Data Association for Multi-Target-Tracking

Michael Grinberg

Vision and Fusion Laboratory
Institute for Anthropomatics
Karlsruhe Institute of Technology (KIT), Germany
michael.grinberg@kit.edu

Technical Report IES-20xx-xx

Abstract: This report describes problems of Multi-Target-Tracking and gives an introduction to the state-of-the-art methods of dynamic state estimation and data association in cluttered environments. A detailed derivation of the Probabilistic Data Association Filter and Joint Probabilistic Data Association Filter is given.

1 Introduction

Object detection and tracking is an important task in modern environment perception and surveillance systems. Its aim is continuous localization of people or objects in an environment by processing data of environment perception sensors such as sonars, radars, lidars, or video cameras. In general, the problem of object tracking can be divided into three subtasks: data association (re-identification), dynamic state estimation (filtering), and track management. The first subtask is responsible for the correct interpretation of the collected observations, i.e., assignment of sensor measurements to the tracked objects (tracks). The second subtask deals with estimation of the dynamic state of the objects (e.g., kinematics) from a sequence of noisy measurements. Finally, the third subtask is responsible for a consistent internal representation of the tracked objects, which includes initiation of new and deletion of obsolete tracks. The following sections give an introduction to the basic state-of-the-art methods for dynamic state estimation and data association in cluttered environments.

2 Dynamic State Estimation

Each measurement process contains sources of noise. Thus, obtained measurements may differ from the expected values. The aim of a dynamic state estimator is the determination of the real value of a not known system state from the obtained (noisy) measurements. This is done by the so-called filtering algorithms that aim at minimization of the noise effects. There exists a variety of such methods. Most of the modern tracking systems use statistical filters that are based on the Bayesian approach. They model system state and noise as random variables and estimate their statistics using certain assumptions about their nature.

Applications with real-time requirements often cannot consider the entire measurement history for achieving the best estimation result. Thus, they proceed recursively using only the last estimated system state and the current measurements. The underlying assumption is that all previous measurements are incorporated in the estimated state and are not required to be processed again in each time step (Markov property). The system evolution is thus modeled by means of a Markov process.

A system state $\mathbf{x}_k$ at discrete time point k is modeled as a realization of a random variable $\mathbf{X}$ in the state space $\mathcal{X}$. The system state between two discrete points in time k and $k+1$ is assumed to behave according to a known system evolution function f (***system model***):

$$\mathbf{X}_{k+1} = f(\mathbf{X}_k, \mathbf{u}_k, \mathbf{W}_k)\,,$$

where $\mathbf{u}_k$ represents the (known) system control parameters and $\mathbf{W}_k$ represents the stochastic component which cannot be modeled analytically (system noise).

The observations $\mathbf{z}_k$ are modeled as a realization of a random variable $\mathbf{Z}$ in the measurement space $\mathcal{Z}$. The measurement process is modeled by means of a ***measurement model*** $h(\mathbf{X}_k, \mathbf{V}_k)$:

$$\mathbf{Z}_k = h(\mathbf{X}_k, \mathbf{V}_k)\,,$$

where $\mathbf{V}_k$ represents the stochastic component of the measurement process (measurement noise). Since the system state can not be observed directly, one speaks of the ***Hidden Markov Model (HMM)***. The relation between system states and observations of a Hidden Markov Model for the case of discrete states is shown in Figure 2.1.

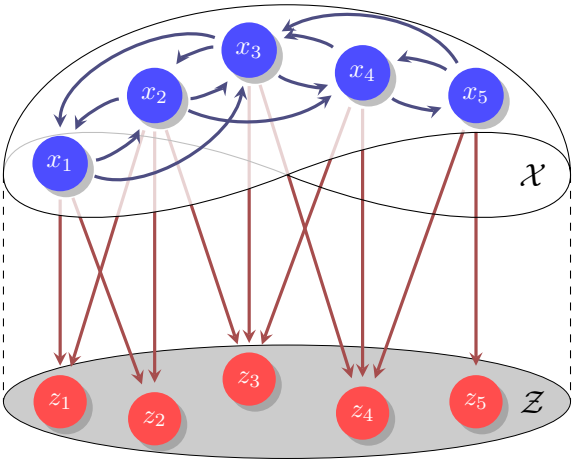


Figure 2.1: Relation between system states and observations in a Hidden Markov Model. Possible state transitions are represented by blue arrows, emission probabilities are indicated as purple arrows. Here, the subscripts of the states and observations are used not for indicating the time index k but serve for enumeration of the both sets.

The state estimation is done using the so-called Predictor-Corrector cycle, which consist of two steps:

Prediction of the probability density functions of the new system state and expected measurements based on the latest state estimate by using the system model and the measurement model.

Correction of the estimated system state and adaptation of the both models based on the actually obtained measurements. It is also called ***Innovation***, ***Update*** or ***Filtering***.

The basic principle of a recursive statistical filter is shown in Figure 2.2. The filter works recursively in a predictor-corrector cycle starting with an initial system state estimate $\hat{\mathbf{x}}_0$. Given a state estimation at time step $k-1$ the filter propagates it in the time using the system model. In this way, an a-priori estimate of the current system state $\hat{\mathbf{x}}_k^-$ is obtained. Then, the measurement model is used for estimating the expected measurement $\hat{\mathbf{z}}_k$. After having obtained the actual measurement $\mathbf{z}_k$, a correction step is performed, in which both the current state and the uncertainties of the both models are updated based on the difference (residuals) between the predicted and actually obtained measurements.

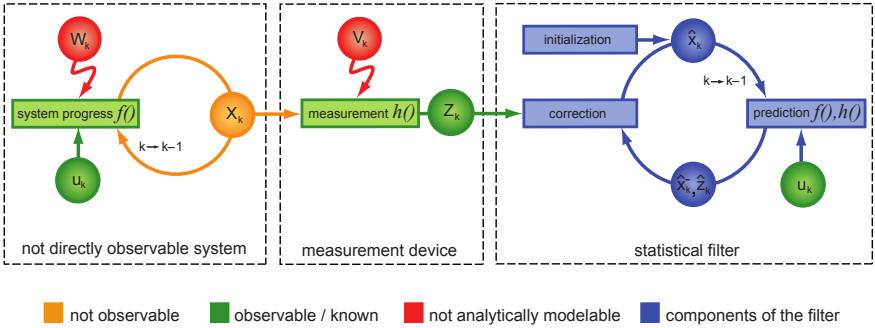


Figure 2.2: Illustration of the dependencies between the observed system and the filtering process of a statistical filter.

One of the simplest statistical dynamic state estimators is the ***Kalman Filter*** introduced by R. E. Kalman in 1960 [Kal60]. It assumes Gaussian distributions of both the state and the noise variables and provides equations for propagation of those distributions using linear system and measurement models. For the case of $\mathbf{W}_k$ and $\mathbf{V}_k$ being uncorrelated and having white Gaussian distribution with zero mean, the Kalman Filter is an optimal estimator in the sense of the least square errors and Bayesian filtering.

A Gaussian distribution can be represented by the two first moments (mean and covariance matrix) and is easily propagated through a linear system resulting in another Gaussian distribution. In case of non-linearities in at least one of the both models, this is not the case anymore. For coping with this problem two different approaches have been proposed. The first one aims at approximation of the non-linear function by using the Taylor series expansion around the mean of the Gaussian distribution (***Extended Kalman Filter (EKF)***, ***Iterative Extended Kalman Filter (IEKF)***). The second approach aims at approximation of the distribution by means of a set of points that can be propagated through the non-linear functions and serve for determination of the new distribution parameters (***Unscented Kalman Filter (UKF)***, ***Central Difference Kalman Filter (CDKF)*** etc.). A generalization of this approach leads to the family of the ***Sequential Monte Carlo Methods (SMCM)*** also known as ***Particle Filters (PF)***. An overview over different dynamic state estimators can be found in [BSL93, Bro98, Sim06]. Following subsections recapitulate the basics of the linear Kalman Filter, Extended Kalman Filter and Iterative Extended Kalman Filter since they build the basement for the data association methods presented in this paper.

2.1 Linear Kalman Filter

In the case of the linear Kalman Filter, system model and measurement model are given as linear equations

$$\begin{aligned} \mathbf{X}_{k+1} &= \mathbf{F}\mathbf{X}_k + \mathbf{G}\mathbf{u}_k + \mathbf{W}_k && \text{(state equation)} \\ \mathbf{Z}_k &= \mathbf{H}\mathbf{X}_k + \mathbf{V}_k && \text{(measurement equation)} \end{aligned} \tag{2.1}$$

with $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ being the system matrix, the control matrix and the measurement matrix, respectively, and

$$\begin{aligned} &\mathbf{X}_k \sim \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_{\mathbf{X}_k\mathbf{X}_k}) \\ &\text{with} \\ &\mathbf{P}_{\mathbf{X}_k\mathbf{X}_k} := \mathrm{Cov}(\mathbf{X}_k, \mathbf{X}_k) = \mathbb{E}\big[(\mathbf{X}_k - \hat{\mathbf{x}}_k)(\mathbf{X}_k - \hat{\mathbf{x}}_k)\big] \,. \end{aligned} \tag{2.2}$$

The noise components $\mathbf{W}$ and $\mathbf{V}$ are assumed to be uncorrelated and to have white Gaussian distribution with zero mean and known covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$:

$$\mathbf{W}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k) \,, \quad \mathbf{V}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k) \,,$$

$$\mathbf{P}_{\mathbf{W}_k\mathbf{W}_t} := \mathrm{Cov}(\mathbf{W}_k, \mathbf{W}_t) = \mathbb{E}\big[\mathbf{W}_k\mathbf{W}_t^T\big] = \begin{cases} \mathbf{Q}_k & \text{for } t = k \\ \mathbf{0} & \text{for } t \neq k \,, \end{cases}$$

$$\mathbf{P}_{\mathbf{V}_k\mathbf{V}_t} := \mathrm{Cov}(\mathbf{V}_k, \mathbf{V}_t) = \mathbb{E}\big[\mathbf{V}_k\mathbf{V}_t^T\big] = \begin{cases} \mathbf{R}_k & \text{for } t = k \\ \mathbf{0} & \text{for } t \neq k \,, \end{cases}$$

$$\mathbf{P}_{\mathbf{W}_k\mathbf{V}_t} := \mathrm{Cov}(\mathbf{W}_k, \mathbf{V}_t) = \mathbb{E}\big[\mathbf{W}_k\mathbf{V}_t^T\big] = \mathbf{0} \quad \text{for all } t \text{ and } k \,,$$

$$\mathbf{P}_{\mathbf{X}_k\mathbf{W}_t} := \mathrm{Cov}(\mathbf{X}_k, \mathbf{W}_t) = \mathbb{E}\big[\mathbf{X}_k\mathbf{W}_t^T\big] = \mathbf{0} \quad \text{for all } t \text{ and } k \,,$$

$$\mathbf{P}_{\mathbf{X}_k\mathbf{V}_t} := \mathrm{Cov}(\mathbf{X}_k, \mathbf{V}_t) = \mathbb{E}\big[\mathbf{X}_k\mathbf{V}_t^T\big] = \mathbf{0} \quad \text{for all } t \text{ and } k \,.$$

In the considered application there is no possibility to influence the observed system. Hence, the control parameter vector $\mathbf{u}$ will be omitted in the following.

As mentioned above, the Kalman Filter gives estimates of the two first moments $\hat{\mathbf{x}}_k$ and $\mathbf{P}_{\mathbf{X}_k\mathbf{X}_k}$ of the distribution of the true state $\mathbf{x}_k$. It is initialized at time step $k = 0$ with initial state estimate $\hat{\mathbf{x}}_0$ and covariance matrix $\mathbf{P}_{\mathbf{X}_0\mathbf{X}_0}$. The recursive expression for the calculation of the a-priori estimates $\hat{\mathbf{x}}_k^-$ and $\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}$ at time step k from the a-posteriori estimates $\hat{\mathbf{x}}_{k-1}$ and $\mathbf{P}_{\mathbf{X}_{k-1}\mathbf{X}_{k-1}}$ at the previous time step $k-1$ is derived by using the state equation (2.1) in the expectation value computation:

$$\hat{\mathbf{x}}_k^- = \mathbb{E}\big[\mathbf{X}_k|\mathbf{z}_{1:k-1}\big] = \mathbb{E}\big[\mathbf{F}\mathbf{X}_{k-1} + \mathbf{W}_{k-1}|\mathbf{z}_{1:k-1}\big] = \mathbf{F}\hat{\mathbf{x}}_{k-1}$$

and

$$\begin{aligned}\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-} =&\mathbb{E}\big[(\mathbf{X}_k-\hat{\mathbf{x}}_k^-)(\mathbf{X}_k-\hat{\mathbf{x}}_k^-)^T\big]\\ =&\mathbb{E}\big[(\mathbf{F}\mathbf{X}_{k-1}+\mathbf{W}_{k-1}-\mathbf{F}\hat{\mathbf{x}}_{k-1})(\mathbf{F}\mathbf{X}_{k-1}+\mathbf{W}_{k-1}-\mathbf{F}\hat{\mathbf{x}}_{k-1})^T\big]\\ =&\mathbf{F}\mathbf{P}_{\mathbf{X}_{k-1}\mathbf{X}_{k-1}}\mathbf{F}^T+\mathbf{Q}_k\,.\end{aligned}\tag{2.3}$$

In equation (2.3), $\mathbf{Q}_k$ represents the unpredictable noise component. The uncertainty $\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}$ of the state grows in each time step by this expression. The counteraction is achieved by integrating new information about the system state that is contained in the new measurements. This is done in the correction step using innovation $\tilde{\mathbf{z}}_k=\mathbf{z}_k-\hat{\mathbf{z}}_k$:

$$\begin{aligned}\hat{\mathbf{x}}_k&=\hat{\mathbf{x}}_k^-+\mathbf{K}_k(\tilde{\mathbf{z}}_k)=\hat{\mathbf{x}}_k^-+\mathbf{K}_k(\mathbf{z}_k-\hat{\mathbf{z}}_k)=\hat{\mathbf{x}}_k^-+\mathbf{K}_k(\mathbf{z}_k-\mathbf{H}\hat{\mathbf{x}}_k^-)\\ &=(\mathbf{I}-\mathbf{K}_k\mathbf{H})\hat{\mathbf{x}}_k^-+\mathbf{K}_k\mathbf{z}_k\,,\end{aligned}$$

and

$$\mathbf{P}_{\mathbf{X}_k\mathbf{X}_k}=\mathbb{E}\big[(\mathbf{X}_k-\hat{\mathbf{x}}_k)(\mathbf{X}_k-\hat{\mathbf{x}}_k)\big]=\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}+\mathbf{K}_k\mathbf{H}\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}$$

with Kalman gain $\mathbf{K}_k=\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}\mathbf{H}^T\mathbf{P}^{-1}_{\tilde{\mathbf{Z}}_k\tilde{\mathbf{Z}}_k}$ and innovation covariance

$$\begin{aligned}\mathbf{P}_{\tilde{\mathbf{Z}}_k\tilde{\mathbf{Z}}_k}&=\mathbb{E}\big[\tilde{\mathbf{Z}}_k\tilde{\mathbf{Z}}_k^T\big]=\mathbb{E}\big[\big((\mathbf{Z}_k-\hat{\mathbf{z}}_k)-0\big)\big((\mathbf{Z}_k-\hat{\mathbf{z}}_k)-0\big)^T\big](=\mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k})\\ &=\mathbb{E}\big[(\mathbf{H}(\mathbf{X}_k-\hat{\mathbf{x}}_k)+\mathbf{V}_k)(\mathbf{H}(\mathbf{X}_k-\hat{\mathbf{x}}_k)+\mathbf{V}_k)^T\big]\\ &=\mathbf{H}\mathbf{P}_{\mathbf{X}_k\mathbf{X}_k}\mathbf{H}^T+\mathbf{R}_k\,.\end{aligned}$$

2.2 Extended Kalman Filter

In the case of non-linearities in the system and measurement models, the state and measurement equation are given by

$$\begin{aligned}\mathbf{X}_{k+1}&=f(\mathbf{X}_k,\mathbf{W}_k)\\ \mathbf{Z}_k&=h(\mathbf{X}_k,\mathbf{V}_k)\,.\end{aligned}$$

In the most of the cases, additive noise model is assumed so that

$$\begin{aligned}\mathbf{X}_{k+1}&=f(\mathbf{X}_k)+\mathbf{W}_k\\ \mathbf{Z}_k&=h(\mathbf{X}_k)+\mathbf{V}_k\,.\end{aligned}$$

The Extended Kalman Filter approximates the non-linear functions f and h using Taylor series expansion around the current mean estimate. Truncation of the Taylor series after the first element leads to a linear function, which can be used for propagation of the Gaussian distribution as in the linear case. The a-priori estimate for the system state and expected measurement can be obtained directly using both nonlinear functions. When propagating state covariance and computing Kalman gain, Jacobians $\mathbf{F}_{k-1}$ and $\mathbf{H}_k$ are used:

Prediction: $\hat{\mathbf{x}}_k^- = f(\hat{\mathbf{x}}_k), \quad \mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-} = \mathbf{F}_{k-1}\mathbf{P}_{\mathbf{X}_{k-1}\mathbf{X}_{k-1}}\mathbf{F}_{k-1}^T + \mathbf{Q}_k$

Correction: $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - h(\hat{\mathbf{x}}_k^-)), \quad \mathbf{P}_{\mathbf{X}_k\mathbf{X}_k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}$

with

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}\mathbf{H}_k^T\mathbf{P}_{\tilde{\mathbf{Z}}_k\tilde{\mathbf{Z}}_k}^{-1}$$
$$\mathbf{P}_{\tilde{\mathbf{Z}}_k\tilde{\mathbf{Z}}_k} = \mathbf{H}_k\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}\mathbf{H}_k^T + \mathbf{R}_k$$

and the Jacobians

$$\mathbf{F}_{k-1} = \left.\frac{\mathrm{d}f}{\mathrm{d}\mathbf{x}}\right|_{\hat{\mathbf{x}}_{k-1}} \quad \text{and} \quad \mathbf{H}_k = \left.\frac{\mathrm{d}h}{\mathrm{d}\mathbf{x}}\right|_{\hat{\mathbf{x}}_k^-}.$$

2.3 Iterative Extended Kalman Filter

The Extended Kalman Filter linearizes the measurement function around the a-priori state estimate $\hat{\mathbf{x}}_k^-$ although a better state estimate is given after the integration of the current measurement. Linearization around the a-posteriori state estimate $\hat{\mathbf{x}}_k$ may improve the estimation. This potential is exploited in the iterative version of the EKF, the Iterative Extended Kalman Filter (IEKF). IEKF iteratively repeats the correction step with the recalculated linearization of the measurement model until a termination constraint is fulfilled. For ensuring non-recurrent integration of the measurement $\mathbf{z}_k$ during the iterations, a correction term $\mathbf{H}_k^{(i)}(\hat{\mathbf{x}}_k^- - \hat{\mathbf{x}}_k^{(i)})$ is used in each iteration i:

$$\hat{\mathbf{x}}_k^{(i+1)} = \hat{\mathbf{x}}_k^- + \mathbf{K}_k^{(i)}(\mathbf{z}_k - h(\hat{\mathbf{x}}_k^{(i)}) - \mathbf{H}_k^{(i)}(\hat{\mathbf{x}}_k^- - \hat{\mathbf{x}}_k^{(i)}))$$
$$\mathbf{K}_k^{(i)} = \mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}(\mathbf{H}_k^{(i)})^T(\mathbf{P}_{\tilde{\mathbf{Z}}_k\tilde{\mathbf{Z}}_k}^{(i)})^{-1}$$

with

$$\mathbf{H}_k^{(i)} = \left.\frac{\mathrm{d}h}{\mathrm{d}\mathbf{x}}\right|_{\hat{\mathbf{x}}_k^{(i)}}$$

and start value $\hat{\mathbf{x}}_k^{(0)} = \hat{\mathbf{x}}_k^-$.

3 Data association

In order to correctly perform the update step, statistical state estimators such as the Kalman Filter assume a correct assignment of measurements to tracks. A correct assignment means that in each time step each track is associated with a single measurement that has been originated from the corresponding object. The problem of assigning measurements to the existing tracks is called the data association problem. Data association is not always a trivial process. Given multiple active tracks and multiple detections, there are often several assignment possibilities being more or less probable. Figure 3.1 illustrates the data association ambiguity in case of three objects and four measurements.

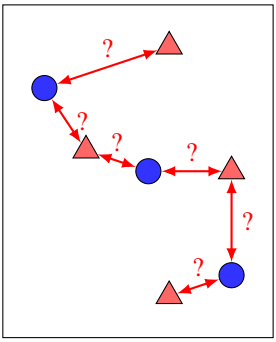


Figure 3.1: Illustration of a possible data association ambiguity in case of three tracks and four measurements. The three expected measurements are visualized by blue circles, the actually obtained – by red triangles.

Further uncertainties are introduced through the fact that a measurement may be evoked not only by a real object but may emerge due to concentration of noise in the data (clutter) or may be missing due to weaknesses of the sensors or of the subsequent data processing algorithms. And, finally, in some systems an object may evoke multiple measurements and several objects may give a joint measurement. This makes unambiguous assignments difficult or even impossible. In case of extended targets this is even worse since object observability represents another uncertainty source. Partial and full occlusions result in incomplete and missing detections and make data association even more challenging.

There exists a number of algorithms for solving the data association problem in multi-target applications. Hereby, a differentiation between the so-called ***single scan algorithms*** (also referred to as ***Single Hypotheses Tracking (SHT)***) and ***multiscan algorithms*** (also referred to as ***Multi-Hypotheses Tracking (MHT)***) is done.

While single scan algorithms consider only data of the current frame (scan), multiscan algorithms simultaneously evaluate multiple hypotheses maintaining them throughout several frames in anticipation that the new data will allow to resolve emerging conflicts [Rei79, CH96]. In practice, single scan algorithms are often preferred due to their simplicity and low computational cost. In the following, algorithms assuming that the number of tracks is known and a detection corresponds to a single track and vice versa will be presented.

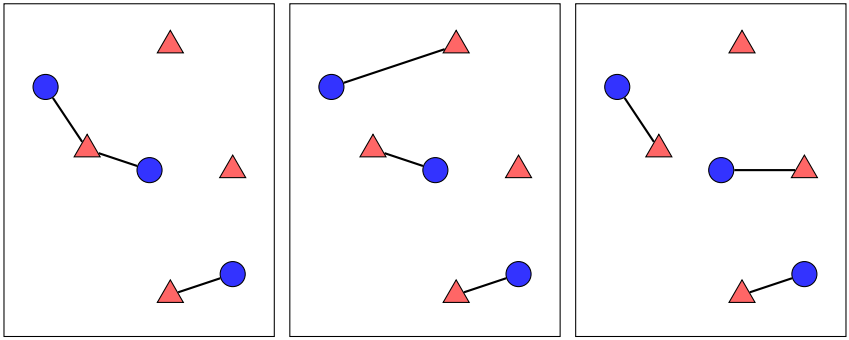
3.1 Nearest Neighbor Algorithms

One of the simplest data association algorithms is the ***Nearest Neighbor algorithm (NN)***. It is a typical single scan algorithm since it considers only measurements belonging to the current data frame (scan). NN algorithm considers only one data association hypothesis, assigning for each track the closest measurement. As shown in Figure 3.2 (a), in multi-target tracking scenarios, the NN algorithm is not optimal since it might assign a single measurement to multiple tracks despite the presence of other measurements. There exists an iterative version of the NN algorithm which prohibits multiple selections. It sequentially choses track-measurement pairs with the closest distance and excludes them from further consideration. This algorithm is suboptimal too, since it minimizes the track-to-measurement distances sequentially and thus may miss the global minimum as shown in Figure 3.2 (b). This problem can be solved by the ***Global Nearest Neighbor algorithm (GNN)*** which seeks for the globally optimal solution with respect to track-to-measurement distances (Figure 3.2 (c)).

3.2 Probabilistic Data Association (PDA)

Nearest Neighbor algorithms make a hard decision by minimizing distances between the predicted and real measurements. This decision might be optimal with respect to the distances in the current frame, however it may be still suboptimal with respect to the whole measurement sequence. Especially in applications where missing detections or obtaining clutter-based detections is possible, the Global Nearest Neighbor algorithm may lead to severe tracking errors. This problem was studied thoroughly in the radar tracking literature and led to development of statistical methods based on the idea of the probabilistic data association.

The main idea of such methods is weighting of different association hypotheses according to their probabilities thus minimizing the association error. Similarly to NN and GNN, PDA-based methods consider at each point in time only currently incoming measurements, i.e., they are single scan algorithms. However,



(a) Association by means of the simple Nearest Neighbor algorithm

(b) Association by means of the iterative NN algorithm

(c) Association by means of the Global Nearest Neighbor algorithm

Figure 3.2: Illustration of Nearest Neighbor data association.

when updating a track, instead of choosing a single measurement with highest association probability they evaluate multiple association hypotheses and use all neighboring measurements weighting them according to the probabilities of the corresponding hypotheses (***All-Neighbours Data Association***). Due to this ***soft decision*** approach, PDA-based methods suffer less from data association errors and are thus better suitable for applications with clutter-based and missing detections. Although PDA-based methods work with multiple association hypotheses they are also referred to as single hypotheses tracking algorithms since the hypotheses are combined to a single hypothesis prior to innovation. The remainder of this section addresses basics of the ***Probabilistic Data Association algorithm (PDA)*** proposed by Bar-Shalom et. al. [BST75, BS78].

The PDA considers each track separately. Let the considered track be denoted by $\mathbf{x}$ with $\mathbf{X}_k \sim \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_{\mathbf{X}_k\mathbf{X}_k})$ as in (2.2). Under Gaussian distribution assumption, the a-priori probability density of the predicted measurement position is given by

$$f_{\mathbf{Z}_k^{\mathbf{x}}} = f(\mathbf{z}_k^{\mathbf{x}} | \mathcal{Z}_{1:(k-1)}^{\mathbf{x}}) = \mathcal{N}(\mathbf{z}_k; \hat{\mathbf{z}}_k^{\mathbf{x}}, \mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}}) \,,$$

with

$$\hat{\mathbf{z}}_k^{\mathbf{x}} = \mathbf{H}\hat{\mathbf{x}}_k^- \quad \text{and} \quad \mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}} = \mathbf{H}\mathbf{P}_{\mathbf{X}_k^-\mathbf{X}_k^-}\mathbf{H}^T + \mathbf{R} \,.$$

For preventing associations with too far lying and thus too improbable measurements a selection region referred to as the ***gating region*** or ***validation gate*** $\Gamma_k^{\mathbf{x}}$ is

defined around $\hat{\mathbf{z}}_k$ with volume $V_k^{\mathbf{x}}$. Associations are only performed with measurements falling inside the gating region. The probability of the correct measurement $\mathbf{z}_k$ to lie inside the gating region is given by

$$P_G^{\mathbf{x}_k} = P(\mathbf{z}_k \in \Gamma_k^{\mathbf{x}}) = \int_{V_k^{\mathbf{x}}} f_{\mathbf{Z}_k^{\mathbf{x}}} \, \mathrm{d}\mathbf{z}_k \,. \tag{3.1}$$

The a-priori probability density function that accounts for gating is thus defined as:

$$p(\mathbf{z}_k^{\mathbf{x}} | \mathcal{Z}_{1:(k-1)}) := \begin{cases} \frac{1}{P_G} f_{\mathbf{Z}_k^{\mathbf{x}}} & \text{for} \quad \mathbf{z}_k^{\mathbf{x}} \in \Gamma_k^{\mathbf{x}} \\ 0 & \text{for} \quad \mathbf{z}_k^{\mathbf{x}} \notin \Gamma_k^{\mathbf{x}} \end{cases} .$$

Often, validation gates are defined as hyper-ellipsoidal regions around $\hat{\mathbf{z}}_k^{\mathbf{x}}$ such that $P_G^{\mathbf{x}_k} = P_G$ is a constant. This is done by choosing

$$\Gamma_k^{\mathbf{x}}(\gamma) = \{\mathbf{z} : (\mathbf{z} - \hat{\mathbf{z}}_k^{\mathbf{x}})^T (\mathbf{P}_{\mathbf{Z}_k \mathbf{Z}_k}^{\mathbf{x}})^{-1} (\mathbf{z} - \hat{\mathbf{z}}_k^{\mathbf{x}}) \leq \gamma\} \tag{3.2}$$

with a constant parameter γ. As the measurements are normally distributed, it holds that

$$(\mathbf{Z} - \hat{\mathbf{z}}_k^{\mathbf{x}})^T (\mathbf{P}_{\mathbf{Z}_k \mathbf{Z}_k}^{\mathbf{x}})^{-1} (\mathbf{Z} - \hat{\mathbf{z}}_k^{\mathbf{x}}) \sim \chi^2_{n_{\mathbf{z}}} \quad \Rightarrow \quad P_G = P(\mathbf{z}_k \in \Gamma_k^{\mathbf{x}}(\gamma)) = \chi^2_{n_{\mathbf{z}}}(\gamma) \,,$$

with $n_{\mathbf{z}}$ being the dimension of the measurement $\mathbf{z}$. Defining a constant P_G leads to certain γ, which can be obtained from the quantile tables of the $n_{\mathbf{z}}$-dimensional chi-square distribution ($\gamma = \chi^2_{n_{\mathbf{z}}, P_G}$). This allows for determination of $\Gamma_k^{\mathbf{x}} = \Gamma_k^{\mathbf{x}}(\gamma)$ as well as $V_k^{\mathbf{x}} = V_k^{\mathbf{x}}(\gamma)$ which is given as

$$V_k^{\mathbf{x}}(\gamma) = c_{n_{\mathbf{z}}} \gamma^{\frac{n_{\mathbf{z}}}{2}} \left| \mathbf{P}_{\mathbf{Z}_k \mathbf{Z}_k}^{\mathbf{x}} \right|^{\frac{1}{2}} ,$$

where $c_{n_{\mathbf{z}}}$ is the volume of the $n_{\mathbf{z}}$-dimensional unity sphere ($c_1 = 2, c_2 = \pi, c_3 = \frac{4}{3}\pi, c_4 = \frac{1}{2}\pi^2, \cdots$).

The set of $m_k^{\mathbf{x}}$ measurements falling into the gating region of a track $\mathbf{x}$ at time step k is denoted by $\mathcal{Z}_k^{\mathbf{x}}$: $\mathcal{Z}_k^{\mathbf{x}} = \{\mathbf{z}_{k,1}, \ldots, \mathbf{z}_{k,m_k^{\mathbf{x}}}\} \in \Gamma_k^{\mathbf{x}}$. For better readability, the superscript $\mathbf{x}$ in $m_k^{\mathbf{x}}$ will be omitted in the following. For each measurement $\mathbf{z}_{k,j} \in \mathcal{Z}_k^{\mathbf{x}}$ a hypothesis is formed, where this measurement is assumed being correct while all other $m_k - 1$ measurements in the gate are assumed to be caused by clutter. This hypothesis is denoted as $\theta_k^{\mathbf{x} \to \mathbf{z}_j}$ with $j \in \{1 \ldots m_k\}$. $\theta_k^{\mathbf{x} \to \mathbf{z}_0}$ denotes the hypothesis of none of the m_k measurements in gate being correct, i.e., that all of them stem from clutter or are false alarms.

In the innovation step of the Bayesian state estimator, estimates produced by each hypothesis are weighted with the weighting factors $\beta_k^{\mathbf{x}\to\mathbf{z}_j}$ (with $j\in\{0\dots m_k\}$) that are defined as

$$\beta_k^{\mathbf{x}\to\mathbf{z}_j}=P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|\mathcal{Z}_{1:k}^{\mathbf{x}})$$

with $\mathcal{Z}_{1:k}^{\mathbf{x}}=\{\mathcal{Z}_1^{\mathbf{x}},\dots,\mathcal{Z}_k^{\mathbf{x}}\}$ and $\sum_{j=0}^{m_k}\beta_k^{\mathbf{x}\to\mathbf{z}_j}=1$. The weighting factors are calculated using Bayes' theorem:

$$\begin{aligned}\beta_k^{\mathbf{x}\to\mathbf{z}_j}&=P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|\mathcal{Z}_{1:k}^{\mathbf{x}})=P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|\mathcal{Z}_k^{\mathbf{x}},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})\\&=\frac{1}{c_k}p(\mathcal{Z}_k^{\mathbf{x}}|\theta_k^{\mathbf{x}\to\mathbf{z}_j},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})\end{aligned}$$

with c_k being the normalization factor. Assuming a Gaussian measurement distribution, the likelihood of the true measurement $\mathbf{z}_{k,j}$ ($j\neq 0$) is given by

$$\begin{aligned}p(\mathbf{z}_{k,j}|\theta_k^{\mathbf{x}\to\mathbf{z}_j},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})&=\tfrac{1}{P_G}f_{\mathbf{z}_k^{\mathbf{x}}}=\tfrac{1}{P_G}\mathcal{N}(\mathbf{z}_{k,j};\hat{\mathbf{z}}_k^{\mathbf{x}},\mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}})\\&=\tfrac{1}{P_G}\mathcal{N}(\mathbf{z}_{k,j}-\hat{\mathbf{z}}_k^{\mathbf{x}};0,\mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}})=\tfrac{1}{P_G}\mathcal{N}(\tilde{\mathbf{z}}_{k,j}^{\mathbf{x}};0,\mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}})\\&=\tfrac{1}{P_G}\cdot|2\pi\cdot\mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}}|^{-\frac{1}{2}}\cdot e^{-\frac{1}{2}(\tilde{\mathbf{z}}_{k,j}^{\mathbf{x}})^T(\mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}})^{-1}\tilde{\mathbf{z}}_{k,j}^{\mathbf{x}}}\end{aligned}$$

with innovation $\tilde{\mathbf{z}}_{k,j}^{\mathbf{x}}=\mathbf{z}_{k,j}-\hat{\mathbf{z}}_k^{\mathbf{x}}$.

Clutter measurements are assumed to be independent from the correct measurement. Their position is assumed to be independent and identically distributed over the whole gating region with uniform distribution on $\Gamma_k^{\mathbf{x}}$. Under these assumptions,

$$p(\mathbf{z}_{k,i}|\theta_k^{\mathbf{x}\to\mathbf{z}_j},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})=\frac{1}{V_k}\quad j\neq i\,.$$

The likelihood of the entire measurement set $\mathcal{Z}_k^{\mathbf{x}}$ falling into the gating region of the track $\mathbf{x}$ at time step k given that either all of them are false alarms ($\theta_k^{\mathbf{x}\to\mathbf{z}_0}$) or the measurement j is the correct measurement and all other measurements are false alarms ($\theta_k^{\mathbf{x}\to\mathbf{z}_j}$, $j=1...m_k$) is given by

$$\begin{aligned}p(\mathcal{Z}_k^{\mathbf{x}}|\theta_k^{\mathbf{x}\to\mathbf{z}_0},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})&=\prod_{i=1}^{m_k}p(\mathbf{z}_{k,i}|\theta_k^{\mathbf{x}\to\mathbf{z}_0},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})=\frac{1}{V_k^{m_k}}\,,\\p(\mathcal{Z}_k^{\mathbf{x}}|\theta_k^{\mathbf{x}\to\mathbf{z}_j},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})&=\prod_{i=1}^{m_k}p(\mathbf{z}_{k,i}|\theta_k^{\mathbf{x}\to\mathbf{z}_j},m_k,\mathcal{Z}_{1:(k-1)}^{\mathbf{x}})\\&=\frac{1}{V_k^{m_k-1}}\tfrac{1}{P_G}\mathcal{N}(\tilde{\mathbf{z}}_{k,j};0,\mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{\mathbf{x}})\,,\quad j=1,\cdots,m_k\,.\end{aligned}$$

The probability mass function of the hypothesis $\theta_k^{\mathbf{x}\to\mathbf{z}_j}$ conditioned on m_k and $\mathcal{Z}^{\mathbf{x}}_{1:(k-1)}$ is given by

$$\begin{aligned} P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|m_k, \mathcal{Z}^{\mathbf{x}}_{1:(k-1)}) &= P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|m_k) \\ &= \frac{P(m_k|\theta_k^{\mathbf{x}\to\mathbf{z}_j})P(\theta_k^{\mathbf{x}\to\mathbf{z}_j})}{\sum_{j=0}^{m_k} P(m_k|\theta_k^{\mathbf{x}\to\mathbf{z}_j})P(\theta_k^{\mathbf{x}\to\mathbf{z}_j})}\,, \end{aligned} \tag{3.3}$$

where $P(\theta_k^{\mathbf{x}\to\mathbf{z}_j})$ (with $j = 1...m_k$) denotes the a-priori probability that the measurement $\mathbf{z}_j$ originated from track $\mathbf{x}$, $P(\theta_k^{\mathbf{x}\to\mathbf{z}_0})$ denotes the a-priori probability that none of the measurements in the gate has been evoked by track $\mathbf{x}$ and $P(m_k|\theta_k^{\mathbf{x}\to\mathbf{z}_0})$ and $P(m_k|\theta_k^{\mathbf{x}\to\mathbf{z}_j})$ denote the probabilities for receiving m_k measurements given that either none or one of them stems from track $\mathbf{x}$.

$P(\theta_k^{\mathbf{x}\to\mathbf{z}_0})$ is obviously given by

$$P(\theta_k^{\mathbf{x}\to\mathbf{z}_0}) = 1 - P_D P_G\,, \tag{3.4}$$

where P_D is the probability that the track evokes a measurement (detection probability), and P_G is the probability of the measurement to fall into the gating region as defined in (3.1).

Under the assumption that each of the m_k measurements in the gate has equal probability of being evoked by track $\mathbf{x}$, the a-priori association probability $P(\theta_k^{\mathbf{x}\to\mathbf{z}_j})$ for $j = 1...m_k$ is given by

$$P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}) = \frac{1}{m_k} P_D P_G \quad \forall j = 1, ..., m_k\,.$$

The probability of the number of measurements being m_k given one of the association hypotheses $\theta_k^{\mathbf{x}\to\mathbf{z}_0}$ or $\theta_k^{\mathbf{x}\to\mathbf{z}_j}$ is equivalent to the probability of the number of false measurements being m_k or $m_k - 1$ correspondingly:

$$\begin{aligned} P(m_k|\theta_k^{\mathbf{x}\to\mathbf{z}_0}) &= \mu_F(m_k)\,, \\ P(m_k|\theta_k^{\mathbf{x}\to\mathbf{z}_j}) &= \mu_F(m_k - 1) \end{aligned} \tag{3.5}$$

with $\mu_F(m)$ being the probability mass function for the number of clutter-based measurements. $\mu_F(m)$ can be modeled in different ways. The number of the clutter-based measurements can be assumed either to have Poisson distribution (***parametric model***) or to be equally distributed over the set $\{0, \cdots, N-1\}$ with N being the maximal number of clutter-based measurements (***non-parametric model***).

Parametric model: Poisson distribution

$$\mu_F(m) = e^{-\hat{m}_k}\frac{\hat{m}_k^m}{m!} = e^{-\lambda V_k}\frac{(\lambda V_k)^m}{m!}\,, \qquad m \in \mathcal{N}_0\,,$$

with λ being the mean clutter density and $\hat{m}_k := \lambda V_k$ being the expected number of clutter measurements in the gating region. If λ is a-priori not known, $\hat{m}_k$ can be estimated by using $\hat{m}_k = m_k - P_D P_G$.

Non-parametric model: Uniform distribution

$$\mu_F(m) = \frac{1}{N}\,, \qquad m = 0, 1, \ldots, N-1\,,$$

where N can be chosen as a great enough arbitrary number since it will be canceled in $P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|m_k, \mathcal{Z}^{\mathbf{x}}_{1:(k-1)})$.

Using (3.4) - (3.5) in (3.3) leads to

$$P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|m_k, \mathcal{Z}^{\mathbf{x}}_{1:(k-1)}) = \begin{cases} \frac{\mu_F(m_k)(1-P_D P_G)}{\mu_F(m_k)(1-P_D P_G)+m_k\cdot\mu_F(m_k-1)\frac{P_D P_G}{m_k}} & j=0 \\ \\ \frac{\mu_F(m_k-1)\frac{P_D P_G}{m_k}}{\mu_F(m_k)(1-P_D P_G)+m_k\cdot\mu_F(m_k-1)\frac{P_D P_G}{m_k}} & j=1,...,m_k \end{cases}$$

and thus to

$$P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|m_k, \mathcal{Z}^{\mathbf{x}}_{1:(k-1)}) = \begin{cases} \frac{(1-P_D P_G)\lambda V_k}{P_D P_G m_k+(1-P_D P_G)\lambda V_k} & j=0 \\ \\ \frac{P_D P_G}{P_D P_G m_k+(1-P_D P_G)\lambda V_k} & j=1,\ldots,m_k \end{cases}$$

for the *parametric model* and to

$$P(\theta_k^{\mathbf{x}\to\mathbf{z}_j}|m_k, \mathcal{Z}^{\mathbf{x}}_{1:(k-1)}) = \begin{cases} (1-P_D P_G) & j=0 \\ \\ \frac{1}{m_k}P_D P_G & j=1,\ldots,m_k \end{cases}$$

for the *non-parametric model*.

This leads to the following weighting factors $\beta_k^{x\to z_j}$:

$$\beta_k^{x\to z_j} = \begin{cases} \frac{b}{b+\sum_{i=1}^{m_k} e_i} & j=0 \\ \\ \frac{e_j}{b+\sum_{i=1}^{m_k} e_i} & j=1,\ldots,m_k \end{cases}$$

with

$$e_j = e^{-\frac{1}{2}(\tilde{\mathbf{z}}^{\mathbf{x}}_{k,j})^T (\mathbf{P}^{\mathbf{x}}_{\mathbf{Z}_k\mathbf{Z}_k})^{-1}\tilde{\mathbf{z}}^{\mathbf{x}}_{k,j}}$$

and

$$b = \lambda|2\pi\mathbf{P}^{\mathbf{x}}_{\mathbf{Z}_k\mathbf{Z}_k}|^{\frac{1}{2}}\frac{(1-P_D P_G)}{P_D} = \begin{cases} \left(\frac{2\pi}{\gamma}\right)^{\frac{n_z}{2}}\lambda V_k c_{n_z}\frac{(1-P_D P_G)}{P_D}\,, & \text{parametric model} \\ \left(\frac{2\pi}{\gamma}\right)^{\frac{n_z}{2}} m_k c_{n_z}\frac{(1-P_D P_G)}{P_D}\,, & \text{non-parametric model.} \end{cases}$$

For each hypothesis, the corresponding state estimate is given by

$$\hat{\mathbf{x}}^{\mathbf{z}_j}_k = \mathbb{E}[\mathbf{X}_k|\theta^{\mathbf{x}\to\mathbf{z}_j}_k, \mathcal{Z}^{\mathbf{x}}_{1:k}] = \begin{cases} \hat{\mathbf{x}}^-_k & j = 0 \\ \hat{\mathbf{x}}^-_k + \mathbf{K}_k(\mathbf{z}_{k,j} - \hat{\mathbf{z}}^{\mathbf{x}}_k) & j = 1,\dots,m_k\,. \end{cases}$$

When considering all hypotheses, this leads to the following composite state estimate for the track $\mathbf{x}$:

$$\begin{aligned} \hat{\mathbf{x}}_k =& \mathbb{E}[\mathbf{X}_k|\mathcal{Z}^{\mathbf{x}}_{1:k}] = \sum_{j=0}^{m_k}\mathbb{E}[\mathbf{X}_k|\theta^{\mathbf{x}\to\mathbf{z}_j}_k, \mathcal{Z}^{\mathbf{x}}_{1:k}]\cdot P(\theta^{\mathbf{x}\to\mathbf{z}_j}_k|\mathcal{Z}^{\mathbf{x}}_{1:k}) \\ =& \sum_{j=0}^{m_k}\hat{\mathbf{x}}^{\mathbf{z}_j}_k\beta^{\mathbf{x}\to\mathbf{z}_j}_k = \sum_{j=0}^{m_k}\beta^{\mathbf{x}\to\mathbf{z}_j}_k\hat{\mathbf{x}}^-_k + \mathbf{K}_k\sum_{j=1}^{m_k}\beta^{\mathbf{x}\to\mathbf{z}_j}_k(\mathbf{z}_{k,j} - \hat{\mathbf{z}}^{\mathbf{x}}_k) \\ =& \hat{\mathbf{x}}^-_k + \mathbf{K}_k\sum_{j=1}^{m_k}\beta^{\mathbf{x}\to\mathbf{z}_j}_k\tilde{\mathbf{z}}^{\mathbf{x}}_{k,j} \end{aligned} \tag{3.6}$$

with composite innovation $\tilde{\mathbf{z}}^{\mathbf{x}}_{k,\text{Comp}} := \sum_{j=1}^{m_k}\beta^{\mathbf{x}\to\mathbf{z}_j}_k\tilde{\mathbf{z}}^{\mathbf{x}}_{k,j}$.

Although the equation (3.6) seems to be linear, this is not the case as the weighting factors $\beta^{\mathbf{x}\to\mathbf{z}_j}_k$ depend on $\tilde{\mathbf{z}}^{\mathbf{x}}_{k,j}$.

The covariance matrix $\mathbf{P}_{\mathbf{X}_k\mathbf{X}_k}$ is calculated according to

$$\mathbf{P}_{\mathbf{X}_k\mathbf{X}_k} = \beta^{\mathbf{x}\to\mathbf{z}_0}_k\mathbf{P}_{\mathbf{X}^-_k\mathbf{X}^-_k} + (1-\beta^{\mathbf{x}\to\mathbf{z}_0}_k)\mathbf{P}^c_{\mathbf{X}_k\mathbf{X}_k} + \tilde{\mathbf{P}}_k$$

with

$$\mathbf{P}^c_{\mathbf{X}_k\mathbf{X}_k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{\mathbf{X}^-_k\mathbf{X}^-_k}$$

and

$$\tilde{\mathbf{P}}_k = \mathbf{K}_k\Big(\sum_{j=1}^{m_k}\beta^{\mathbf{x}\to\mathbf{z}_j}_k\tilde{\mathbf{z}}^{\mathbf{x}}_{k,j}(\tilde{\mathbf{z}}^{\mathbf{x}}_{k,j})^T - \tilde{\mathbf{z}}^{\mathbf{x}}_{k,\text{Comp}}(\tilde{\mathbf{z}}^{\mathbf{x}}_{k,\text{Comp}})^T\Big)\mathbf{K}^T_k\,.$$

Hereby, the predicted covariance matrix $\mathbf{P}_{\mathbf{X}_k\mathbf{X}_k}$ is weighted with the factor $\beta_k^{\mathbf{x}\to\mathbf{z}_0}$, which is related to the case of none of the obtained measurement being correct. $\mathbf{P}^c_{\mathbf{X}_k\mathbf{X}_k}$ is the covariance matrix calculated under the assumption that the innovation is performed with the correct measurement, i.e., without association error. It is weighted with the factor $(1-\beta_k^{\mathbf{x}\to\mathbf{z}_0})$. Since it is not known which of the m_k measurements is the correct one, the state covariance is increased by means of the matrix $\tilde{\mathbf{P}}_k$ which incorporates the measurement association errors.

3.3 Joint Probabilistic Data Association (JPDA)

In the PDA, each track is considered separately. This justifies the assumption that either all or all but one measurements falling into the gating region of a track are due to clutter. In the presence of multiple closely spaced targets this assumption may be invalid since true measurements of one target may fall into the gating region of another target causing permanent non-random interference. This issue is accounted for in the extension of the PDA called ***Joint Probabilistic Data Association (JPDA)*** proposed by Bar Shalom et al. [FBSS83]. Instead of considering each track separately, JPDA considers association configurations, the so-called ***joint events***. A joint event $\Theta_k(T)$ is defined as an conjunction of associations $\theta_k^{\mathbf{x}_{t_j}\to\mathbf{z}_j}$ between measurements $\mathbf{z}_j$ and possible causes $\mathbf{x}_{t_j}$ that can be given by either an existing track ($t_j \neq 0$) or clutter ($t_j = 0$):

$$\Theta_k(T) = \bigcap_{j=1}^{m_k} \theta_k^{\mathbf{x}_{t_j}\to\mathbf{z}_j}, \quad T = (t_1, ..., t_{m_k}), \quad t_j \in \{0, \dots, n_k\},$$

with n_k being number of currently tracked targets. T are ordered sets of m_k (possibly repeating) track numbers including 0, which represents the clutter source.

For reduction of complexity, tracks are partitioned into independent clusters and joint events are built for each cluster separately. A cluster is defined as a set of tracks which share no measurements with tracks that do not belong to the cluster.

For easier clutter handling, clutter measurements are considered to be identically distributed over the whole cluster volume V independently of the gating regions of the tracks. This implies that each measurement should be able to be associated with each track in the cluster and hence $P_G = 1$. However, this would also imply usage of too far lying measurements for update of a track. In order to avoid this, a binary ***validation matrix*** Ω_k is defined:

$$\Omega_k = [\omega_{ji}]_k, \quad j = 1, \cdots, m_k; \quad i = 0, 1, \cdots, n_k$$

with

$$\omega_{ji} = \begin{cases} 0 & \text{if} \quad \mathbf{z}_j \notin \Gamma^{\mathbf{x}_i}(\gamma) \\ 1 & \text{if} \quad \mathbf{z}_j \in \Gamma^{\mathbf{x}_i}(\gamma)\,, \end{cases}$$

and gating regions $\Gamma^{\mathbf{x}_i}(\gamma)$ as defined in (3.2). The first column ($i = 0$) of Ω_k stands for association with no track, i.e., indicates that a measurement j stems from clutter. As mentioned above, this can be applicable to each measurement in cluster, hence $\forall j : \omega_{j0} = 1$.

Each joint event $\Theta_k(T)$ can be represented through a binary matrix $\hat{\Omega}(\Theta_k(T))$ with

$$\hat{\Omega}(\Theta_k(T)) = [\hat{\omega}_{ji}(\Theta_k(T))], \quad j = 1, \cdots, m_k; \quad i = 0, 1, \cdots, n_k$$

and

$$\hat{\omega}_{ji}(\Theta_k(T)) = \begin{cases} 1 & \text{if} \quad \theta_k^{\mathbf{x}_i \to \mathbf{z}_j} \subset \Theta_k(T) \\ 0 & \text{else}\,. \end{cases}$$

In JPDA, a joint event $\Theta_k(T)$ is considered to be ***"feasible"*** under following conditions:

- A measurement may have only one origin:

$$\sum_{i=0}^{n_k} \hat{\omega}_{ji}(\Theta_k(T)) = 1, \quad j = 1, \cdots, m_k$$

- A track may evoke at most one measurement:

$$\sum_{j=1}^{m_k} \hat{\omega}_{ji}(\Theta_k(T)) \leq 1, \quad i = 1, \cdots, n_k\,.$$

A matrix $\hat{\Omega}$ defining a feasible event $\Theta_k(T)$ can be built from the validation matrix Ω_k by picking out elements in a way such that each row and each column contains at most one "1". The only exception is made for the first column which may contain multiple non-zero entries since more than one measurement may be due to clutter. The set of all feasible joint events in the following is denoted by Ξ_k with

$$\sum_{\{\Theta_k(T) \in \Xi_k\}} P(\Theta_k(T)) = 1\,.$$

For better readability, in the following three auxiliary entities $\vartheta_i(\Theta_k(T))$, $\tau_j(\Theta_k(T))$ and $\phi(\Theta_k(T))$ are defined for a joint event $\Theta_k(T)$:

$$\vartheta_i(\Theta_k(T)) := \sum_{j=1}^{m_k} \hat{\omega}_{ji}(\Theta_k(T)), \quad i = 1, \cdots, n_k$$
$$\tau_j(\Theta_k(T)) := \sum_{i=1}^{n_k} \hat{\omega}_{ji}(\Theta_k(T)), \quad j = 1, \cdots, m_k$$
$$\phi(\Theta_k(T)) := \sum_{j=1}^{m_k} (1 - \tau_j(\Theta_k(t)))$$

$\vartheta_i(\Theta_k(T))$ indicates whether in $\Theta_k(T)$ the ith track has been assigned a measurement. $\tau_j(\Theta_k(T))$ indicates whether the jth measurement has been assigned to a track. Finally, $\phi(\Theta_k(T))$ specifies the number of the clutter based measurements in $\Theta_k(T)$.

The weighting factors $\beta_k^{\mathbf{x}_i \to \mathbf{z}_j}$ $(i = 1, \cdots, n_k; \quad j = 0, \cdots, m_k)$ can be calculated as follows:

$$\beta_k^{\mathbf{x}_i \to \mathbf{z}_j} := P(\theta_k^{\mathbf{x}_i \to \mathbf{z}_j} | \mathcal{Z}_{1:k}) = \sum_{\Theta_k(T) \in \Xi_k} P(\Theta_k(T) | \mathcal{Z}_{1:k})\, \hat{\omega}_{ji}(\Theta_k(T))\,.$$

The a-posteriori probability of a joint event $\Theta_k(T)$ conditioned on all received measurements including the current measurement set can be calculated using the Bayes' rule:

$$\begin{aligned} P(\Theta_k(T)|\mathcal{Z}_{1:k}) &= P(\Theta_k(T)|\mathcal{Z}_k, m_k, \mathcal{Z}_{1:(k-1)}) \\ &= \tfrac{1}{c_k} p(\mathcal{Z}_k|\Theta_k(T), m_k, \mathcal{Z}_{1:(k-1)}) P(\Theta_k(T)|m_k, \mathcal{Z}_{1:(k-1)}) \end{aligned} \tag{3.7}$$

with c_k being the normalization constant.

Similar to the calculations in the PDA, the likelihood of a measurement $\mathbf{z}_{k,j}$ given that it stems from a track $\mathbf{x}_{t_j}$ with $t_j \neq 0$ or from clutter ($t_j = 0$) is given by

$$p(\mathbf{z}_{k,j}|\theta_k^{\mathbf{x}_{t_j} \to \mathbf{z}_j}, \mathcal{Z}_{1:(k-1)}) = \begin{cases} \mathcal{N}(\mathbf{z}_{k,j}; \hat{\mathbf{z}}_k^{\mathbf{x}_{t_j}}, \mathbf{P}_{\mathbf{Z}_k \mathbf{Z}_k}^{x_{t_j}}) & \text{for } t_j \neq 0\,, \\ \frac{1}{V} & \text{for } t_j = 0\,. \end{cases}$$

Under the previously mentioned independence assumption of the clutter-based measurements and the true measurements, this leads to the following expression for the likelihood of the current measurement set $\mathcal{Z}_k$ conditioned on a joint event

$\Theta_k(T)$ and number of measurements being m_k:

$$p(\mathcal{Z}_k|\Theta_k(T), m_k, \mathcal{Z}_{1:(k-1)}) = \prod_{j=1}^{m_k} p(\mathbf{z}_{k,j}|\theta_k^{\mathbf{x}_{t_j}\to\mathbf{z}_j}, \mathcal{Z}_{1:(k-1)})$$
$$= \frac{1}{V^{\phi(\Theta_k(T))}} \prod_{j=1}^{m_k} \left(\mathcal{N}(\mathbf{z}_{k,j}; \hat{\mathbf{z}}_k^{\mathbf{x}_{t_j}}, \mathbf{P}_{\mathbf{z}_k\mathbf{z}_k}^{x_{t_j}})\right)^{\tau_j(\Theta_k(T))}$$

where $\phi(\Theta_k(T))$ is the number of the clutter-based measurements in the joint event $\Theta_k(T)$ and $\tau_j(\Theta_k(T))$ serves for picking out the likelihoods of the measurements, that in $\Theta_k(T)$ have been declared as being non-clutter.

The a-priori probability $P(\Theta_k(T)|m_k, \mathcal{Z}_{1:(k-1)})$ of a joint event $\Theta_k(T) \in \Xi_k$ in (3.7), conditioned on the number of received measurements is equivalent to the probability of assigning the tracks according to $\vartheta_i(\Theta_k(T))$ and getting additionally $\phi(\Theta_k(T))$ clutter-based measurements:

$$\begin{aligned}
&P(\Theta_k(T)|m_k, \mathcal{Z}_{1:(k-1)}) \\
&\quad = P\big(\Theta_k(T),\, \vartheta(\Theta_k(T))_1, \ldots, \vartheta_{n_k}(\Theta_k(T)),\, \phi(\Theta_k(T))|\mathcal{Z}_{1:(k-1)}\big) \\
&\quad\quad = P\big(\Theta_k(T)\,|\,\vartheta(\Theta_k(T))_1, \ldots, \vartheta_{n_k}(\Theta_k(T)),\, \phi(\Theta_k(T))\big) \\
&\quad\quad\quad \cdot P\big(\vartheta_1(\Theta_k(T)), \ldots, \vartheta_{n_k}(\Theta_k(T)),\, \phi(\Theta_k(T))|\mathcal{Z}_{1:(k-1)}\big)\,.
\end{aligned}$$

An expression for the computation of the first factor follows from combinatorics with an assumption that each of the joint events $\Theta_k(T)$ has equal a-priori probability. It is given as a reciprocal of the number of all events that assign measurements to the tracks as defined by $\vartheta_i(\Theta_k(T))$ for $i = 1, \ldots, n_k$ and have $\phi(\Theta_k(T))$ clutter measurements:

$$P(\Theta_k(T)|\vartheta_1(\Theta_k(T)), ..., \vartheta_{n_k}(\Theta_k(T)), \phi(\Theta_k(T))) = \left(\frac{m_k!}{\phi(\Theta_k(T))!}\right)^{-1}$$
$$= \frac{\phi(\Theta_k(T))!}{m_k!}$$

The second factor is given by

$$P(\vartheta_1(\Theta_k(T)), \ldots, \vartheta_{n_k}(\Theta_k(T)), \phi(\Theta_k(T))|\mathcal{Z}_{1:(k-1)})$$
$$= \prod_{i=1}^{n_k} \left((P_D^{\mathbf{x}_i})^{\vartheta_i(\Theta_k(T))} \cdot (1 - P_D^{\mathbf{x}_i})^{1-\vartheta_i(\Theta_k(T))}\right) \cdot \mu_F(\phi(\Theta_k(T)))$$

with $P_D^{\mathbf{x}_i}$ being the probability for the track $\mathbf{x}_i$ to be detected and $\mu_F(\phi(\Theta_k(T)))$ being the probability mass function for the number of clutter-based measurements that can be modeled as described in Section 3.2 (see page 13).

This leads to

$$P(\Theta_k(T)|\mathcal{Z}_{1:k}) = \frac{\phi(\Theta_k(T))!}{c_k m_k!} \frac{\mu_F(\phi(\Theta_k(T)))}{V^{\phi(\Theta_k(T))}} \prod_{j=1}^{m_k} \left(\mathcal{N}(\mathbf{z}_{k,j}; \hat{\mathbf{z}}_k^{\mathbf{x}_{t_j}}, \mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{x_{t_j}})\right)^{\tau_j(\Theta_k(T))} \cdot \prod_{i=1}^{n_k} \left((P_D^{\mathbf{x}_i})^{\vartheta_i(\Theta_k(T))} (1 - P_D^{\mathbf{x}_i})^{1-\vartheta_i(\Theta_k(T))}\right)$$

and hence to

$$P(\Theta_k(T)|\mathcal{Z}_{1:k}) = \lambda^{\phi(\Theta_k(T))} \frac{e^{-\lambda V}}{c_k \cdot m_k!} \prod_{j=1}^{m_k} \left(\mathcal{N}(\mathbf{z}_{k,j}; \hat{\mathbf{z}}_k^{\mathbf{x}_{t_j}}, \mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{x_{t_j}})\right)^{\tau_j(\Theta_k(T))} \cdot \prod_{i=1}^{n_k} \left((P_D^{\mathbf{x}_i})^{\vartheta_i(\Theta_k(T))} (1 - P_D^{\mathbf{x}_i})^{1-\vartheta_i(\Theta_k(T))}\right)$$

for the *parametric model* of clutter distribution and to

$$P(\Theta_k(T)|\mathcal{Z}_{1:k}) = \frac{1}{\tilde{c}_k} \frac{\phi(\Theta_k(T))!}{V^{\phi(\Theta_k(T))}} \prod_{j=1}^{m_k} \left(\mathcal{N}(\mathbf{z}_{k,j}; \hat{\mathbf{z}}_k^{\mathbf{x}_{t_j}}, \mathbf{P}_{\mathbf{Z}_k\mathbf{Z}_k}^{x_{t_j}})\right)^{\tau_j(\Theta_k(T))} \cdot \prod_{i=1}^{n_k} \left((P_D^{\mathbf{x}_i})^{\vartheta_i(\Theta_k(T))} (1 - P_D^{\mathbf{x}_i})^{1-\vartheta_i(\Theta_k(T))}\right)$$

for the *nonparametric model* of clutter distribution.

4 Conclusion and Outlook

This report has presented basics of the state-of-the-art methods for tracking of multiple objects in cluttered environments. An overview and a detailed description of the basic state-of-the-art approaches for data association and dynamic state estimation has been given. However, all described data association approaches consider existence of the tracked targets as given. Track initiation and maintenance has to be done outside of scope of the tracking algorithms.

In practice, target existence is often subject to uncertainties due to great amount of clutter and missing detections. An elegant way of modeling those uncertainties has been proposed by Mušicki et al. The ***Integrated Probabilistic Data Association (IPDA)*** and ***Joint Integrated Probabilistic Data Association (JIPDA)*** algorithms proposed in [MES94, ME02] are extensions of the PDA and JPDA algorithms respectively. Additionally to the expressions for data association probabilities they

provide expressions for computation of the track existence probabilities that are directly accounted for (integrated) when computing the association probabilities.

Track existence is modeled as a Markov process with the constant state transition probabilities between the states “track exists” and “track does not exist”. Observability aspect can be also accounted for by using three states (“track exists and is observable”, “track exists but is not observable” and “track does not exist”).

Estimation of the track existence probability offers a solid basis for track initiations and terminations and allows for better handling of clutter and missing detections. However, in some applications such as vision-based object tracking, corrupted measurements due to split, merged and incomplete detections bear an additional source for problems. Here, the above-mentioned approaches have to be extended in order to be able to cope with the introduced effects. Low-level information which can be obtained by robust re-identification and tracking of dedicated feature points in the image offers great potential for solving such problems. An approach which utilizes such information and allows for handling of incomplete, split and merged detections has been proposed in [GOB09]. It is called Feature-Based Probabilistic Data Association and Tracking Algorithm (FBPDA).

Bibliography

[Bro98] Eli Brookner. *Tracking and Kalman filtering made easy*. John Wiley & Sons, New York, 1998.

[BS78] Yaakov Bar-Shalom. Tracking Methods in a Multitarget Environment. In *IEEE Transactions on Automatic Control*, volume 23, pages 618–626, 1978.

[BSL93] Yaakov Bar-Shalom and Xiao-Rong Li. *Estimation and Tracking: Principles, techniques, and software*. Artech House, Boston, London, 1993.

[BST75] Yaakov Bar-Shalom and Edison Tse. Tracking in a Cluttered Environment with Probability Data Association. *Automatica*, 11:451–460, 1975.

[CH96] Ingemar J. Cox and Sunita L. Hingorani. An efficient implementation of Reid’s multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 18, pages 138–150, 1996.

[FBSS83] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe. Sonar Tracking of Multiple Targets Using Joint Probabilistic Data Association. *IEEE Journal of Oceanic Engineering*, 8(3):173–183, 1983.

[GOB09] Michael Grinberg, Florian Ohr, and Jürgen Beyerer. Feature-Based Probabilistic Data Association (FBPDA) for Visual Multi-Target Detection and Tracking under Occlusions and Split and Merge Effects. In *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems*, pages 291–298, 2009.

[Kal60] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[ME02] Darko Mušicki and Rob Evans. Joint Integrated Probabilistic Data Association - JIPDA. In *Proceedings of the 5th International Conference on Information Fusion*, pages 1120–1125, 2002.

[MES94] Darko Mušicki, Rob Evans, and S. Stankovic. Integrated Probabilistic Data Association. In *IEEE Transactions on Automatic Control*, volume 39, 1994.

[Rei79] Donald B. Reid. An Algorithm for Tracking Multiple Targets. In *IEEE Transactions on Automatic Control*, volume AC-24, pages 843–854, 1979.

[Sim06] Dan Simon. *Optimal state estimation: Kalman, H∞, and nonlinear approaches*. Wiley-Interscience, Hoboken, NJ, 2006.