EUSIPCO 2013 1569746757

PROPOSAL DISTRIBUTION FOR PARTICLE FILTERING APPLIED TO TERRAIN NAVIGATION

Achille Murangira, Christian Musso

Onera - The French Aerospace Lab
Chemin de la Hunière et des Joncherettes
BP 80100
FR-91123 PALAISEAU CEDEX
France

ABSTRACT

This article provides a methodology for designing a proposal distribution in the context of particle filtering for terrain navigation. The suggested method is based on the use of an importance distribution centered around an estimate of the maximum a posteriori (MAP). By assuming a Gaussian prior, we show that the computation of the MAP can be reduced to an optimization problem in a space of lower state dimension. Furthermore, we introduce a new method for choosing the covariance of the proposal. In this case, numerical experiments show that the method can improve upon classical sampling methods.

Index Terms— Bayesian filtering, particle filtering, importance sampling, proposal distribution, maximum a posteriori, terrain navigation, radar-altimeter, digital terrain elevation data

1. INTRODUCTION

The problem of terrain navigation has been addressed in recent years via the use of non-linear filters such as the point mass filter [1] and particle filters [2] [3]. The latter is a sequential Monte Carlo method based on importance sampling that can in theory approximate any posterior distribution thanks to a finite number of particles. The main ingredients of particle filtering are a prediction step where particles are sampled according to a prior distribution that incorporates the dynamic model and a correction step that weights these particles according to an observation provided by a sensor through the use of a likelihood function. However, in the case of a strong mismatch between the sampling distribution and the weighting function, particle filters can diverge. To avoid this, it is advocated in the filtering literature to sample from an importance distribution that takes into account the measurement in the prediction step.

The goal of this article is to provide a methodology for choosing this importance distribution in the context of terrain navigation. This distribution is based on the maximum a posteriori (MAP). Since most navigation filters estimate at least six components (position and velocity), we show how the computation of the MAP can be reduced to finding a maximum in a two-dimensional space with the assumption that the prior distribution is Gaussian. The performance is illustrated on a specific scenario where severe divergence may occur. Section 2 recalls the basics of particle of filtering in the context of terrain navigation. Section 3 outlines the principle of importance sampling and provides three different versions of importance distributions. Section 4 is devoted to the derivation of the MAP under the hypothesis of a Gaussian a priori and a partially linear observation equation. Finally, section 5 illustrates the benefits of a MAP based proposal in a static setting as well as on a navigation filter scenario.

2. PARTICLE FILTERING IN THE CONTEXT OF TERRAIN NAVIGATION

Consider an aircraft equipped with an inertial navigation system (INS) which provides geographical position and velocity coordinates estimates as well as attitude components. The systems suffer from increasing estimation errors or drifts over time : it is essential to use an external sensor as a means of correcting these errors. We will assume that the aircraft uses a radar-altimeter as the external sensor and we are interested in the problem of estimating the position and velocity given the altimeter measurements. The radar-altimeter yields ground clearance measurements that can be compared to a database stored on board which will be referred to as the digital terrain elevation database (DTED).

2.1. Dynamic and observation model

For the sake of simplicity we will not detail the hybridization of the INS and state space representation. We refer the reader to [2] for more information. Let $x_k = \begin{pmatrix} x_{k,1} & x_{k,2} & x_{k,3} & \dot{x}_{k,1} & \dot{x}_{k,2} & \dot{x}_{k,3} \end{pmatrix}^T$ be the state vector composed of the position and velocity coordinates. We assume a near constant velocity model.

$$x_k = Fx_{k-1} + G_k w_{k-1} \tag{1}$$

where

$$F = \begin{pmatrix} I_3 & \Delta I_3 \\ 0_3 & I_3 \end{pmatrix} \quad G_k = \begin{pmatrix} \frac{\Delta^2}{2} I_3 \\ \Delta I_3 \end{pmatrix} \tag{2}$$

Δ is the sampling period, I_3 is the identity matrix and 0_3 is the 3×3 null matrix. The dynamic noise w_k is a vector of $\mathbb{R}^3$.

At each time step, the aircraft's radar-altimeter measures the ground clearance y_k i.e the distance between the aircraft and the ground according to the following equation

$$y_k = x_{k,3} - h_{DTED}(x_{k,1}, x_{k,2}) + v_k = h(x_k) + v_k \tag{3}$$

The term $x_{k,3}$ is the absolute altitude while $x_{k,1}, x_{k,2}$ are the horizontal position components and $h_{DTED}(x_{k,1}, x_{k,2})$ is the terrain height. v_k is an additive measurement noise which models the sensor imperfection. It is assumed that v_k follows zero mean Gaussian distribution with standard deviation σ_v.

In Bayesian estimation an estimator of the state x_k is obtained by approximating the filtering distribution $p_k = p(x_k|y_{0:k})$ where $y_{0:k}$ stands for the vector of measurements from time 0 up to time index k. Particle filtering is a class of sequential Monte Carlo algorithms that approximates p_k via a weighted mixture of Dirac measures.

$$p_k \approx \sum_{i=1}^{N} \omega_k^i \delta_{x_k^i} \tag{4}$$

$\left(\omega_k^i\right)_{i=1\ldots N}$ are the particle weights and the $\left(x_k^i\right)_{i=1\ldots N}$ are the so-called particles. To obtain the previous representation, a prediction step allows the particles to evolve according to the dynamics (1) and a correction step weights each x_k^i according to the likelihood function $g(x_k) = p(y_k|x_k)$. Suppose a particle approximation given by $\left(\omega_{k-1}^i, x_{k-1}^i\right)_{i=1\ldots N}$

1. *prediction* : at time k sample x_k^i according to $p(x_k|x_{k-1}^i)$
2. *correction* : compute the new weights according to $\omega_k^i = \omega_{k-1}^i g(x_k^i)$
3. normalize the weights $\omega_k^i = \frac{\omega_k^i}{\sum_{j=1}^N \omega_k^j}$
4. compute $N_{eff} = \frac{1}{\sum_{i=1}^N {\omega_k^i}^2}$
5. if $N_{eff} < N_{th}$, resample the particle set according to high/low weights

The above algorithm constitutes the SIR (Sequential Importance Resampling) recursion. However, in practice after a while particles tend to have a weight close to zero. The resampling step is necessary to avoid this particle degeneracy and consists in selecting the particles proportionally to their weight w_k^i so that only significant particles are duplicated and those with negligible weight are discarded. The decision to resample is made whenever the effective sample size N_{eff} falls below a certain predefined threshold N_{th}.

In practical applications, the SIR may not be able to track x_k successfully. Whenever the measurement noise is low, sampling according to the prior proposal $p(x_k|x_{k-1})$ may be inefficient since most particles will fall in regions where the likelihood $g(x_k)$ is close to zero. A solution is to choose a sampling distribution close to the posterior : the next section recalls basics of importance sampling and suggests a new proposal distribution.

3. CHOICE OF THE PROPOSAL DENSITY

In this section we consider a static setting where the unknown state x is distributed according to a prior q and is observed through a measurement $y = h(x) + v$, h being a non-linear function.

As in the previous section, the likelihood is denoted g so that the posterior $p(x|y) \propto g(x)q(x)$. Direct sampling from $p(x|y)$ generally not possible. Let P be the covariance of q.

3.1. Importance sampling

We wish to compute an estimate of the posterior $p(x|y)$. The importance sampling (I.S.) estimator $\hat{p}$ of $p(x|y)$ is obtained by drawing samples x^i from an alternative distribution $\tilde{q}$ so that

$$p \approx \sum_{i=1}^{N} w^i \delta_{x^i} \triangleq \hat{p} \tag{5}$$

where $\tilde{w}^i = \frac{g(x^i)q(x^i)}{\tilde{q}(x^i)}$ are the importance weights and $w^i = \frac{\tilde{w}^i}{\sum_{j=1}^N \tilde{w}^j}$.

Choosing a good importance distribution is crucial. Indeed, the Monte Carlo error of the estimator $\hat{p}$ is related to the asymptotic variance $V_{\tilde{q}}$ of the normalized importance weights w^i [4].

$$V_{\tilde{q}} = \frac{1}{N} \left(\frac{\int \frac{g^2(x)q^2(x)}{\tilde{q}(x)} dx}{\left(\int g(x)q(x)dx\right)^2} - 1 \right) \tag{6}$$

In case the prior $\tilde{q}$ and g have little overlap, most samples x^i have negligible weight and the above variance is high. Conversely, this variance is zero when $q = p(x|y)$.

3.2. Design of an importance distribution based on the maximum a posteriori

In light of the previous observation, we wish to design a proposal distribution $\tilde{q}$ close to the posterior. One way to do this is to use a Gaussian distribution centered on the mode $\hat{x}_{MAP}$

of $p(x|y)$ and with a covariance close to the posterior covariance $\mathrm{Cov}(X|Y)$. Two methods can approximate $\mathrm{Cov}(X|Y)$: Laplace's method [4], the inverse of the observed Fisher information matrix evaluated at the MAP [5](see eq. (7) below) denoted $J^{-1}(\hat{x}_{MAP})$ which is an approximation of Laplace's method. A necessary hypothesis is that the posterior be unimodal.

$$J(x) = -\frac{\partial^2 \log g}{\partial x^2}(x) - \frac{\partial^2 \log q}{\partial x^2}(x) \tag{7}$$

Laplace's method involves fourth order derivatives of h whereas J^{-1} uses second order derivatives. To simplify this even further we suggest the use of Cramer-Rao's lower bound where only the derivative $\frac{\partial h}{\partial x}$ is needed : indeed, the Cramer-Rao bound CRB is linked to J via $CRB^{-1} = \mathbb{E}_Y(J)$.
However, the use of $J^{-1}(\hat{x}_{MAP})$ can result in bad estimates. Indeed, if $J^{-1}(\hat{x}_{MAP})$ is too small w.r.t to P, $\frac{q}{\tilde{q}}$ is unbounded and the variance (6) can become infinite. To avoid this, we suggest a Gaussian proposal $\tilde{q}$ centered around the MAP with a covariance K that has eigenvalues greater than the prior covariance eigenvalues. To design a proposal even closer to the posterior, we propose that K be a rotated version of the prior covariance P, i.e. such that the corresponding ellipsoid has the orientation of that of $J^{-1}(\hat{x}_{MAP})$. In short, K has the shape of P and the orientation of $J^{-1}(\hat{x}_{MAP})$.
If E_J is the matrix of eigenvector of $J^{-1}(\hat{x}_{MAP})$ and Λ_P is the matrix of eigenvalues of P, the orientation adjustment is obtained by

$$K = E_J \Lambda_P E_J^T \triangleq P^r \tag{8}$$

The main difficulty is then to compute efficiently an estimate of the MAP of a six-dimensional state (3 position components, 3 velocity components).

4. MAXIMUM A POSTERIORI ESTIMATION UNDER GAUSSIAN PRIOR ASSUMPTION

In this section we will show how it is possible to compute an estimate of the MAP by assuming the a priori q density is Gaussian.

The maximum a posteriori is defined by :

$$\hat{x}_{MAP} = \arg\max p(x|y)$$

where the a posteriori density $p(x|y) \propto p(y|x)q(x)$

In some applications, the likelihood $p(y|x)$ only depends on a smaller part of x and can be broken down into a non-linear part and linear part. Assume the state vector is decomposed into 2 sub-states x_1 and x_2 i.e. $x = (x_1^T, x_2^T)^T$. Then the measurement equation becomes

$$y = Ax_2 - h_l(x_1) + v \tag{9}$$

where h_l is a non-linear function of x_1.
In map aided terrain navigation, x_1 is a two-dimensional vector of horizontal components of x and x_2 includes the altitude component as well as the three velocity components. The measurement equation (3) does indeed have the form of equation (9) where $A = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$ and $h_l = h_{DTED}$.

In the following, we will assume that the prior $q(x_1, x_2)$ is Gaussian with mean $\mu = (\mu_1^T, \mu_2^T)^T$ and covariance matrix

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}$$

and the likelihood $p(y|x)$ is a zero mean Gaussian distribution with covariance R. Notice that the posterior may be expressed as

$$\begin{aligned} p(x|y) \quad &\alpha \quad p(y|x_1, x_2)q(x_1, x_2) && (10) \\ &= \quad p(y|x_1, x_2)q(x_2|x_1)q(x_1) && (11) \end{aligned}$$

Using this factorization the maximization of the posterior leads to :

$$\max_{x_1, x_2} p(x_1, x_2|y) \propto \max_{x_1} q(x_1) \max_{x_2} \{p(y|x_1, x_2)q(x_2|x_1)\} \tag{12}$$

The maximization can then be carried out by first maximizing $F_1(x_1, x_2) = p(y|x_1, x_2)q(x_2|x_1)$ w.r.t x_2 which yields a function of x_1.
Maximizing F_1 w.r.t x_2 boils down to the minimization of the following function w.r.t x_2 :

$$\begin{aligned} Q(x_1, x_2) = &(y + h_l(x_1) - Ax_2) R^{-1} (y + h_l(x_1) - Ax_2)^T \\ &+ (x_2 - \mathbb{E}_1(x_2)) \Sigma_{2|1}^{-1} (x_2 - \mathbb{E}_1(x_2))^T \end{aligned}$$

where $\mathbb{E}_1(x_2) = \mathbb{E}(x_2|x_1)$ and $\Sigma_{2|1} = \mathrm{Cov}(x_2|x_1)$. We can reorder the terms above so that :

$$Q(x_1, x_2) = \Gamma(x_1) + (x_2 - \gamma(x_1)) \Omega^{-1} (x_2 - \gamma(x_1))^T \tag{13}$$

where

$$\begin{aligned} \Gamma(x_1) \quad &= \quad (y + h_l(x_1) - A\gamma(x_1)) R^{-1} (y + h_l(x_1) - A\gamma(x_1))^T \\ &+ \quad (\gamma(x_1) - \mathbb{E}_1(x_2)) \Sigma_{2|1}^{-1} (\gamma(x_1) - \mathbb{E}_1(x_2))^T \\ \gamma(x_1) \quad &= \quad \Omega \left[A^T R^{-1} (y + h_l(x_1)) + \Sigma_{2|1}^{-1} \mathbb{E}_1(x_2) \right] \\ \Omega \quad &= \quad \left(A^T R^{-1} A + \Sigma_{2|1}^{-1} \right)^{-1} \end{aligned}$$

The conditional expectation $\mathbb{E}_1(x_2)$ and covariance matrix $\Sigma_{2|1}$ can be worked out as :

$$\begin{aligned} \mathbb{E}_1(x_2) \quad &= \quad \mathbb{E}(x_2) + P_{21} P_{11}^{-1} (x_1 - \mathbb{E}(x_1)) && (14) \\ \Sigma_{2|1} \quad &= \quad P_{22} - P_{21} P_{11}^{-1} P_{12} && (15) \end{aligned}$$

Since the function Γ does not depend on x_2, the minimum of Q w.r.t x_2 is achieved in $\hat{x}_2 = \gamma(x_1)$. The next step to maximizing the posterior $p(x_1, x_2|y)$ is to substitute $\gamma(x_1)$ to x_2 in (13).

$$\begin{aligned} &\max_{x_1} q(x_1) \max_{x_2} \{p(y|x_1, x_2)q(x_2|x_1)\} \\ &\qquad = \max_{x_1} q(x_1) \exp(-\frac{1}{2}\Gamma(x_1)) \end{aligned} \tag{16}$$

Maximizing $q(x_1)\exp(-\frac{1}{2}\Gamma(x_1))$ can be achieved through numerical optimization methods. In the problem at hand, x_1 is a two-dimensional vector representing the horizontal position of the aircraft and the task of optimizing a function of two dimension has a significantly inferior computation cost than that of optimizing a 6-D function.
In a filtering setting, i.e. where we wish to estimate $p(x_k|y_{0:k})$ sequentially, $q = p(x_k|y_{0:k-1})$. In order to evaluate the importance weights $w_k^i \propto \frac{g(x_k^i)q(x_k^i)}{\tilde{q}(x_k{}^i)}$, computing $q(x_k^i)$ can be done by using a kernel density estimate of $p(x_k|y_{0:k-1})$. In our filtering simulations, for simplicity we assumed a Gaussian density q. The use of a Gaussian proposal has been investigated in [6] where the moments are estimated by an unscented filter. An other method is based on the local linearisation of the optimal proposal density $p(x_k|x_{k-1}, y_k)$ [7] but it requires the computation of a covariance for each particle.

5. SIMULATION RESULTS

5.1. Static scenario

We consider here the problem of estimating a vector $x = (x_1, x_2)$ in $\mathbb{R}^2$. The measurement function is given by a Gaussian pdf $h(x) = \frac{1}{2\pi|C|^{1/2}}\exp\left(-\frac{1}{2}(x-m)^T C^{-1}(x-m)\right)$ where $m = (1,1)^T$ and $C = \begin{pmatrix} 0.0075 & 0.0025 \\ 0.0025 & 0.0075 \end{pmatrix}$.
As usual, we assume that the observation equation takes the form $y = h(x) + v$, where $v \sim \mathcal{N}(0, \sigma_v^2)$ and that $x \sim \mathcal{N}(\mu, P)$.
The following values were used : $\mu = (0.94, 0.94)^T$, $P = \begin{pmatrix} 0.005^2 & 0.001^2 \\ 0.001^2 & 0.002^2 \end{pmatrix}$, $\sigma_v = 0.01$.

We wish to compare the importance sampling estimates of the posterior mean denoted by $\hat{x}^N = \sum_{i=1}^N w^i x^i$ obtained by using the following proposals $\tilde{q}$:

1. $\tilde{q}_0 = q$
2. $\tilde{q}_1 = \mathcal{N}(\hat{x}_{MAP}, P)$ where P is the prior covariance
3. $\tilde{q}_2 = \mathcal{N}(\hat{x}_{MAP}, J^{-1})$ where J defined in equation 7
4. $\tilde{q}_3 = \mathcal{N}(\hat{x}_{MAP}, P^r)$ where P^r is a rotated version of P (see eq.(8))

The comparison of these choices of importance distribution is carried out by computing the root mean squared errors (RMSE) $\mathbb{E}(\|\hat{x}^N - \hat{x}_{true}\|^2)^{1/2}$ as a function of the sample size N, using 100 independent runs. The value $\hat{x}_{true} = \mathbb{E}(X|Y)$ was computed by numerical integration. The results are summarized in figure 1. The light blue line corresponds to the distance between the MAP estimate and $\hat{x}_{true}$. The value $\sigma_v = 0.01$ causes a mismatch between the proposal and the posterior. It appears that the best estimator is $\tilde{q}_3$, centered on the MAP and with covariance obtained by tilting the prior covariance P in the direction of the posterior covariance which we approximate via (7) as can be seen in figure 1. In contrast, the use of $\tilde{q}_2$ does not guarantee the convergence of the I.S. estimator as evidenced by the RMSE plot.

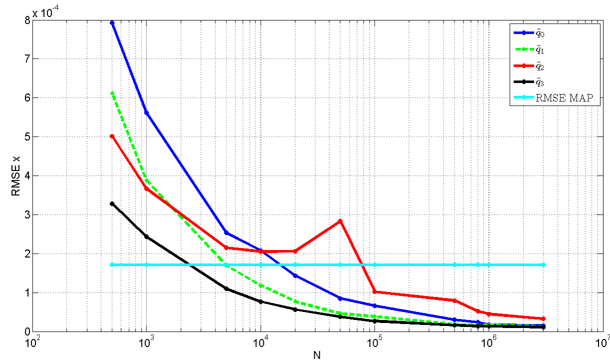
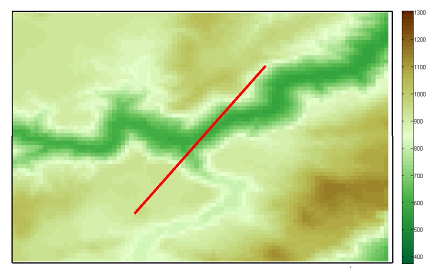
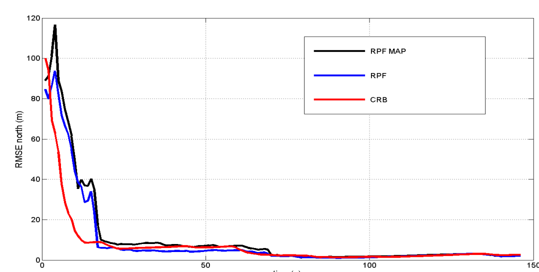


Fig. 1. Comparison of proposals $\tilde{q}$ ($\sigma_v = 0.01$)

5.2. Application to terrain navigation



(a) Terrain elevation and aircraft trajectory (red line)



(b) RMSE of $\hat{x}_1$ vs time

Fig. 2. Navigation filter : comparison of the standard RPF and RPF-MAP

We consider a navigation scenario where the aircraft will suddenly encounter a canyon. In this case, as explained previously the information contained in the measurement function is very high and most samples fall in a region where the likelihood is low. Instead of only sampling from the prior and weighting by the likelihood, we propose the following adaptation of the correction step which consists in sampling from the proposal $\tilde{q}_3$ defined in the previous section:

1. normalize the weights $\omega_k^i = \frac{\omega_k^i}{\sum_{j=1}^N \omega_k^j}$
2. compute $N_{eff} = \frac{1}{\sum_{i=1}^N {\omega_k^i}^2}$. If $N_{eff} < N_{min}$
   - compute the empirical mean x_{k^-} and covariance P_{k^-} of the predictive distribution $\hat{P}$ using $\left(\omega_{k-1}^i, x_k^i\right)$, $i = 1 \dots N$
   - compute $\hat{x}_{MAP}$ as in 4 using a Gaussian approximation $\mathcal{N}(x_{k^-}, P_{k^-})$ for q.
   - sample $x_k^i \sim \tilde{q}_3 = \mathcal{N}(\hat{x}_{MAP}, P_{k^-}^r)$ where $P_{k^-}^r$ is obtained by tilting P_{k^-} as in 3.2 and compute the weights $w_k^i = \frac{g(x_k^i)q(x_k^i)}{\tilde{q}_3(x_k^i)}$
3. resample if necessary

N_{min} is a threshold that triggers the use of the MAP in the correction step and can be different from the resampling threshold.

The navigation filter implemented here estimates position and velocity vector x_k defined in section 2. The initial state x_0 follows a Gaussian distribution $\mathcal{N}\left(\bar{x}_0, P_0\right)$ where $P_0 = \operatorname{diag}\begin{pmatrix}100^2 & 100^2 & 100^2 & 3^2 & 3^2 & 1^2\end{pmatrix}$. The prior mean $\bar{x}_0$ is sampled from a Gaussian distribution centered around the true state with covariance P_0. The dynamic noise is a zero mean Gaussian distribution with covariance $Q = \operatorname{diag}\begin{pmatrix}1^2 & 1^2 & 0.5^2\end{pmatrix}$. The radar altimeter measurement noise is also a zero mean Gaussian with standard deviation $\sigma_v = 1\,m$ and the sampling period is $\Delta = 0.1\,s$. The true aircraft's trajectory is straight and its duration is 14.5 seconds (145 measurements).
We compared a standard regularized particle filter (RPF) [8] with a version corresponding to the algorithmic description above which we denote RPF-MAP. The comparison was carried out using 150 independent runs using $N = 2000$ particles for the RPF-MAP and $N = 4000$ for the RPF which insures comparable computational load.Each filter was initialized with the same particle cloud. The RMSE of the convergent runs was computed and the results are available in figure 2. Out of the 150 runs, 67% were convergent for the standard RPF whereas 85% converged for the RPF-MAP. This illustrates the usefulness of the MAP and the covariance rotation according to the orientation of $\operatorname{Cov}(X|Y)$ as a way of sampling in the region of interest to mitigate the divergence phenomenon.

The algorithm, as currently formulated, assumes that the particle cloud has a single mode at each time step. In early stages of a typical navigation filter, this is generally not true. However, it is possible to use a mixture model as in [9]: $p(x_k|y_{0:k}) = \sum_{l=1}^L \alpha_{l,k} p_l(x_k|y_{0:k})$ where each component is approximately unimodal and make a Gaussian approximation for each p_l . In this case, a local importance distribution $\tilde{q}_l$ based on the local MAP can be used to sample more efficiently.

6. CONCLUSION

In this article, we have provided a methodology for designing an importance sampling method based on the maximum a posteriori. The main contribution is the reduction of the dimension of the underlying optimization problem and the use of a proposal centered on the MAP with a suitably chosen covariance. Simulations in a difficult terrain navigation setting show the interest of this method. Still, it remains that in order to take advantage of the developed method, it is necessary to partition the filter into unimodal components .

7. REFERENCES

[1] N. Bergman, L. Ljung, and F. Gustafsson, “Terrain navigation using bayesian statistics,” *Control Systems, IEEE*, vol. 19, no. 3, pp. 33–40, 1999.

[2] P.-J. Nordlund and F. Gustafsson, “Marginalized particle filter for accurate and reliable terrain-aided navigation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1385 –1399, October 2009.

[3] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, “Novel approach to nonlinear/non-gaussian bayesian state estimation,” *Radar and Signal Processing, IEE Proceedings F*, vol. 140, no. 2, pp. 107 –113, April 1993.

[4] P. Bui Quang, C. Musso, and F. Le Gland, “Multidimensional Laplace Formulas for Nonlinear Bayesian Estimation,” in *Proceedings of the 20th European Signal Processing Conference (EUSIPCO 2012),* Bucharest, Romania, August 27–31, 2012.

[5] B. Jungbacker and S. J. Koopman, “Monte carlo estimation for nonlinear non-gaussian state space models,” *Biometrika*, vol. 94, no. 4, pp. 827–839, 2007.

[6] R. Van der Merwe, N. de Freitas, and E. Doucet, A.and Wan, “The Unscented Particle Filter,” in *Advances in Neural Information Processing Systems 13*, Nov. 2000.

[7] A. Doucet, S. Godsill, and Andrieu C., “On sequential monte carlo sampling methods for bayesian filtering,” *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[8] A. Doucet, N. De Freitas, and N. Gordon, Eds., *Sequential Monte Carlo methods in practice*, Springer, 2001.

[9] A. Murangira, C. Musso, K. Dahia, and J. Allard, “Robust regularized particle filter for terrain navigation,” in *Proceedings of the 14th International Conference on Information Fusion (FUSION)*, July 2011.