

Track Association Problems In AD

Dr. Waqas Afzal

November 28, 2022

Nomenclature

C Clutter measurement matrix.

\mathbf{c} Clutter measurement vector.

θ Data association vector.

Z Measurement matrix.

\mathbf{z} Measurement vector.

m Number of measurements.

n Number of objects.

X Object state matrix.

O Object measurement matrix.

\mathbf{o} Object measurement vector.

\mathbf{x} Object state vector.

Θ Set of all data association vectors.

k Time step.

Acronyms

CAP	Combat Air Patrol.
CDF	Cumulative Density Function.
EM	Electromagnetic.
ESM	Electronic Support Measures.
GNN	Global Nearest Neighbor.
HMM	Hidden Markov Model.
JPDA	Joint Probabilistic Data Association.
MHT	Multi Hypothesis Tracker.
MMSE	Minimum Mean Squared Error.
MOT	Multiple Object Tracking.
MSE	Mean Squared Error.

NLOS	Non Line-of-Sight.
PDF	Probability Density Function.
PMF	Probability Mass Function.
SNR	Signal-to-Noise Ratio.
SOT	Single Object Tracking.
TBD	Track-Before-Detect.

Abstract

This article provides a survey of the problems arising in multi-target tracking in multi-sensor scenarios in aircraft tracking.

1 Introduction

In scenarios, where multiple overlapping sensors are employed to track multiple aircraft targets, track association becomes of paramount importance. Errors in track association can lead to consequent errors in decision support. This can result in higher level errors in strategy making.

With the increase in the number of sensors being employed to track targets, a common problem faced by air defense operators is that the number of tracks is inconsistent with the number of targets. This affects the operators ability to analyze formations and subsequently, the intent of the aircraft.

This paper provides an introductory survey of the track association problem in the multi-target, multi-sensor scenario. An attempt is made to exhaustively present the basics of this problem.

We shall start off with a couple of definitions:

- **Single Object Tracking (SOT):** Single Object Tracking (SOT) is a filtering problem. It can be described as the sequential processing of noisy sensor measurements to determine the object's state:
 - Position
 - Other properties like kinematics, etc.
 - Other attributes related to the detected object like Electronic Support Measures (ESM) data.
- **Multiple Object Tracking (MOT):** Multiple Object Tracking (MOT) is the sequential processing of noisy sensor measurements to determine
 - the number of dynamic objects,
 - and each dynamic object's state.

Simply put, MOT, jointly detects object and applies SOT to each of these objects.

Typically, in tracking problems, we can have multiple sensors like Radars or Cameras, streaming data like radio signals or image pixels to a detector or processor. The detector then obtains the objects of interest from the data and provides measurements corresponding to each dynamic object to the MOT module (fig. 1). The MOT module processes the measurements and computes the probability densities for the object states.

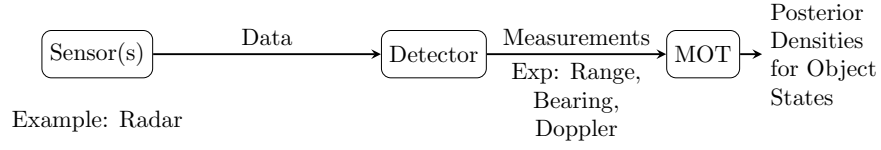


Figure 1: Multiple Object Tracking (MOT)

Scenarios where the detected objects are not easily 'discernible' from the data, the detector can output many false detections. An example of this is a radar detecting objects that are very far away leading to a very low Signal-to-Noise Ratio (SNR) of their received signal. In such, scenarios a Track-Before-Detect (TBD) approach can be employed where MOT is applied directly to the sensors' data (fig. 2).

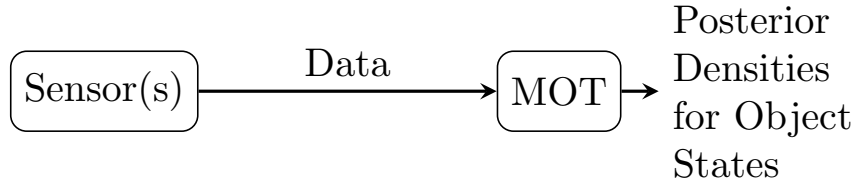


Figure 2: Track-Before-Detect (TBD)

2 Types of Tracking

Dynamic object tracking can be categorized as follows:

1. **Point Object Tracking:** a single detection is returned per object per time step. Therefore, a one-on-one association can be assumed for each object per measurement.
2. **Extended Object Tracking:** multiple detections are returned per object per time step. This can happen due to the size and geometry of objects. For example, a radar can return multiple detection for a single object due to multiple reflections from the same object.
3. **Group Object Tracking:** multiple resolvable objects are treated as a group per time step. The aim is to track the group and not the objects constituting the group. Combat Air Patrols (CAPs) can be taken as an example of group object tracking. It has to be noted here that each object in the group returns separate detections.
4. **Tracking With Multi-Path Propagation:** tracking Non Line-of-Sight (NLOS) dynamic objects is known as tracking with multi-path propagation. Radars are mostly used to track Non Line-of-Sight (NLOS) objects by capturing Electromagnetic (EM) reflected energy from a surface. Therefore, signals are received from multiple 'paths' per object per time step. This mostly results in multiple detections per object per time step.
5. **Tracking with Unresolved Measurements:** in this case, multiple objects return a single detection. This can be due to the objects moving very closely in a synchronized manner.

It can be noted here that extended object, group object and multipath based tracking receive multiple detections per object. Tracking with unresolved measurements, on the hand, can be described as multiple objects per measurement tracking.

3 Challenges in Multiple Object Tracking (MOT)

MOT is a complex problem where multiple problems need to be solved simultaneously. The challenges faced in MOT can be summarized as below.

- The first challenge is detecting objects. Since the number of dynamic object is unknown, detection needs to be done at each time step.
- The states of the detected objects are also unknown, thus their states need to be estimated at each time step.
- The objects are mostly dynamic, thus, their states are constantly changing.
- New objects can enter the sensor's range and objects detected in the previous time step can leave the sensor's range. This is also termed as birth and death of dynamic objects.

- Objects can occlude one another. Therefore, the occluded objects cannot be detected.
- Errors in detection. The sensors used to detect objects have inherent deficiencies and therefore, this can result in false detections (false alarms) or missed detections (errors). MOT needs to take into consideration these errors as well.
- Finally, the most significant challenge in MOT is data association. This problem arises due to:
 - which detections are from objects detected previously.
 - which detections are from new objects in the sensors range.
 - false detections.

4 Bayesian Filtering

The evolution of an object's states over time can be modeled as a Hidden Markov Model (HMM). MOT is basically detection and estimation of the dynamic object's hidden states. The multiple sources of uncertainty mentioned in the previous section makes this an extremely complex task.

Bayesian filters have been shown to perform well in estimating hidden states in the presence of uncertainty. All Bayesian filters successively follow a two step process.

- **Motion Update:** Use a motion model to predict the state of the objects in the next time-step.
- **Measurement Update:** Update the object state's densities using a measurement model based on the measurement obtained in the current time-step.

Let the state of an object at time step k be given by \mathbf{x}_k and a measurement at k be given by \mathbf{z}_k . Assuming we have sequence of measurements given by:

$$\mathbf{z}_{1:\tau} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_\tau)$$

then the posterior is given by

$$p_{\mathbf{x}_k|\mathbf{z}_{1:\tau}}(\mathbf{x}_k|\mathbf{z}_{1:\tau}) = p(\mathbf{x}_k|\mathbf{z}_{1:\tau})$$

Assuming that the posterior is Gaussian and the expected value of an object state at time step k and τ measurements is given as:

$$\bar{\mathbf{x}}_{k|\tau}$$

then the Gaussian posterior is given by

$$p(\mathbf{x}_k|\mathbf{z}_{1:\tau}) = \mathcal{N}(\mathbf{x}_k; \bar{\mathbf{x}}_{k|\tau}, P_{k|\tau})$$

where $P_{k|\tau}$ is the covariance matrix.

Bayesian filtering alternatively applies the prediction and measurement updates to update the posterior of the object state at each time step.

Initial: $p(\mathbf{x}_1)$
 Update: $p(\mathbf{x}_1|\mathbf{z}_1)$
 Prediction: $p(\mathbf{x}_2|\mathbf{z}_1)$
 Update: $p(\mathbf{x}_2|\mathbf{z}_{1:2})$
 Prediction: $p(\mathbf{x}_3|\mathbf{z}_{1:2})$
 Update: $p(\mathbf{x}_3|\mathbf{z}_{1:3})$
 Prediction: $p(\mathbf{x}_4|\mathbf{z}_{1:3})$
 Update: $p(\mathbf{x}_4|\mathbf{z}_{1:4})$
 Prediction: $p(\mathbf{x}_5|\mathbf{z}_{1:4})$

Due to their simplicity, Bayesian filters are widely employed in tracking problems.

Motion Modeling: The prediction step in Bayesian filtering uses a motion model of the object being tracked to predict the next state of the object based on its current state. The motion model is a function f that maps the current state to the next state of the object

$$f : \mathbf{x}_k \rightarrow \mathbf{x}_{k+1}$$

accurate prediction helps in data association. Motion models can be categorized as

- **Linear**
 - Constant Velocity
 - Constant Acceleration
- **Non-linear**
 - Coordinated Turn

Using a motion model is dependent on the current object phase. If the object is moving in a straight line, then constant velocity or acceleration would be good model. However, if an object is turning then the coordinated turn model would be better suited to predict its next state.

Typically, real-life motion is affected by many random factors .e.g. tires can slip when moving. These random factors need to be included in the motion model. Mostly additive Gaussian noise is used to model the random factors affecting the motion.

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}) + q_{k-1}$$

where f can be a linear or non-linear mapping depending on the motion model. This can be described probabilistically using transition densities

$$p(\mathbf{x}_k|\mathbf{x}_{k-1})$$

so for additive Gaussian noise with zero mean and covariance matrix Q the transition density is given by

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; f_{k-1}(\mathbf{x}_{k-1}), Q_{k-1})$$

Measurement Modeling: The measurement update step in Bayesian filtering uses a measurement model based on the sensor being used. The data from the sensor is used in the measurement model to generate the measurement update.

The measurement model h describes the relation between the object state and the measurement.

$$h : \mathbf{x}_k \rightarrow \mathbf{z}_k$$

The measurement model is different for different sensors. Different objects may have different models for the same sensor as well. Typically, sensor measurement are affected by random factors and usually modeled as additive random noise v .

$$\mathbf{z}_k = h_k(\mathbf{x}_k) + v_k$$

where $h()$ can be a linear or non-linear mapping. The measurement likelihood can then be defined as the density

$$p(\mathbf{z}_k | \mathbf{x}_k)$$

if the measurement noise is assumed to be Gaussian with zero mean and covariance R then the measurement likelihood is given as follows

$$p(\mathbf{z}_k | \mathbf{x}_k) = \mathcal{N}(\mathbf{z}_k; h(\mathbf{x}_k), R_k)$$

It has to be noted here that Bayesian Filtering does not require the motion and measurement model to be very accurate. The models that provide sufficient prediction and measurement updates are enough for tracking using Bayesian filtering. The design of the models is therefore, open to the designers choice. More accurate modeling would require more computational power.

4.1 Bayesian Filtering Recursion

The Bayesian filtering recursion of prediction and update can be described using the motion and measurement models as follows:

- **Prediction (Chapman-Kolmogorov Equation):** The motion prediction update is computed as follows

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \quad (1)$$

- **Update (Bayes Rule):** Using eq. (1) as the prior, the measurement update (posterior) can be computed by using the measurement likelihood as follows:

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})} \quad (2)$$

- **Predicted Likelihood (Distribution/Probability of Evidence)**

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) d\mathbf{x}_k \quad (3)$$

4.2 Estimators & Performance Measure

Using distributions can give a good intuition regarding the state of an object. However, plotting distributions of multiple object can lead to confusions. Therefore estimates of the objects' states are obtained from the distributions for a clearer visualization.

Estimates are also needed for quantitatively evaluating the performance of the Bayesian Filtering algorithm. For a given posterior $p(\mathbf{x}_k|\mathbf{z}_{1:\tau})$, the estimate $\hat{\mathbf{x}}$ at time instant k given measurements up to and including time instant τ is given as:

$$\hat{\mathbf{x}}_{k|\tau}$$

4.2.1 Performance Measure

An objective performance measure is needed to gauge the performance of the filtering algorithm. The performance measure provides a way to compare the result with the ground truth. One of the most commonly used measure is the Mean Squared Error (MSE)

$$\text{MSE}(\hat{\mathbf{x}}_{k|\tau}) = \text{E}\{(\hat{\mathbf{x}}_{k|\tau} - \mathbf{x}_k)^T(\hat{\mathbf{x}}_{k|\tau} - \mathbf{x}_k)\} \quad (4)$$

Other performance measures can be obtained depending on the application.

Estimates of the object states can also be obtained using the performance measures as well. For example, an estimate of the object states can be obtained by minimizing the Mean Squared Error (MSE). Commonly known as Minimum Mean Squared Error (MMSE)

$$\hat{\mathbf{x}}_{k|\tau}^{MMSE} = \arg \min_{\hat{\mathbf{x}}_{k|\tau}} \text{MSE}(\hat{\mathbf{x}}_{k|\tau})$$

4.3 Bayesian Filtering Example: Kalman Filter

Kalman filter is a widely used Bayesian filter. If the motion and measurement models are linear with additive Gaussian noise and the prior is also Gaussian, then the Kalman filter gives a closed form solution. In fact if all models are linear and all noise is Gaussian, then the Kalman filter is the Minimum Mean Squared Error (MMSE) estimator. Using the following definitions,

$$\begin{aligned} \text{Motion Model: } p(\mathbf{x}_k|\mathbf{x}_{k-1}) &= \mathcal{N}(\mathbf{x}_k; F_{k-1}\mathbf{x}_{k-1}, Q_{k-1}) \\ \text{Measurement Model: } p(\mathbf{z}_k|\mathbf{x}_k) &= \mathcal{N}(\mathbf{z}_k; H_k\mathbf{x}_k, R_k) \\ \text{Initial Density: } p(\mathbf{x}_0) &= \mathcal{N}(\mathbf{x}_0; \bar{\mathbf{x}}_0, P_0) \\ \text{Posterior Density: } p(\mathbf{x}_k|\mathbf{z}_{1:\tau}) &= \mathcal{N}(\mathbf{x}_k; \bar{\mathbf{x}}_{k|\tau}, P_{k|\tau}) \\ \text{Expected Value: } \bar{\mathbf{x}}_{k|\tau} & \\ \text{Covariance: } P_{k|\tau} & \end{aligned}$$

the Kalman filter updates are given as:

$$\begin{array}{ll}
\text{Prediction:} & \begin{cases} \bar{\mathbf{x}}_{k|k-1} = F_{k-1}\bar{\mathbf{x}}_{k-1|k-1} & \text{Use model to predict} \\ P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^T + Q_{k-1} & \text{Increase uncertainty} \end{cases} \\
\text{Update:} & \begin{cases} \bar{\mathbf{z}}_k = H_k\bar{\mathbf{x}}_{k|k-1} & \text{Predicted Measurement} \\ \epsilon_k = \mathbf{z}_k - \bar{\mathbf{z}}_k & \text{Innovation} \\ S_k = H_kP_{k|k-1}H_k^T + R_k & \text{Innovation Covariance} \\ K_k = P_{k|k-1}H_k^T S_k^{-1} & \text{Kalman gain} \\ \bar{\mathbf{x}}_{k|k} = \bar{\mathbf{x}}_{k|k-1} + K_k\epsilon_k & \text{Weighted average} \\ P_{k|k} = P_{k|k-1} - K_kH_kP_{k|k-1} & \text{Decrease uncertainty} \end{cases}
\end{array}$$

5 Tracking

In object tracking, it is desired to recursively estimate an object's state vector that varies stochastically across time using noisy measurements. This can be achieved using Bayesian Filtering. However, tracking is complicated further by:

- missed detections. In this scenario the object being tracked is not detected.
- clutter detections. Usually we can ignore clutter detections.
- unknown data-associations. Perhaps the most involved challenge as there is no information regarding measurements and their associations.

The states of the n objects at time instant k are given by the matrix,

$$X_k = [\mathbf{x}_k^1, \mathbf{x}_k^2, \dots, \mathbf{x}_k^n]$$

These states need to be tracked in the presence of missed detections and clutter. Therefore, the following is needed to track the object:

- n -object measurement model.
- n -object motion model.
- n -object states prior.
- Methods for handling the data association. Handling multiple measurements association to multiple objects is a challenging problem.

Once the aforementioned data is obtained and consolidated, then the n objects can be tracked using the following algorithms.

- Global Nearest Neighbor (GNN)
- Joint Probabilistic Data Association (JPDA)
- Multi Hypothesis Tracker (MHT)

5.1 Multiple Object Measurement Model

The probability of detection provides a measure of the chances of detecting an object. At each time instant, an object is either detected or missed. Bernoulli distributions are used to model experiments with binary outcomes. A discrete random variable which takes a true value with probability p and a false value with probability $1 - p$ is said to be Bernoulli distributed. If a discrete random variable b is said to be Bernoulli distributed then it is given by:

$$b = \begin{cases} true & \text{with probability } p \\ false & \text{with probability } 1 - p \end{cases} \quad (5)$$

The Binomial distribution models a sequence of trials that are Bernoulli distributed. If j Bernoulli experiments are conducted with probability of getting a true value p then the corresponding Binomial random variable b is given as:

$$Pr[b = i] = \binom{j}{i} p^i (1 - p)^{j-i} \quad (6)$$

It can be noted that $E[b] = pj$ for the above Binomial distribution.

The Poisson distribution is a discrete probability distribution that models a given number of independent events occurring in a fixed time or space interval with a known constant mean occurrence rate. The distribution of a Poisson random variable b with the mean interval occurrence rate λ is given as:

$$Pr[b = i] = Po[i; \lambda] = \frac{\lambda^i \exp(-\lambda)}{i!} \quad (7)$$

It can be noted that $E[p] = \text{Var}[p] = \lambda$.

The measurement model can be updated using the probability of detection of objects. Assuming there are n objects that need to be tracked and the object measurement matrix be denoted by O ,

$$O_k = [O_k^1, O_k^2, \dots, O_k^n] \quad (8)$$

where, O_k^i , corresponds to the i^{th} object and is given as

$$O_k^i = \begin{cases} \emptyset & \text{with probability } 1 - P^D(\mathbf{x}_k^i) \\ \mathbf{o}_k^i & \text{with probability } P^D(\mathbf{x}_k^i) \end{cases} \quad (9)$$

The likelihood for the measurement vector \mathbf{o}_k^i is given by

$$p(\mathbf{o}_k^i | \mathbf{x}_k^i) = g_k(\mathbf{o}_k^i | \mathbf{x}_k^i) \quad (10)$$

The measurement model is then given by

$$p(O_k^i | \mathbf{x}_k^i) = \begin{cases} 1 - P^D(\mathbf{x}_k^i) & \text{if } O_k^i = \emptyset \\ P^D(\mathbf{x}_k^i) g_k(\mathbf{o}_k^i | \mathbf{x}_k^i) & \text{if } O_k^i = \mathbf{o}_k^i \end{cases} \quad (11)$$

Equation (11) captures both the probability of detection and if detected, the distribution of detection.

in the presence of clutter, measurement from clutter will also be received. Let the matrix C represent the clutter measurements. The clutter measurements are given as

$$C_k = [\mathbf{c}_k^1, \mathbf{c}_k^2, \dots, \mathbf{c}_k^{m_k^c}] \quad (12)$$

where m_k^c represents the number of clutter measurements. The clutter C_k is Poisson with intensity,

$$\lambda_c = \bar{\lambda}_c f_c(c) \quad (13)$$

The measurement matrix at time instant k Z_k is then given by

$$Z_k = \Pi(O_k, C_k) \quad (14)$$

where Π randomly shuffles the columns of O_k and C_k .

5.2 Data Association

When tracking multiple objects, a sensor will return multiple measurements. Associating this data with each object is a challenging problem. Techniques need to be developed to correctly associate the data to objects as incorrect association will lead to faulty tracking.

The prior for n objects given by,

$$p(X_0) = p(\mathbf{x}_0^1, \mathbf{x}_0^2, \dots, \mathbf{x}_0^n)$$

If the state random variables are independent then the joint prior density can be written as the product of individual densities. Typically in tracking, the initial prior is assumed to be independent. Therefore it can be written as

$$p(X_0) = \prod_{i=1}^n p^i(\mathbf{x}_0^i)$$

We need to compute an n -dimensional measurement likelihood.

$$p(Z_k|X_k) \quad (15)$$

s This can be expressed as the sum over data associations (time indexing dropped for brevity)

$$p_{Z|X}(Z|X) = \sum_{\theta \in \Theta} p_{Z, \theta|X}(Z, \theta|X) \quad (16)$$

$$= \sum_{\theta \in \Theta} p_{Z|X, \theta, m}(Z|X, \theta, m) p_{\theta, m|X}(\theta, m|X) \quad (17)$$

The number of measurements m represents the width of Z , thus introducing it will not affect the distribution. θ is the data association vector that associates measurements with detected objects or clutter and Θ is the set of all data association vectors. For n objects, and m_k measurements, the data association vector θ is n -dimensional and each component of θ can take one of m_k values from the following set.

$$\{0, 1, 2, \dots, m_k\} \quad (18)$$

where 0 represents a misdetected object and the values $1, 2, \dots, m_k$ are indices representing the measurements. For example, if we have 2 objects ($X = [\mathbf{x}_1, \mathbf{x}_2]$) and 2 measurements ($Z = [\mathbf{z}_1, \mathbf{z}_2]$), then we can have the following combination of values for the data association vector.

$$\begin{aligned}
[0, 0] &\rightarrow \text{obj 1: misdetected,} & \text{obj 2: misdetected} & (19) \\
[0, 1] &\rightarrow \text{obj 1: misdetected,} & \text{obj 2: measurement 1} & \\
[0, 2] &\rightarrow \text{obj 1: misdetected,} & \text{obj 2: measurement 1} & \\
[1, 0] &\rightarrow \text{obj 1: measurement 1,} & \text{obj 2: misdetected} & \\
[1, 2] &\rightarrow \text{obj 1: measurement 1,} & \text{obj 2: measurement 2} & \\
[2, 0] &\rightarrow \text{obj 1: measurement 2,} & \text{obj 2: misdetected} & \\
[2, 1] &\rightarrow \text{obj 1: measurement 2,} & \text{obj 2: measurement 1} & \\
\end{aligned} \tag{20}$$

it has to be noted that the associations $[1, 1]$ and $[2, 2]$ are missing. This is because the same measurement cannot be associated to two different objects. This arises from the assumption that the object is a point object. So, at most one measurement would represent one object.

The misdetection in the data association represents a measurement with no detection. Therefore, the measurement is associated with clutter. For the associations shown in eq. (19), the corresponding object and clutter measurements are given in table 1. m^o and m^c are the number of detected objects and the

θ	O^1	O^2	C	m^o	m^c
$[0, 0]$	$[\]$	$[\]$	$[\mathbf{z}_1, \mathbf{z}_2]$	0	2
$[0, 1]$	$[\]$	\mathbf{z}_1	\mathbf{z}_2	1	1
$[0, 2]$	$[\]$	\mathbf{z}_2	\mathbf{z}_1	1	1
$[1, 0]$	\mathbf{z}_1	$[\]$	\mathbf{z}_2	1	1
$[1, 2]$	\mathbf{z}_1	\mathbf{z}_2	$[\]$	2	0
$[2, 0]$	\mathbf{z}_2	$[\]$	\mathbf{z}_1	1	1
$[2, 1]$	\mathbf{z}_2	\mathbf{z}_1	$[\]$	2	0

Table 1: Object and Clutter Measurements

clutter detections respectively.

Appendix A Definitions

Probability Distributions: Probability distributions can be provided in different forms depending on whether the random variables are discrete or continuous.

- **Discrete Random Variables**

- **Probability Mass Function (PMF):** a probability mass function is a function that gives the probability that a discrete random variable is exactly equal to some value.

- **Absolutely Continuous Random Variables**

- **Probability Density Function (PDF):** describes the infinitesimal probability of any given value, and the probability that the outcome lies in a given interval can be computed by integrating the probability density function over that interval.
- **Cumulative Density Function (CDF):** describes the probability that the random variable is no larger than a given value (i.e., $P(X < x)$ for some x). The Cumulative Density Function (CDF) is the area under the probability density function from $-\infty$ to x

Probability Distributions: A mixture distribution is a distribution formed from the weighted combination of two or more component distributions. The component distributions can be univariate or multivariate. For example, a highway is populated by 67% cars and 33% trucks. The weight of cars is distributed as Normal(3200, 300) pounds, whereas the weight of trucks is distributed as Normal(12K, 4K). The variability distribution for all vehicles is a mixture of these, with the Normal(3200, 300) distribution given a weight of 0.67, the Normal(12K, 4K) given a weight of 0.33.

Prior: a prior probability distribution, often simply called the prior, of an uncertain quantity is the probability distribution that would express one's beliefs about this quantity before some evidence is taken into account. For example, the prior could be the probability distribution representing the chances that the next day could be sunny.

Posterior: the probability distribution of the proposition (hypothesis) after taking the evidence (measurement) into account.

Likelihood: the probability distribution of the evidence (measurement) given the hypothesis. The likelihood quantifies the extent to which the evidence or measurement supports the proposition or hypothesis.