Restrictive Information and Kalman-type Filters for Hybrid Positioning

Simo Ali-Löytty
Institute of Mathematics, Tampere University of Technology
simo.ali-loytty@tut.fi

Abstract
This paper presents a new algorithm whereby restrictive information $Bx \leq \beta$ can be used in a Kalman-type filter. We also present hybrid positioning simulations. Simulations show that this new algorithm is fast to compute and gives almost the same accuracy as the particle filter with millions of particles. We found that in some cases restrictive information such as mobile phone network sector and maximum range information dramatically improve filter accuracy.

1 Introduction
Hybrid positioning means positioning using measurements from different sources e.g. GPS and a cellular network. Range, pseudorange, deltarange, altitude, restrictive and compass measurements are examples of typical measurements in hybrid positioning. By restrictive information, we mean the knowledge that the state $x$ e.g. position is inside some area, which is the most often a polyhedron $Bx \leq \beta$. In the simplest case, the area is a half-space. Base station sector and maximum range information (Cell ID) are examples of restrictive information [9].

Filters are used to compute an estimate of the state using current and past measurement data. Filters also give an approximation of the error covariance matrix. Kalman-type filters approximate the probability distribution of the state as a Gaussian. Extended Kalman Filter (EKF) [5, 7], Second Order Extended Kalman Filter (EKF2) [3] and Unscented Kalman Filter (UKF) [6] are examples of Kalman-type filters.

An outline of the paper is as follows. After Bayesian filter basics and problem formulation in Section 2, the basic idea of the new algorithm is presented in Section 3. Section 4 provides the mathematical fundamentals of the algorithm. In Section 5, we concentrate on hybrid positioning and illustrate how the new algorithm works in practice with sector and maximum range information. Finally, simulation results are given in Section 6.

2 Bayesian Filtering
We consider the discrete-time non-linear non-Gaussian system
\begin{align}
    x_k &= f_{k-1}(x_{k-1}) + w_{k-1}, \quad (1) \\
    y_k &= h_k(x_k) + v_k, \quad (2)
\end{align}
where the vectors $x_k \in \mathbb{R}^{n_x}$ and $y_k \in \mathbb{R}^{n_y}$ represent the state of the system and the measurement at time $t_k$, $k \in \mathbb{N}$, respectively. We assume that errors $w_k$ and $v_k$ are white, mutually independent and independent of the initial state $x_0$. The aim of the filtering is to find conditional probability density function (posterior)

\[ p(x_k|y_{1:k}) , \]
where $y_{1:k} \triangleq \{y_1, \ldots, y_k\}$. Posterior can be determined recursively according to the following relations.

\textit{Prediction (prior):}

\[ p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1}; \quad (3) \]
\[
\begin{align*}
\text{Update (posterior)}: \\
p(x_k|y_{1:k}) = & \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{\int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k},
\end{align*}
\]

where the transition pdf \(p(x_k|x_{k-1})\) can be derived from (1) and the likelihood \(p(y_k|x_k)\) can be derived from (2). The initial condition for the recursion is given by the pdf of the initial state \(p(x_0|y_{1:0}) = p(x_0)\). Knowledge of the posterior distribution enables one to compute an optimal state estimate with respect to any criterion. For example, the minimum mean-square error (MMSE) estimate is the conditional mean of \(x_k\) [3, 8], in general and in our case the conditional probability density function cannot be determined analytically. Extended Kalman Filter (EKF) is a traditional approximation of Bayesian filter. EKF is Kalman filtering applied to a linearization of system (1), (2). EKF is very commonly used in satellite-based positioning [7]. The EKF algorithm is described for example in [3, 5].

### 2.1 Restrictive information

Restrictive information is a measurement with measurement function (2)

\[
h(x) = \begin{cases} 
1, & \text{if } x \in A \\
0, & \text{if } x \notin A 
\end{cases},
\]

where \(A \subset \mathbb{R}^{n_x}\). In this paper, we consider only a case where restrictive information does not have error, so then error term \(v = 0\) is a constant random variable. The likelihood function of restriction information is the characteristic function

\[
p(1|x) = \chi_A(x) = \begin{cases} 
1, & \text{if } x \in A \\
0, & \text{if } x \notin A 
\end{cases},
\]

and respectively \(p(0|x) = \chi_{\neg A}(x)\), where \(\neg\) is the complement of \(A\). Unfortunately, it is not straightforward to use restrictive information with Kalman-type filters because measurement is a dichotomy variable and the measurement function is non-linear.

### 3 New algorithm in a nutshell

The main idea of the new algorithm is that we use some Kalman-type filter and extend it to use restrictive information.

1. Compute the posterior without restrictive information using Kalman-type filter.
2. If there is new restrictive information, model restrictive information as a likelihood function that is one inside a certain polyhedron and zero outside.
3. Compute new mean and covariance estimates.
4. Approximate the posterior distribution with a Gaussian, with the new mean and covariance.
5. Repeat this every timestep.

The heart of this algorithm is in stages (2-3), which present a fast way to approximate integrals

\[
\int_A xp(x)dx \text{ and } \int_A (x - \mu)(x - \mu)^T p(x)dx,
\]

where \(p(x)\) is multinormal density function. Notice that this algorithm does not require any other measurements, so the algorithm works also if there is only restriction information.

### 4 Derivation of the algorithm

In this section, we concentrate on the mathematical fundamentals of the new algorithm. First of all, the algorithm changes only the update part (4) so prediction part (3) remains the same as in the ordinary filter.

1. **Compute the posterior without restrictive information using Kalman-type filter.**

Because of the independence, we can write likelihood function as the product

\[
p(y_k|x_k) = \chi_A(x_k)p(y'_k|x_k),
\]

where \(A\) is the intersection of all new restrictive information, this means that state is inside \(A\) at time instant \(t_k\) and \(y'_k\) are other measurements than restrictive information at time instant \(t_k\). Using Eqs. (4) and (6) we find that

\[
p(x_k|y_{1:k}) \propto \chi_A(x_k)p(x_k|y'_{1:k}),
\]
where

\[ p(x_k|y'_{1:k}) = \frac{p(y'_k|x_k)p(x_k|y_{1:k-1})}{\int p(y'_k|x_k)p(x_k|y_{1:k-1})dx_k} \]

is posterior without current restrictive information. This can be computed using the ordinary Kalman-type filter. We use notation that \( \mu_\text{old} \) and \( \Sigma_\text{old} \) represent the approximations of the mean and the covariance matrix of this distribution, respectively.

We also approximate \( p(x_k|y'_{1:k}) \) with a Gaussian.

2. If there is new restrictive information, model restrictive information as a likelihood function that is one inside a certain polyhedron (see Fig 1)

\[ p(x_k|y_{1:k}) \]

and zero outside.

We use the same restrictive information only once within a certain time period, because otherwise our approximation error can accumulate and cause unwanted phenomena. Let the area \( \mathcal{A} \) be inscribed in the polyhedron (see Fig 1)

\[ B = \left\{ x \mid \begin{bmatrix} b_1^T \\ \vdots \\ b_n^T \end{bmatrix} x \leq \beta \right\} = \{ x | Bx \leq \beta \}, \tag{8} \]

where matrix \( B \) is organized so that the first inequality reduces probability the most, that is

\[ \frac{\beta_1 - b_1^T \mu_\text{old}}{\sqrt{b_1^T \Sigma_\text{old} b_1}} \leq \cdots \leq \frac{\beta_n - b_n^T \mu_\text{old}}{\sqrt{b_n^T \Sigma_\text{old} b_n}}. \]

We use Gram-Schmidt orthonormalization on \( \{b_1, \ldots, b_n\} \) with inner product

\[ < x, y > = x^T \Sigma_\text{old} y, \]

and get matrix \( A \) whose row space is same than the row space of matrix \( B \) and \( A \Sigma_\text{old} A^T = I \). Now we approximate the likelihood of restrictive information with one inside the polyhedron (see Fig 1)

\[ A' = \{ x \mid |Ax - Ax_{\text{mid}}| \leq \alpha \} \tag{9} \]

and zero elsewhere. So we replace \( \chi_A \) in Eq. (7) with \( \chi_{A'} \). Vectors \( Ax_{\text{mid}} \) and \( \alpha \) we select so that \( B \) is subset of \( A' \) and probability that state is inside \( A' \) is as small as possible. Then also \( A \) is subset of \( A' \). If

\[ x = [x_1, \ldots, x_N] \]

are vertices of polyhedron \( B \), then

\[ \alpha = \frac{up - low}{2} \]

and

\[ Ax_{\text{mid}} = A\mu_\text{old} - \frac{up + low}{2}, \tag{11} \]

where

\[ up = A\mu_\text{old} - \min_{i \in \{1, \ldots, N\}} (AX_{e_i}) \]

and

\[ low = A\mu_\text{old} - \max_{i \in \{1, \ldots, N\}} (AX_{e_i}). \]

Operators min and max are taken separately with every components. We use this same notation through the paper.

3. Compute new mean and covariance estimates. In this stage, we compute the mean \( \mu_\text{new} \) and the covariance matrix \( \Sigma_\text{new} \) of the posterior approximation

\[ p(x_k|y_{1:k}) \approx \frac{\chi_{A'}(x_k)p(x_k|y'_{1:k})}{p_{\text{new}}}, \]

where \( A \) is replaced with \( A' \). This approximation is accurate if the old posterior \( p(x_k|y'_{1:k}) \) is Gaussian and \( A = A' \). Here \( p_{\text{new}} \) is probability that state is inside polyhedron \( A' \),

\[ p_{\text{new}} = \int \chi_{A'}(x_k)p(x_k|y'_{1:k})dx_k \]

\[ = \int_{|x| \leq \alpha} p_\text{old}(z)dz \tag{12} \]

\[ = \prod_{i=1}^n (\Phi(u_{i}) - \Phi(low_{i})) \]

where \( \Phi(x) \) is cumulative density function (cdf) of \( x \), when \( x \sim N(0,1) \). The idea of the solution is that with the change variables,

\[ z = Ax_k - Ax_{\text{mid}}, \]

\[ E(z) = \mu_z = A\mu_\text{old} - Ax_{\text{mid}}, \tag{13} \]

\[ V(z) = \Sigma_z = A\Sigma_\text{old}A^T = I, \]

we can compute the integral iteratively. We use this same idea also in the following integrals, when computing mean and covariance matrix. If posterior is
Gaussian mixture then the new weight is proportional to old weight times the probability $p_{new}$.

The approximation of the mean of the posterior distribution is

$$
\mu_{\text{new}} = \int x_k \frac{\chi_{A}(x_k)p(x_k | y_{1:k})}{p_{\text{new}}} dx_k = \frac{A^{-1}}{p_{\text{new}}} \int_{|z| \leq \alpha} z p_{\alpha}(z) dz + x_{\text{mid}}
$$

$$
= \frac{A^{-1}}{p_{\text{new}}} \int z p_{\alpha'}(z) dz + x_{\text{mid}}
$$

$$
= \mu_{\text{old}} + \Sigma_{\text{old}} A^T \epsilon
$$

where

$$
p_{\alpha'}(z) = \frac{\chi_{|z| \leq \alpha}(z)p_{\alpha}(z)}{p_{\text{new}}}
$$

and

$$
\epsilon = \sum_{i=1}^{n} \frac{\exp \left( -\frac{u_i^2}{2} \right) - \exp \left( -\frac{low_i^2}{2} \right)}{\sqrt{2\pi} (\Phi (up_i) - \Phi (low_i))}.
$$

Details of the computation are given in Appendix A, Eq. (17). This equation also works when matrix $A$ is not a square matrix. Proof goes similarly than previously.

4. Approximate the posterior distribution with a Gaussian, with the new mean and covariance.

This stage usually produces the most approximation errors in this algorithm. This is the reason why we use each restrictive information only once.

5. Repeat this every timestep.

5 Example of the new algorithm

In this section, we illustrate how the new algorithm works. In this example, Fig 1, restrictive information consists of two base station Cell IDs, which means that we know the base stations sector and maximum range information. In the figure, dark area (area $A$ in Eq. (7)) is the intersection of this information and so we know that user is inside this area. Black dashed polyhedron represents an polyhedron approximation of the true restrictive area (8). Black solid polyhedron represents the certain polyhedron approximation of the polyhedron approximation of the true restrictive area (9). Red man stands on prior mean and a red dashed ellipse represent prior covariance. The covariances ($\Sigma$) are visualized with ellipses that satisfy the equation

$$(x - \mu)^T \Sigma^{-1} (x - \mu) = 2.2173.$$  

The constant was chosen so that if the distribution is Gaussian then there is 67% probability mass inside the ellipse. Respectively, blue man stands on posterior mean and a blue solid ellipse represent posterior covariance.

It is quite evident that using restrictive information improves estimates notably only when there is much probability mass outside the restrictive area. Usually, this happens two different cases. The first case is that we have only a few measurement, which
Figure 1: Red means prior distribution. Dashed black polyhedron represents the first approximation of restrictive area (8). Solid black polyhedron represents the certain approximation of the true restrictive area (9). Blue represents the mean and covariance approximations after applying the restrictive information.

causes large posterior covariance without restrictive information, but we have quite good restrictive information such as sector information. Other possibly is that Kalman-type filters works totally wrong, this means that filter is inconsistent so that predicted errors are smaller than actual errors [2]. Both of these cases are possible in hybrid positioning, so it is reasonable to use restrictive information in hybrid positioning. In Section 6, we see more specifically how restrictive information improves Kalman-type filter performance in hybrid positioning.

6 Simulations

In the simulations, we use the position-velocity model, so the state \( \mathbf{x} = \begin{bmatrix} \mathbf{r}_u \\ \mathbf{v}_u \end{bmatrix} \) consists of user position vector \( \mathbf{r}_u \) and user velocity vector \( \mathbf{v}_u \) and state-dynamic Eq. (1) is linear (same as in paper [1]). We use base station range measurement and altitude measurement
\[
\begin{align*}
y_b &= \| \mathbf{r}_b - \mathbf{r}_u \| + \epsilon_b, \\
y_a &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{r}_u + \epsilon_a,
\end{align*}
\]
where \( \mathbf{r}_b \) is base station position vector and \( \epsilon \)s are error terms. Restrictive information that we use are base station sector information and maximum range information.

Figure 2: This figure illustrate how the new algorithm, EKF with restrictive information (EKF-BOX), improve EKF in mobile phone positioning. In this case EKF and EKFBOX mean of the position errors are 442 m and 93 m, respectively.

From Fig. 2, we get an idea of how the new algorithm works. These simulations use only a few (one
or two at the same time instant) base station range measurements with variance \((100 \text{ m})^2\) and altitude measurement with variance \((300 \text{ m})^2\). In this case, restrictive information keeps the filter always consistent (not inconsistent). In this paper, we use the general inconsistency test, with risk level 5% [2]. Restrictive information also decreases the mean error from 442 meters to 93 meters.

In Table 1, we have listed a summary of a hundred 300 second simulations. The simulations use only a few (one or two at the same time instant) base station range measurements with variance \((80 \text{ m})^2\) and altitude measurements with variance \((300 \text{ m})^2\). Summary consist of following columns: \(\text{Time}\) is computation time using Matlab in our implementation, scaled so that computation time of EKF is 1. This gives a rough idea of the time complexity of each algorithm. \(\text{Err. } \mu\) is 2D position error mean. \(\text{Err. } 95\%\) gives a radius containing 95% of the 2D errors. \(\text{Err. ref.}\) is 2D error to reference posterior mean, which is computed using particle filter with \(10^6\) particles. \(\text{Inc. } \%\) is how many percentage of time filter is inconsistent with respect to the general inconsistency test, with risk level 5% [2]. Solvers are sorted so that mean positioning errors are in descending order. We also test positioning using only restrictive information, we call this the BOX-solver. We also use "BOX" suffix when we use the new algorithm for incorporating restrictive information.

Table 1: Summary of 100 different simulations with very poor geometry. Simulations only use a few base station range measurements and very inaccurate altitude measurements; an example of this geometry is in Fig. 2.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EKF</td>
<td>1</td>
<td>393</td>
<td>1356</td>
<td>358</td>
<td>28.1</td>
</tr>
<tr>
<td>BOX</td>
<td>1</td>
<td>314</td>
<td>669</td>
<td>253</td>
<td>0.1</td>
</tr>
<tr>
<td>EKFBOX</td>
<td>2</td>
<td>192</td>
<td>625</td>
<td>126</td>
<td>7.7</td>
</tr>
<tr>
<td>SMC</td>
<td>74</td>
<td>145</td>
<td>436</td>
<td>17</td>
<td>0.2</td>
</tr>
<tr>
<td>Ref</td>
<td>(\infty)</td>
<td>145</td>
<td>423</td>
<td>0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

From Table 1, we see that restrictive information improves greatly EKF performance. Every statistic, except CPU time, of the EKF is improved by the new algorithm. However EKFBOX-solver’s computation time is much less than the particle filter (SMC) [4, 8]. SMC and reference solvers are the same solver, but SMC solver uses \(10^4\) particles, when reference solves uses \(10^6\) particles. Note that the geometry is quite bad and traditional solvers like EKF give even worse results than BOX-solver, which uses only restrictive information and not range measurements or altitude measurements at all.

More simulation results can be found in [1].

7 Conclusions

We presented a new hybrid positioning algorithm, which makes possible to use restrictive information with Kalman-type filters. The new algorithm needs clearly less computation time than the general nonlinear Bayesian filters such as the particle filter with \(10^4\) particles. Simulations show that the new algorithm is very well suited to hybrid positioning application, where we have sector and maximum range restrictive information. In this application, the new algorithm improves EKF much and gives almost same accuracy as reference particle filters.

A Auxiliary integrals

In this section, we compute some auxiliary integrals. Here \(z\) is \(n_z\)-dimensional Gaussian random variable with mean \(\mu_z\) and covariance matrix is identity matrix. So the density function of \(z\) is

\[
p_z(z) = \prod_{i=1}^{n_z} p_{z_i}(z_i) = \prod_{i=1}^{n_z} \exp\left(\frac{-(z_i-\mu_z_i)^2}{2}\right)
\]

and so

\[
p_z^r(z) = \frac{\chi_{|z|\leq \alpha}(z)p_z(z)}{p_{\text{new}}} = \prod_{i=1}^{n_z} \chi_{|z_i|\leq \alpha_i}(z_i) \exp\left(\frac{-(z_i-\mu_z_i)^2}{2}\right)
\]

\[
= \prod_{i=1}^{n_z} \left(\Phi(u_{pi}) - \Phi(l_{oi})\right) \sqrt{2\pi}
\]

\[
= \prod_{i=1}^{n_z} p_{z_i}(z_i),
\]

which is also a density function.
Auxiliary integral of derivation of Eq. (14)

\[ \mu_{x'} = \int z p_{x'}(z)dz \]

\[ = \int (z - \mu_x) \prod_{i=1}^{n_x} \phi(z_i)dz + \mu_x \]

\[ = \sum_{j=1}^{n_x} \epsilon_j \int (z_j - \mu_{z_j}) \prod_{i=1}^{n_x} \phi(z_i)dz_j + \mu_x \]

\[ = \sum_{j=1}^{n_x} \epsilon_j \int \frac{\phi(z_j - \mu_{z_j})}{\Phi(\epsilon_{up_j} - \Phi(\mu_{low_j}))} + \mu_x \]

\[ = \sum_{j=1}^{n_x} \epsilon_j \sqrt{\frac{\epsilon_j^2}{\pi}} - \exp \left( -\frac{\epsilon_j^2}{2} \right) + \mu_x \]

\[ = \epsilon + \mu_x, \]

where \( p_{x_{\text{new}}} \) is defined in Eq. (12) and from Eqs. (10), (11) and (13) we get that

\[ up = \mu_x + \alpha \quad \text{and} \quad low = \mu_x - \alpha. \]

Auxiliary integral of derivation of Eq. (15)

\[ \Sigma_{x'} = \int (z - \mu_{x'})(z - \mu_{x'})^T p_{x'}(z)dz \]

\[ = \sum_{j=1}^{n_x} \epsilon_j \epsilon_j^T \int (z_j - \mu_{x_j})^2 \phi(z_j)dz_j \]

\[ = \sum_{j=1}^{n_x} \epsilon_j \epsilon_j^T \int (z_j - \mu_{z_j} - \epsilon_j)^2 \phi(z_j)dz_j \]

\[ = \sum_{j=1}^{n_x} \epsilon_j \epsilon_j^T \int (z_j - \mu_{z_j})^2 \phi(z_j)dz_j - \text{diag}(\epsilon)^2 \]

\[ = I - \text{diag}(\delta) - \text{diag}(\epsilon) \text{diag}(A_{\text{mid}} - A_{x_{\text{mid}}}) \]

\[ (17) \]

Acknowledgments

This study was partly funded by Nokia Corporation. Simo Ali-Löytty acknowledges the financial support of the Nokia Foundation and the Tampere Graduate School in Information Science and Engineering. The particle filter we used was implemented by Duane Petrovich.

References


