TTK4250

Week 6

Single-target tracking

Edmund Førland Brekke

27. September 2024

Recap from last week

Hybrid systems

$$
\begin{aligned}
\mathbf{y}_k &= \begin{bmatrix} \mathbf{x}_k \\ s_k \end{bmatrix} \\
\mathbf{x}_k &= \mathbf{f}^{(s_k)}(\mathbf{x}_{k-1}) + \mathbf{v}_k, , \quad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{(s_k)}) \\
\mathbf{z}_k &= \mathbf{h}^{(s_k)}(\mathbf{x}_k) + \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{(s_k)}) \\
\pi^{ij} &= \Pr\{s_k = j \mid s_{k-1} = i\}
\end{aligned}
$$

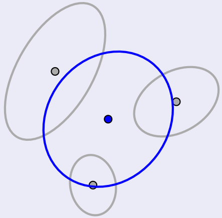
Generic Bayesian procedure

- The models involved.
- The previous posterior.
- The prediction.
- The posterior.

The IMM method

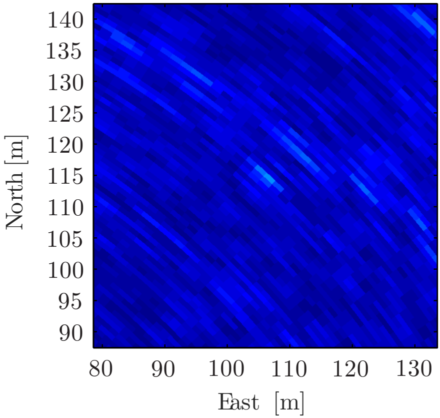
- M filters in parallel.
- Approximate $p(\mathbf{x}_{k-1}|s_k, \mathbf{z}_{1:k-1})$ as a single Gaussian.
- Perform an (E)KF cycle for each value of s_k.

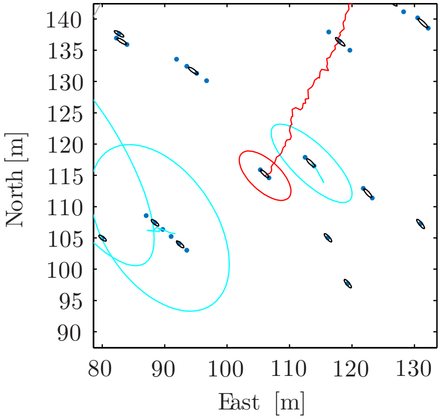
Mixture reduction



An example of target tracking

Tracking a human rebreather diver in 90kHz active sonar data.[1]





[1] Brekke et al. (2010) "The Signal-to-Noise Ratio of Human Divers", Proc. OCEANS.

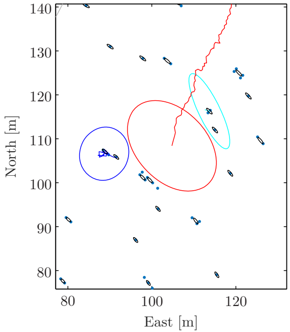
Why is single-target tracking more than just filtering?

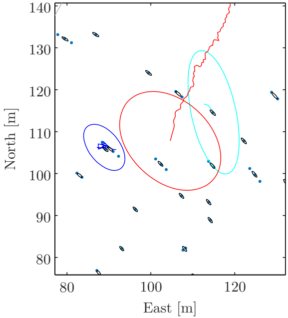
Misdetections.

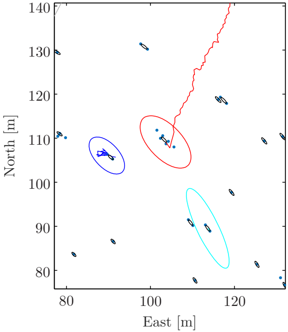
A target does not necessarily give a measurement in each sensor scan.

Clutter.

A measurement does not necessarily come from a target.







As a consequence, **data association** is an issue: We need to make inference on which, if any, measurement comes from the target.

Assumptions of (standard) single-target tracking

1. At time step $k-1$ one and only one target exists in the surveillance region $\mathcal{S}$ with state vector $\mathbf{x}_{k-1}$.
2. The prior density of $\mathbf{x}_{k-1}$ is given as $p_{k-1}(\mathbf{x}_{k-1})$.
3. The state vector of the target evolves from time step $k-1$ to k according to a Markov model of the form $f_{\mathbf{x}}(\mathbf{x}_k|\mathbf{x}_{k-1}) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$.
4. A measurement of the target is detected with probability P_{D}.
5. If a target-originating measurement exists, then it is related to $\mathbf{x}_k$ according to a likelihood of the form $f_{\mathbf{z}}(\mathbf{z}_k|\mathbf{x}_k) = p(\mathbf{z}_k|\mathbf{x}_k)$.
6. An unknown number of φ_k measurements originate from clutter, where the discrete-valued random variable φ_k is distributed according to $\mu(\varphi)$.
7. If $\mathbf{z}$ is a clutter measurement, then it is distributed according to a pdf $c(\mathbf{z})$, independently of all other measurements.

Based on all this the goal is to estimate the predicted pdf $p_{k|k-1}(\mathbf{x}_k) = p(\mathbf{x}_k|Z_{1:k-1})$ and the posterior pdf $p_k(\mathbf{x}_k) = p(\mathbf{x}_k|Z_{1:k})$.

Here Z_k is the set of measurement at time k, and $Z_{1:k}$ is the collection of all such sets so far received.

The prior density

Gaussian prior

In most cases, the prior density, a.k.a. previous posterior density, is a Gaussian

$$p_{k-1}(\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{\mathbf{x}}_{k-1}, \mathbf{P}_{k-1}).$$

But most of the general formulas that we will derive also hold if $p_{k-1}(\mathbf{x}_{k-1})$ is something more complicated.

Gaussian mixture prior

Towards the end of Chapter 7 we shall encounter a situation where the prior density is a Gaussian mixture

$$p_{k-1}(\mathbf{x}_{k-1}) = \sum_{s_k} w_k^{(s_k)} \mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{\mathbf{x}}_{k-1}^{(s_k)}, \mathbf{P}_{k-1}^{(s_k)}).$$

The process model and the predicted density

We have covered this already, but let us repeat it.

Prediction and the product identity

- Let the process model be Gaussian-linear:

$$f_{\mathbf{x}}(\mathbf{x}_k \mid \mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k\,;\,\mathbf{F}\mathbf{x}_{k-1}, \mathbf{Q})$$

- The predicted density then becomes

$$\begin{aligned}
p_{k|k-1}(\mathbf{x}_k) &= \int f_{\mathbf{x}}(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p_{k-1}(\mathbf{x}_{k-1}) \mathrm{d}\mathbf{x}_{k-1} \\
&= \int \mathcal{N}(\mathbf{x}_k\,;\,\mathbf{F}\mathbf{x}_{k-1}, \mathbf{Q}) \mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{\mathbf{x}}_{k-1}, \mathbf{P}_{k-1}) \mathrm{d}\mathbf{x}_{k-1} \\
&= \int \mathcal{N}(\mathbf{x}_k\,;\,\mathbf{F}\hat{\mathbf{x}}_{k-1}, \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^{\mathsf{T}} + \mathbf{Q})\, \mathcal{N}(\mathbf{x}_{k-1}\,;\,\ldots) \mathrm{d}\mathbf{x}_{k-1} \\
&= \mathcal{N}(\mathbf{x}_k\,;\,\mathbf{F}\hat{\mathbf{x}}_{k-1}, \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^{\mathsf{T}} + \mathbf{Q}) \int \mathcal{N}(\mathbf{x}_{k-1}\,;\,\ldots) \mathrm{d}\mathbf{x}_{k-1} \\
&= \mathcal{N}(\mathbf{x}_k\,;\,\mathbf{F}\hat{\mathbf{x}}_{k-1}, \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^{\mathsf{T}} + \mathbf{Q}).
\end{aligned}$$

Detection probability

Why detection probability is an issue

No detector can in general guarantee target detection.

- The target could be occluded.
- Noise from the surroundings (e.g., wave crests) could be detected instead.
- The target could be near invisible (e.g., optical camera at night or stealth airplane in radar data).

Modeling the misdetection problem

- We can model target detection as a Bernoulli random variable δ with the distribution

$$\Pr\{\delta\} = \begin{cases} P_{\mathrm{D}} & \text{if } \delta = 1 \\ 1 - P_{\mathrm{D}} & \text{if } \delta = 0. \end{cases}$$

- For now we assume that the detection probability P_{D} is fixed, but sometimes it is modeled as something that depends on the target state, i.e., $P_{\mathrm{D}}(\mathbf{x}_k)$.
- Typical values for P_{D} may lie in the area between 0.5 and 0.95.

Clutter

Causes of false alarms

No detector can in general guarantee that false alarms will not occur.

- Electric noise in the sensor itself can cause false alarms.
- Electromagnetic or acoustic beams may be reflected by arbitrary stones, waves, birds etc.
- Interference effects or deliberate jammings can cause false alarms.

The false alarm rate P_{FA} and its relationship to P_{D}

- We define P_{FA} as the probability that a given sensor cell contains a false alarm.
- Most detector can be tuned to decrease or increase P_{FA}, ...
- but unfortunately this leads to a corresponding decrease or increase of P_{D}.

Beware: A single sensor scan should never be used to declare the presence or absence of an object. This is a task for a complete tracking system (e.g., M/N-logic or IPDA). P_{D} and P_{FA} on this slide on refers to the measurement extraction process.

The Poisson clutter model

Physical motivation.

We are given a sensor array with a large number N of resolution cells, and assume that no target is present.

- For each resolution cell, there is a probability P_{FA} that it will yield a false alarm.
- The spatial distribution of any given false alarm is therefore uniform over the sensor image.
- The number φ of false alarms will have a binomial distribution:

$$\Pr(\varphi) = \binom{N}{\varphi} P_{\mathrm{FA}}^{\varphi} (1 - P_{\mathrm{FA}})^{N-\varphi}$$

- Let $\Lambda = NP_{\mathrm{FA}}$. If N is sufficiently large and P_{FA} is sufficiently small, then the binomial is well approximated by the Poisson distribution

$$\mu(\varphi) = \Pr(\varphi) = \mathrm{Poiss}(\varphi; \Lambda) = \exp(-\Lambda)\frac{\Lambda^{\varphi}}{\varphi!}.$$

We refer to Λ as the rate parameter. Also notice that Λ is the expectation of the Poisson distribution.

The Poisson clutter model continued

The Poisson model described on the previous slide is an example of a **point process**.

Definition: Point process.

A point process describes how a random number of points are distributed randomly in some space (say, $\mathbb{R}^2$).

Point processes are often referred to as random finite sets in the tracking literature.[2]

The Poisson process

- It is characterized by the following property: Given two disjoint regions $A \subset \mathbb{R}^2$ and $B \subset \mathbb{R}^2$, the number of points in A is independent of the number of points in B.
- In general, it is specified by an **intensity function** $\lambda(\mathbf{z}) = \Lambda c(\mathbf{z})$.
  Here $c(\mathbf{z})$ is the pdf of an individual point of the point process.
- We can find the expected number of points according to $\Lambda = \int \lambda(\mathbf{z})\,\mathrm{d}\mathbf{z}$.

[2] Mahler (2007): "Statistical Multisource-Multitarget Information Fusion".

An alternative: The diffuse clutter model

- In theory, P_{FA} is a tunable parameter, and the Poisson intensity λ will be directly given by P_{FA} and the sensor resolution.
- In practice, the actual false alarm rate is likely to be significantly different from the desired false alarm rate due to blurring, target extent and unmodeled noise artefacts.
- It may therefore be necessary to estimate λ, but such estimates are likely to be very inaccurate.

The diffuse clutter model

We let all clutter cardinalities have the same probability

$$\mu(\phi = 0) = \mu(\phi = 1) = \mu(\phi = 2) = \ldots = \epsilon \tag{1}$$

- The spatial density of clutter measurements is again assumed uniform within the region of interest.
- To ensure that (1) is a valid pmf, simply say that all cardinalities above some upper limit have probability zero.

The data association problem

- We receive m_k measurements $Z_k = \{\mathbf{z}_k^1, \ldots, \mathbf{z}_k^{m_k}\}$ at time k.
- Denote the cumulative measurement set by $Z_{1:k} = (Z_1, \ldots, Z_k)$.
- **At most one of the measurements in Z_k can come from the target. The other ones must be clutter.**
- The outcome space can be divided into $m_k + 1$ mutually exclusive events:

$$\begin{aligned} a_k &= 0 && \text{No measurement originates from the target} \\ a_k &= 1 && \text{Measurement 1 originates from the target} \\ &\vdots \\ a_k &= m_k && \text{Measurement } m_k \text{ originates from the target.} \end{aligned} \tag{2}$$

- The total probability theorem gives the posterior

$$p_k(\mathbf{x}_k) = \sum_{a_k} p(\mathbf{x}_k \mid a_k, Z_{1:k}) \Pr\{a_k \mid Z_{1:k}\}. \tag{3}$$

The goal in single-target tracking is typically to find $p(\mathbf{x}_k \mid a_k, Z_{1:k})$ and $\Pr\{a_k \mid Z_{1:k}\}$, and then simplify the mixture (3) so that we get a tractable approximation of the true posterior.

The measurement update in general terms

The event-conditional posteriors

The pdf terms in (3) are found according to Bayes rule:

$$p(\mathbf{x}_k \mid a_k, Z_{1:k}) \propto \begin{cases} p_{k|k-1}(\mathbf{x}_k) & \text{if} \quad a_k = 0 \\ f_{\mathbf{z}}(\mathbf{z}_k^i \mid \mathbf{x}_k) p_{k|k-1}(\mathbf{x}_k) & \text{if} \quad a_k > 0 \end{cases}$$

Under Gaussian-linear assumptions this boils down to a Kalman filter.

The posterior association probabilities.

The probabilities $P\{a_k \mid Z_{1:k}\}$ require some more work. Let us first define the "measurement likelihood"

$$l^{a_k} = \int f_{\mathbf{z}}(\mathbf{z}_k^{a_k} \mid \mathbf{x}_k) p_{k|k-1}(\mathbf{x}_k) \, \mathrm{d}\mathbf{x}_k.$$

It can then be shown (Theorem 7.3.1 in the book) that the posterior association probabilities are

$$\Pr\{a_k \mid Z_{1:k}\} \propto \begin{cases} (1 - P_{\mathrm{D}})\mu(m_k) & \text{if} \quad a_k = 0 \\ l^{a_k} P_{\mathrm{D}} \mu(m_k - 1)/m_k c(\mathbf{z}_k^{a_k}) & \text{if} \quad a_k > 0. \end{cases} \tag{4}$$

The PDAF: Assuming Gaussian-linearity

The Probabilistic Data Association Filter (PDAF) is one of the simplest and most popular tracking algorithms. The key underlying assumption is

$$p_{k|k-1}(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}), \tag{5}$$

together with Gaussian-linearity of the process model and measurement model:

$$\begin{aligned} f_{\mathbf{x}}(\mathbf{x}_k|\mathbf{x}_{k-1}) &= \mathcal{N}(\mathbf{x}_k\,;\,\mathbf{F}\mathbf{x}_{k-1}, Q) \\ f_{\mathbf{z}}(\mathbf{z}_k|\mathbf{x}_k) &= \mathcal{N}(\mathbf{z}_k\,;\,\mathbf{H}\mathbf{x}_k, \mathbf{R}). \end{aligned}$$

The event-conditional posteriors

For each measurement $\mathbf{z}_k^{a_k}$ we have $p(\hat{\mathbf{x}}_k \mid a_k, Z_{1:k}) = \mathcal{N}(\mathbf{x}_k\,;\,\hat{\mathbf{x}}_k^{a_k}, \mathbf{P}_k^{a_k})$ where

$$\begin{aligned} \hat{\mathbf{x}}_k^{a_k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{W}_k(\mathbf{z}_k^{a_k} - \mathbf{H}\hat{\mathbf{x}}_{k|k-1}) \\ \mathbf{P}_k^{a_k} &= (\mathbf{I} - \mathbf{W}_k\mathbf{H})\mathbf{P}_{k|k-1} \\ \mathbf{W}_k &= \mathbf{P}_{k|k-1}\mathbf{H}^\mathsf{T}(\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\mathsf{T} + \mathbf{R})^{-1} \end{aligned}$$

For $a_k = 0$ we have $p(\mathbf{x}_k \mid a_k, Z_{1:k}) = p_{k|k-1}(\mathbf{x})$.

The validation gate

We also assume that only measurements within some gate $\mathcal{G}$, given g standard deviations around $\hat{\mathbf{z}}_{k|k-1} = \mathbf{H}\hat{\mathbf{x}}_{k|k-1}$ according to $\mathbf{S}_k = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\mathsf{T} + \mathbf{R}$ are validated:

$$\mathcal{G} = \left\{ \mathbf{z} \;\text{ such that }\; (\mathbf{z} - \hat{\mathbf{z}}_{k|k-1})^\mathsf{T}\mathbf{S}_k^{-1}(\mathbf{z} - \hat{\mathbf{z}}_{k|k-1}) < g^2 \right\}.$$

Clutter measurements are assumed spatially uniform with pdf $c(\mathbf{z}) = 1/V_k$ where **V_k is the volume of the gate.** Since there is a chance the true measurement is outside the gate, we introduce the gate probability P_G and replace P_D with $P_\mathrm{D}P_\mathrm{G}$, and also divide the likelihood by $f_\mathbf{z}(\cdot)$ by P_G in the weight calculation.

- We have the following expression for the gate volume:

  $$V_k = c_{n_z}|g^2\mathbf{S}_k|^{1/2}, \;\; c_1 = 2, \;\; c_2 = \pi, \;\; c_3 = 4\pi/3, \ldots$$

  where c_{n_z} is the dimension of the measurement space (e.g., 2 for range-bearing measurements).
- The gate probability P_G is more difficult to calculate.[3] It is as good as 1 if $g \geq 3$.

[3] Bar-Shalom & Li (1995): “Multitarget-Multisensor Tracking: Principles and Techniques”, p. 96.

Association weights for parametric and diffuse clutter models

Association weights with the Poisson model.

$$\beta_k^i = \Pr\{a_k = i \mid Z_{1:k}\} \propto \begin{cases} \lambda \dfrac{1 - P_\mathrm{D} P_\mathrm{G}}{P_\mathrm{D}} & \text{if} \quad i = 0 \\ \mathcal{N}(\mathbf{z}_k^i\,;\, \hat{\mathbf{z}}_{k|k-1}, \mathbf{S}_k) & \text{if} \quad i > 0 \end{cases}$$

Association weights with the diffuse model.

$$\beta_k^i = \Pr\{a_k = i \mid Z_{1:k}\} \propto \begin{cases} \dfrac{m_k}{V_k} \dfrac{1 - P_\mathrm{D} P_\mathrm{G}}{P_\mathrm{D}} & \text{if} \quad i = 0 \\ \mathcal{N}(\mathbf{z}_k^i\,;\, \hat{\mathbf{z}}_{k|k-1}, \mathbf{S}_k) & \text{if} \quad i > 0 \end{cases}$$

We see that m_k / V_k takes the role of a clutter estimate when prior knowledge about the clutter density is not provided.

The PDAF: Mixture reduction

- In order to ensure Gaussianity at the next cycle we need to reduce the Gaussian mixture to a single Gaussian.
- Moment matching: We construct a Gaussian with the same first- and second-order moments (i.e., expectation and covariance) as the mixture.

The expectation.

$$\hat{\mathbf{x}}_k = \beta_k^0 \, \hat{\mathbf{x}}_{k|k-1} + \sum_{i=1}^{m_k} \beta_k^i \, \hat{\mathbf{x}}_k^i \tag{6}$$

The covariance.

First we find the "spread-of-the-innovations":

$$\tilde{\mathbf{P}}_k = \mathbf{W}_k \left[\sum_{i=1}^{m_k} \beta_k^i \boldsymbol{\nu}_k^i (\boldsymbol{\nu}_k^i)^\mathsf{T} - \boldsymbol{\nu}_k \boldsymbol{\nu}_k^\mathsf{T} \right] \mathbf{W}_k^\mathsf{T}. \tag{7}$$

The covariance of the mixture is then

$$\mathbf{P}_k = \mathbf{P}_{k|k-1} - (1 - \beta_k^0) \mathbf{W}_k \mathbf{S}_k \mathbf{W}_k^\mathsf{T} + \tilde{\mathbf{P}}_k. \tag{8}$$

Illustration of the PDAF

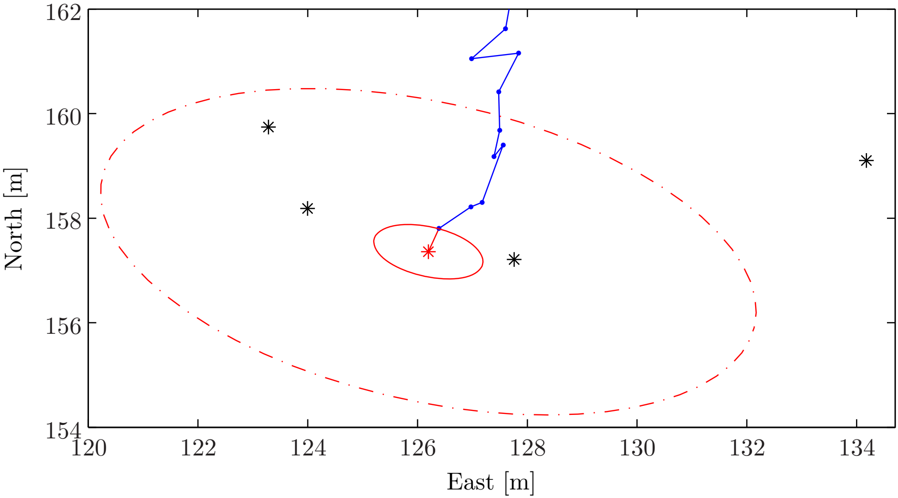


Illustration of the PDAF

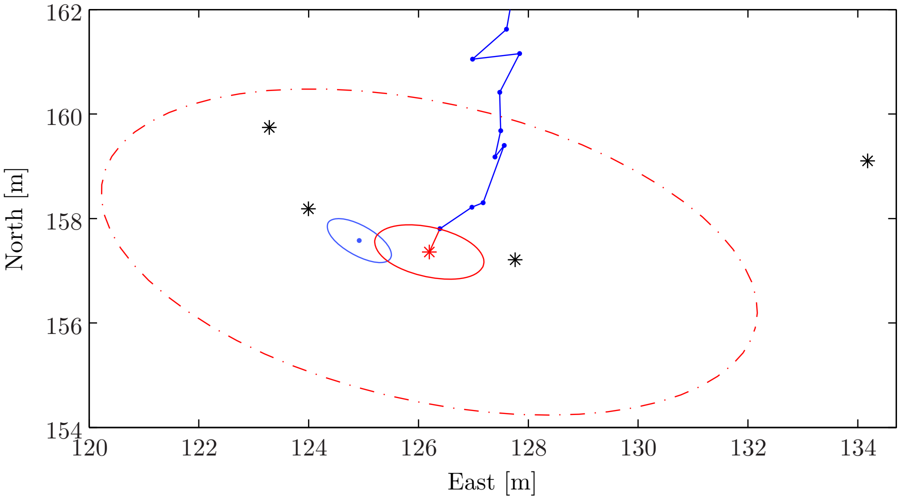


Illustration of the PDAF

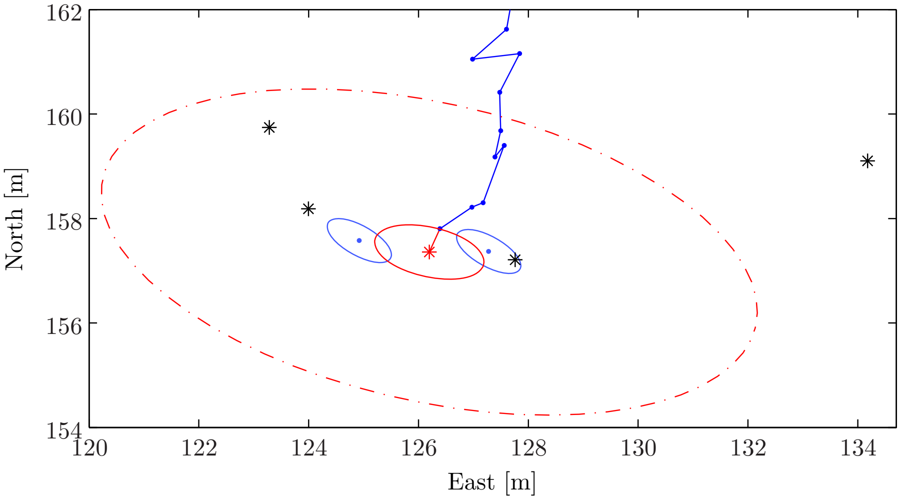


Illustration of the PDAF

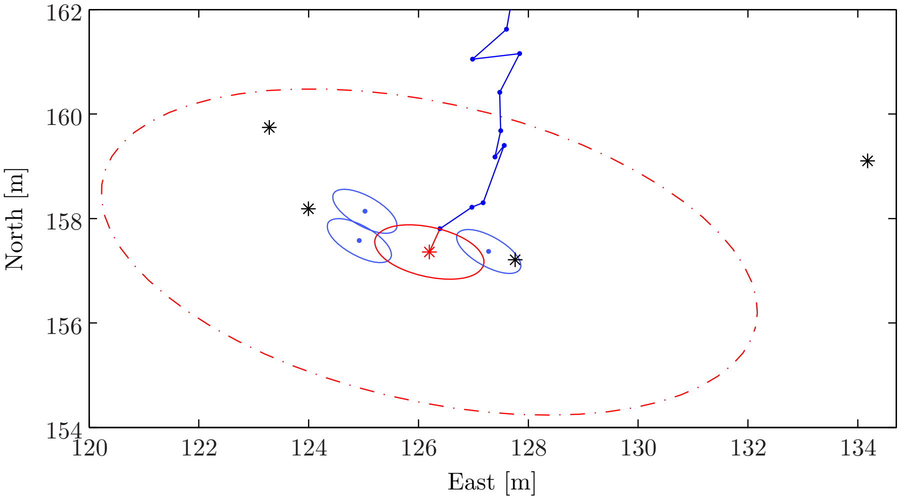


Illustration of the PDAF

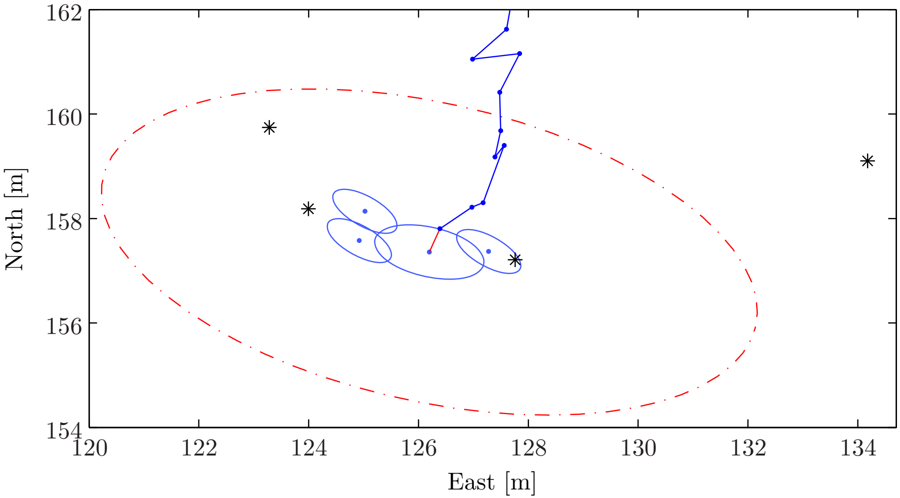


Illustration of the PDAF

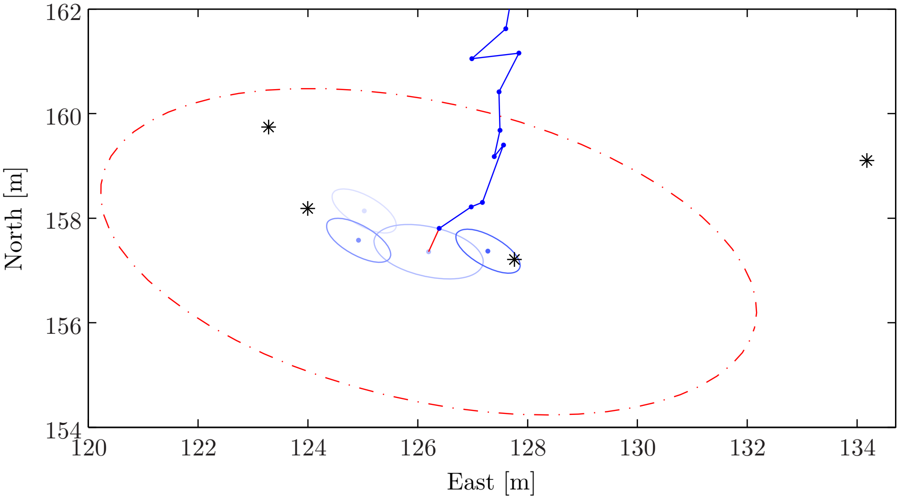


Illustration of the PDAF

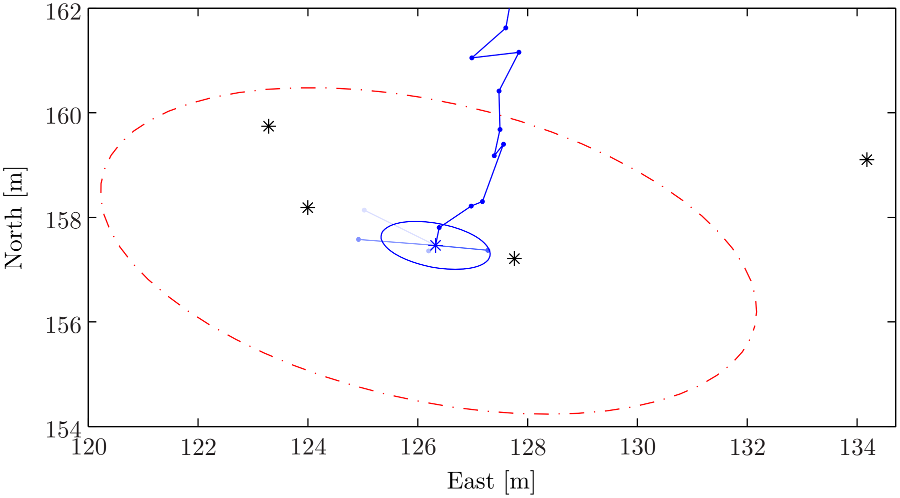
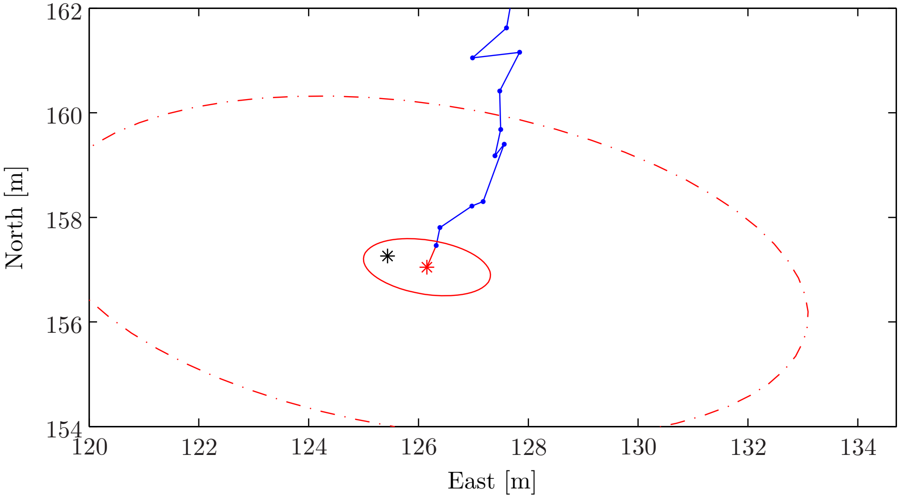
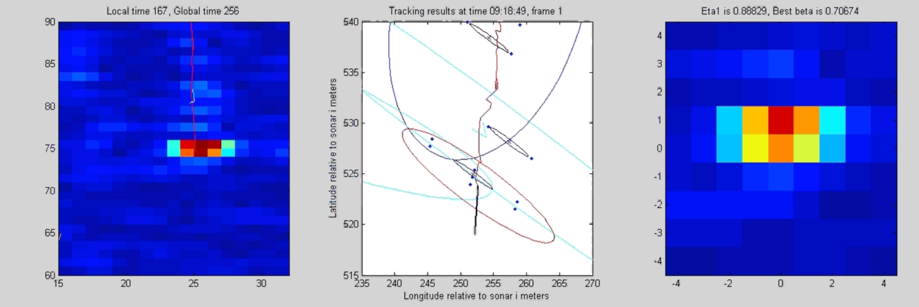


Illustration of the PDAF



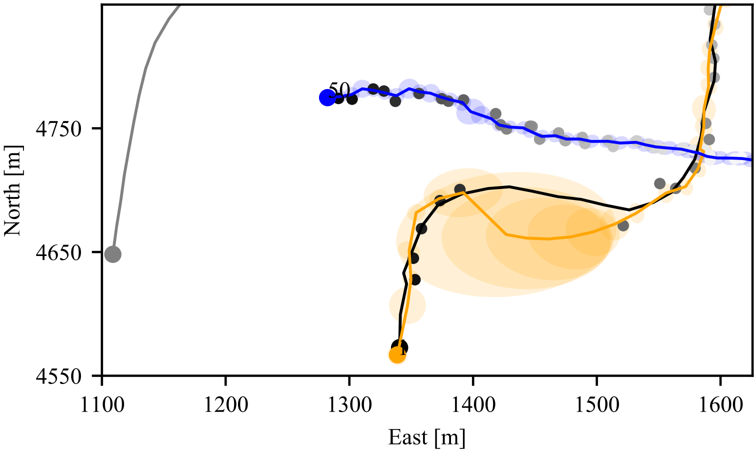
Visualization ideas

- Display measurements, tracks, covariance ellipses and validation gates.
- It can be nice to superpose the tracks on the raw sensor data.
- Experiment with different colors for tracks at different stages or of different quality.
- Zoom.



Visualization ideas[4]

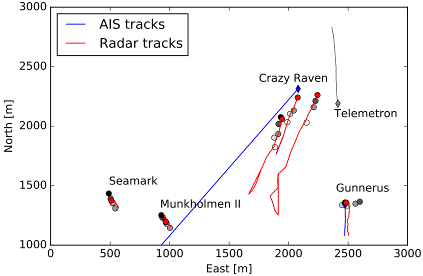
- Make movies.
- Alternatively: Display temporal information using color shades.



[4] Brekke, Hem & Tokle (2020): "Multitarget Tracking With Multiple Models and Visibility: Derivation and Verification on Maritime Radar Data", IEEE-JOE, 2021.

Visualization ideas[5]

- Compare with actual video footage.
- “Ground truth” (AIS) and radar tracks in different colors.



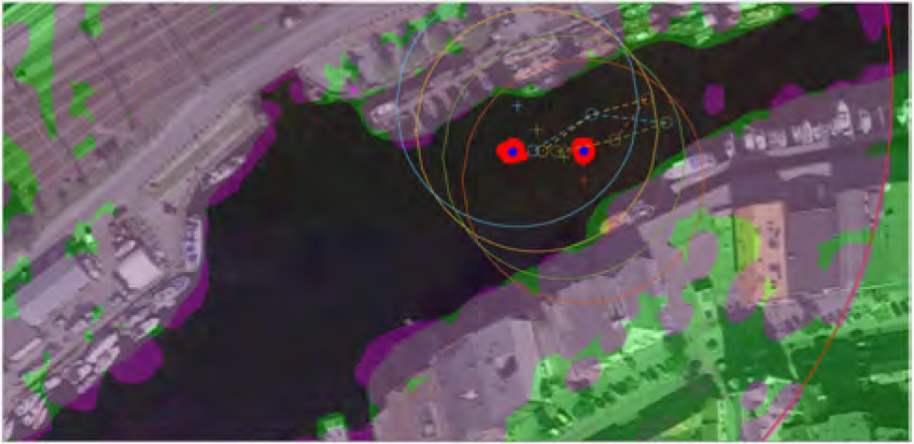
(a) Radar tracks from PDAF.

(b) The same scenario seen from drone video.

[5]Brekke, Wilthil, Eriksen, Kufoalor, Helgesen, Hagen, Breivik & Johansen (2019): “The Autosea project: Developing closed-loop target tracking and collision avoidance systems”, Proc. ICMASS.

Visualization ideas[6]

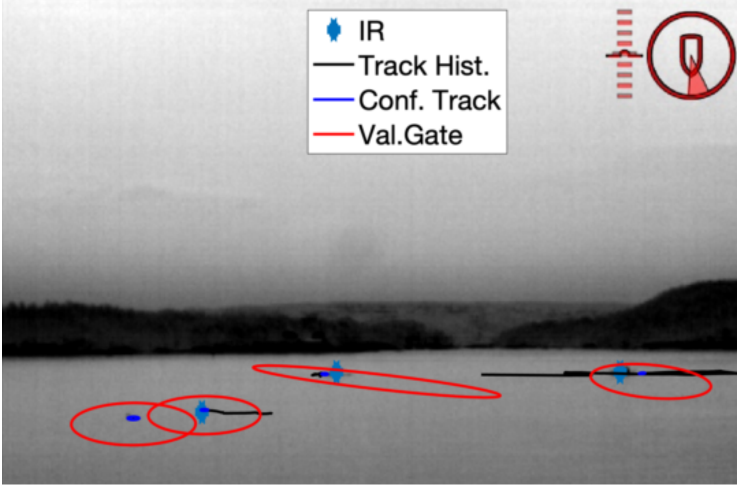
- Visualize together with geographical surroundings.
- Visualize together with processed sensor data.



[6] Jesper Pedersen (2018): "Harbor Surveillance with a K-best, Track Terminating, Hypothesis-Oriented MHT", MSc thesis, NTNU.

Visualization ideas[7]

- Compare with video footage (IR in this case).
- Project the tracking results into the video.



[7] Helgesen, Brekke, Stahl & Engelhardtsen (2020): "Low Altitude Georeferencing for Imaging Sensors in Maritime Tracking", Proc. IFAC World Congress.

Reflections

Roles of tracking: Safety critical or not so critical?

- Collision avoidance. **Please don't miss a target.**
- Defence applications. **Please don't establish false tracks.**
- Surveillance. **Please don't make so many mistakes that operator gets jaded.**
- Post-processing and data interpretation. **Please don't learn the wrong things.**

Alternatives to the Bayesian paradigm

- Nearest neighbor techniques (e.g. SORT).
- Counting techniques (e.g. Recursive RANSAC).
- Signal processing techniques (e.g. Hough transform for track-before-detect).
- Machine learning.
  - Typically used to extract detections (e.g. bounding boxes) that are fed to a tracker.
  - Can also be used to learn how to do data association (transformers).
  - Can learn to recognize motion and distance.
- Occupancy grid + vector field.

The Bayesian approach typically works well insofar as it is easy to make simple and reliable models.