Chapter 4

Tracking and Localisation Using Radio Frequency

4.1 Introduction

In the previous chapter Radio Frequency (RF) sensor models as used for communication purposes or for localisation in robotics were presented. This chapter presents the topics of localisation and tracking in the context of RF based range-only localisation, and discusses the different components needed. A major emphasis is placed on the quality of sensor models.

The main issues presented in this chapter are summarised as:

- Probabilistic tracking filters provide a method to estimate the state of a process. Three major filtering types, namely the Kalman Filter (KF), the Particle Filter (PF) and the Histogram Filter (HF) are presented. These filters are used to implement tracking with RF sensors and their suitability is discussed.
- Process models describing the motion of a person are shown. These models are used in the tracking filters to describe the motion of a person when no measurements are available.
- The importance of observation models is discussed and it is shown how the previously named filtering algorithms can approximate RF measurements.

4.2 Probabilistic filtering algorithms

4.2.1 Recursive Bayesian Estimation

In Recursive Bayesian Estimation the Probability Density Function (PDF) of a random vector is tracked over time in a two-step process.[1] At each time step k, a model describing the evolution of the random vector, as well as an observation model describing how observations are related to the state, are present. The state transition function serves as the model of how the state vector $\mathbf{x}_k$ evolves over time. It is a function of the state vector at the previous time step, and some additional noise.

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{q}_k) \tag{4.1}$$

The quantity $\mathbf{q}_k$ represents an independent noise sequence with known PDF $p(\mathbf{q}_k)$ and denotes the uncertainties, i.e. the parts of the system not modelled. The measurement function $\mathbf{h}_k$ is a model for the measurements $\mathbf{z}_k$. These measurements observe the state vector $\mathbf{x}_k$.

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{r}_k) \tag{4.2}$$

Here again, $\mathbf{r}_k$ is an independent noise sequence with known PDF $p(\mathbf{r}_k)$, this time describing the uncertainties in the observation model.

The first step in recursive Bayesian estimation consists of predicting the change in the PDF of the state vector $\mathbf{x}_k$ from time step $k-1$ to k. At time $k-1$ the PDF $p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$ is known, where $\mathbf{Z}_{k-1} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{k-1}\}$. Using the state transition model $\mathbf{f}_k$ the *prior* PDF $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ is obtained.

$$p(\mathbf{x}_k|\mathbf{Z}_{k-1}) = \int_{-\infty}^{\infty} p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})d\mathbf{x}_{k-1} \tag{4.3}$$

The update step is applied once a measurement is available to obtain the *posterior* PDF $p(\mathbf{x}_k|\mathbf{Z}_k)$. Using Bayes theorem the posterior is given by

$$p(\mathbf{x}_k|\mathbf{Z}_k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{k-1})}{\int_{-\infty}^{\infty} p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{k-1})d\mathbf{x}_k} \tag{4.4}$$

This posterior distribution is then used at the next time step to find the new prior.

For the derivation of the state transition density $p(\mathbf{x}_k|\mathbf{x}_{k-1})$, the assumption is made that the state transition function $\mathbf{f}_k$ is a Markov process. This implies that the state at time k depends only on the state at time $k-1$, and is independent of all other previous states.

[1] Appendix A introduces the basics of probability distributions.

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \int_{-\infty}^{\infty} \delta(\mathbf{x}_k - \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{q}_{k-1}))p(\mathbf{q}_{k-1})d\mathbf{q}_{k-1} \tag{4.5}$$

The observation likelihood $p(\mathbf{z}_k|\mathbf{x}_k)$ can be found from the observation model $\mathbf{h}_k$

$$p(\mathbf{z}_k|\mathbf{x}_k) = \int_{-\infty}^{\infty} \delta(\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k, \mathbf{r}_k))p(\mathbf{r}_k)d\mathbf{r}_k \tag{4.6}$$

It is possible to implement a recursive Bayes filter for certain simple estimation problems, however, for the majority of estimation tasks the integration in the prediction step and the multiplication in the update step will be difficult to obtain in a closed form. Therefore it is common to restrict the problem to finite state spaces and consequently turn the integral into a finite sum [91]. The application presented here uses highly multi-modal observation functions, and it is easier to use filtering methods that do not require the calculations in a closed form. For this reason, the use of filtering algorithms as described in the following subsections is considered appropriate. The algorithms either restrict the estimation problems to certain assumptions (e.g. Gaussian distributions), or make approximations.

4.2.2 Kalman Filter

The KF is arguably one of the most important and popular filtering algorithms used for state estimation. It is a recursive filter used to calculate the state of a dynamic system. The filter provides the best state estimate given models describing the process evolution and how measurements relate to the process. Furthermore, it propagates the states and the associated covariances. This subsection provides a description of the algorithm and the interested reader is referred to the literature [10, 40, 56] for the derivation of the KF.

As the KF is a recursive estimator, only the previous state and the current measurement are required to calculate the new (current) estimate of the state. The following assumptions have to be met for the filter to operate consistently:

- A linear model $\mathbf{f}(\mathbf{x}, \mathbf{u})$ of the system describing its evolution must be present. This model is a function of the previous state and the inputs to the system.

$$\mathbf{x}_k^- = \mathbf{F}_k\mathbf{x}_{k-1} + \mathbf{B}_k\mathbf{u}_k + \mathbf{v}_k \tag{4.7}$$

- A linear measurement model $\mathbf{h}(\mathbf{z})$, relating observations $\mathbf{z}$ of the system to the state variables is also required.

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k \tag{4.8}$$

- The noises present in both the system and the measurement model, must be zero mean

Gaussian random variables. Furthermore, they need to have independent probability distributions

$$\mathcal{E}(\mathbf{v}_k) = \mathcal{E}(\mathbf{w}_k) = 0 \tag{4.9}$$

which are uncorrelated

$$\mathcal{E}(\mathbf{v}_i \mathbf{w}_j^T) = 0 \tag{4.10}$$

Their covariances are given by:

$$\mathcal{E}(\mathbf{v}_i \mathbf{v}_j^T) = \delta_{ij} \mathbf{Q}_i \tag{4.11}$$

$$\mathcal{E}(\mathbf{w}_i \mathbf{w}_j^T) = \delta_{ij} \mathbf{R}_i \tag{4.12}$$

Under these assumptions, the KF generates estimates for the state $\mathbf{x}_k$, and the corresponding covariance $\mathbf{P}_k$, based on the predicted state $\mathbf{x}_k^-$ and covariance $\mathbf{P}_k^-$ as

$$\mathbf{x}_k^+ = \mathbf{x}_k^- + \mathbf{W}_k \left[\mathbf{z}_k + \mathbf{H}_k \mathbf{x}_k^- \right] \tag{4.13}$$

$$\mathbf{P}_k^+ = \mathbf{P}_k^- - \mathbf{W}_k \mathbf{S}_k \mathbf{W}_k^T \tag{4.14}$$

with $\mathbf{W}$ called the gain matrix and $\mathbf{S}$ being the innovation covariance matrix. These can be calculated using

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \tag{4.15}$$

$$\mathbf{W}_k = \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{S}_k^{-1} \tag{4.16}$$

The quantity $\mathbf{R}$ is the estimated observation noise covariance, which describes the uncertainties of the sensor when making observations. This quantity must be provided by the designer of the filter.

The Extended Kalman Filter (EKF) is the extension of the KF and is capable of handling non-linear Gaussian processes to a certain extent, by linearising models around the prior state. It also relies on the assumption that the probability distribution is uni-modal and near Gaussian. Sum of Gaussians (SOG) filters [4] are a further method for approximating non-linear functions as a weighted sum of multiple Gaussians, allowing these filters to represent multi-modal distributions. SOG filters are also known as Gaussian Mixture Models (GMMs).

4.2.3 Particle Filter

Particle Filters [5, 37, 39, 74] are a sample based approach for tracking the evolution of a state vector over time. They are well suited for problems where the state transition model and/or the observation model are highly non-linear. Samples are defined as a set of support points $\mathbf{x}^i$ that can be used to describe a PDF. Each support point is associated with a corresponding weight w^i. Using n samples and the corresponding weights, a PDF can be written (approximated) in the following form:

$$p(\mathbf{x}) \approx \sum_{i=1}^{n} w^i \delta(\mathbf{x} - \mathbf{x}^i) \tag{4.17}$$

Importance sampling

Often the true distribution of the state f (or any other distribution one wants to sample from) is not available, and in this case it is easier to draw samples from a *proposal distribution g*. This distribution resembles the true distribution up to some proportional constant. To compensate for this, the weight attached to the sample, is chosen to be

$$w^i = \frac{f(\mathbf{x}^i)}{g(\mathbf{x}^i)} \qquad i = 1, \ldots, n \tag{4.18}$$

This process is termed *Importance Sampling*.

Filtering process

Starting with a known distribution of the state vector, a PF cycle requires the following steps:

Prediction: Each of the n samples from the previous step PDF are passed through the state transition model

$$\mathbf{x}_k^-(i) = \mathbf{f}_k(\mathbf{x}_{k-1}(i), \mathbf{q}_{k-1}(i)) \qquad i = 1, \ldots, n \tag{4.19}$$

to obtain a new set of particles representing the prior PDF $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ of the state.

Incorporation of observations: When a new observation $\mathbf{z}_k$ is made, a likelihood weighting $\Lambda_k(i)$ for each of the prior samples $\mathbf{x}_k^-(i)$ is calculated. This weighting is obtained from the PDF $p(\mathbf{z}_k|\mathbf{x}_k(i))$ of the observation model and is proportional to the model.

$$\Lambda_k(i) = \frac{p(\mathbf{z}_k|\mathbf{x}_k^-(i))}{\sum_{j=1}^{n} p(\mathbf{z}_k|\mathbf{x}_k^-(j))} \qquad i = 1, \ldots, n \tag{4.20}$$

The likelihood weighting is a scalar value because the measurement, and the prior states are specified. Once again the denominator serves to normalize the sum of the likelihoods to one.

Calculation of the posterior and resampling: After assigning a weight to each particle, it is now possible to resample from the prior distribution (or the corresponding sample set) according to the assigned weights. Consequently a sample based estimate of the posterior distribution is obtained. To achieve this the cumulative likelihoods $\Upsilon_k(i)$ are calculated

$$\Upsilon_k(i) = \begin{cases} 0, & i = 0 \\ \sum_{j=1}^{i} \Lambda_k(j), & otherwise \end{cases} \tag{4.21}$$

To find the *posterior sample* $\mathbf{x}_k^+$, a value u_i is drawn from the uniform distribution $\mathcal{U}(0,1]$. The sample from the prior set is chosen according to this scalar.

$$\mathbf{x}_k^+(i) = \mathbf{x}_k^-(j), \qquad where \qquad \Upsilon_k(j-1) < u_i \leq \Upsilon_k(j) \tag{4.22}$$

The posterior distribution $p(\mathbf{x}_k|\mathbf{Z}_k)$ is then represented by the set of the posterior samples $\{\mathbf{x}_k^+(1), \ldots, \mathbf{x}_k^+(n)\}$.

A note on resampling: when performing the resampling step, particles with lower weights are moved to areas where particles with higher weights are located. Over time, this tends to concentrate particles in areas where the PDF has higher probabilities, i.e. modes. This effect is known as sample impoverishment, where the particles no longer represent the true distribution with its tails and areas of lower probability. Instead, only the areas with high probabilities are represented. Often, this is not desired, and adequate measures must be taken to prevent this effect from occurring. Alternatively, other filtering techniques have to be used if the effect of sample impoverishment cannot be overcome.

4.2.4 Histogram Filter

Occupancy Grids [30] offer the possibility to represent an environment using a matrix of cells, i.e. a grid with fixed resolution. A value in the range (0,1) denotes the probability that an individual cell C_i in the grid is occupied. Values close to one indicate a high probability that a particular cell is occupied, whereas values close to zero indicate that the cell is most likely not occupied. Such a grid is usually initialised with the probability for all cells being occupied of 0.5. This value denotes the maximum uncertainty about the state of a cell. Occupancy Grids (OGs) are successfully employed for localisation as well as for mapping purposes. Indeed, they can be used for a variety of purposes where it is appropriate to represent the state of a cell using stochastic estimates. The state of the cells is updated using Bayes formula

$$\mathcal{P}(C_i = Occ|\mathbf{x}_k, \mathbf{Z}_k) \propto \mathcal{P}(\mathbf{z}_k|C_i = Occ)\mathcal{P}(C_i = Occ|\mathbf{Z}_{k-1}) \tag{4.23}$$

The posterior probability of a cell being occupied is proportional to the prior probability of that cell being occupied, multiplied by the sensor-likelihood (or observation) for the particular cell. As mentioned in [9], the rectangular grid representation is not ideally suited for non-rectangular environments. Also, to be able to represent information at a detailed level, the size of the cells must be very small, consequently increasing the storage and computational requirements.

OG are well suited to hold information that has a regular (spatial) distribution. If the required granularity, memory and computational requirements are well defined, OGs can be implemented and used successfully. Care has to be taken not to unnecessarily waste computing resources if a lot of the cells do not change their state of occupancy, but are nonetheless updated with each observation. It should also be noted that grid based representations generally lose information through the discretisation of a true continuous distribution.

In the scope of this thesis OGs are used to represent the position of an agent (or more precisely the state PDF describing the position), which generally changes over time. Incorporating the element of motion into OGs leads to the HF. One method to account for the motion between observations is by convolving the OG with a suitable kernel. This convolution has the effect of blurring the OG (similar to a de-noising filter in computer vision [36]), consequently broadening the range of possible positions of the agent. Figure 4.2 shows the comparison of a position estimate immediately after an observation, and after a period of time without any observations. Other possibilities to incorporate motion exist [21, 32, 91], but are not used in the scope of this thesis. Using a suitable symmetrical Gaussian convolution kernel, the motion of a person can be adequately modelled. Such a kernel is appropriate as a person can very quickly change direction, and the Gaussian kernel smoothes the estimate evenly in all directions.

4.3 Process models for a person's motion

Motion models are used in filters to predict the evolution of the state vector in the absence of observations. These models are important because the evolution of the state must be propagated consistently when no observations are available. The following two sections describe the Constant Velocity Model and Gaussian Kernel Convolution, used in this thesis to propagate the motion of a person.

4.3.1 Constant Velocity Model

The *Constant Velocity Model* [10] assumes that in the absence of measurements the position tracking process evolves at constant speed. If the state vector $\mathbf{x}$ contains the x and y positions and the corresponding velocities $\dot{x}$ and $\dot{y}$,

$$\mathbf{x} = [x, \dot{x}, y, \dot{y}]^T \tag{4.24}$$

the model will propagate the change in position under the assumption that the velocities are constant during the time step T. The corresponding State-Transition Matrix for this case is given as

$$F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.25}$$

This model may be unsuitable for certain tracking algorithms if the dynamics of the underlying process are faster compared to the time interval at which observations are guaranteed to be available. In the case of localising a person, it is necessary to consider how fast a person can change the direction of motion. This can be a dynamic process, with time intervals of less than one second. If the person is observed only at 5 second intervals, and the constant velocity model is used, a completely incorrect position may be predicted. The person might have changed direction straight after the last observation, and walked in a completely different direction. In this case the algorithm would have predicted that the person continued walking in the initial direction.

4.3.2 Gaussian Kernel Convolution

Using a suitable symmetrical Gaussian kernel the motion of a person can be adequately modelled by convolving the state PDF with the kernel. Such a kernel is appropriate as a person can very quickly change direction, and the chosen Gaussian kernel smoothes the estimate represented by an OG evenly in all directions.

$$k(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2}\left(\frac{x^2 + y^2}{\sigma^2}\right)\right] \tag{4.26}$$

An example of such a kernel is shown in figure 4.1. It makes the reasonable assumption that the motion of a person in the x and the y directions can be modelled as uncorrelated. This is true for small velocities and not highly unreasonable for higher velocities. Using the

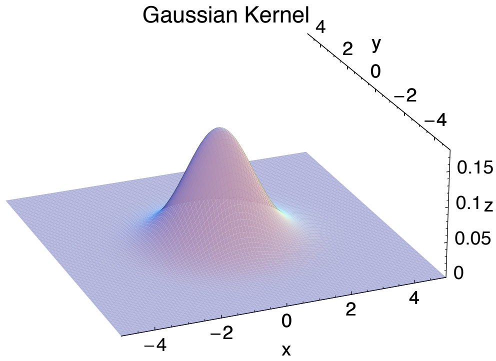


Figure 4.1: Gaussian kernel — This symmetric kernel is suitable to model the unpredictable movements of a person. When used in the prediction step, all directions are equally likely and do not depend on other variables as for example orientation etc.

variance σ, this motion model can be adapted. A large variance can model a fast movement, while a smaller variance would be suitable for modelling a slower movement.

Figure 4.2 shows an example of the application of a Gaussian kernel. A state PDF, as presented in figure 4.2(a), might be obtained during the simulation of a HF that tracks the position of an agent. The figure shows the PDF immediately after the incorporation of a measurement. As no subsequent measurements are available, the state PDF is convolved with a kernel as shown in figure 4.1, resulting in the state PDF at a later time step. The PDF is shown in figure 4.2(b). It can be clearly seen that the application of the convolution using the kernel has a smoothing effect on the state PDF, as is the case in the prediction step in a recursive Bayesian framework. It can also be seen that the kernel does not favour any direction, and the smoothing occurs with the same magnitude in all directions.

4.4 Observation models

4.4.1 On the importance of a correct observation model

The precise characterisation of the sensors used is a key component for the successful implementation of a tracking algorithm. Even an ideal algorithm cannot compensate for an incorrectly characterised sensor, and localisation accuracy strongly depends on the availability of a correct and accurate sensor model for the sensors used. In this context a correct sensor model implies that it appropriately reflects the physical behaviour of the sensor, for example, if the sensor exhibits a multi-modal behaviour, the model should incorporate this behaviour and not a uni-modal one. An accurate model is a sensor model that incorporates

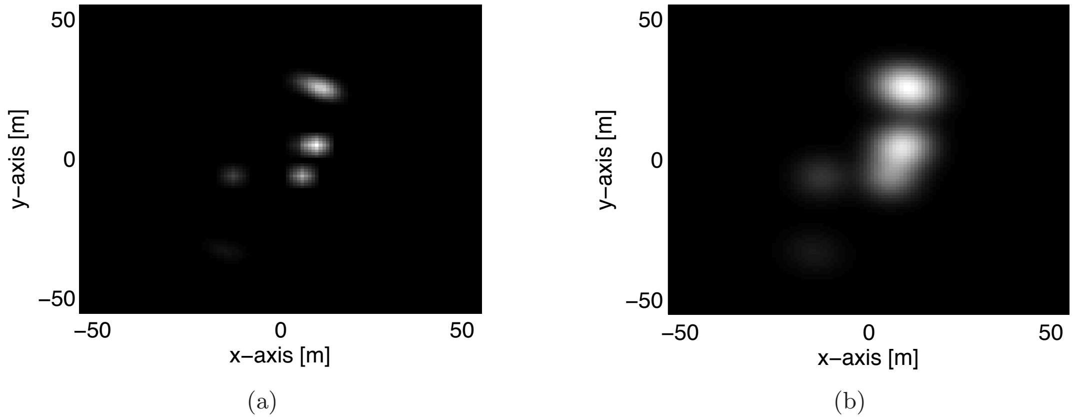


Figure 4.2: Picture (a) shows an OG representation of the possible positions of an agent. Between observations, the estimate tends to become more uncertain, and is modelled by convolving the state PDF with a Gaussian kernel. This is shown in (b), immediately before the incorporation of the next measurement. Here, two previously distinctly separated hypotheses have merged to one large area representing where the agent could be located.

the measurement uncertainties at a magnitude that truly corresponds to the real occurring uncertainties and does not under- or overestimate these. Figure 4.3 shows a multi-modal sensor likelihood function typically encountered when dealing with RF sensors. One measured value reported by such a sensor can correspond to multiple function values, which could be interpreted as the distance between a transmitter and a receiver. If, for this case, the sensor likelihood function is chosen to be more simple (or even incorrect), it will not accurately reflect the true physical behaviour of the sensor (the dashed line in figure 4.3), and will lead to a loss of hypotheses. The simplified model will reduce the three previous hypotheses to one. The problem is accentuated because the remaining hypothesis can be often incorrect. Nevertheless, simpler models are still useful to some extent. The simpler model can be used if a lower accuracy is acceptable. It is noted here, and shown later in this chapter, that the KF is only capable of representing one hypothesis.

It is usually very tedious to obtain a good sensor representation. This often involves numerous, repetitive measurements under several operational modes. For this reason it is necessary to have an application requiring a very accurate model. The work involved in obtaining it should be justifiable with a real need for accuracy in the application. It is not feasible to develop a sensor model enabling localisation with millimetre accuracy, when the real application demands only metre accuracy.

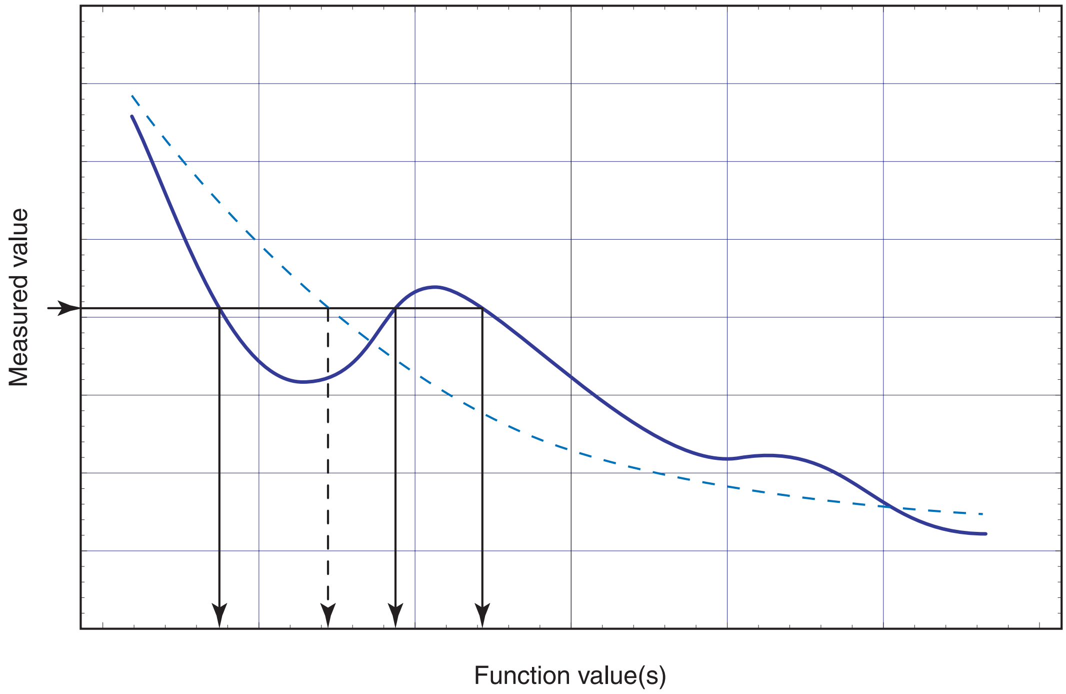


Figure 4.3: Need for correct sensor model — A sensor exhibits a multi-modal behaviour (solid line). The same sensor could also be modelled with a (simpler) uni-modal approximation (dashed line).

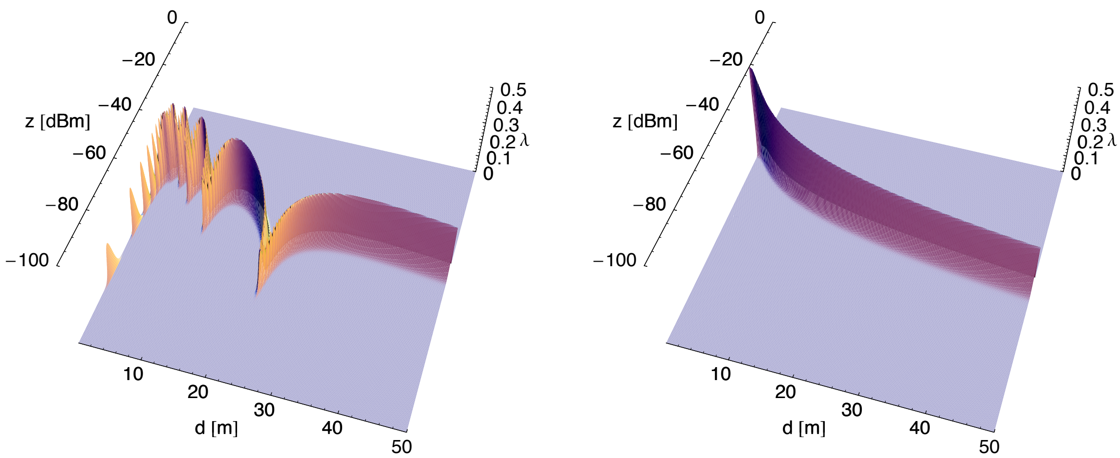
4.4.2 Multi-modal observations

The KF, the PF and the HF each approximate multi-modal distributions from RF sensors in their own particular way, though the quality varies.

KF approximation

The KF can only incorporate measurements that have a Gaussian distribution. RF sensors may provide highly non-Gaussian observation likelihood functions as seen in chapter 3. The reader is reminded of the two-ray model with the oscillating behaviour of the signal path loss. It is necessary to convert the non-Gaussian observations into Gaussian observations if they are to be used in a KF approach. Two possible approaches for this using RF measurements are shown below:

Ignoring the reflective components: A large number of tracking algorithms are implemented ignoring the multi-modal properties of the RF propagation. These algorithms tend to use a simplified model, the nth power model. The model is an approximated function, exponentially decaying with the inverse of the distance. The effects of reflections etc. are ignored, and to compensate for the uncertainties, large variances are



(a) (b)

Figure 4.4: Sensor models used for KF measurement approximation — (a) shows the two-ray model based sensor likelihood function. (b) shows the nth power model. It should be noted that the nth power model shown here is not a good approximation of the two-ray model as a very small variance for the random variable has been used. This leads to the situation that the *spikes* of the two-ray model are not correctly represented by the narrow nth power model.

often used for the random variable describing the noise in the model. Figure 4.4(a) shows the sensor likelihood functions for the two-ray model and figure 4.4(b) shows the corresponding nth power model. These models are now used to demonstrate how multi-modal measurements can be approximated so that they can be incorporated in a KF.

Assuming that the two-ray model represents reality, the likelihood function corresponding to an observation will be as shown in figure 4.4(a). The KF observation will be generated from figure 4.4(b) as done often in mobile robotics applications. Figure 4.5 depicts this situation. For a sample measurement of -35 dBm, the true likelihood function is shown. The likelihood as obtained from the nth power model with a corresponding Gaussian function fitted is also shown. The variance σ^2 of the fitted Gaussian function has been chosen to closely represent the underlying nth power model. It can be seen that this Gaussian function is not a good representation of the multi-modal likelihood function. The probability mass is highly concentrated around the mean due to the chosen variance, and the area covered by the Gaussian function is smaller than the area covered by the true likelihood. It is also of interest to note that the mean of the Gaussian does not correspond to the Maximum Likelihood (ML) location of the true likelihood.

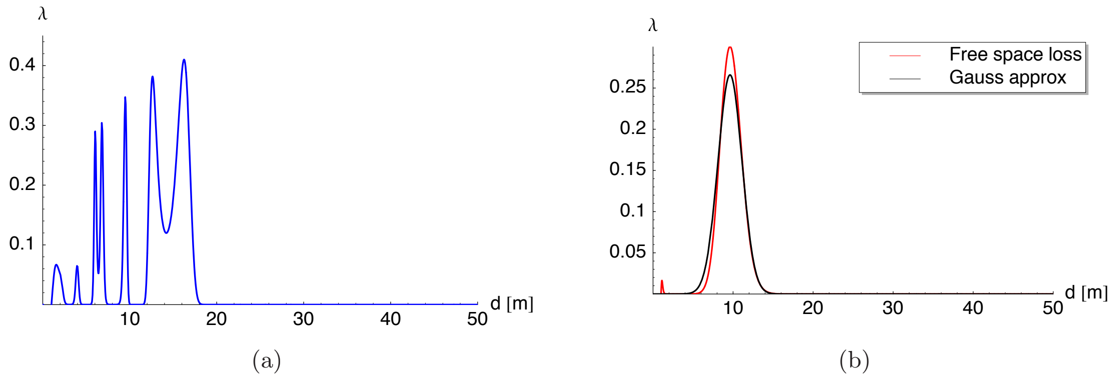


Figure 4.5: Approximating a multi-modal distribution with a Gaussian — For a measurement of -35 dBm (a) shows the true multi-modal distribution. (b) shows the distribution for the same measurement as it is obtained if the nth power model is used. A Gaussian distribution fitted to this measurement and which can be used in a KF is also shown. It is interesting to note, that although the conditionals $p(z|x)$ describing the probability of a measurement given the distance are of Gaussian distribution in both models, the sensor likelihoods $\lambda(x \mid z)$ describing the likelihood of a distance given a measurement even in the case of the nth power model only are not of Gaussian distribution and therefore have to be approximated for the KF.

Approximating the multi-modal distribution with a single Gaussian function: A second possible method for approximating the multi-modal distribution using a Gaussian function is to calculate the first two moments, the mean and the standard deviation, of the entire multi-modal sensor distribution

$$\mu = \sum_x x p(x) \tag{4.27}$$

$$\sigma = \sqrt{\mathcal{E}((x-\mu)^2)} \tag{4.28}$$

In practice this is often best done numerically, as the sensor model may be very complicated, if there exists a closed form expression. Figure 4.6 shows the likelihood function for a measurement of -45 dBm obtained from the theoretical two-ray model. Several dominant peaks can be seen, with the true location being most likely close to the ML. Also shown is a Gaussian approximation to this likelihood function with the mean and standard deviation calculated using equations (4.27), and (4.28). As the likelihood has several peaks spread along the distance domain, the variance, or alternatively the standard deviation, of this Gaussian is high. Nevertheless, the probability mass of the likelihood at the right edge is almost not contained by the Gaussian approximation. It is clearly seen, that the Gaussian approximation with a single mode at the mean, is again not an appropriate representation of the multiple modes contained in the

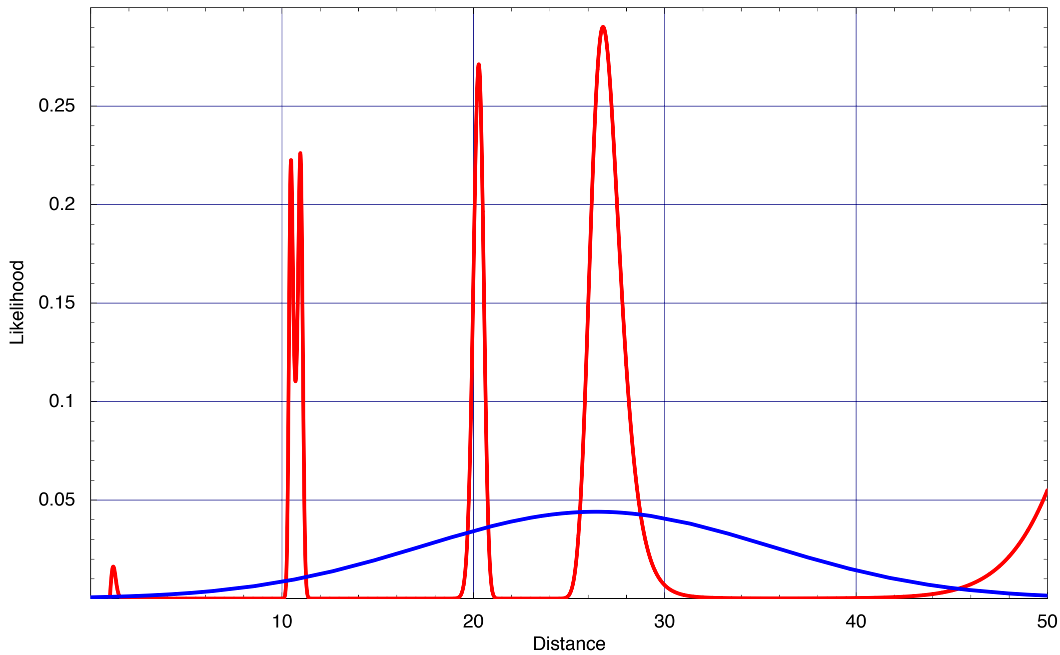


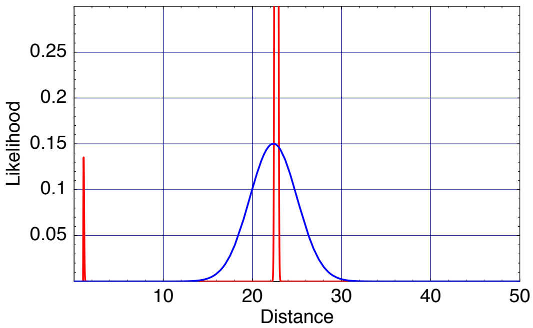
Figure 4.6: Gaussian approximation of a multi-modal distribution — A highly non Gaussian distribution is poorly approximated by a Gaussian distribution.

original likelihood function. In this example, the mode of the Gaussian, and the mode of the likelihood are almost at the same position. This is not necessarily the case, as can be seen in figure 4.7 (c) and (d), for a different measurement of -35 dBm.

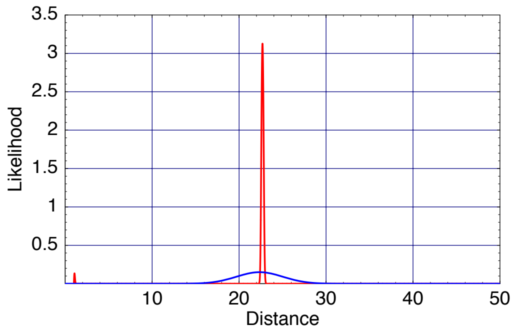
The method presented to approximate a multi-modal distribution using a uni-modal distribution is related to the question of how a sensor likelihood function is interpreted. Clearly, if a sensor likelihood function is multi-modal and is interpreted in the way that such a sensor measurement is not able to distinguish between individual hypotheses (modes), a uni-modal approximation simplifies this to the extent that it removes all but one hypothesis. This introduces the risk that the remaining hypothesis is not the correct hypothesis, as already mentioned in subsection 4.4.1.

PF approximation

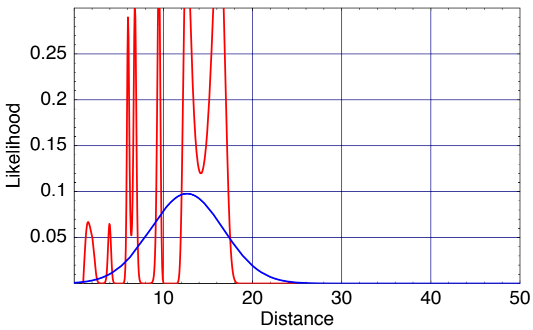
The PF handles observations by weighting the particles. Using the likelihood function obtained from a measurement, the filter modifies the weighting of the particles. This occurs by multiplying the previous weight by the likelihood obtained from the measurement. Using this method it is theoretically possible to capture the whole distribution, instead of the first two moments. In practice, the quality of the approximation is limited by the number of particles used. Figure 4.8 shows an example for a measurement obtained using the two-ray



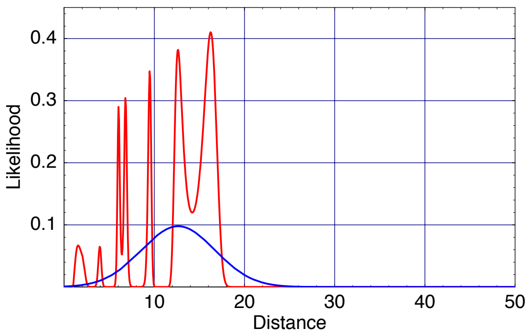
(a)



(b)



(c)



(d)

Figure 4.7: Gaussian approximation of two multi-modal distributions — Measurements of -65 dBm, (a) zoomed and (b) complete distribution and -35 dBm, (c) zoomed and (d) the complete distribution.

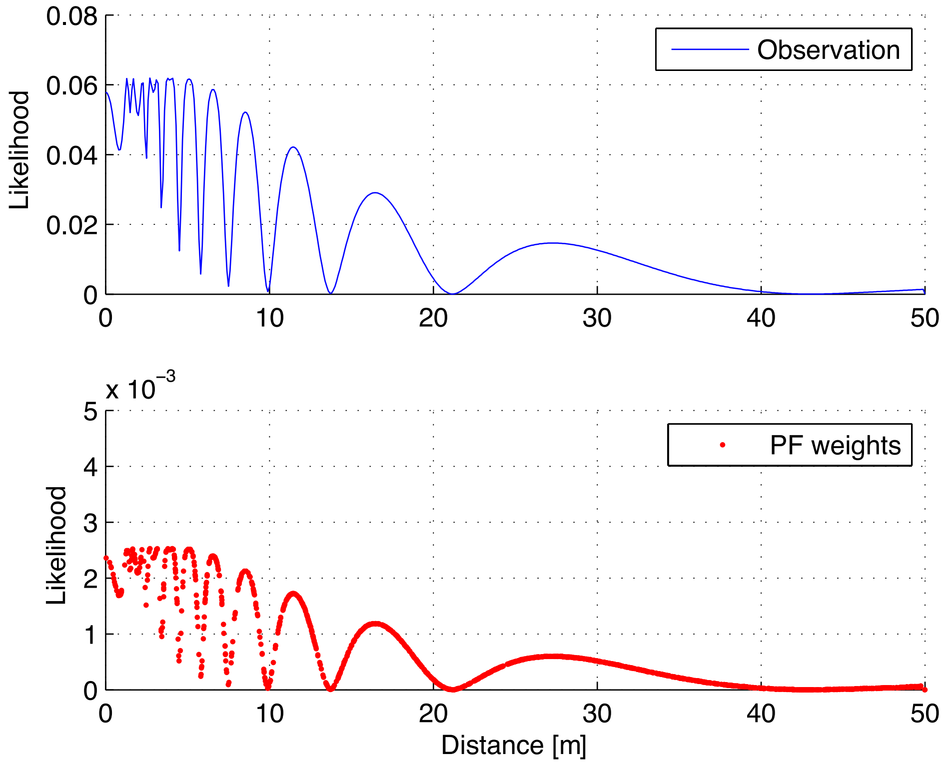


Figure 4.8: Observation with PF — A multi-modal sensor measurement (top). Assuming that the particles had uniform weights and were uniformly distributed before the measurement occurred, their weights after will correspond exactly to the measurement likelihood function values (bottom).

model. It can be seen that it is possible to approximate the original distribution closely. The number of particles used in the filtering process is an important variable that has a major influence on the quality of the approximation. Low numbers of particles are not useful for approximating highly complex distributions, but do allow for fast computation.

HF approximation

Multi-modal distributions can also be approximated for use in the HF. This is achieved through a decomposition of the state into finite non-overlapping regions. Often the regions are of the same size, and for each region a value is assigned to represent the probability of the state in this region. HFs can be considered as piecewise constant approximations of a continuous density [91]. The quality of the approximation depends on the chosen size of the regions. As in the case of the PF approximation, the need for an accurate representation has to be weighed against a probably competing need for fast computation. Figure 4.9

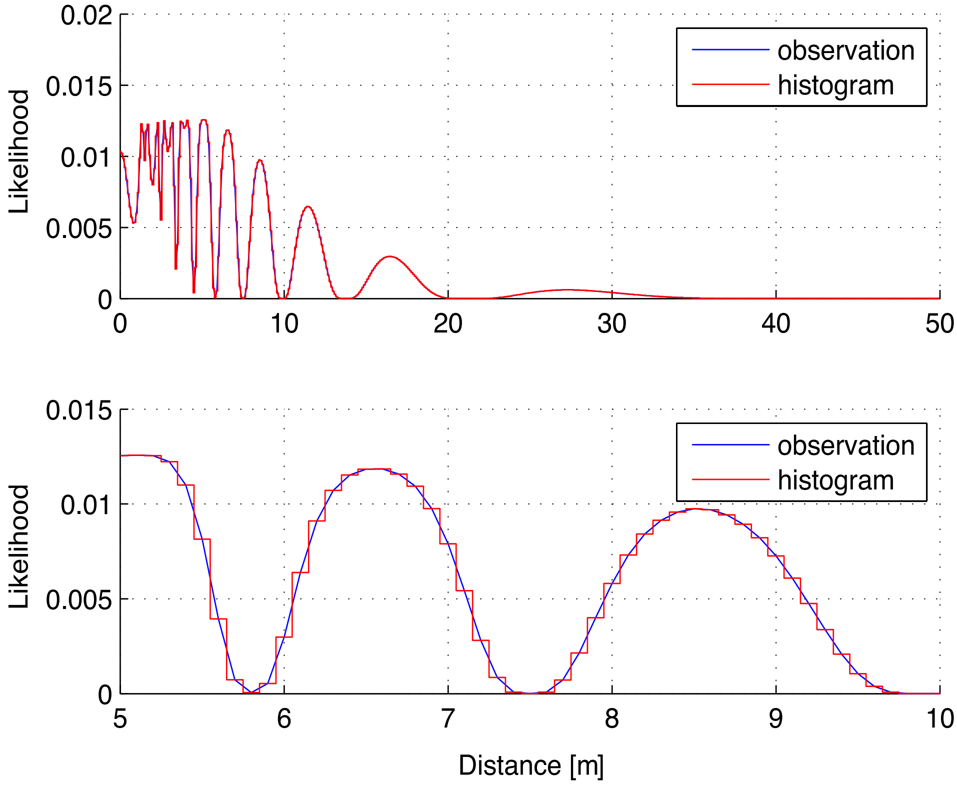


Figure 4.9: Observation with HF — A multi-modal sensor measurement with the histogram representation overlayed (top). In the zoomed image (bottom), the constant piecewise approximation can be seen.

shows an example of an HF approximation for a RF measurement, obtained from a two-ray model.

It has been demonstrated that the KF is not capable of handling multi-modal measurement likelihood functions without significantly simplifying the distribution. In contrast the PF and the HF are both able to approximate multi-modal distributions. Depending on the complexity of the distribution computational needs however, may be significant.

4.4.3 Bias of observation models

Sensors are sometimes modelled in a way that the variance σ_z of the measured signal z is a function of the mean measured signal. This is done to describe the physical behaviour of the sensor under consideration, and is necessary to be included. It can be concluded that the sensor exhibits this behaviour rather than being a property of the sensor model.

It is of interest to analyse the behaviour of such a sensor model when used, and to verify that, if a certain measurement value is observed, the corresponding true value of the state can be recovered. If this is not the case an explanation needs to be found and the consequences for the application have to be described.

For a uni-modal likelihood-function, the effect of the variance on the measurement likelihood function will be analysed for two cases:

1. For a single measurement $\mathbf{z}_{true}$, corresponding to the true value $\mathbf{x}_{true}$.
2. For multiple observations (and in the limit for infinite observations) of the true value of the state $\mathbf{x}_{true}$ with a real sensor with noise contamination.

The ML approach is adopted here for the interpretation of a multi-modal likelihood function. Other measures, such as the mean, are not suitable here as seen previously when examining multi-modal distributions for the KF.

Single observation of the true value: Ideally, the true value of the state $\mathbf{x}_{true}$ should be obtained for the case where the measurement $\mathbf{z}$ corresponds to the true value $\mathbf{z}_{true}$. In the context of ML this can be written as

$$\begin{aligned} \arg\max \lambda(\mathbf{z}_{true}) &= \mathbf{x}_{true} \\ \hat{\mathbf{x}}_{\text{ML}} &= \mathbf{x}_{true} \end{aligned} \tag{4.29}$$

This equation is interpreted to mean that the true value is the value that maximises the likelihood function obtained for the measurement corresponding to the true value.

Multiple noise corrupted observations of the true value: Similar to the single observation case, the ideal result would also be to recover the true value, when observing the true value multiple times corrupted with noise. At the very least in the limit, when the number of observations n approaches infinity, the result should be the true value. In the context of ML this can be formulated as

$$\arg\max \left(\prod_{n\to\infty} \lambda(\mathbf{z}_n) \right) = \mathbf{x}_{true} \tag{4.30}$$

The conditions for uni-modal sensor likelihood-functions can be extended to the multi-modal case, by realising that a muliti-modal function can have multiple modes with the same function-value. In this case the ML or Maximum A Posteriori (MAP) estimator would return a set $\mathcal{X}$ of possible true positions $\hat{\mathbf{x}}_{true,1\ldots m}$. The number m, of elements in this set, corresponds to the number of modes with the same maximum function value of the resulting likelihood function. The true position $\mathbf{x}_{true}$ must be a member of this set.

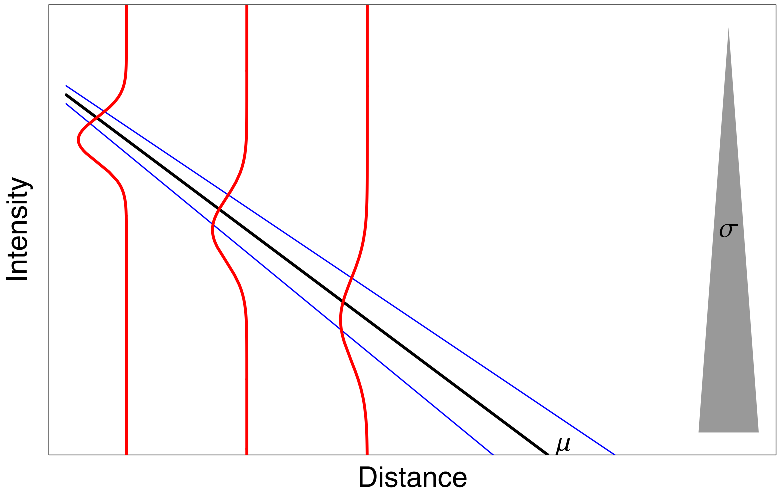


Figure 4.10: Signal variance as a function of the signal mean — The signal mean decreases with distance, while the signal variance increases (indicated by the blue lines) with decreasing signal mean (black line).

$$
\begin{aligned}
\mathcal{X} &= \{\hat{\mathbf{x}}_{true,1} \dots \hat{\mathbf{x}}_{true,m}\} \\
\mathbf{x}_{true} &\in \mathcal{X}
\end{aligned}
\tag{4.31}
$$

The true value is preserved, if the above condition is satisfied, though hidden amongst other hypotheses.

Figure 4.10 illustrates the signal mean of a distance sensor. The distribution is Gaussian for a given distance. The signal mean intensity decreases with increasing distance, while the signal variance increases with decreasing signal mean intensity. This is a behaviour as it occurs for example in the RF model as developed later in chapter 5.

For the single true measurement case, and the case where the number of measurements approaches infinity, the corresponding likelihood function is a non-Gaussian function. The function has its ML not on the true position corresponding to the true measurement. This is depicted in figure 4.11. The ML estimate of the resulting likelihood function for both cases tends to be shifted towards areas with smaller signal variance.

The magnitude of this shift (bias) is influenced by three factors:

- The dependency of the variance on the signal mean: A function where the variance changes a lot with respect to the signal mean may lead to a larger shift as opposed to a function where the variance changes are small.

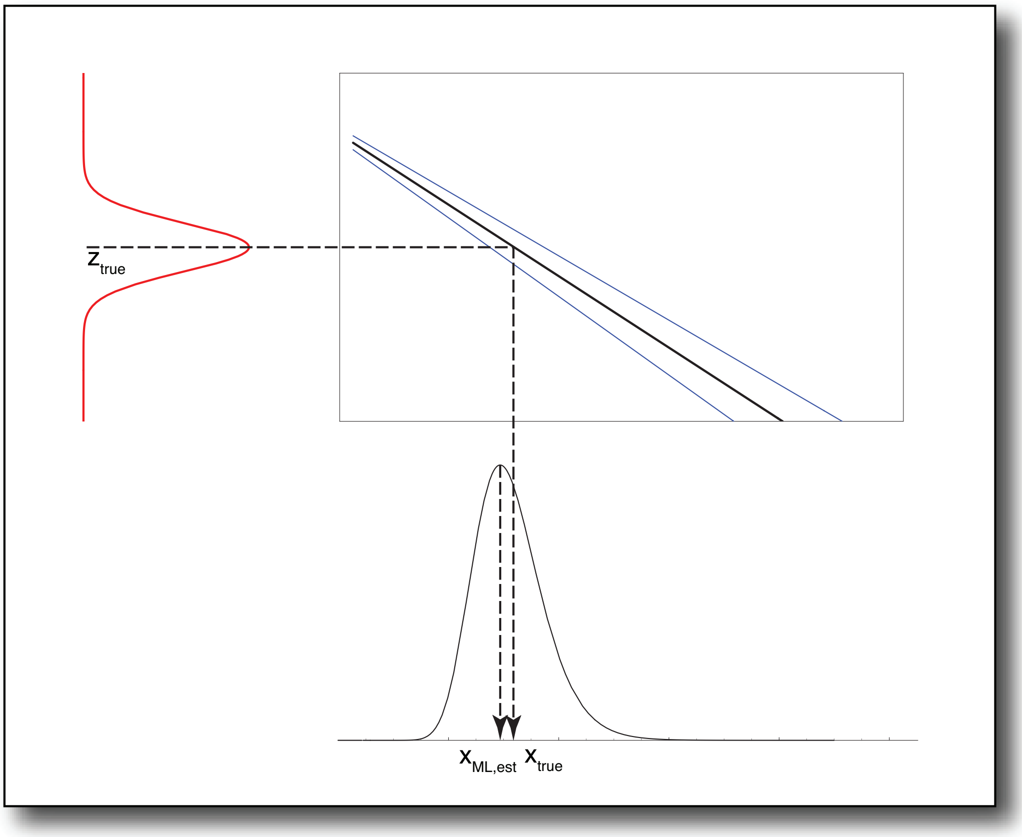


Figure 4.11: Sensor bias — When observing the true values z_{true} it would be expected to infer the true value x_{true}. In the case of a non-constant variance that depends on the signal mean the ML estimate $x_{ML,est}$ does not coincide with the true value x_{true}, but is shifted. The likelihood function is obtained either from the single measurement of z_{true}, or as the product of an infinite number of noise corrupted measurements with mean z_{true} and variance $\sigma_{sensornoise}$.

- The form of the underlying signal mean function and the magnitude of the gradient (derivative) in the region around the true measurement: For small gradients (derivatives) at the true measurement the function describing the signal mean varies a lot given only small changes in the measurement (as can occur due to measurement noise). In such a case the shift would also be larger. Conversely, large gradients will lead to smaller changes.
- The magnitude of the variance of the sensor noise at the true measurement: For small variances of the sensor noise, the shift will be also small. This is based on the fact that for very small changes in the measurement, only small changes in the corresponding distance value can occur. Larger variations in the measurement can lead to larger changes in the distance.

For the case treating an infinite number of measurements an interpretation of this phenomenon in a Bayesian way is possible. If the smaller variance and consequently a higher mode is interpreted as the sensor being more sure of the location, it is not surprising that the product of an infinite number of measurements tends to be shifted towards the areas where the variance is small.

Sensor bias for multi-modal sensor models

It is not difficult to extend the considerations made for a uni-modal model to the case of a multi-modal sensor model and describe the resulting effects. As figure 4.12 shows, for a multi-modal function there will be regions where, depending on the underlying signal mean, the variance increases or decreases. These regions are separated at points where the underlying mean function has either a local minimum or a maximum, and at the boundary points. Within each region there is a bias in the direction of smaller variance.

Remark on sensor bias

It is important to note that the behaviour and the consequences presented in this section are not specific to RF sensors or other sensors used for distance measurements. Neither is it a result of the non-linearity of the underlying function describing the mean. This behaviour depends only on the fact that the signal variance is a function of the measured value itself. Such a behaviour might also occur for other sensors, or in other applications with different observation models.

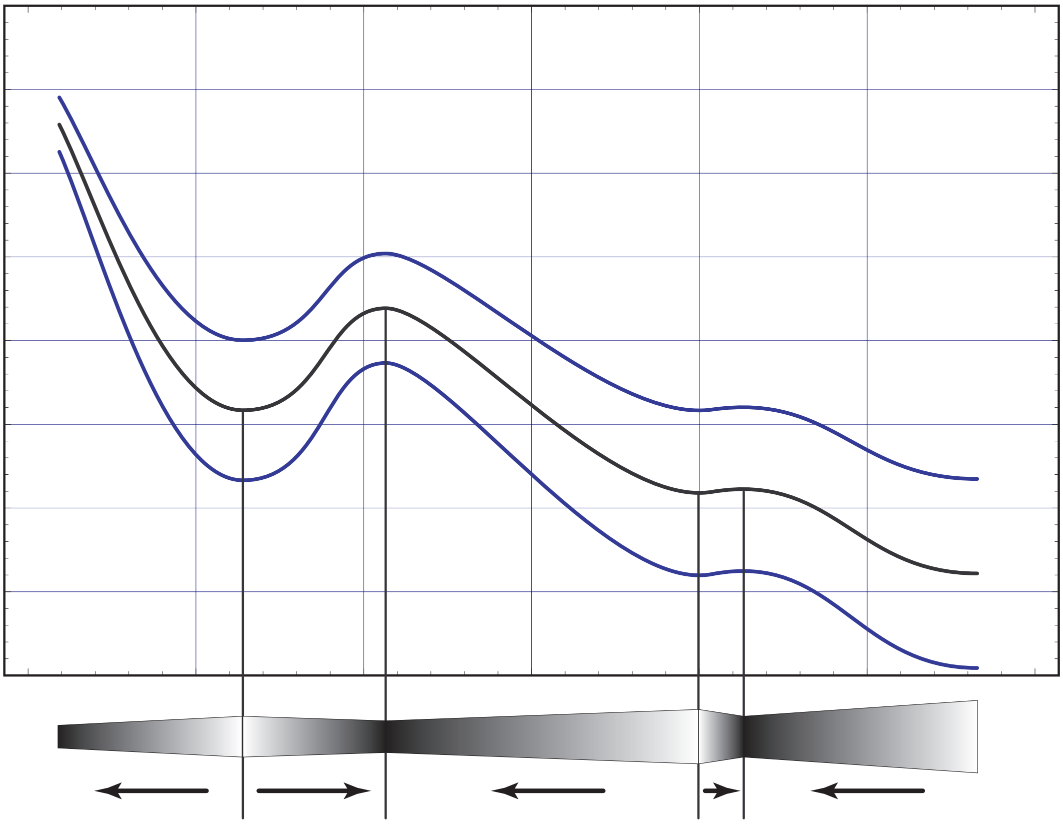


Figure 4.12: Sensor bias for multi-modal functions — When observing the true value z_{true} it would be expected to infer the true value x_{true}. In the case of a non-constant variance depending on the signal mean the ML estimate $x_{ML,est}$ does not coincide with the true value x_{true}, but is shifted. This figure again shows the signal mean (black) and the variance (blue). The arrows indicate the direction of the bias, but not the magnitude.

4.5 Summary

This chapter brought together various aspects of localisation and tracking using RF technology, and multi-modal range-only sensors. Recursive Bayesian estimation and three popular probabilistic tracking algorithms, the KF, the PF and the HF, were introduced. These filters were investigated with respect to their suitability of using multi-modal observations. It was found that the KF is not suitable for such applications as the multi-modal distribution needs to be approximated by a Gaussian distribution. This approximation does not reflect the true occurring sensor likelihood function.

A sensor bias inherent to the way some sensors are modelled was described. This bias leads to incorrect observations. The important conclusion was drawn that this bias is inherent to some sensor models and that the bias will affect the estimation process. For the localisation task addressed in this thesis, the bias will most likely lead to incorrect position estimates. This was shown when evaluating the effect of ranging errors.

As part of the task of tracking and localisation, two process models describing the motion of a person were shown. The constant velocity model and a Gaussian kernel were presented. Both models reflect the difficulty of modelling the motion of a person and tend to propagate the process in rather a simple way. This does not allow for a high-accuracy state propagation. This property of the motion models has to be counteracted by the observation model if high accuracy localisation is to be achieved.