of a Gaussian noise and outliers. Consequently, they employ robust estimation techniques to preprocess (clean) the radar data. Here, we use a different approach. We assume that the glint noise can be modeled by some non-Gaussian distribution. We then seek an optimal filter for the non-Gaussian noise.

There have been a number of researchers who consider the problem of the Kalman filtering in non-Gaussian environment [8–18]. One of the most effective schemes is proposed by Masreliez [11,12]. He introduced a nonlinear score function as the correction term in the state estimate and the results are often nearly optimal. While this approach seems promising, it encounters the difficulty of implementing the convolution operation involved in the evaluation of the score function. This precludes the practical applications of the method.

The score function implementation problem is recently solved by Wu and Kundu [19]. The method employs an adaptive normal expansion to expand the score function and truncates the higher order terms in the expanded series. Consequently, the score function can be approximated by a few central moments of the observation prediction density. The normal expansion is made adaptive by using the concept of conjugate recentering and the saddle point method. It is shown that the approximation is satisfactory and the method is simple and practically feasible. We employ this method in the tracking problem.

The approach developed in [19] is easy to implement for the scalar observation. However, the radar observation is often not scalar. We can tackle this problem by using a spherical modeling of the target [1, 2]. The state is formulated as the range, the elevation and the bearing angle of the target. After the model linearization, the linear state equations are obtained. Since the state now is defined in the

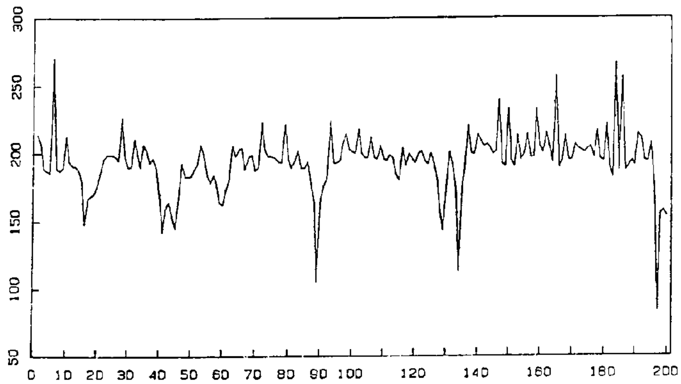

Fig. 1. Glint noise record.

spherical coordinate, the observation are decoupled into three sets of independent components. By doing so, the scalar score function approximation scheme can be applied. Spherical modeling can be made more attractive by some simplification. In this case, the state equations are further decoupled into three channels. Thus, three Kalman filters of smaller dimension can be run in parallel and the computational complexity is greatly reduced. However, since the target model is linearized in spherical coordinate, dynamic error will occur [5]. We use a simple scheme to solve the problem. Simulations show that by using this scheme, the tracking performance is almost as good as that in the rectangular modeling.

The organization of the paper is as follows. In Section II, the Masreliez's approach is briefly reviewed. The implementation of the score function is described in Section III. Tracking with glint noise is then considered in Section IV using the algorithms described in Sections II and III. In Section V, we present some simulation results and the conclusion is drawn in Section VI.

II. SCORE FUNCTION APPROACH

A. General Filtering Problem

The general filter problem can be formulated as the estimation of the state given all the past history of the observation. Consider a linear system described as follows:

$$x_{k+1} = \phi_k x_k + w_k \tag{1}$$

$$z_k = H_k x_k + v_k \tag{2}$$

where x_k is the state vector, w_k and v_k represent white noise sequences and are assumed to be mutually independent. The basic problem is to estimate the state x_k from the noisy observation $(z_1,\ldots,z_k)$. The probability density of the state conditioned on all the available observation data is called the *a posteriori* density. If this density is known, an estimation for any type of performance criterion can be found. Thus, the estimation problem can be viewed as the problem of determining the *a posteriori* density. In addition, one is frequently interested in performing the filtering recursively. The recursive determination of the *a posteriori* density is generally referred to as the Bayesian approach. Denote $f(\cdot)$ as a density and $Z^k = \{z_0, z_1, \ldots, z_k\}$. The Bayesian approach is described by the following relations [8].

$$f(x_k \mid Z^k) = \frac{f(x_k \mid Z^{k-1}) f(z_k \mid x_k)}{f(z_k \mid Z^{k-1})} \tag{3}$$

$$f(x_k \mid Z^{k-1}) = \int f(x_{k-1} \mid Z^{k-1}) f(x_k \mid x_{k-1})\, dx_{k-1} \tag{4}$$

where the normalizing constant $f(z_k \mid Z^{k-1})$ is given by

$$f(z_k \mid Z^{k-1}) = \int f(x_k \mid Z^{k-1}) f(z_k \mid x_k)\, dx_k. \tag{5}$$

The $f(z_k \mid x_k)$ in (3) is determined by the observation noise density $f(v_k)$ and (2). Similarly, $f(x_k \mid x_{k-1})$ in (4) is determined by the state noise density $f(w_k)$ and (1). Theoretically, knowing these densities, we can determine the *a posteriori* density $f(x_k \mid Z^k)$. However, it is generally impossible to carry out the integration in (4) for every instant. Consequently, the *a posteriori* density cannot be determined for most applications. There is only one exception, i.e., when the initial state and all the noise sequences are Gaussian. In this case, (3) and (4) are reduced to the standard Kalman filter equations, namely

$$\hat{x}_k = \overline{x}_k + M_k H_k^t (H_k M_k H_k^t + R_k)^{-1} (z_k - H_k \overline{x}_k) \tag{6}$$

$$P_k = M_k - M_k H_k^t (H_k M_k H_k^t + R_k)^{-1} H_k M_k \tag{7}$$

$$\overline{x}_{k+1} = \phi_k \hat{x}_k \tag{8}$$

$$M_{k+1} = \phi_k P_k \phi_k^t + Q_k \tag{9}$$

where $f(x_k \mid Z^{k-1}) \sim N(\overline{x}_k, M_k)$, $f(x_k \mid Z^k) \sim N(\hat{x}_k, P_k)$, $E\{w_k w_j\} = Q_k \delta_{kj}$ and $E\{v_k v_j\} = R_k \delta_{kj}$.

B. Score Function Approach

In this section we briefly review the Masreliez's algorithm. Consider a linear system described in (1) and (2). The variables w_k and v_k can be non-Gaussian. The density of z_k conditioned on the *prior* observations is denoted by $f(z_k \mid Z^{k-1})$. We name $f(z_k \mid Z^{k-1})$ the *observation prediction density* and assume that it is twice differentiable. Similarly, $f(x_k \mid Z^{k-1})$ is the density of x_k conditioned on *prior* observations and is named the *state prediction density*. The filtering problem is to estimate the state vector x_k from the noisy observations Z^k. Assuming that $f(x_k \mid Z^{k-1})$ is a Gaussian density with mean $\overline{x}_k$, and covariance matrix M_k, Masreliez has shown that the minimum variance state estimation $\hat{x}_k$, and its covariance matrix $P_k = E\{(x_k - \hat{x}_k)(x_k - \hat{x}_k)^t \mid Z^k\}$ can be recursively calculated as follows [12]:

$$\hat{x}_k = \overline{x}_k + M_k H_k^t g_k(z_k) \quad (10)$$

$$P_k = M_k - M_k H_k^t G_k(z_k) H_k M_k \quad (11)$$

$$\overline{x}_{k+1} = \phi_k \hat{x}_k \quad (12)$$

$$M_{k+1} = \phi_k P_k \phi_k^t + Q_k \quad (13)$$

where $g_k(\cdot)$ is a column vector with components:

$$\{g_k(z_k)\}_i = -\left[\frac{\partial f(z_k \mid Z^{k-1})}{\partial (z_k)_i}\right][f(z_k \mid Z^{k-1})]^{-1} \quad (14)$$

and $G_k(z_k)$ is a matrix with elements

$$\{G_k(z_k)\}_{ij} = \frac{\partial \{g_k(z_k)\}_i}{\partial (z_k)_j}. \quad (15)$$

The function $g_k(\cdot)$ is called the score function of $f(z_k \mid Z^{k-1})$. It is $g_k(\cdot)$ that suggests how to modify the Kalman filter in the non-Gaussian noise. Assuming that w_k is Gaussian and v_k is non-Gaussian, we can see that the score function $g(\cdot)$ operating on the residual $z_k - H_k\overline{x}_k$ will deemphasize the influence of large residuals when the observation prediction density is long tailed, and, on the other hand, emphasize the large residuals when the observation density is short tailed. This is intuitively appealing. It is easy to check that the filter is reduced to the standard Kalman filter if the initial state, w_k, and v_k (for all ks) are Gaussian.

The following procedure summarizes the implementation of the filter.

Step 0: Assume that at stage $k-1$, $\hat{x}_{k-1}$ and P_{k-1} are known.

Step 1: Calculate $M_k = \phi_{k-1} P_{k-1} \phi_{k-1}^t + Q_{k-1}$.

Step 2: Approximate the state prediction density $f(x_k \mid Z^{k-1})$ by a Gaussian distribution with mean $\overline{x}_k = \phi_{k-1}\hat{x}_{k-1}$ and covariance matrix M_k.

Step 3: Find the observation prediction density $f(z_k \mid Z^{k-1})$ by convolving $f(H_k x_k \mid Z^{k-1})$ with $f_{v_k}(\cdot)$.

Step 4: Find $g_k(z_k)$ and $G_k(z_k)$.

Step 5: Apply (10) and (11) to find $\hat{x}_k$ and P_k.

Step 6: Let $k \to k+1$ and start all over from step 1.

The procedure outlined above is straightforward in principle. However, the convolution operation in step 3 is difficult to implement in general except for very simple cases. Also, in step 4, the differentiation operations involved in the evaluation of the score function and its derivative are not trivial.

III. EVALUATION OF SCORE FUNCTION

In this section, we describe an adaptive method to evaluate the score function using the concept of normal expansion, conjugate density and the saddle point method. We assume that all the distribution are univariate. But, the idea developed here can be extended to multivariate ones also.

A. Adaptive Normal Approximation of Distribution

Assume that $f(x)$ is a continuous density funtion. The polynomial $p_n(x)$ is called the *orthogonal polynomial* associated with the distribution $f(x)$ if $p_n(x)$ is a polynomial of degree n, the coefficient of x^n is positive and $p_n(x)$ satisfies the orthogonality conditions

$$\int_{-\infty}^{+\infty} p_m(x)p_n(x)f(x)\,dx = \begin{cases} 1, & \text{if } m = n \\ 0, & \text{if } m \neq n \end{cases}. \quad (16)$$

It has been shown that [25] under certain regularity conditions a distribution can be expanded by a family of orthogonal polynomials and its associated distribution. The most popular orthogonal polynomials may be the *Hermite* polynomials which are associated with the normal distribution. The Hermite polynomial $H_n(\cdot)$ is defined as

$$\left(\frac{\partial}{\partial x}\right)^n e^{-x^2/2} = (-1)^n H_n(x) e^{-x^2/2}. \quad (17)$$

Then, we have

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x, \quad H_4(x) = x^4 - 6x^2 + 3, \ldots. \quad (18)$$

Let $g(\cdot)$ be a distribution with zero mean and unit variance and $M(t)$ be its moment generating function (MGF). The normal expansion is then described as follows [21].

$$g(x) = \phi(x)\left\{1 + \frac{\rho_3}{3!}H_3(x) + \frac{\rho_4}{4!}H_4(x) + \frac{\rho_5}{5!}H_5(x) + \frac{(\rho_6 + 10\rho_3^2)}{6!}H_6(x) + \cdots\right\} \quad (19)$$

where ρ_n is the standardized cumulant which is defined as $K^{(n)}(0)/K^{(2)}(0)$ where $K(T) = \ln[M(T)]$ and $K^{(n)}(\cdot)$ denotes the nth derivative of $K(\cdot)$. An approximation is made by retaining several terms in the expanded series given by (19). This is called the normal approximation. Since the series expands the distribution at its mean, if the distribution being approximated is close to a Gaussian one, we expect that the error is smaller around the mean. Based on this observation, we develop an adaptive scheme which will enhance the approximation accuracy.

The basic idea is to use a low-order expansion at each point of the distribution rather than to use a high-order expansion for the whole distribution at a single point. Suppose that we want to approximate the distribution at a point, say, x_0. We first transform the original distribution to a distribution which has its mean at x_0. Then, we apply the normal approximation on the transformed distribution and evaluate it at x_0. Since the normal approximation is good around the mean, this approach yields a better result than the straightforward normal approximation. The procedure for the transformation is called *recentering* [22] and the transformed distribution is called the *conjugate density* [23]. We now formally state the definition.

Definition. A density $g(z)$ is called the conjugate density of $f(x)$ at a point x_0 if there are constants α and T_0 such that

1) $$z = x - x_0 \tag{20}$$

2) $$g(z) = \alpha e^{T_0 z} f(z + x_0) \tag{21}$$

3) $$\int_{-\infty}^{+\infty} g(z) = 1; \qquad \int_{-\infty}^{+\infty} z g(z)\,dz = 0. \tag{22}$$

If, for the time being, we assume that the conjugate density is known, $g(z)$ can be normalized and expanded as in (19).

$$g(z) = \frac{1}{\sigma_z}\phi\left(\frac{z}{\sigma_z}\right)\left\{1 + \frac{\rho_{z,3}}{3!}H_3\left(\frac{z}{\sigma_z}\right) + \frac{\rho_{z,4}}{4!}H_4\left(\frac{z}{\sigma_z}\right) + \cdots\right\} \tag{23}$$

where σ_z and $\rho_{z,i}$ are the standard deviation and the ith standardized cumulant of $g(z)$, respectively. From condition 2, we know that $f(z + x_0) = \alpha^{-1}e^{-T_0 z}g(z)$. Then, $f(x_0) = \alpha^{-1}g(0)$.

$$g(0) = \frac{\phi(0)}{\sigma_z}\left\{1 + \tfrac{1}{8}\rho_{z,4} - \tfrac{1}{48}(\rho_{z,6} + 10\rho_3^2) + \cdots\right\}. \tag{24}$$

If the distribution doesn't deviate too much from Gaussianity, only the first term in the series provides good approximation, i.e.,

$$f(x_0) = \alpha^{-1}g(0) \approx \frac{\phi(0)}{\alpha\sigma_z}. \tag{25}$$

In order to use (25), we have to find the conjugate density. This is carried out as follows. Let the MGF of $f(x)$ be $M(T)$ and $K(T) = \ln[M(T)]$. Then,

$$M(T) = e^{K(T)} = \int_{-\infty}^{+\infty} e^{Tx} f(x)\,dx. \tag{26}$$

Let T be real. Multiplying both sides of (26) by e^{-Tx_0}, changing the variable x to s where $s = x - x_0$, and differentiating both sides with respect to T, we have

$$(K'(T) - x_0)e^{K(T) - Tx_0} = \int_{-\infty}^{+\infty} s e^{Ts} f(s + x_0)\,ds. \tag{27}$$

Now, if we choose T_0 and α such that

$$K'(T_0) - x_0 = 0 \tag{28}$$

$$1/\alpha = \int_{-\infty}^{+\infty} e^{T_0 s} f(s + x_0)\,ds = e^{K(T_0) - T_0 x_0} \tag{29}$$

then

$$\int_{-\infty}^{+\infty} \alpha e^{T_0 s} f(s + x_0)\,ds = 1 \tag{30}$$

$$\int_{-\infty}^{+\infty} s\alpha e^{T_0 s} f(s + x_0)\,ds = 0. \tag{31}$$

Hence, $g(s) = \alpha e^{T_0 s} f(s + x_0)$ is the conjugate density.

The solution of $K'(T) - x_0 = 0$ is referred to as the *saddle point* of $e^{-Tx_0}M(T)$. It is shown in [20] that under some regularity condition a unique real saddle point exists. Once the conjugate density is found, its moments can also be found. This is done by differentiating (26) n times and using (28), (29). Denote the nth moment of $g(z)$ by $\mu_{z,n}$, then

$$\mu_{z,n} = K^{(n)}(T_0). \tag{32}$$

The relation of moments and cumulants (denoted by $\kappa_{z,n}$) can be found in [24]. In particular, we have $\kappa_{z,n} = \mu_{z,n}$ for $n = 1,2,3$.

The MGT of a distribution is nothing but the Laplace transform of the distribution. One of the most important properties of Laplace transform is that the convolution in time domain or spatial domain can be transformed into multiplication in frequency domain. This property is directly applicable to the MGFs. This is also the key concept that we can avoid the convolution operation in the estimation of the score function as required in Masreliez's approach.

B. Adaptive Normal Approximation of Score Functions

The distribution approximation technique discussed above can also be extended to find the approximation of score function. The idea is to find the expansion of $f(\cdot)$ and $f'(\cdot)$ via the conjugate recentering. From that, we find the expansion of $f'(\cdot)/f(\cdot)$ and truncate it to obtain the approximation. Here, we make the assumption that $f'(\cdot)$ possesses a series expansion and therfore can be obtained by a term-by-term differentiation of $f(\cdot)$. First, we construct the conjugate density $g(z)$ at x_0 and expressed $f(x)$ in terms of $g(z)$.

$$f(x) = f(z + x_0) = \alpha^{-1}e^{-T_0 z}g(z). \tag{33}$$

Then

$$f'(x) = -T_0\alpha^{-1}e^{-T_0 z}g(z) + \alpha^{-1}e^{-T_0 z}g'(z) \tag{34}$$

$$\left.\frac{f'(x)}{f(x)}\right|_{x=x_0} = -T_0 + \left.\frac{g'(z)}{g(z)}\right|_{z=0}. \tag{35}$$

The expansion of $g(z)$ using (23) is given by

$$g(z) = \frac{1}{\sigma_z}\phi\left(\frac{z}{\sigma_z}\right)\left\{1 + \frac{\rho_{z,3}}{3!}H_3\left(\frac{z}{\sigma_z}\right) + \frac{\rho_{z,4}}{4!}H_4\left(\frac{z}{\sigma_z}\right) + \cdots\right\}. \tag{36}$$

Then

$$g'(z) = -\frac{1}{\sigma_z^2}\phi\left(\frac{z}{\sigma_z}\right)\left\{H_1\left(\frac{z}{\sigma_z}\right) + \frac{\rho_{z,3}}{3!}H_4\left(\frac{z}{\sigma_z}\right) + \frac{\rho_{z,4}}{4!}H_5\left(\frac{z}{\sigma_z}\right) + \cdots\right\} \tag{37}$$

$$g(0) = \frac{\phi(0)}{\sigma_z}\left\{1 + \tfrac{1}{8}\rho_{z,4} - \tfrac{1}{48}\left(\rho_{z,6} + 10\rho_3^2\right) + \cdots\right\} \tag{38}$$

$$g'(0) = -\frac{\phi(0)}{\sigma_z^2}\left\{\tfrac{1}{2}\rho_{z,3} - \tfrac{1}{8}\rho_{z,5} + \cdots\right\}. \tag{39}$$

Hence,

$$\frac{g'(0)}{g(0)} = -\frac{\rho_{z,3}}{2\sigma_z} - \frac{1}{\sigma_z}\left\{\tfrac{1}{8}\left(\rho_{z,5} - \tfrac{1}{2}\rho_{z,3}\rho_{z,4}\right) + \cdots\right\}. \tag{40}$$

Retaining the first term in (40), we obtain the approximation of the score funtion.

$$\frac{f'(x_0)}{f(x_0)} \approx -T_0 - \frac{\rho_{z,3}}{2\sigma_z} = -T_0 - \frac{\mu_{z,3}}{2\sigma_z^4}. \tag{41}$$

Here T_0 is obtained from (28). To implement Masreliez's filter, we also need the derivative of the score function. This is done in the following manner. The relationship between x and T is established through the equation $K'(T) = x$. Thus,

$$K''(T)\frac{dT}{dx} = 1 \tag{42}$$

$$\frac{dT}{dx} = \frac{1}{K''(T)}. \tag{43}$$

Taking the derivative of the score function (41) with respect to x, we obtain

$$\left.\frac{d}{dx}\left(\frac{f'}{f}\right)\right|_{x=x_0} \approx \left.\left\{-\frac{dT}{dx} - \frac{dT}{dx}\frac{d}{dT}\left[\frac{K^{(3)}(T)}{2K''(T)^2}\right]\right\}\right|_{T=T_0} \approx -\frac{1}{\sigma_z^2}\left[1 + \frac{\mu_{z,4}}{2\sigma_z^4} - \frac{\mu_{z,3}^2}{\sigma_z^6}\right]. \tag{44}$$

In the above derivation, we assume that the distribution under consideration is univariate. It turns out that the score function approximation scheme is simple and computationally efficient. It is not difficult to extend the result to the multivariate distributions. However, in this case, the approximation formula becomes very complicated. The computational complexity is then greatly increased. The observation distributions in a radar system are usually multivariate. In the following section, we model the target dynamics in a spherical coordinate. This makes the observation distributions univariate. Thus, the efficient scalar score function approximation scheme can be applied.

IV. TRACKING WITH GLINT NOISE

A. Non-Gaussian Glint Noise

As we have seen in Fig. 1, the glint noise is clearly non-Gaussian and long tailed. Now, the problem is how to model it. Borden and Mumford [27] consider the distribution of the glint as a student's t distribution with two degrees of freedom and develop a method to produce glint-like signals. From the empirical studies, Hewer, Martin, and Zeh [6] argue that the glint can be modeled as a mixture of a Gaussian noise and outliers. Their results are based on the analysis of normal QQ-plots of glint noise records. Indeed, from [6, Fig. 2], we can see that the QQ-plot is fairly linear around the origin. This indicates that the distribution is Gaussian-like around its mean. But, in the tail region, the plot deviates the linearity and indicates a non-Gaussian long-tailed character. The data in the tail region is essentially associated with the glint spikes and are considered to be outliers. They are modeled as a Gaussian noise with large variance. This leads to the Gaussian-mixture noise model. Although this model is simple, it is not suitable for our use. It is easy to see that the score function of the Gaussain mixture is not robust. The score function of a Gaussian mixture increases without bound as the observation goes to infinity. Here, we remedy this problem by modeling the glint as a mixture of a Gaussian and a Laplacian noise, i.e.,

$$f_t(x) = (1-\epsilon)f_g(x) + \epsilon f_l(x)$$

where $f_t(\cdot)$, $f_g(\cdot)$, and $f_l(\cdot)$ represent the glint, the Gaussian, and the Laplacian distribution, respectively, and ϵ is a small positive number less than one. The variance of $f_l(\cdot)$ is larger than that of $f_g(x)$. One can show that the score function of this distribution is robust. It is also interesting to note that this score function is very similar to some of the psi-function in the robust M-estimator. In fact, the score function does do the similar work as the psi-function does.

B. System Model

Assume that the dynamics of a maneuvering target in the rectangular coordinate is described by the following differential equations.

$$\ddot{x} = -\alpha\dot{x} + w_x(t) + u_x(t) \tag{45}$$

$$\ddot{y} = -\alpha\dot{y} + w_y(t) + u_y(t) \tag{46}$$

$$\ddot{z} = -\alpha\dot{z} + w_z(t) + u_z(t) \tag{47}$$

where α is the viscous drag coefficient, $w_x(t)$, $w_y(t)$, and $w_z(t)$ are the driving noise, and $u_x(t)$, $u_y(t)$, and $u_z(t)$ are some deterministic inputs. Let us assume that the driving noise and the inputs are piecewisely constant. Representing the differential equations as a set of state equations and discretizing them, we obtain

$$x_{k+1} = \phi x_k + \Gamma(w_k + u_k) \tag{48}$$

where

$$x_k = (x_k, \dot{x}_k, y_k, \dot{y}_k, z_k, \dot{z}_k)^t \tag{49}$$

$$\phi = \begin{pmatrix} 1 & p_1 & 0 & 0 & 0 & 0 \\ 0 & p_1' & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & p_1 & 0 & 0 \\ 0 & 0 & 0 & p_1' & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & p_1 \\ 0 & 0 & 0 & 0 & 0 & p_1' \end{pmatrix} \tag{50}$$

$$\Gamma = \begin{pmatrix} p_2 & 0 & 0 \\ p_2' & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & p_2' & 0 \\ 0 & 0 & p_2 \\ 0 & 0 & p_2' \end{pmatrix} \tag{51}$$

with w_k a zero mean Gaussian driving noise vector and u_k a deterministic input vector at sampling instant k. It can be shown that

$$p_1 = (1 - e^{-\alpha T})/\alpha$$

$$p_1' = e^{-\alpha T} \tag{52}$$

$$p_2 = (e^{-\alpha T} - 1 + \alpha T)/\alpha^2 \tag{53}$$

$$p_2' = (1 - e^{-\alpha T})/\alpha \tag{54}$$

where T is the sampling interval. The observation equations are described as

$$r_k = \sqrt{x_k^2 + y_k^2 + z_k^2} + v_k^r \tag{55}$$

$$b_k = \tan^{-1}\left(\frac{y_k}{x_k}\right) + v_k^b \tag{56}$$

$$e_k = \sin^{-1}\frac{z_k}{\sqrt{x_k^2 + y_k^2}} + v_k^e \tag{57}$$

where $v_k = (v_k^r, v_k^b, v_k^e)^t$ is the observation noise, r_k represents range, and e_k and b_k represent elevation and bearing measurements, respectively. Here, we assume that v_k is the glint noise and is non-Gaussian. It is readily seen that radar observations are obtained in the spherical coordinate while the target dynamics are described in the rectangular coordinate. In order to apply Kalman filter, the nonlinear relations must be linearized. Expanding the nonlinear functions in the observation equation using the Taylor series and truncating high-order terms, we obtain a set of linear observation equations.

$$z_k = H_k x_k + v_k \tag{58}$$

where $z_k = (r_k, b_k, e_k)^t$ and

$$H_k = \begin{pmatrix} \frac{x_k}{r_k} & 0 & \frac{y_k}{r_k} & 0 & \frac{z_k}{r_k} & 0 \\ \frac{-y_k}{x_k^2 + y_k^2} & 0 & \frac{-x_k}{x_k^2 + y_k^2} & 0 & 0 & 0 \\ \frac{-x_k z_k}{r_k^2\sqrt{x_k^2 + y_k^2}} & 0 & \frac{-y_k z_k}{r_k^2\sqrt{x_k^2 + y_k^2}} & 0 & \frac{\sqrt{x_k^2 + y_k^2}}{r_k^2} & 0 \end{pmatrix}. \tag{59}$$

The application of an extended Kalman filter is adequate if the observation is Gaussian. Unfortunately, the glint noise is non-Gaussian. Here, we propose to apply the score function based filtering scheme as described in Section II and III to tackle the problem. As we mentioned in Section III, the score function approximation scheme is most efficient for univariate densities. However, the distribution of the observation in (58) is multivariate. Thus, the filtering in the rectangular coordinate is not appropriate for our purpose.

If the state can be defined in the spherical coordinate, the three components of the observation will become independent. However, modeling target dynamics in spherical coordinates (r, b, e) usually involves the simultaneous solution of three very complex nonlinear differential equations. One way to overcome this problem is to linearize the target dynamics. This enables us to obtain the linear state equations in terms of r, $\dot{r}$, b, $\dot{b}$, e, and $\dot{e}$. If, in addition, some simplifying assumption is made, the target model can be reduced to three independent channels. This method is proposed by Gholson and Moose in [1] and is called the approximate spherical modeling. Using this modeling, we cannot only apply the score function algorithm but also achieve computational saving. It is described in [1] that the new state equations of the spherical modeling are as follows.

Range Channel:

$$\begin{pmatrix} r \\ \dot{r} \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & p_1 \\ 0 & p_1' \end{pmatrix} \begin{pmatrix} r \\ \dot{r} \end{pmatrix}_k + \begin{pmatrix} p_2 & 0 \\ p_2' & 0 \end{pmatrix} (u_r + w_r) \tag{60}$$

$$z_k^r = r_k + v_k^r. \tag{61}$$

Elevation Channel:

$$\begin{pmatrix} e \\ \dot{e} \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & p_1 \\ 0 & p_1' \end{pmatrix} \begin{pmatrix} e \\ \dot{e} \end{pmatrix}_k + \begin{pmatrix} p_2/d_1 & 0 \\ p_2'/d_1 & 0 \end{pmatrix} (u_e + w_e) \tag{62}$$

$$z_k^e = e_k + v_k^e. \tag{63}$$

Bearing Channel:

$$\begin{pmatrix} b \\ \dot{b} \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & p_1 \\ 0 & p_1' \end{pmatrix} \begin{pmatrix} b \\ \dot{b} \end{pmatrix}_k + \begin{pmatrix} p_2/d_1 & 0 \\ p_2'/d_2 & 0 \end{pmatrix} (u_b + w_b) \tag{64}$$

$$z_k^b = b_k + v_k^b \tag{65}$$

where $d_1 = \bar{r}_k$, $d_2 = \cos\bar{e}_k$, $u_r = u_x$, $w_r = w_x$, $u_e = u_z$, $w_e = w_z$, $u_b = u_y$, and $w_b = w_y$. Here $\bar{x}_k$ represents the one step prediction of x at instant $k-1$. Thus, we obtain three independent channels with scalar observations. Now the filtering can be processed in parallel with smaller dimensionality. The computation is then simplified and the score function of the observation prediction density can be efficiently implemented.

Although the approximate spherical modeling is simple, there is one problem. It has been shown [5] that the linearization of the target dynamics will lead to large dynamic errors. Here, we use a simple algorithm which can significantly reduce the error. Observing the Kalman equations descibed in (6)–(9), we find that the Kalman filtering can be divided into two parts, namely, filtering and prediction. The target model is only used in the prediction. To reduce the effect of the model linearization, we can separate the operation of filtering and prediction. We can perform filtering in the spherical coordinate and prediction in the rectangular coordinate. At every stage, we transform the state from the spherical coordinate to the rectangular coordinate after filtering. Then, we perform the prediction in the rectangular coordinate. To continue the filtering in next stage, the state is transformed back to the spherical coordinate. The transformation from the spherical to the rectangular coordinate is described in the following equations.

$$x_k = r_k \cos e_k \cos b_k \tag{66}$$

$$y_k = r_k \cos e_k \sin b_k \tag{67}$$

$$z_k = r_k \sin e_k \tag{68}$$

and the transformation from the rectangular coordinate to the spherical coordinate is described in (55)–(57).

Simulations show that by using the spherical modeling and only performing the state prediction in the rectangular coordinate, the result is almost as good as that in the rectangular modeling.

Since the approximate spherical modeling decouples the system into three independent channels. In the following, we summarize the implementation of the score function for one channel only. Let x_k be the state and z_k is the observation of one channel in the spherical coordinate. We assume that the MGF of $f(v_k)$ is known. Since $f(x_k|Z^{k-1})$ is Gaussian with known mean and variance, the MGF of $f(H_k x_k|Z^{k-1})$ can be easily found. Let the MGF of $f(v_k)$ and $f(H_k x_k|Z^{k-1})$ be $M_v(T)$ and $M_x(T)$, respectively.

1) *Find the MGF of* $f(z_k|Z^{k-1})$: Since $f(z_k|Z^{k-1})$ is obtained from the convolution of $f(v_k)$ and $f(H_k x_k|Z^{k-1})$, the MGF of $f(z_k|Z^{k-1})$ is $M_v(T)M_x(T)$.

2) *Find the conjugate density of* $f(z_k|Z^{k-1})$ *at* z_k: Let $K(T) = \ln(M_v(T)M_x(T))$. The conjugate density of $f(z_k|Z^{k-1})$ at z_k is constructed as $g(s) = \alpha_k e^{T_k s} f(s + z_k)$ where T_k is chosen as the saddle point of $\{M_v(T)M_x(T)e^{-Tz_k}\}$, i.e., $K'(T_k) - z_k = 0$, and $1/\alpha_k$ as $\{M_v(T_k)M_x(T_k)e^{-T_k z_k}\}$.

3) *Find the second, the third, and the fourth moment of* $g(s)$: σ_s^2, $\mu_{s,3}$, and $\mu_{s,4}$ can be found by

$$\sigma_s^2 = K^{(2)}(T_k); \qquad \mu_{s,3} = K^{(3)}(T_k); \qquad \mu_{s,4} = K^{(4)}(T_k) \tag{69}$$

where $K^{(i)}(T)$ denotes the ith derivative of $K(T)$.

4) *Approximate the score function of* $f(z_k|Z^{k-1})$ *and its derivative*:

$$-\frac{f'(z_k|Z^{k-1})}{f(z_k|Z^{k-1})} \approx T_k + \frac{\mu_{s,3}}{2\sigma_s^2} \tag{70}$$

$$\frac{d}{dz_k}\left(\frac{f'}{f}\right) = \frac{1}{\sigma_s^2}\left[1 + \frac{\mu_{s,4}}{2\sigma_s^4} - \frac{\mu_{s,3}^2}{\sigma_s^6}\right] \tag{71}$$

To use this scheme, the MGF of the noise distribution has to be known. If the close form of the MGF does not exist, some estimation method has to be applied. The whole tracking algorithm can now be summarized as follows.

1) Find the score functions using the procedure described above for three channels (the state are defined in the spherical coordinate).

2) Use (10)–(13) to accomplish the filtering for three channels.

3) Transform all the states to the rectangular coordinate using (66)–(68).

4) Predict the state in the next stage using (1).

5) Transform the states to the spherical coordinate using (55)–(57).

6) Go to step 1).

V. SIMULATIONS

In this section, we perform some simulations to examine the performance of our new scheme. We first see how well the score function approximation can be. Specifically, we use a non-Gaussian distribution model for experiments, namely,

$$f(x) = pf_i(x-\mu_i) + (1-p)f_j(x-\mu_j). \quad (72)$$

The subscript of f (i and j) stands for type of distribution and μ for the mean of the distribution. We only consider the three most common types of distributions, namely, uniform, Gaussian, and Laplacian which are denoted by u, g, and l, respectively. The definitions of these distributions are given as follows:

Uniform:

$$f_u(x) = \frac{1}{2m} \qquad -m \le x \le m. \quad (73)$$

Gaussian:

$$f_g(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-x^2/2\sigma^2}. \quad (74)$$

Laplacian:

$$f_l(x) = \frac{1}{2\eta}e^{-|x|/\eta}. \quad (75)$$

Note that the model in (72) can have either one distribution or a mixture of two distributions.

We show three examples of score function approximation. The distributions we used results from the convolution of non-Gaussian distributions with a Gaussian distribution. This type of distribution is what we encounter in the filtering problem. The non-Gaussian distributions are 1) uniform distribution, 2) uniform and Laplacian mixed distribution, and 3) Laplacian and Laplacian mixed distribution. The Gaussian distribution is assumed to have zero mean and variance σ^2. Using the model in (72), we specify the parameters as follows:

1) $p = 1,\ i = u,\ m = 1,\ (\sigma^2 = 0.1)$.

2) $p = 0.99;\ i = u,\ m = 1;\ j = l,\ \eta = 5;$
$\mu_i = \mu_j = 0;\ (\sigma^2 = 0.2)$.

3) $p = 0.5;\ i = j = l,\ \eta_i = \eta_j = 1,$
$\mu_i = -\mu_j = 3;\ (\sigma^2 = 1)$.

The MGF of these distributions and the moments of their conjugate densities can be found in the Appendix. Figs. 2–4 are the plots of the score functions of the distributions 1)–3). It can be seen clearly that the approximation scheme is quite satisfactory. It is also easy to find that the approximation error mainly concentrates on the area where the slope of the score function is steep.

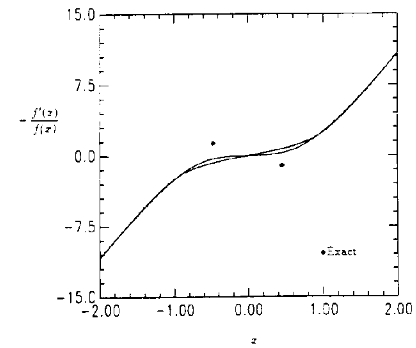

Fig. 2. Approximated and exact score functions: uniform convolved with Gaussian distribution.

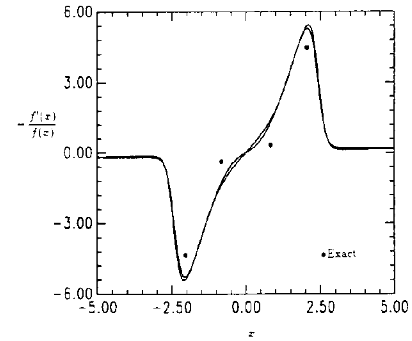

Fig. 3. Approximated and exact score functions: mixture of uniform and Laplacian convolved with Gaussian distribution.

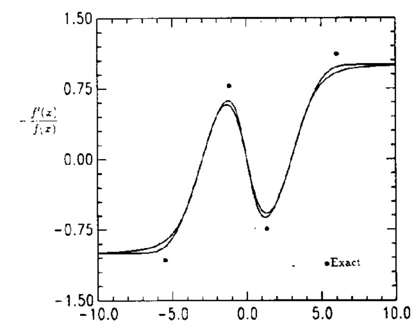

Fig. 4. Approximated and exact score functions: mixture of two Laplacian convolved with Gaussian distribution.

Next, we examine the tracking performance of the approximate spherical modeling. We perform the filtering in the spherical coordinate and the state prediction in the rectangular coordinate. The parameters of the simulated systems are listed as follows.

$\alpha = 0.5$ s, $T = 0.5$ s

$Q = \mathrm{diag}(4\times10^{-4}\ \mathrm{km}^2/\mathrm{s}^4,\ 1\times10^{-4}\ \mathrm{km}^2/\mathrm{s}^4,\ 5\times10^{-5}\ \mathrm{km}^2/\mathrm{s}^4)$

$R = \mathrm{diag}(2.5\times10^{-3}\ \mathrm{km}^2,\ 1\times10^{-5}\ \mathrm{rad}^2,\ 1\times10^{-5}\ \mathrm{rad}^2)$

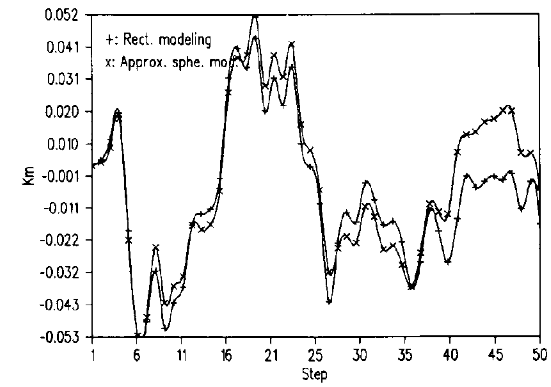

Fig. 5. Range error in x direction.

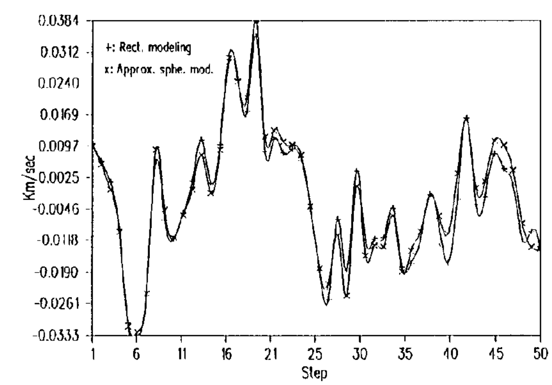

Fig. 6. Speed error in x direction.

$U = (-0.15\ \text{km/s}^2,\ 0.04\ \text{km/s}^2,\ 0.02\ \text{km/s}^2)^t$

$Q_i = \text{diag}(2.4 \times 10^{-3}\ \text{km}^2/\text{s}^4,\ 6 \times 10^{-4}\ \text{km}^2/\text{s}^4,\ 6 \times 10^{-4}\ \text{km}^2/\text{s}^4)$

$x_i = (10\ \text{km},\ -0.3\ \text{km/s},\ 1\ \text{km},\ 0.06\ \text{km/s},\ 2\ \text{km},\ 0.03\ \text{km/s})^t$

where Q_i is the initial estimated state covariance matrix and x_i is the initial estimate of the state. Figs. 5 and 6 show the comparison of the tracking errors (range and velocity) in x direction for the rectangular modeling and the approximate spherical modeling. The error signals are obtained by a subtraction of the true states from the estimated states. It can be seen from these figures that the error due to the approximate spherical approximation is rather small. It also shows that only the state prediction in rectangular coordinate is enough.

Finally, we compare the tracking performance of the linear Kalman filter (extend Kalman filter) and the nonlinear Kalman filter (the score function based filter). As we described in Section IV, the glint measurement noise is assumed to be a mixture of a Gaussian noise with smaller variance and a Laplacian noise with larger variance. In our simulation, the occurrence probability of the Laplacian noise is assumed to be 0.05. The parameters of the Laplacian

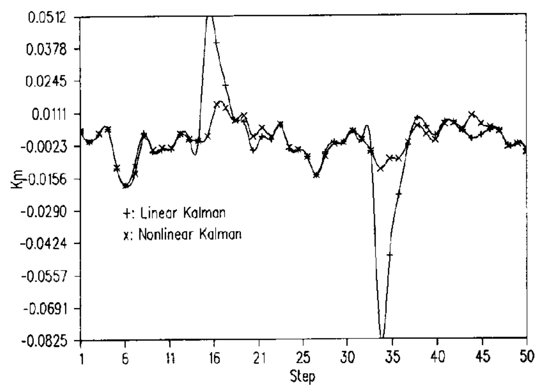

Fig. 7. Range error in x direction.

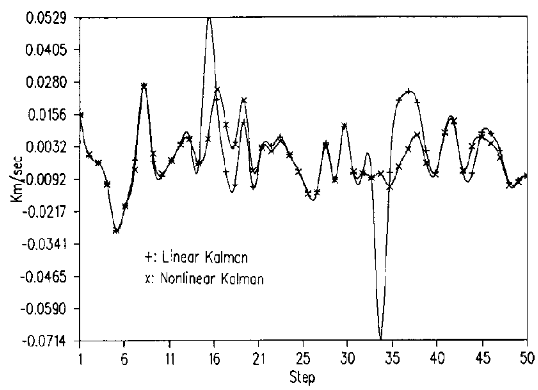

Fig. 8. Speed error in x direction.

noise are $\eta_r = 0.05$ km, $\eta_b = 0.003$ rad, $\eta_e = 0.003$ rad. The parameters of the system are as follows.

$\alpha = 0.5/\text{s},\ T = 0.5\ \text{s}$

$Q = \text{diag}(4 \times 10^{-4}\ \text{km}^2/\text{s}^4,\ 1 \times 10^{-4}\ \text{km}^2/\text{s}^4,\ 5 \times 10^{-5}\ \text{km}^2/\text{s}^4)$

$R = \text{diag}(5 \times 10^{-5}\ \text{km}^2,\ 4 \times 10^{-7}\ \text{rad}^2,\ 4 \times 10^{-7}\ \text{rad}^2)$

$U = (-0.15\ \text{km/s}^2,\ 0.04\ \text{km/s}^2,\ 0.02\ \text{km/s}^2)^t$

$Q_i = \text{diag}(2.4 \times 10^{-3}\ \text{km}^2/\text{s}^4,\ 6 \times 10^{-4}\ \text{km}^2/\text{s}^4,\ 6 \times 10^{-4}\ \text{km}^2/\text{s}^4)$

$x_i = (10\ \text{km},\ -0.3\ \text{km/s},\ 1\ \text{km},\ 0.06\ \text{km/s},\ 2\ \text{km},\ 0.03\ \text{km/s})^t$

where Q_i is the initial state estimate covariance matrix and x_i is the initial estimate of the state. Two cases are simulated. In the first case, we assume that the Laplacian noise is ignored for the linear Kalman filter. Consequently, the variances of the observation noise are just those of the Gaussian components. In the second case, we include the variance of the Laplacian noise. Figs. 7 and 8 illustrate the tracking error (range and velocity) in x direction for the first case. We can see that the difference between the linear Kalman filter and the nonlinear Kalman

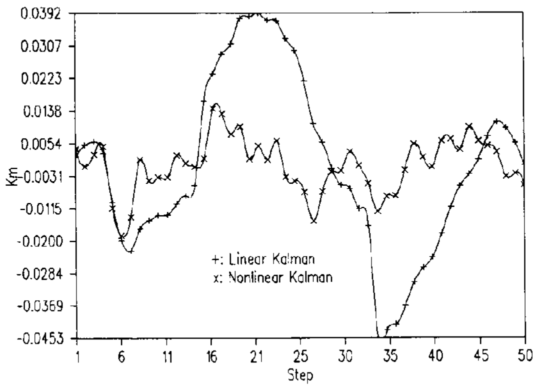

Fig. 9. Range error in x direction.

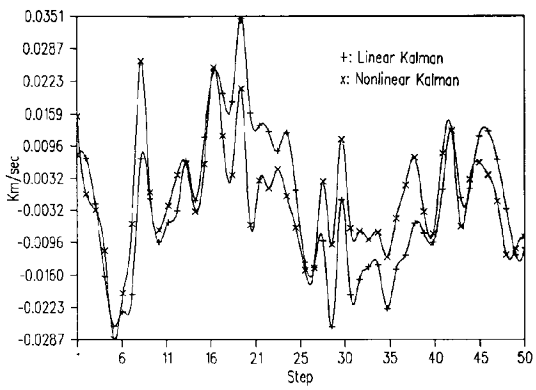

Fig. 10. Speed error in x direction.

filter is significant. In the presence of the Gaussian noise, the performance of both filters are almost the same. However, in the presence of the high variance Laplacian noise, the linear Kalman filter fails. This is because the linear Kalman filter uses a constant gain and cannot distinguish two different kinds of noises. On the contrary, the nonlinear Kalman filter adaptively changes the gain and is able to estimate the states nicely. Figs. 9 and 10 show the tracking errors (range and velocity) in x direction for the second case. For the linear Kalman filter, the variance of the Laplacian noise is included and this results in too much smoothing in the presence of Gaussian noise. This is not so obviously shown in Fig. 10. In Fig. 11, we show the comparison of the velocity tracking. From that, we can clearly see the oversmoothing effect. Based on these results, we conclude that for the non-Gaussian glint noise, the linear Kalman is insufficient. A more sophisticated filtering scheme is needed. The nonlinear algorithm we developed here provides a simple and feasible way to solve the problem.

VI. CONCLUSION

Conventional tracking algorithms are all based on an idealized assumption, i.e., Gaussian observation

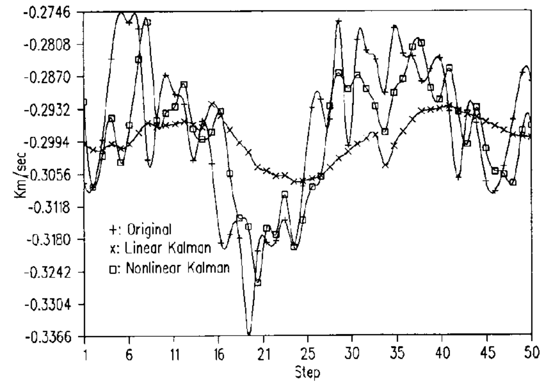

Fig. 11. Target speed in x direction.

noise. In real applications, however, the noise is not necessarily Gaussian. In those cases, the conventional algorithms are not sufficient. Due to the target glint, the observation noise which is referred to as glint noise in a radar system, is non-Gaussian. In this paper, we develop a new tracking algorithm based on the score function approach. The algorithm is a Kalman type of recursive filtering scheme which can work nearly optimally with the glint noise. An approximate spherical modeling is adopted in order to efficiently use the algorithm. Using this model, we are able to decouple the target model into three independent channels and the computational complexity is reduced dramatically. The main cost of the computation for the score function approach is on the search of the saddle point. For a scalar case, this computational requirement is small. Simulations show that using the new algorithm the tracking performance is greatly improved. The scheme developed here can be extended to some other kind of tracking algorithms such as multiple model and input estimate, etc. Research in this direction is now underway.

APPENDIX

1) *The Uniform Distribution Convolved with the Gaussian Distribution*:

$$f(x) = f_u(x) * f_g(x) \tag{76}$$

where $f_u(x)$ is a zero mean uniform distribution defined as

$$f_u(x) = \frac{1}{2m}, \qquad -m \leq x \leq m \tag{77}$$

and $f_g(x)$ is a zero mean Gaussian distribution defined as

$$f_g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}, \qquad -\infty \leq x \leq +\infty. \tag{78}$$

Let $M(T)$, $M_u(T)$, and $M_g(T)$ be the MGFs of $f(x)$, $f_u(x)$, and $f_g(x)$, respectively. Then,

$$M(T) = M_u(T)M_g(T) \tag{79}$$

$$M_u(T) = \int_{-\infty}^{+\infty} f_u(x)e^{Tx}\,dx = \frac{1}{2mT}(e^{mT} - e^{-mT}) = \frac{1}{mT}\sinh(mT) \tag{80}$$

$$M_g(T) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{-(x^2/2\sigma^2)+Tx}\,dx = e^{(\sigma^2/2)T^2} \tag{81}$$

$$M(T) = \frac{1}{mT}\sinh(mT)e^{(\sigma^2/2)T^2} \tag{82}$$

$$K(T) = \ln(M(T)) = -\ln(mT) + \ln[\sinh(mT)] + \frac{\sigma^2}{2}T^2 \tag{83}$$

$$K'(T) = -\frac{1}{T} + m\coth(mT) + \sigma^2 T. \tag{84}$$

The saddle point can be found by solving the equation $K'(T) - x = 0$.

$$K^{(2)}(T) = \frac{1}{T^2} - m^2\operatorname{csch}^2(mT) + \sigma^2$$

$$K^{(3)}(T) = -\frac{2}{T^3} - 2m^3\operatorname{csch}^2(mT)\coth(mT)$$

$$K^{(4)}(T) = \frac{6}{T^4} - 2m^4\operatorname{csch}^2(mT) \times [2\coth^2(mT) + \operatorname{csch}^2(mT)]. \tag{85}$$

2) *The Laplacian Distribution Convolved with the Gaussian Distribution*:

$$f(x) = f_l(x) * f_g(x) \tag{86}$$

where $f_l(x)$ is a Laplacian distribution defined as

$$f_l(x) = \frac{1}{2\mu}e^{-|x|/\mu}, \qquad -\infty \le x \le +\infty \tag{87}$$

$$M_l(T) = \int_{-\infty}^{+\infty} e^{(-|x|/\mu)+Tx}dx = \frac{1}{1-\mu^2T^2} \tag{88}$$

Then,

$$M(T) = M_l(T)M_g(T) = \frac{1}{1-\mu^2T^2}e^{(\sigma^2/2)T^2} \tag{89}$$

$$K(T) = -\ln(1-\mu^2T^2) + \frac{\sigma^2}{2}T^2 \tag{90}$$

$$K'(T) = \frac{2\mu^2T}{1-\mu^2T^2} + \sigma^2T \tag{91}$$

$$K^{(2)}(T) = \frac{4\mu^4T^2}{(1-\mu^2T^2)^2} + \frac{2\mu^2}{(1-\mu^2T^2)} + \sigma^2 \tag{92}$$

$$K^{(3)}(T) = \frac{16\mu^6T^3}{(1-\mu^2T^2)^3} + \frac{8\mu^4T}{(1-\mu^2T^2)^2} + \frac{4\mu^4T}{(1-\mu^2T^2)^2} \tag{93}$$

$$K^{(4)}(T) = \frac{96\mu^8T^4}{(1-\mu^2T^2)^4} + \frac{96\mu^6T^2}{(1-\mu^2T^2)^3} + \frac{12\mu^4}{(1-\mu^2T^2)^2}. \tag{94}$$

REFERENCES

[1] Gholson, N. H., and Moose, R. L. (1977)
Maneuvering target tracking using adaptive state estimation.
IEEE Transactions on Aerospace and Electronic Systems, **AES-13** (May 1977), 310–317.

[2] Moose, R. L., et al. (1979)
Modeling and estimation for tracking maneuvering targets.
IEEE Transactions on Aerospace and Electronic Systems, **AES-15** (May 1979), 448–456.

[3] Singer, R. A. (1970)
Estimating optimal tracking filter performance for manned maneuvering targets.
IEEE Transactions on Aerospace and Electronic Systems, (July 1970), 473–483.

[4] Pearson, J. B., and Stear, F. B. (1974)
Kalman filter applications in airborne radar tracking.
IEEE Transactions on Aerospace and Electronic Systems, **AES-10** (May 1974), 319–329.

[5] Spingarn, K., and Weidemann, H. L. (1972)
Linear regression filtering and prediction for tracking maneuvering aircraft targets.
IEEE Transactions on Aerospace and Electronic Systems, **AES-8** (Nov. 1972), 800–810.

[6] Hewer, G. A., Martin, R. D., and Zeh, J. (1987)
Robust preprocessing for Kalman filtering of glint noise.
IEEE Transactions on Aerospace and Electronic Systems, **AES-23** (Jan. 1987), 120–128.

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

[8] Sorenson, H. W., and Stubberud, A. R. (1968)
Non-linear filtering by approximation of the *a posteriori* density.
International Journal of Control, **18** (1968), 33–51.

[9] Sorenson, H. W., and Alspach, D. L. (1971)
Recursive Bayesian Estimation using Gaussian sums.
Automatica, 7 (1971), 465–479.

[10] Alspach, D. L., and Sorenson, H. W. (1972)
Nonlinear Bayesian estimation using Gaussian sum approximations.
IEEE Transactions on Automatic Control, **AC-17**, 4 (Aug. 1972), 465–478.

[11] Masreliez, C. J. (1972)
Robust recursive estimation and filtering,
Ph.D. dissertation, University of Washington, Seattle, 1972.

[12] Masreliez, C. J. (1975)
Approximate non-Gaussian filtering with linear state and observaton relations.
IEEE Transactions on Automatic Control, **AC-20** (1975), 107–110.

[13] Tollet, I. H. (1975)
Robust forecasting.
Ph.D. dissertation, University of Washington, Seattle, 1975.

[14] Ershov, A. A., and Lipster, R. S. (1978)
A robust Kalman filter in discrete time.
Automatic Remote Control, **39** (1978), 359–367.

[15] Ershov, A. A. (1978)
Robust filtering algorithms.
Automatic Remote Control, **39** (July 1978), 992–996.

[16] West, M. (1981)
Robust sequential approximate Bayesian estimation.
Journal of the Royal Statistics Society B, **2** (1981), 157–166.

[17] Tsai, C., and Kurz, L. (1983)
An adaptive robustizing approach to Kalman filtering.
Automatica, **19** (1983), 279–299.

[18] Subba Rao, T., and Yar, M. (1984)
Linear and non-linear filters for linear, but not Gaussian processes.
International Journal of Control, (1984), 235–246.

[19] Wu, W.-R. (1989)
Kalman filtering in non-Gaussian environment using efficient score function approximation.
In *Proceedings of 1989 IEEE International Symposium on Circuits and Systems*, 1989, 413–416.

[20] Daniels, H. E. (1954)
Saddle point approximations in statistics.
Annals of Mathematical Statistics, **25** (1954), 631–650.

[21] Barndorff-Nielsen, O., and Cox, D. R. (1979)
Edgeworth and saddle point approximations with statistical applications.
Journal of the Royal Statistical Society B, **41** (1979), 279–312.

[22] Hampel, F. R. (1973)
Some small sample asymptotics.
In *Proceedings of the Prague Symposium on Asymptotic Statistics*, J. Hajek (Ed.), Vol. II, 1973, 109–126; held in Prague: Charles University, 1973.

[23] Khinchin, A. I. (1949)
Mathematic Foundation of Statistical Mechanics.
New York: Dover, 1949.

[24] Kendall, M. G., and Stuart, A. (1958)
Advanced Theory of Statistics, Vol. 1.
London: Griffin, 1958.

[25] Cramer, H. (1946)
Mathematical Methods of Statistics.
Princeton, NJ: Princeton University Press, 1946.

[26] Masreliez, D. J., and Martin, R. D. (1977)
Robust Bayesian estimation for the linear model and robustifying the Kalman filter.
IEEE Transactions on Automatic Control, **AC-22** (1977), 361–371.

[27] Borden, B. H., and Mumford, M. L. (1983)
A statistical glint/radar cross section target model.
IEEE Transactions on Aerospace and Electronic Systems, **AES-19** (Sept. 1983), 781–785.

Wen-Rong Wu was born in Taiwan, R.O.C., in 1958. He received his B.S. degree in mechanical engineering from Tatung Institute of Technology, Taiwan, in 1980, the M.S. degrees in mechanical and electrical engineering, and Ph.D. degree in electrical engineering from State University of New York at Buffalo in 1985, 1986, and 1989, respectively.

Since August 1989 he has been a faculty member in the Department of Communication Engineering in National Chiao Tung University, Taiwan. His research interests include estimation theory, digital signal processing, and image processing.