Extended Target Tracking using Principal Components

Johan Degerman, Johannes Wintenby
Electronic Defense Systems
Saab AB
Göteborg, Sweden
Email: johan.degerman@saabgroup.com

Daniel Svensson
Dept. of Signals and Systems
Chalmers University of Technology
Göteborg, Sweden
Email: daniel.svensson@chalmers.se

***Abstract*—The increased resolution in today's radar systems enables tracking of small targets. However, tracking both small and large targets in a dense target scenario raises considerable challenges. The data association of tracks to measurement groups is highly dependent on good target extension models for filtering and likelihood computation. In our attempt to design a tracker for extended targets, we start by adopting the results from the technique referred to as random matrices, which enables us to separate the filtering into an extension and a kinematical part. We re-define the measurement model and discard the assumption of independent Gaussian-distributed plots. Instead we assume the principal components to be Gaussian distributed. Then, through a heuristic approach, we create a two-stage Kalman filter, where the first stage estimates the principal components, and the second stage estimates the centre of gravity, using the output from the first stage as measurement uncertainty. The advantage of having a Kalman filter with data-driven measurement noise over a standard Kalman filter is demonstrated using simulated data, where a significant improvement in kinematical accuracy is shown.**

Keywords: Target tracking, extended targets, random matrices, principal components, Kalman filtering.

I. Introduction

Extended target tracking is increasingly studied in radar signal processing research. The main reason is the desire to track both small and large objects. To distinguish between small objects, the resolution is increased. That causes problems when tracking large objects since the point-target assumption no longer holds, i.e. each target is not represented by only one plot. Given the point-target assumption, the tracking problem is limited to association and filtering of target kinematics. By relaxing this constraint, we add the problem of determining which plots make up the target. Unless special care is taken, large objects will produce erroneous tracks.

One situation when the tracking of both small and large objects is important is sea or harbour surveillance, particularly for seaborne radar. We may require the ability to track small rafts at large distances, while having larger freighters passing nearby. Also, we may want to detect small boats docking or leaving large ships, which means that the resolution problem coincides with the dense target tracking problem. Typically, large ships will either produce several tracks, or be represented by erroneous tracks updated with plots alternating between fore and aft, thus yielding highly unreliable velocity vectors. Either way, we will need to improve the maritime picture if we require reliable situational awareness.

The literature contains many publications addressing the problem of tracking extended objects, as the image analysis community has dealt with this problem since the dawn of video tracking [1], [2], [3], [4]. In optical video the update rate is high, making the problem of finding *what is* the object as the primary concern, and hence making motion modelling less important. Radar trackers, on the other hand, traditionally have the opposite problem of finding *where is* the target since the update rate as well as the resolution is low in comparison. Tracking of extended targets, or target groups, is sometimes regarded as cluster tracking, and Bayesian solutions have recently been developed using random matrices [5]. Our work is very influenced by [5], which contains an in-depth description of the theory concerning Bayesian filtering for extended targets. It does not cover the data association part of tracking, however. Interesting examples concerning the combined association of plot-to-cluster and cluster-to-track can be found in [6], [8], [7], which use the multi-hypothesis framework to handle clustering ambiguities.

The outline of the paper is as follows. In Section II, we describe the random matrices approach and how our approach differs from [5]. Further, we propose an alternative measurement model, and postulate that the centre of gravity and extension should be independent, rather than the plots themselves. In Section III, we describe our extension to the random matrices using principal components, and discuss how it transforms the problem into the tracking of the physical entities length, width, and orientation, in the ordinary state vector format. From this we develop a two-stage Kalman filter that is evaluated using synthetic data in Section V. The results show improved performance compared to an ordinary Kalman filter, and to the approach in [5].

II. Random matrices

The basic idea of having a random matrix is to be able to adapt the uncertainty of target position to incoming data. When targets cannot be modelled as points, the concept of measurement uncertainty has to be redefined. We will assume that, in the standard Kalman filter terminology, the measurement noise for the centre of gravity is directly proportional to the

978-0-9824438-2-8/11/$26.00 ©2011 IEEE

extension. Thus, we are trying to find a way to recursively estimate the measurement noise for the centre of gravity. However, data-driven noise estimation is not a trivial task, since it affects the two-way coupling between track filtering to plot association. While adaption of the process noise will primarily affect the prediction and gating, adaption of the measurement noise will affect the measurement likelihood and hence primarily influence the data association. Naturally, both the process noise and the measurement noise will be involved in both track filtering and association, but what we want to stress here is the risk for instabilities when adding data-driven measurement noise to a tracker already utilizing data-driven process noise in some form.

We will denote the set of n_k measurements (plots) at time t_k by $\mathbf{z}_k$, and all sets of measurements up to time t_k by $\mathbf{Z}_k$. Further, the kinematic states of a target are captured by the state vector $\mathbf{x}_k$, and the target extent is modeled by a positive semi-definite random matrix $\mathbf{X}_k$. The problem that we study is hence the calculation (approximation) of the joint posterior density $p\big(\mathbf{x}, \mathbf{X} \big| \mathbf{z}_k, \mathbf{Z}_{k-1}\big)$.

Just as in [5], we rewrite the joint posterior density using Bayes' rule as

$$p\left(\mathbf{x}_k, \mathbf{X}_k | \mathbf{Z}_k\right) = p\left(\mathbf{x}_k, \mathbf{X}_k | \mathbf{z}_k, \mathbf{Z}_{k-1}\right) = \frac{p\left(\mathbf{z}_k | \mathbf{x}_k, \mathbf{X}_k\right) p\left(\mathbf{x}_k, \mathbf{X}_k | \mathbf{Z}_{k-1}\right)}{p\left(\mathbf{z}_k | \mathbf{Z}_{k-1}\right)}. \quad (1)$$

Further, we make the same simplification as [5] on the joint prior density, by stating that the temporal change of the extension does not affect the predicted density of the kinematical state. It is shown in [5] that we can use this approximation to compute the prior densities for $\mathbf{x}_k$ and for $\mathbf{X}_k$ separately

$$p\left(\mathbf{x}_k, \mathbf{X}_k, \mathbf{Z}_{k-1}\right) \approx \int \underbrace{p\left(\mathbf{x}_k | \mathbf{X}_k, \mathbf{x}_{k-1}\right)}_{\text{motion model}} \underbrace{p\left(\mathbf{x}_{k-1} | \mathbf{X}_k, \mathbf{Z}_{k-1}\right)}_{\approx\text{previous update}} d\mathbf{x}_{k-1} \cdot \int \underbrace{p\left(\mathbf{X}_k | \mathbf{X}_{k-1}\right)}_{\text{transition model}} \underbrace{p\left(\mathbf{X}_{k-1} | \mathbf{Z}_{k-1}\right)}_{\text{previous update}} d\mathbf{X}_{k-1}. \quad (2)$$

This approximation is visualized in the Bayesian network in Figure 1, where the diagonal dependence $\mathbf{X}_{k-1} \rightarrow \mathbf{x}_k$ is moved to $\mathbf{X}_k \rightarrow \mathbf{x}_k$.

The difference between the approach of this paper and the one in [5] starts with the likelihood function $p\left(\mathbf{z}_k | \mathbf{x}_k, \mathbf{X}_k\right)$. In [5] it is assumed that the n_k plots in $\mathbf{z}_k$ at time t_k are independently Gaussian-distributed around the centre position $\mathbf{x}_k$, with covariance $\mathbf{X}_k$. They rewrite the product of Gaussians in terms of the measured centre of gravity $\bar{z}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} \mathbf{z}_k^{(i)}$ and the scatter matrix $Z_k = \sum_{i=1}^{n_k} \left(\mathbf{z}_k^{(i)} - \bar{z}_k\right)^2$ as

$$p\left(\mathbf{z}_k | \mathbf{x}_k, \mathbf{X}_k\right) = \prod_{i=1}^{n_k} \frac{1}{\left|4\pi^2 \mathbf{X}_k\right|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left(z_i - \mathbf{x}_k\right)^\intercal \mathbf{X}_k^{-1}\left(z_i - \mathbf{x}_k\right)\right) = \frac{1}{\left|4\pi^2 \Sigma\right|^{\frac{1}{2}}} \operatorname{exptr}\left(-\frac{1}{2}\left(\left(\bar{z} - \mathbf{x}_k\right)\left(\bar{z} - \mathbf{x}_k\right)^\intercal\right)\left(\frac{\mathbf{X}_k}{n_k}\right)^{-1}\right) \cdot \frac{1}{\left|4\pi^2 \Sigma\right|^{\frac{n_k - 1}{2}}} \operatorname{exptr}\left(-\frac{1}{2} Z \mathbf{X}_k^{-1}\right), \quad (3)$$

which means that the centre of gravity is Gaussian with mean $\mathbf{x}_k$ and covariance $\frac{\mathbf{X}_k}{n_k}$, and that the scatter matrix has a distribution proportional to the Wishart distribution with $n_k - 1$ degrees of freedom. We view n_k as a deterministic parameter, hence it is not being conditioned on in the declaration of the likelihood function. By looking at equation (3) we can identify the measurement equation as a product of the likelihood for two independent observables $\bar{z}_k$ and Z_k, where the probability density for the scatter matrix Z_k is not dependent on the kinematical state $\mathbf{x}_k$. The likelihood is hence

$$p\left(\mathbf{z}_k | \mathbf{x}_k, \mathbf{X}_k\right) = p\left(\bar{z}_k | \mathbf{x}_k, \mathbf{X}_k\right) p\left(Z_k | \mathbf{X}_k\right). \quad (4)$$

As in [5] we consider Gaussian-distributed plots, but instead of assuming that each plot $\mathbf{z}_k^{(i)}$ at time t_k is independent we assume that *the measured position and extension are independent*. Here, by position we mean centre of gravity. In fact, we do not consider the individual plots at all. The position measurement consists of an estimated centre of gravity $\bar{z}_k$, which is modelled as a Gaussian with mean $\mathbf{x}_k$ and covariance $\frac{\mathbf{X}_k}{n_k}$. The extension measurement is the scatter matrix Z_k,

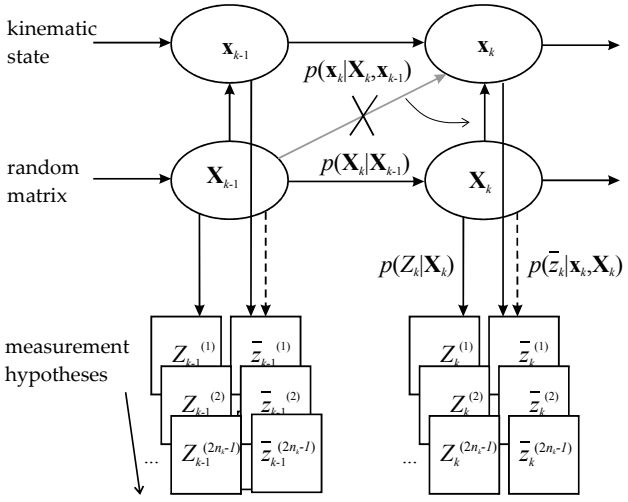


Figure 1. Bayesian network for the random matrix and state vector. State and random matrix transition models and measurement likelihood models are shown. The approximation in (2) is visualized by the move of the diagonal arrow. The reason for using a dashed line for the coupling from random matrix $\mathbf{X}_k$ to measurement $\bar{z}_k$ is explained in Section IV.

whose likelihood is represented by a Wishart distribution with covariance matrix $\mathbf{X}_k$ and $n_k - 1$ degrees of freedom, instead of the Wishart proportional likelihood (3) as in [5]. Centering the scatter matrix to $\bar{z}_k$ reduces the degrees of freedom from n_k to $n_k - 1$ [9]. Both the Gaussian distribution for $\bar{z}_k$ and Wishart distribution for Z_k is consistent with having Gaussian-distributed plots, however. In Figure 1 we show the random matrices approach in a Bayesian network, with the dependencies between the random variables and the measurements.

The main motivation for assuming independent position and extension is to modify the extension part and devise a likelihood function that is less sensitive to the number of plots n_k. When having a model with independent scatterers, the likelihood causes the track-to-object association to favor single measurements over large clusters. This is usually compensated by including a false alarm density parameter in the likelihood, which penalizes the non-assigned plots. We believe that, given a model for the object size, it should be possible to assign the correct plots without being dependent on this parameter. Furthermore, we would also like to be able to filter the extension with other information than just the position scattering of a collection of plots. By modifying the likelihood function we are able loosen the inherent dependence on the number of plots and their scattering. We will continue focusing on the likelihood function and association probabilities also in the next section, before returning to the filtering aspects.

III. Two-stage tracking filter for extended targets

In the previous section we have assumed independent position and extension measurements by formulating the likelihood as a function of measured centre of gravity and scatter matrix (4). The scatter matrix parameterizes the extension by an ellipse in two dimensions. This ellipse can also be expressed in terms of two principal axes and their orientation. We denote the principal axes and orientation l_1,l_2,α for the normalized scatter matrix $\frac{Z_k}{n_k}$, and λ_1,λ_2,θ for the random matrix $\mathbf{X}_k$. The axes are ordered such that $l_1 > l_2$ and $\lambda_1 > \lambda_2$. For clarity reasons we have skipped the temporal index k. Now we can write the eigen-decomposition as

$$\frac{Z_k}{n_k} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} l_1^2 & 0 \\ 0 & l_2^2 \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \quad (5)$$

$$\mathbf{X}_k = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad (6)$$

It is reasonable to believe that the lengths of the principal axes and their orientation are independent, thus we postulate the likelihood for the extension and orientation to be

$$p\left(Z_k|\mathbf{X}_k\right) = p\left(l_1^2, l_2^2|\mathbf{X}_k\right) p\left(\alpha|\mathbf{X}_k\right) \quad (7)$$

Now, given that the scatter matrix Z_k is Wishart distributed with covariance $\mathbf{X}_k$ we can derive the joint distribution of the eigenvalues $p\left(l_1^2, l_2^2|\mathbf{X}_k\right)$ and the distribution of the orientation $p\left(\alpha|\mathbf{X}_k\right)$ analytically. They becomes surprisingly complex and contain hypergeometric series which can only be evaluated through approximations, see Appendix A.

Even if we could use those formulas for evaluating the likelihood of the extension, it is unfeasible to find an approximate Bayesian solution for updating the stochastic parameters λ_1, λ_2, and θ, given Wishart distributed scatter matrices, i.e. Gaussian-distributed plots. But is the Gaussian model for plots justifiable from a physical point of view? If we take a naval ship as an example, the scatterers are more likely to be situated at the edges, and less likely to appear at the centre. Hence, instead of approximating the distributions, *we suggest redefining the model assumption by stating that the principal components are Gaussian-distributed*. Tracking the random matrix $\mathbf{X}_k$ is then replaced by tracking the state vector $\mathbf{\Lambda}_k = [\lambda_1, \lambda_2, \theta]_k^T$, which includes the principal components of $\mathbf{X}_k$ and its rotation.

Our approach, which we call two-stage filtering, is based on the fact that we can separate the extension from the kinematics. And thus, we perform the filtering of the extension in the first stage, and then use the extension estimate in the second stage, in order to achieve better filtering of the kinematics. This setup is visualized in Figure 2.

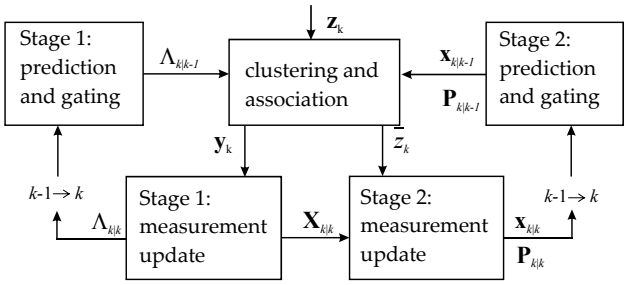


Figure 2. The two-stage filter where the matrix $\mathbf{X}_{k|k}$ is being transferred from the extension filter to the kinematical filter.

For updating the first stage filter, we use a measurement $\mathbf{y}_k$ that includes information about the length, width and rotation of the targets. We thus have a general approach, which is not limited to plot scattering solely as information about the extension, but can utilize any type of information representing the length, width, and orientation. If plots are the source of information about the target extent, the normalized scatter matrix provides the measurement vector $\mathbf{y}_k = [l_1, l_2, \alpha]_k^T$, comprising the principal components and the orientation of the scatter matrix. With linear-Gaussian models, the prediction and measurement update equations are

$$\mathbf{\Lambda}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{\Lambda}_{k-1} + \begin{bmatrix} \sigma_\lambda^2 & 0 & 0 \\ 0 & \sigma_\lambda^2 & 0 \\ 0 & 0 & \sigma_\theta^2 \end{bmatrix} \quad (8)$$

$$\mathbf{y}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{\Lambda}_k + \begin{bmatrix} \sigma_l^2 & 0 & 0 \\ 0 & \sigma_l^2 & 0 \\ 0 & 0 & \sigma_\alpha^2 \end{bmatrix}, \quad (9)$$

where the process noise σ_λ, σ_θ and measurement noise σ_l, σ_α are design parameters.

We have seen that the assumption of Gaussian-distributed plots leads to complicated equations (14)-(16). The reversed situation, i.e. assuming Gaussian-distributed principal components l_1, l_2, α, corresponds to a complicated non-Gaussian distribution for the plots, and thus non-Gaussian estimated centre of gravity (unless we have *many* plots). Nevertheless, we approximate the estimated centre of gravity as Gaussian although it may be theoretically incompatible with Gaussian principal components. Along with this heuristic we will also make a crucial simplifying assumption by *neglecting the uncertainty in the extension when filtering the kinematical state* $\mathbf{x}_k$. Hence, we assume that $\mathbf{X}_k$ is stochastic between measurements, but deterministic when used in the update of position.

$$p\left(\mathbf{x}_k|\mathbf{X}_k, \mathbf{Z}_k\right) \approx p\left(\mathbf{x}_k|\mathbf{X}_{k|k}, \mathbf{Z}_k\right) \tag{10}$$

This assumption is indicated in the Bayesian network in Figure 1 by making the connection between the random matrix and the position measurement dashed. To put all this in other words, we build a Kalman filter for estimating the centre of gravity, with data-driven measurement noise that is estimated through the principal components. This is a heuristic approach that enables us to build a two-stage Kalman filter for the tracking of extended objects. The first stage is a Kalman filter for the principal components, and the second stage is a Kalman filter for the centre of gravity (position). It produces the extension-filter state vector $\mathbf{\Lambda}_{k|k}$, which is used in equation (6) to form $\mathbf{X}_{k|k}$. The second filter is also linear Gaussian using $\mathbf{X}_{k|k}$ as measurement noise. In this paper we use the following linear-Gaussian process and measurement models for the centre of gravity:

$$\begin{aligned}\mathbf{x}_k = \begin{bmatrix} x_k \\ y_k \\ \dot{x}_k \\ \dot{y}_k \end{bmatrix} &= \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} \\ &+ \begin{bmatrix} \frac{\Delta T^4}{4} & \frac{\Delta T^3}{2} \\ \frac{\Delta T^3}{2} & \Delta T^2 \end{bmatrix} \otimes \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_a^2 \end{bmatrix} \end{aligned} \tag{11}$$

$$\begin{bmatrix} \bar{z}_x^k \\ \bar{z}_y^k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{x}_k + \mathbf{R}_k. \tag{12}$$

We will consider three different measurement noise models for the second stage filter: one fixed and two data driven.

$$1)\, \mathbf{R}_k = \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix},\ 2)\, \mathbf{R}_k = \mathbf{X}_{k|k},\ 3)\, \mathbf{R}_k = \frac{\mathbf{X}_{k|k}}{n_k} \tag{13}$$

We have assumed that the distribution of the centre of gravity is Gaussian, and if the plots are Gaussian the estimated measurement uncertainty would be $\frac{\mathbf{X}_{k|k}}{n_k}$. However, in the evaluation we will also examine the larger measurement uncertainty $\mathbf{X}_{k|k}$. Imagine that the plots are not independently distributed, and that we for example get a cluster of measurements from only the fore or the aft of a large ship. Then, the estimated mean should not decrease with n_k. Our goal in this study is to both investigate the advantage of having data-driven measurement noise over fixed, and to evaluate different modelling of the measurement noise.

IV. Implementational aspects

In a multiple extended target scenario, or in the presence of background clutter, we need to spawn measurement hypotheses by grouping plots. The problem of constructing cluster hypotheses can be solved in a number of ways. Here, we use a hierarchical clustering algorithm called single linkage clustering, based on the distances between the plots. Unlike the average linkage clustering, this method recomputes the distances between the clusters. It assigns a plot to a cluster based on the smallest distance between the plot and any of the plots in the cluster. Hence, it is computationally efficient. Moreover, it does not result in a combinatorial explosion of hypotheses; n_k plots will generate $2n_k - 1$ cluster hypotheses in total. An example of the clustering is shown in Figure 3, which is taken from the two-target scenario presented in Section V. The cluster dendrogram is displayed along with the plots from the two horizontally aligned elliptical targets.

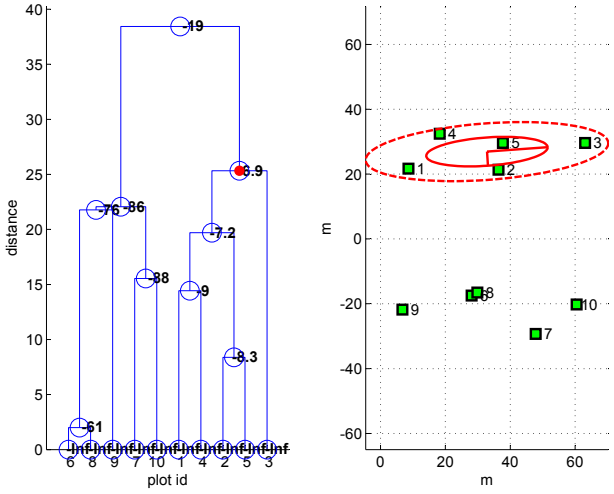


Figure 3. The single linkage cluster dendrogram with association likelihood (for the track representing the uppermost target to the right) at each node using measurement noise model $\mathbf{R}_k = \mathbf{X}_{k|k}$. Each node represents one of the $2n_k - 1$ measurement hypotheses (see Figure 1). The likelihood is the product of the stage one and stage two Kalman filter likelihoods. The maximum log-likelihood of -6.9 is marked in red and represents the clustering hypothesis shown by the ellipses.

Each node corresponds to a cluster hypothesis and the maximum likelihood hypothesis for the uppermost target is marked red. To evaluate the cluster hypotheses we use the likelihood from both the first and the second stage Kalman filters. This means that we have to calculate the number of plots n_k, the centre of gravity $\bar{z}_k$, and the scatter matrix Z_k, for each node. This is performed in a bottom-up fashion using the simple formulas displayed in Appendix B. Then, for all nodes, we calculate the extension measurements l_1, l_2, and α by diagonalizing the normalized scatter matrix $\frac{Z_k}{n_k}$ using the equations found in Appendix C. These are used to update λ_1, λ_2, and θ using the first stage of the Kalman filter, which inserted into equation (6) yields the estimated extension $\mathbf{X}_{k|k}$.

Now, there are a few exceptions when we get less than three measurements. If $n_k = 1$, we simply use the predicted values for $\lambda_1^{k|k-1}$, $\lambda_2^{k|k-1}$, and $\theta_{k|k-1}$ and hence use $\mathbf{X}_{k|k-1}$ in the second stage of the filter. If $n_k = 2$, we use the predicted semi-minor axis $\lambda_2^{k|k-1}$ and the updated semi-major axis $\lambda_1^{k|k}$ and orientation $\theta_2^{k|k}$.

V. Simulations and Results

We will generate uniformly distributed plots from elliptically shaped two-dimensional targets, representing large-sized naval ships. The semi-major axis is set to 40m and the semi-minor axis is 10m, hence the target is 80m long and 20m wide. The target is travelling with a speed of 10m/s and the measurement interval is $\Delta T = 1$s. In the first simulation, we consider a single target moving in a figure-of-eight pattern, see Figure 4, and analyze the difference between simulated centre of gravity and the filtered position (estimated centre of gravity) from the second stage Kalman filter.

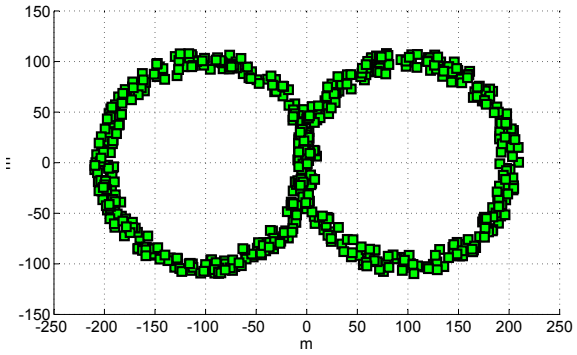


Figure 4. The single target scenario trajectory, displayed using accumulated plots from one lap. This scenario is used for evaluating the filtering performance in terms of position error.

The second simulation is a typical two-target scenario, see Figure 5, for analyzing how the combined likelihood estimate from both the first stage (extension) and the second stage (kinematic) Kalman filter can improve the ability to keep the track identity. Two targets moving at a speed of 10m/s are on collision course, but make a rapid 45° turn when their centres of gravity are 50m apart, and travel side by side for 200m before separating. Resolution aspects and radar measurement errors are not considered here. Also, we do not include background clutter, which is a significant simplification of the general target tracking problem.

A. Filter parameter settings

For the two-stage Kalman filter presented in Section III, we used the following parameter setting. For the first stage extension filter, the process noise for the major and minor semi-axes should be low (target size is relatively constant); we use $\sigma_\lambda = 1$m. Further, the orientation process noise is set to $\sigma_\theta = 10^\circ$. The measurement noise for the extension is set to $\sigma_l = 5$m and $\sigma_\alpha = 30^\circ$. The second stage filter, estimating

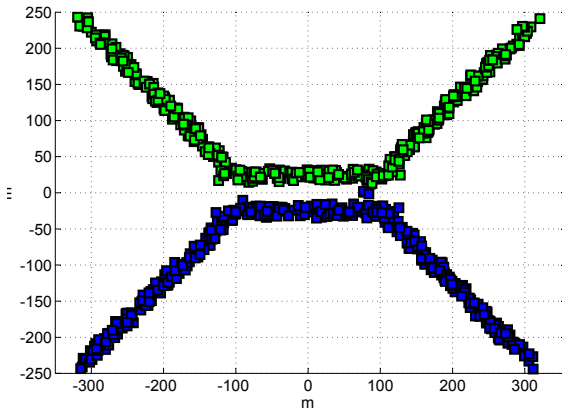


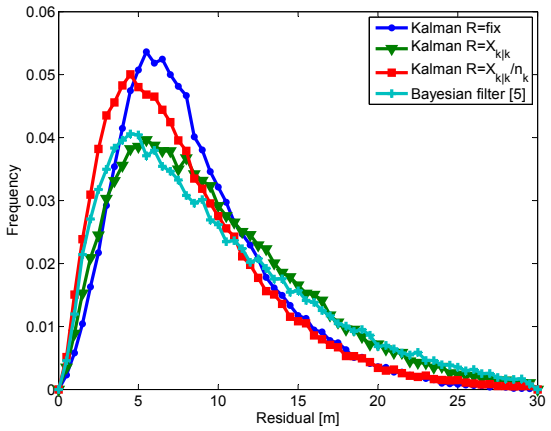
Figure 5. Trajectories of the two-target scenario, displayed using accumulated plots. This scenario is used for evaluating the filtering performance, and the association performance in terms of track loss probability.

the centre of gravity, uses a process noise for the kinematical state vector set according to equation (11) with acceleration uncertainty $\sigma_a = 1\text{m/s}^2$. The fixed measurement noise model in (13) is set to $\sigma_x = \sigma_y = 5$m. This is such that two standard deviations matches the size of the semi-minor axis 10m.

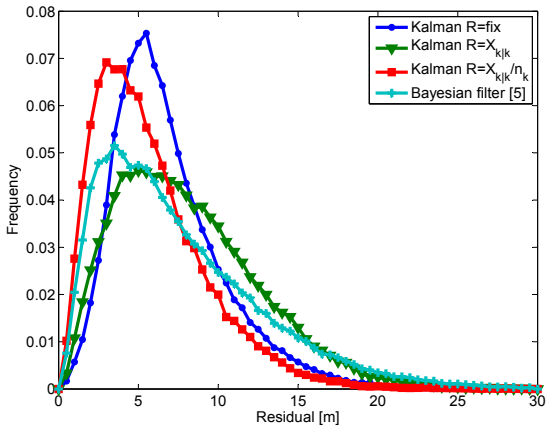
B. Single target filtering

In this part we do not consider the association problem. The simulated target follows the trajectory shown in Figure 4 and generates a number of plots n_k, which are used for updating the states of the first and second-stage filters. The plots are generated from a Poisson distribution with two different values for the expected number of plots: $E[n_k] = 3$ and $E[n_k] = 5$. We compare the error, i.e. the Euclidian norm of difference between simulated position and the position given by the filtered state vector $\mathbf{x}_{k|k}$ for the three different measurement noise models given in (13) and the Bayesian extended target filter given by [5]. To get a good statistical basis we generated 500 laps. The relative frequency of the errors is shown in the histogram in Figure 6.

We see that the smallest errors are obtained when the measurement noise is estimated by $\frac{\mathbf{X}_{k|k}}{n_k}$. This is expected, since the measurement uncertainty decreases with the number of plots for uniformly distributed plots. When using $\mathbf{X}_{k|k}$ as measurement noise, and thus neglecting this information, we lose accuracy by over-estimating the measurement uncertainty. The Bayesian filter has no intrinsic coupling between position and extension, which means that the extension estimate does not aid the filtering of the centre of gravity in the same way as for the proposed two-stage filtering approach. On the other hand, it will not deteriorate the results either, in case the extension estimate is unreliable. Nevertheless, in this evaluation of the estimation of the centre of gravity, where the association is perfect, the two-stage filter with $R = \frac{\mathbf{X}_{k|k}}{n_k}$ performs better.



(a) Error when expected number of plots per target is 3.



(b) Error when expected number of plots per target is 5.

Figure 6. Histogram of the error of the centre-of-gravity estimate of the two-stage Kalman filter for the single-target scenario in Figure 4, for three different measurement noise models (see (13)). The number of plots are Poisson distributed with two different expected values.

C. Two-target filtering and association

In this part we include the data association problem, and will hence not always update the filters with the correct plots. We use the two-target scenario in Figure 5, where two targets come as close as having 30m between them, see Figure 7. In a simplistic approach we assign the target to the cluster having the largest likelihood estimates from the Kalman filters. We sum the log-likelihood from the first and second stage filter, see Figure 2, to get the combined likelihood. A false alarm probability for penalizing unassigned plots is not included. Furthermore, the association is performed locally, meaning that the tracks do not know of each other. This enables for unwanted track merging, and make the association much more difficult. To simplify things we restrict the number of plots to be constant, and use three different values in the simulation $n_k = 3, 5, 10$. Also, we do not allow for assigning tracks to single plots. This is a significant simplification, but the benefit is that the problem of determining the extension likelihood for single plots in the first stage filter is avoided.

The two-target scenario was not evaluated in terms of error in position, but in track loss probability. If the two targets are correctly identified after having travelled the path in Figure 5 we count a correct assignment. If any of the tracks loses the target we count one error. This can be due to a track merge, or if one of the tracks is assigned to the top cluster node when the targets part, i.e. assigned to all plots from both targets. The latter causes the extension to grow while the track continues straight ahead during the separation. If both tracks lose their target we count two errors. This can be due to a switch of target identities, or that both tracks associate with the top cluster node. The biggest challenge for the tracking filter is to handle the situation when the targets come close and make the first rapid $45°$ turn. In Figure 7 we visualize how the data-driven measurement noise adapts the uncertainty to avoid track merge or mix-up.

As mentioned previously, we compute the log-likelihood as the sum of the first and second stage filter log-likelihood. An example of the likelihood for each cluster is shown in Figure 3, where the two targets are traveling side by side. The correct node is the one with the highest log-likelihood, much thanks to the contribution from the first stage filter. Since we do not allow for single-plot assignment (for simplicity) the log likelihood is minus infinity for leaves in the dendrogram. In the evaluation we will compare the results to a standard Kalman filter with fixed measurement noise, utilizing only the second stage filter. Thus the likelihood will only consider the cluster centre of gravity, which leads to an increased risk of false assignments.

A total of 100 iterations were run for the two-target scenario, for each of the three choices of measurement noise model in (13). In Figure 8, the results of the evaluation are presented.

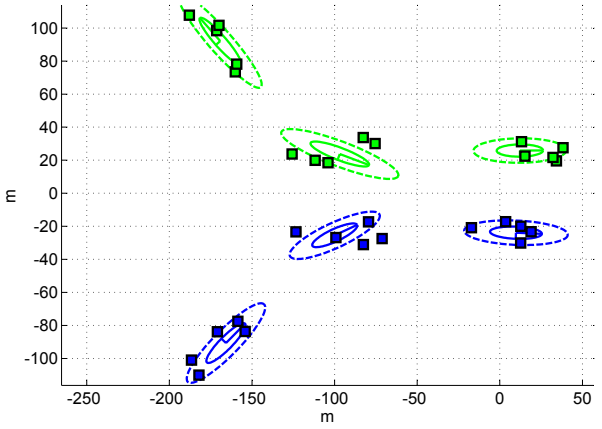
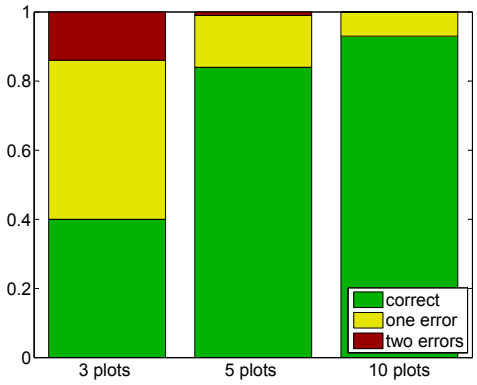
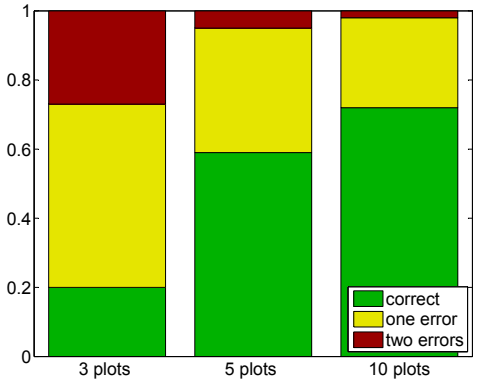


Figure 7. Visualization of the stage two filter measurement noise $\mathbf{R}_k = \mathbf{X}_{k|k}$ by ellipses for every 10th measurement update as the two targets come together. The number of plots is constant, $n_k = 5$. The inner and outer ellipses show the first and second standard deviation, i.e. λ_1, λ_2 and $2\lambda_1, 2\lambda_2$ including the orientation θ.

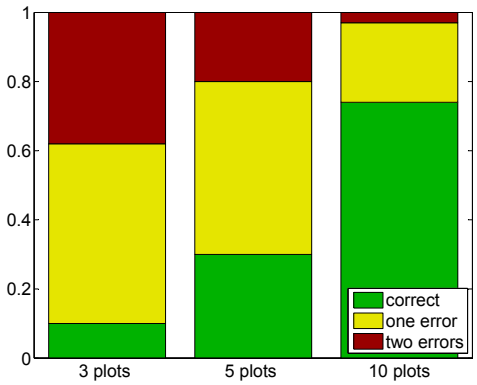
The results indicate, in contrast to the results from the single target filtering, that the coupling between the extension and kinematics is not beneficial. False cluster association leads to incorrect measurement uncertainty which in turn generate a bias for the next association likelihood, and so on. The improvement in kinematical accuracy from the filtering is overshadowed by the increased risk for false associations due to instabilities. This problem can, and should, be handled in a multi-hypothesis framework. Of course, the fact that the maximum likelihood association is not computed globally also affects the performance negatively.



(a) Fixed $\mathbf{R}_k$



(b) Data-driven $\mathbf{R}_k = \mathbf{X}_{k|k}$



(c) Data-driven $\mathbf{R}_k = \dfrac{\mathbf{X}_{k|k}}{n_k}$

Figure 8. The results in terms of the proportion of correct/one error/two errors tracks, for three different choices of measurement noise are displayed in (a)-(c). As we increase the number of plots generated by the target, the proportion of correct tracks increases rapidly. Note that the best results, achived in (b), corresponds to the measurement noise model $\mathbf{R}_k = \mathbf{X}_{k|k}$ that yielded mediocre filtering result in the single target simulation.

VI. Conclusions

We have designed a two-stage Kalman filter for extended target tracking using random matrices. The filter is based on the assumption that the principal components and the centre of gravity of the random matrix that describes the extension are Gaussian distributed. As measurements, we consider the two independent variables of estimated mean and scatter matrix. As a consequence, this filter is not restricted to information only represented by the scattering of the plots. Extension information may be encapsulated in single plots, for example.

The two-stage filter is evaluated on a single-target scenario, without data association problems, and a two-target scenario. The results from the single-target simulation clearly show that the filter using a data-driven measurement noise model $\mathbf{R}_k = \frac{\mathbf{X}_{k|k}}{n_k}$ for the centre of gravity performs better than both the standard Kalman filter and the Bayesian extended target filter derived in [5]. However, in the two target simulation this noise model exhibits worse performance than both the noise model $\mathbf{R}_k = \mathbf{X}_{k|k}$ and the fixed noise model. Hence, the filter that produces the lowest error given a correct association does not necessarily yield the best tracking result when including the track-to-cluster association problem. This relates to the discussion in the beginning of Section II regarding the dynamics between track filtering and data association.

We believe that the identified difference in tracking performance for the three considered measurement models can be equalized when the likelihood is computed for a global hypothesis, including the probability of false alarm. Thus, future work will comprise the development of more sophisticated data association, and will also contain a comparative study with the Bayesian filter including association. Another logical next step for the evaluation of tracking and data association performance would be to incorporate the two-stage filter in a multiple hypothesis framework, where we utilize the cluster nodes to hypothesize over both target movement and measurement clustering. This would mitigate potential instabilities caused by data-driven noise adaption in the case of large clusters and motion ambiguities.

In any case, we have seen in this study that a Kalman filter can be used to filter clustered measurements by utilizing a pre-stage filter on the principal components. The results are promising although much testing and evaluation remain.

Acknowledgement

This research has been supported by the National Aeronautic Research Program (NFFP), which is funded by VINNOVA (The Swedish Governmental Agency for Innovation Systems). The research project is a joint collaboration between the signal processing group at Chalmers University of Technology and Saab Electronic Defense Systems, Saab AB.

APPENDIX A
PROBABILITY DISTRIBUTIONS FOR PRINCIPAL COMPONENTS OF 2D WISHART MATRICES

The distribution of the orientation α in (5) is derived in [10] and is written as

$$p\left(\alpha|\mathbf{X}_k\right) = p\left(\xi(\alpha); \lambda_1^2, \lambda_2^2, n_k\right) = \frac{1}{\pi\left(n_k+1\right)}\left(\frac{\lambda_1^2\lambda_2^2}{\left(\frac{\lambda_1^2+\lambda_2^2}{2}\right)^2}\right)^{\frac{n_k}{2}} {}_2F_1\left(2, n_k, \frac{n_k+3}{2}; \xi(\alpha)\right) \tag{14}$$

where ${}_2F_1$ is Gauss hypergeometric series of scalar argument. The argument $\xi(\alpha)$ is defined as

$$\xi(\alpha) = \left[\begin{array}{cc} \cos\alpha & \sin\alpha \end{array}\right] \frac{\mathbf{X}_k}{\mathrm{tr}\left(\mathbf{X}_k\right)} \left[\begin{array}{c} \cos\alpha \\ \sin\alpha \end{array}\right] \tag{15}$$

In order to evaluate ${}_2F_1$ we recommend using the Laplace approximation for Gauss hypergeometric function on one scalar argument, which is given in [11]. The joint distribution of the eigenvalues is [9]

$$p\left(l_1^2, l_2^2|\mathbf{X}_k\right) = \left(\frac{n_k}{2}\right)^{n_k} \frac{\pi^2}{|\mathbf{X}|^{\frac{n_k}{2}}\,\Gamma_2\left(\frac{n_k}{2}\right)\Gamma_2\left(1\right)} \cdot l_1^{n_k-3} l_2^{n_k-3}\left(l_1^2-l_2^2\right) {}_0F_0\left(-\frac{n_k}{2}\left[\begin{array}{cc} l_1^2 & 0 \\ 0 & l_2^2 \end{array}\right], \mathbf{X}_k^{-1}\right) \tag{16}$$

where ${}_0F_0$ is the hypergeometric function of lowest order, which is basically an exponential trace. However, it is a function of *two* matrix arguments which means that we need to make an approximation in order to evaluate the function [12].

APPENDIX B
COMPUTING THE SCATTER MATRIX

To compute the measurements centre of gravity $\bar{z}_k$ and scatter matrix Z_k for the whole cluster tree bottom up, use the following simple equations. The node indexed 3 is the parent of node 1 and 2.

$$n_3 = n_1 + n_2 \tag{17}$$

$$\bar{z}_3 = \frac{1}{n_1+n_2}\left(n_1\bar{z}_1 + n_2\bar{z}_2\right) \tag{18}$$

$$R_3 = R_1 + R_2 + \frac{n_1 n_2}{n_1+n_2}\left(\bar{z}_1-\bar{z}_2\right)\left(\bar{z}_1-\bar{z}_2\right)^{\mathsf{T}} \tag{19}$$

APPENDIX C
COMPUTING THE PRINCIPAL COMPONENTS

To compute the principal components we use the following eqautions for the eigenvalues

$$l_1 = \frac{r_{11}+r_{22}}{2} + \sqrt{\frac{\left(r_{11}+r_{22}\right)^2}{4} - \left(r_{11}r_{22}-r_{12}^2\right)} \tag{20}$$

$$l_2 = \frac{r_{11}+r_{22}}{2} - \sqrt{\frac{\left(r_{11}+r_{22}\right)^2}{4} - \left(r_{11}r_{22}-r_{12}^2\right)} \tag{21}$$

and eigenvectors

$$\mathbf{q}_1^{r_{11}\geq r_{22}} = \left[\begin{array}{c} \frac{(r_{11}-r_{22})}{2} + \sqrt{\frac{(r_{11}-r_{22})^2}{4}+r_{12}^2} \\ r_{12} \end{array}\right] \tag{22}$$

$$\mathbf{q}_2^{r_{11}\geq r_{22}} = \left[\begin{array}{c} r_{12} \\ -\frac{(r_{11}-r_{22})}{2} - \sqrt{\frac{(r_{11}-r_{22})^2}{4}+r_{12}^2} \end{array}\right] \tag{23}$$

$$\mathbf{q}_1^{r_{11}<r_{22}} = \left[\begin{array}{c} r_{12} \\ -\frac{(r_{11}-r_{22})}{2} + \sqrt{\frac{(r_{11}-r_{22})^2}{4}+r_{12}^2} \end{array}\right] \tag{24}$$

$$\mathbf{q}_2^{r_{11}<r_{22}} = \left[\begin{array}{c} \frac{(r_{11}-r_{22})}{2} - \sqrt{\frac{(r_{11}-r_{22})^2}{4}+r_{12}^2} \\ r_{12} \end{array}\right] \tag{25}$$

$$\mathbf{q} = \frac{\mathbf{q}}{|\mathbf{q}|} \tag{26}$$

REFERENCES

[1] Rasmussen, C.; Hager, G.D.; ,"Probabilistic data association methods for tracking complex visual objects", *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.23, no.6, pp.560-576, Jun 2001

[2] Isard, M.; MacCormick, J.; ,"BraMBLe: a Bayesian multiple-blob tracker", *Computer Vision, 2001. ICCV 2001. Proceedings. Eight IEEE International Conference on* , vol.2, no., pp.34-41 vol.2, 2001

[3] Wei Du; Piater, J.; "Tracking by cluster analysis of feature points using a mixture particle filter", *Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on* , vol., no., pp. 165- 170, 15-16 Sept. 2005

[4] G. Gennari, G. D. Hager, "Probabilistic Data Association Methods in Visual Tracking of Groups", cvpr, vol. 2, pp.876-881, *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR'04) - Volume 2, 2004

[5] Koch, W. "Bayesian approach to extended object and cluster tracking using random matrices", *Aerospace and Electronic Systems, IEEE Transactions on* vol. 44 no.3, pp 1042 - 1059, July 2008

[6] Lau, B. Arras, K. Wolfram, B. "Multi-model hypothesis Group Tracking and Group Size Estimation" *International Journal of Social Robotics Volume 2*, Number 1, 19-30, March 2010

[7] Manuel Mucientes; Wolfram Burgard; "Multiple Hypothesis Tracking of Clusters of People", *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on* , vol., no., pp.692-697, Oct. 2006

[8] Garcia J. Berlanga A. Molina J.M. and de Miguel G. "A machine-learning approach to multiple-detection data association for ASDE radar" *Proceedings of the Seventh International Conference on Information Fusion*, vol I, pp 123-129, Stockholm, Sweden 2004

[9] Muirhead, R.J., *Aspects of Multivariate Statistical Theory*. 1982. Wiley, New York.

[10] T. Sugiyama, "On the Distribution of the Latent Vectors for Principal Component Analysis" Ann. Math. Statist. Volume 36, Number 6 (1965), 1875-1876.

[11] Butler, R.W. Wood, A.T "Laplace approximations for hypergeometric functions with matrix arguments" *The Annals of Statistics* 2002, Vol. 30, No. 4, 1155–1177

[12] Butler, R.W. Wood, A.T "Laplace approximations to hypergeometric functions of two matrix arguments" *Journal of Multivariate Analysis*. Volume 94 Issue 1, May 2005