FREQUENCY MODULATED CONTINUOUS WAVEFORM RADAR FOR COLLISION PREVENTION IN LARGE VEHICLES

A thesis submitted in fulfilment of the requirements for the
degree of Doctor of Philosophy in the ACFR at
The University of Sydney

Bryan Clarke
October 2015

I have examined this thesis and attest that it is in a form suitable for examination for the degree of Doctor of Philosophy.

_____

()   Principal Adviser

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the University or other institute of higher learning, except where due acknowledgement has been made in the text.

Bryan Clarke

31 March 2015

Bryan Clarke   Doctor of Philosophy

The University of Sydney   March 2015

Frequency modulated continuous waveform radar for collision prevention in large vehicles

Large vehicles are used in mining, agriculture, stevedoring and other applications. These are often dynamic, unstructured and crowded environments containing numerous large and light vehicles, pedestrians and static infrastructure operating in close proximity. The drivers of large vehicles can have very limited visibility, which contributes to poor situation awareness and an increased risk of collision with other agents.

This thesis is focused on the development of reliable sensing for this close proximity problem in large vehicles operating in harsh environmental conditions. It emphasises the use of in-depth knowledge of a sensor's physics and performance characteristics to develop effective mathematical models for use in different mapping algorithms.

An analysis of the close proximity problem and the demands it poses on sensing technologies is presented. This guides the design and modelling process for a frequency modulated continuous waveform (FMCW) radar sensor for use in solving the close proximity problem. Radar offers better all-weather performance than other sensing modalities, but its measurement structure is more complex and often degraded by noise and clutter. The commonly used constant false alarm rate (CFAR) threshold approach performs poorly in applications with frequent extended targets and a short measurement vector, as is the case here. Therefore, a static detection threshold is calculated using measurements of clutter made using the radar, allowing clutter measurements to be filtered out in known environments.

The detection threshold is used to develop a heuristic sensor model for occupancy grid mapping. This results in a more reliable representation of the environment than is achieved using the detection threshold alone. A Gaussian mixture extended Kalman probability hypothesis density filter (GM-EK-PHD) is implemented to allow better mapping in dynamic environments using the FMCW radar. These methods are effective in producing maps of the environment that can be displayed to the driver of a large vehicle to enhance their situation awareness and better avoid collisions.

The concepts developed in this thesis are validated using simulated and real data from a low-cost 24GHz FMCW radar developed at the Australian Centre for Field Robotics at the University of Sydney. These results show improvements in the map quality obtained by being involved in the sensor design process and having access to internal sensor parameters.

The approaches used are not specific to this radar and could be used with other FMCW radars and potentially other sensors.

Acknowledgements

I could not have done this on my own and there are many, many people to thank.

Firstly, my primary supervisor, Eduardo Nebot, and my associate supervisors, Graham Brooker and Stewart Worrall. Without your guidance, genius and occasional pep-talks, this never would have been done. You guys are awesome.

My colleagues in the Intelligent Vehicles group, James Ward, Asher Bender, Gabriel Agamennoni, David Orchansky and Andrew Maclean, who provided so much assistance and inspiration over the course of my thesis.

A HUGE thank you to Javier Martinez, Dai Bang Nguyen, Ben Stewart and Pak Hung (Victor) Chan, without who I could not have gotten the radar running and software logging. Thanks for fixing every problem with the cabling, or the antenna, or the gearbox, or the motor, or the FFT, or a million other things.

Everyone at the ACFR who provided help conducting tests, gave me guidance on how to focus my research, provided a surface to bounce ideas off, loaned me books, came on powerlunches, ate Tim-Tams, visited the pub, played board games and engaged in countless other research activities/diversions with me, thank you for helping me get through this without me going crazy . Special thanks to, in no particular order, Daniel Wilson whose desk was beside mine for four years straight, Dan Bongiorno for getting me rockclimbing and swimming, Andrew Palmer for keeping me going rockclimbing and hiking, Zac Taylor for getting me running more, Adrian Ball for getting me into D&D, Chris Brunner for being so keen on triathlons, Donald Dansereau for being a beer connoisseur, and Lachlan Toohey for explaining to me the best pass to take through Wentworth Falls.

I am indebted to many friends outside the ACFR too. You're all legends.

Finally, thank you to my wonderful Mum and Dad, for all the love and support and food and car repairs. You are the best parents ever.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivations

Large vehicles are used in mining, agriculture, construction, stevedoring and other applications. These work environments can be unstructured, dynamic and crowded. Large numbers of large and light vehicles must operate near to one another or interact with one another. The environment may also contain pedestrians, buildings and other fixed infrastructure that the vehicles must avoid.

The consequences of a collision between a large vehicle and other agents can be disastrous, but unfortunately these accidents do occur. Accidents can be classified as occurring at high or low speed. High speed accidents usually happen on the road, when the vehicles are running at speed. Low speed accidents make up the vast majority of collisions involving large vehicles [40], and occur when the large vehicle is starting to move or moving at low speed in close proximity with other objects. These are usually called **close proximity accidents** [27], and range in severity from causing minor equipment damage to injuries and fatalities, as well as having large costs associated with equipment damage and work stoppage. Several close proximity accidents are shown in Fig. 1.1.

A major factor contributing to close proximity accidents is poor driver visibility [22]. The drivers of large vehicles operating in such challenging environments often suffer from very poor visibility. Large vehicles have blind spots due to the design and positioning of the cabin and the bulk of the vehicle [27] [28]. Adverse weather conditions such as rain, fog, snow and dust can make it difficult for the operators to see in the

1

(a) A close proximity accident [27].

(b) A close proximity accident [69].

(c) A close proximity accident [28].

Figure 1.1: Close proximity accidents between large and light vehicles can have disastrous and fatal consequences.

limited field of view available to them [76]. Poor visibility leads to poor **situation awareness**: an inability to accurately understand the state of the environment; predict the future state; and take appropriate action to maintain safe operation [12]. Close proximity accidents are one potential consequence of poor situation awareness. This thesis is focused on describing and addressing this close proximity problem.

1.2 The Close Proximity Problem

1.2.1 Blind spots

There is a large volume of space that is hidden from the sight of the driver of a large vehicle due to the vehicle's size and construction. A diagram of a typical haul truck, a Komatsu 730E is shown in Figure 1.2, viewed from above. The truck is 6.68m wide at the front and 12.8m long [29]. The driver's cabin is located on the left side of the vehicle with the top of the cabin 6.25m above the ground. The bulk of the vehicle, the height of the cabin above the ground, and the pillars of the cabin obscure a large proportion of the space around the truck, especially at close range. This causes a number of blind spots:

- Immediately underneath the truck (which is of concern for maintenance workers) [27];

- Immediately behind the cabin;

- Immediately in front of and beside the truck, where the height of the cabin prevents the driver from looking straight down;

- Beside the vehicle in the regions not monitored by the wing mirrors;

- Behind the pillars of the driver's cabin.

Due to the height of the cabin above the ground, the driver overlooks the area at close range. This is depicted in more detail in Fig. 1.3 and Fig. 1.4 for a slightly smaller truck. In Fig. 1.3, the blue line shows the minimum height for an object to be visible to the driver, as a function of distance from the vehicle. Within the bounds of the solid pink bar beneath the diagram, a ~2m tall object is not visible to the driver. Within the bounds of the solid blue bar, the ground itself and anything on it is not visible.
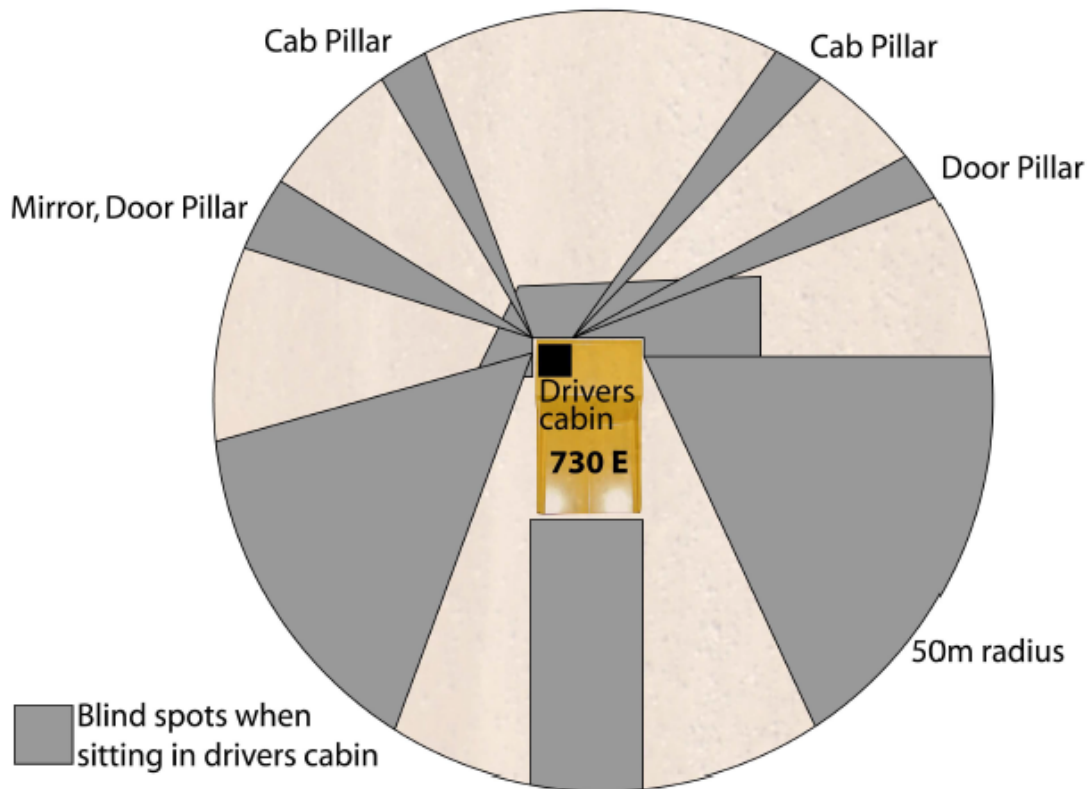
Figure 1.2: Blind spots of a typical mining haul truck (in this case, a Komatsu 730E) haul truck, viewed from above. [27]
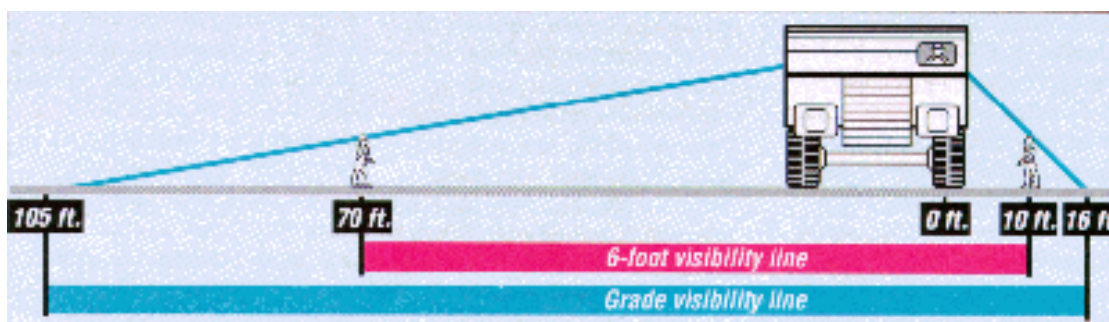
Figure 1.3: Blindspots beside a large vehicle, viewed from in front. Objects beside the vehicle need to be some distance away before they are visible to the driver [39].
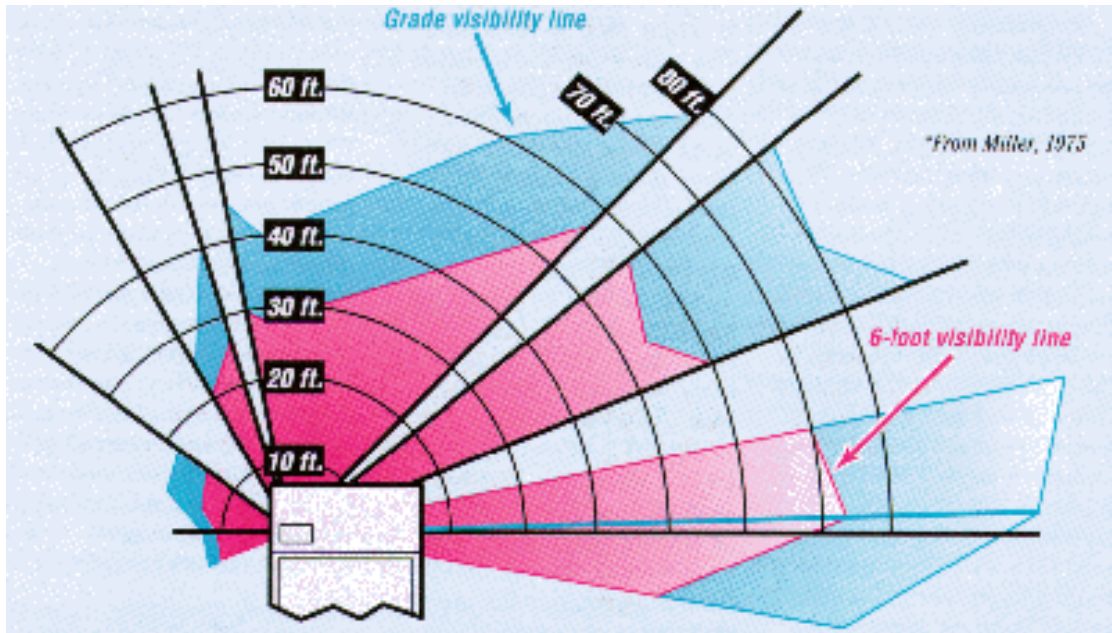
Figure 1.4: Detailed view of blind spots for the driver of a smaller haul truck, as seen from above [39]. An object 2m tall cannot be seen until the solid pink line. The ground cannot be seen until the solid blue line. The uncoloured regions are totally obscured by the pillars of the cabin.

In Fig. 1.4, the solid pink line shows the varying distance encloses the region in which a 2m tall object is invisible to the driver, and the solid blue line encloses the region where the driver cannot see the ground. The uncoloured regions are those obscured by the pillars of the cabin (drawn in solid grey in Fig. 1.2).

These blind spots have been a major contributing factor in numerous collisions involving large vehicles [28] [22]. The regions in front of or behind the truck are each large enough to park a light vehicle in, meaning that a truck driver can start moving with a vehicle in its path and be unaware of the danger until the collision. As Fig. 1.1 shows, this is extremely dangerous to anyone in the light vehicle.

1.2.2 Weather

Large vehicles must operate outdoors in what can be very challenging environments, and operations need to be able to continue even when weather conditions are poor. Rain, fog, snow, dust and low light conditions make it even more difficult for the driver of the vehicle to see in their limited field of view.

In mine environments, dust is a major problem, impeding visibility and settling on

(a) Fog [76]

(b) Dust [76]

(c) Rain [76]

Figure 1.5: Bad weather impairs driver vision and deteriorates camera and laser data.

all exposed surfaces [76]. Mines can be located in extreme environments with frequent rain, snow and fog. Water on the windscreen distorts vision, and precipitation in the air decreases the maximum viewing distance. In combination with dust, these weather phenomena can lead to all exposed surfaces being covered in mud. Night-time and other low-light operations can impede driver visibility, as can situations where the sun is shining into the drivers' eyes [76].

The combination of bad weather and blind spots means that the driver of a large vehicle only has information about a limited region of space around the vehicle, and it is often of poor quality.

1.2.3   Complex, Close-Range Interactions

Avoiding dangerous close-range operations is not an option for large vehicles. They must often operate in close proximity to other large and light vehicles in parking areas,

Figure 1.6: Large vehicles are required to engage in complex, close-range operations as part of their normal operation in a mine. [27]

intersections, areas near equipment like shovels and crushers where vehicles are loaded or unloaded [27]. Typically these operations take place at low speed, but the presence of blind spots described in Section 1.2.1 means that even low speed operations can be dangerous without good situation awareness.

1.2.4 Driver-Related Factors

Other factors contribute to close proximity accidents, such as driver error, fatigue, inexperience, distraction or inattentiveness. These behavioural factors are outside the scope of this thesis.

1.2.5 Situation Awareness

For the driver of a heavy vehicle, limited visibility due to blind spots, poor visibility due to bad weather, and the complex and dynamic nature of the environment lead to poor situational awareness. A comprehensive definition of **situational awareness** is given in [12] as "the perception of the elements in the environment within a volume of

PERCEIVE

the elements in the environment

ACT

to achieve goals while maintaining safety

COMPREHEND

the meaning of the elements in the environment

PROJECT

the current state of the environment
into the future
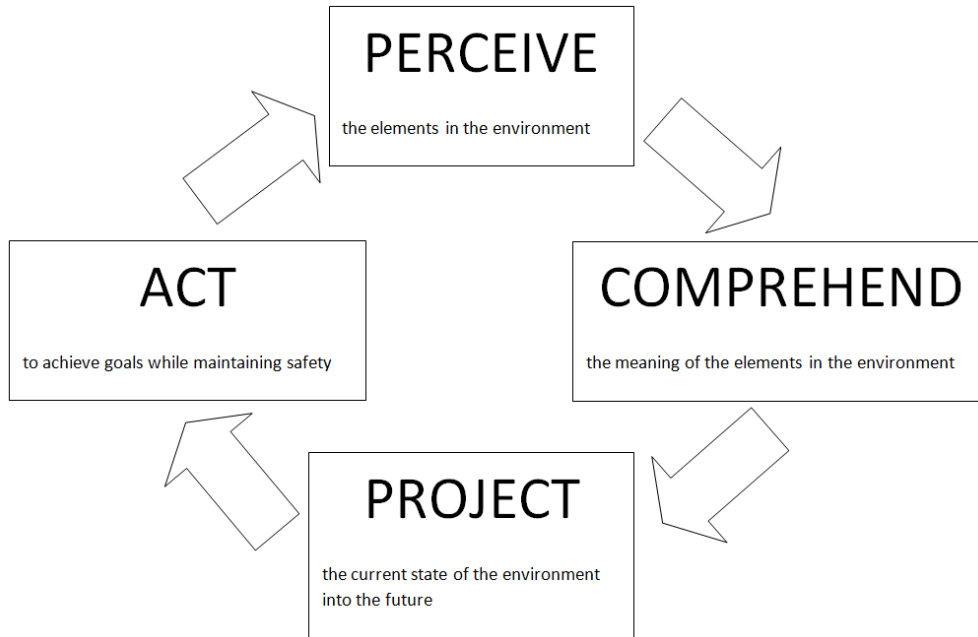
Figure 1.7: Driver situation awareness is important for safety, and relies on the driver continuously perceiving their environment, comprehending the meaning of these percepts, projecting the current environmental state into the future, and then taking safe action.

time and space, the comprehension of their meaning, and the projection of their status in the near future". A driver with good situational awareness can understand what is happening and what is going to happen around them, and take actions to ensure their own safety and the safety of others [48] [2]. This is illustrated in Fig. 1.7. With poor situational awareness, safe operation is much harder and accidents are more likely. Poor driver visibility hampers their perception of the environment, preventing swift comprehension and accurate projection into the future.

1.3   Situation Awareness Technologies

Technology can be used to improve driver awareness of objects in close proximity under these conditions. Two approaches to this can be called active and passive situation awareness systems.

Active situation awareness systems require objects in the environment to actively broadcast their position to be detected by the system. These systems require modification of the environment by attaching transponders or communications units, such as

RFID tags or wireless beacons, to the environmental features and agents that need to be detected. The communications units can communicate with other nearby units or a central base station via a wireless network [27]. Information about the local area can be transmitted between vehicles and displayed to the driver. These systems can allow detection and identification of potential collision hazards at a long range, with a low false-alarm rate and without the need for line-of-sight. But they require a unit to be fitted to every feature in the environment; anything without one will be invisible to the system [76]. They are typically accurate to within a few metres [27] which is a problem for close-range scenarios; a few metres of error can make a big difference when an object is only a few metres away. In longer-range, higher speed scenarios they work well; a few metres of error is less of an issue when there is larger spacing between vehicles and this error is unaffected by vehicle speed.

Passive situation awareness systems do not require any modification of or action by objects in the environment for the system to be able to detect them. These systems achieve this by mounting exteroceptive sensors on the vehicle to directly monitor the state of the environment. This does not require every feature and agent in the environment to be modified with a transponder, and may be done with little or no environmental modification at all. The capabilities of a passive situation awareness system are heavily dependent on those of the sensor being used. Depending on the sensor and the environment, there can be problems with false alarms and missed detections [76]. The sensors are typically limited in range and require line of sight, which is not desirable for long-range high-speed applications but acceptable for short-range low-speed scenarios.

Both active and passive situation awareness systems are already in use [27] [57]. This thesis focuses on passive situation awareness as being best suited to solving the close proximity problem. Not every agent in the environment can be fitted with a transponder, some large objects cannot realistically be fitted with enough transponders to represent their full extent, there is a risk of a transponder failing, and removing the need to make these modifications to the environment can help keep the cost down.

1.4 Sensors for Passive Situation Awareness

Vehicles can be fitted with sensors such as sonar, stereo and monocular cameras, lasers and radar. These can detect potential collision hazards that the driver cannot see and

provide warnings to prevent close proximity accidents, without the need to modify every object in the environment. But the effectiveness of such systems depends on the sensor being able to function reliably under the bad weather conditions when the driver's eyesight cannot.

In a mining environment, the main issues for using sensors are "durability, and the ability to operate in different weather conditions" [76]. For these reasons, this thesis argues that millimetre/microwave radar is the most appropriate sensor for a situation awareness system in these environments. Further comparison of different sensors will be presented in Sec. 2.2.1.

1.5   State of the Art

1.5.1   Radars In Intelligent Transport

Radars are capable of making long-range, high-precision measurements at a high frequency [5]. However, high-precision radars can be very expensive, costing up to $100,000, which can make them uncompetitive with other sensors such as 3D lidar which may have worse performance in bad weather but can be mounted and housed in ways to partially mitigate this problem. Existing radar applications in intelligent transport are restricted to applications such as adaptive cruise-control, collision avoidance and blind spot monitoring for large vehicles or in mapping and localising in static environments such as mines or container ports [1, 5]. Most are a collision-avoidance driver assistance tool rather than as part of a fully automated vehicle, which reduces manufacturers' liability [64] as well as being a much simpler problem to solve than constructing a fully autonomous vehicle system. The ability to use Doppler to detect a target's speed in addition to its position makes it a very useful sensor for transport applications [4]. Automotive radars are typically fairly narrow, unscanned beams covering only a small volume of space in a fixed direction [3] [4], and cannot be cheaply purchased separate from an entire car. Increasing research is being done on wider-scanning automotive radars but at the time this research was commenced, no such sensors existed. Radars with wide beams are used to search larger volumes of space for collision avoidance in large vehicles, but these provide poor cross-range resolution [57]. They can inform the driver of a large vehicle that an obstacle is present somewhere within the beam footprint, but not where it is.

To solve the close proximity problem for large vehicles, a narrow-beam, wide-scanning radar is needed to more accurately detect both the presence and position of collision threats, in all weather conditions at a lower cost.

1.5.2   Radar Processing and Mapping

FMCW radar is able to search a volume of space and detect multiple objects within it in a single measurement. However all measurements are to some extent corrupted by unwanted power from noise or clutter within the beam volume.

To discriminate between objects of interest in the environment and noise or clutter, a threshold detection filter is typically used. These can be fixed or dynamic, with the most popular of the latter being constant false-alarm rate (CFAR) processors [58]. These do not work as well with short measurement vectors or targets that occupy multiple measurement bins, due to the multiple close targets raising the threshold and masking one another out; these conditions are typical of low-cost radars and ground-vehicle operating environments, respectively.

No detection filter will be perfectly reliable; there will be false detections and missed detections. It is also worthwhile to combine measurement information in a map rather than just displaying it onscreenThrun2002. Consecutive measurements can then be fused together probabilistically, providing more accurate estimates of the environment than a single raw measurement.

One of the most prominent mapping approaches with radar is using occupancy grids (also called certainty grids or evidence grids). In an occupancy grid, the environment is divided into a tessellating pattern of uniform size grid cells. Each cell contains an independent random variable representing evidence of some state of the environment (hence the alternative names as evidence or certainty grids). Traditionally, the state being monitored is whether the area covered by the cell is empty or occupied by some object, with multiple measurements being combined to update the estimate.

Occupancy grids are simple to implement and produce robust maps [67], and because they "explicitly model free space" [50], occupancy grids are effective for obstacle avoidance, navigation and path planning. However they cannot represent target position uncertainty well, and it becomes computationally expensive to have a high-resolution grid or a wide-beam radar updating it at high speed. A sensor model is needed to perform the occupancy update, typically a heuristic, which is used to converts from

measurements to probabilities of occupancy.

Another option is a feature-based map, using tools such as a Kalman or particle filter. Feature-based maps model the environment as a list of features, "interesting" and recognisable landmarks that can be defined by a simplified model such as a point, line, corner or circle. The map consists of a list of the estimated positions of these features and their defining properties, and often a representation of the uncertainty of the positions (represented as a covariance matrix in an extended Kalman filter). Map management routines are needed to add new features to the map, or identify which measurements correspond to what known targets (data association) [1]. Feature-based maps can be implemented in different ways, such as using Gaussian noise-approximations (extended Kalman filters [63], unscented Kalman filters [68]) or particle filters [67].

Probability hypothesis density (PHD) filters [74] and cardinalised probability hypothesis density (CPHD) filters [73] are feature-based map representations that are growing in popularity due to their ability to reflect uncertainty in both target position and target existence. They overcome the more traditional mapping approaches' conceptual problems caused by radar's more complex detection characteristics [1, 44] where a detected target may or may not really exist.

1.6 Contributions

This thesis is focused on the development of reliable sensing for the close proximity problem in large vehicles operating in harsh environmental conditions. It emphasises the use of in-depth knowledge of a sensor's physics and performance characteristics to develop effective mathematical models for use in different mapping algorithms.

The contributions of this thesis are:

- An analysis of the close proximity problem and the demands it poses on sensing technologies.

- The modelling for a FMCW radar sensor for use in addressing the close proximity problem

- The development of a method for calculating a static detection threshold for use with FMCW radar, using measurements of clutter made using the radar.

- The development and implementation of a method of mapping in static environments that includes the modelling of a FMCW radar's detection and noise characteristics, using an occupancy grid approach.

- The development and implementation of a method of mapping in dynamic environments, using a Gaussian mixture extended Kalman probability hypothesis density filter, for use with a FMCW radar in solving the close proximity problem.

This thesis includes experimental results using simulated and real data from a low-cost 24GHz FMCW radar developed at the Australian Centre for Field Robotics at the University of Sydney. These results show improvements in the map quality obtained by being involved in the sensor design process and having access to internal sensor parameters.

The approaches used are not specific to this radar and could be used with other FMCW radars and potentially other sensors.

1.7 Structure of the thesis

Chapter 2 covers the background of the close proximity problem, situation awareness, FMCW radar and radar mapping.

Chapter 3 presents the design and development of our close proximity detector (CPD) radar, including the requirements analysis, radar chirp linearisation, radar characterisation and calibration, and the development of a detection filter.

Chapter 4 introduces the challenge of mapping with FMCW radar in static environments, including occupancy grid maps. The detection filter developed in Chapter 3 is used as a base for developing a probabilistic sensor model for occupancy grid mapping with radar.

Chapter 5 is focused on mapping with FMCW radar in dynamic environments, using occupancy grids and probability hypothesis density (PHD) filtering. The strengths of both approaches are compared, and analysis of a radar's performance characteristics are used to select the configuration parameters for an implementation of Gaussian mixture extended Kalman filter PHD filtering (GM-EKF-PHD). Simulated and real experimental data using both static and moving targets is used to analyse the performance of the implemented GM-EKF-PHD radar and the CPD radar in dynamic environments.

Experimental methods, results and discussion are included in each respective chapter.

Chapter 2

Background

2.1 The Close Proximity Problem

2.1.1 Mobile Agents In Mines

Operating any moving vehicle can be dangerous for both the operator and others nearby. Large vehicles are especially dangerous; their greater size poses a greater threat to objects in their vicinity, and the driver's view of the surroundings is very restricted [76]. Large vehicles are used in many industries; this thesis focuses on mining as an example of an application with some of the most challenging environments. A solution for the close proximity problem that works for this application should work in other industries and environments.

Large vehicles used in mining include haul trucks, bulldozers, front-end loaders, excavators, motor graders and mobile cranes [57]. These are used for extracting and transporting material, constructing haul roads and other necessary infrastructure of the mines and other operations. This thesis will focus on haul trucks, as these are numerous and are involved in a large proportion of close proximity accidents [32].

Light vehicles used in mining are typically four-wheel drives, to cope with the uneven terrain. They are used to transport personnel around the mine for tasks such as surveying, repair/maintenance and production [76]. Light vehicles may interact with large vehicles on haul roads and in parking areas. Pedestrians are also present in a mine; staff engaged in tasks like surveying, repair and maintenance, production management or moving between areas [76].

2.1.2   Close Proximity Accidents in Mines

The environments where large vehicles are commonly used can often be unstructured and complex. In an open-cut mine, haul trucks are loaded at the ore face by a mechanised shovel and then driven along haul roads to a stockpile or crusher where the ore is dumped [76]. There may be multiple trucks filled by each shovel, or depositing ore at the same site. Open-cut mines often have narrow haul roads roads with sharp turns, berms and intersections that obscure vision of oncoming vehicles, and high and steep walls. Light vehicles may travel on the same roads as large vehicles or move through the same areas within the mine such as parking areas or driver changeover areas. It is not always possible for vehicles to maintain a large separation from one another and regrettably accidents sometimes occur between them. Government agencies in Australia and internationally monitor and investigate incidences of accidents in mines.

In a survey of 2808 mining incidents in 16 countries, with the majority dating between 1980 and 2008, Macneill found that trucks (including haul, service and on-highway trucks) were involved in 258 of 2808 incidents (9.2%), the highest of any single equipment type [32]. Powered haulage (which includes haul trucks) was involved in 32% of all fatalities in the USA in metal and non-metal mining in 2013 [70]. One quarter of construction worker deaths are due to "collisions, rollovers, struck-by accidents, and a variety of other equipment-related incidents" [22]. In Western Australia 2005-06 to 2009-10, there were 4 fatalities involving vehicle or mobile equipment collision, and 2 from being struck by vehicle or mobile equipment, out of 20 fatalities overall [54].

In the 594 equipment-related fatal construction site accidents in the USA between 1990 and 2007 [22], dump trucks were involved in 173 of the 594 accidents, more than any other vehicle type. Trucks of unspecified type were involved in 73 accidents. 56% of dump truck accidents involved blind spots, and 7% of dump truck accidents involved bad lighting.

From these statistics it becomes clear that large vehicles can be dangerous, and when close proximity accidents do occur, blind spots are often a major factor.

2.1.3   Blind Areas

Although it is anecdotally stated that 90% of the information used by drivers is visual, there is no research to actually justify this number [61]. Vision nonetheless plays a very important role in driver situation awareness.

(a) Blind spots for a large vehicle, ground level [46] (b) Blind spots for a large vehicle, 1.5m above ground level [46]
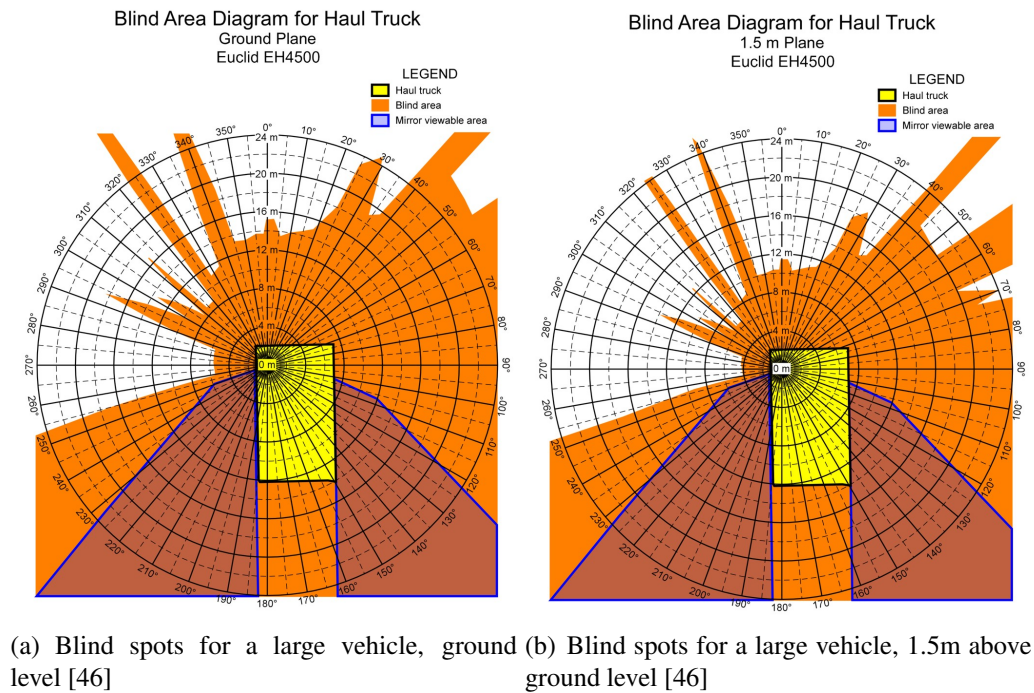
Figure 2.1: Large vehicles have blind spots.

Visibility can be degraded by blind spots, inadequate lighting and bad weather. Of particular concern are **blind spots** or **blind areas**, which the US National Institute for Occupational Safety and Health defines as "those areas where the equipment operator, who is seated in the equipment cab, cannot see an object by direct line of sight or in mirrors" [45].

Blind spots are present to some extent in nearly all vehicles due to the vehicle's size, construction or the driver's limited range of motion. Larger vehicles typically have larger blind spots, and vehicles the size of large haul trucks used in mining can have very large ones. The blind spot is more accurately a blind volume; the non-visible zone is a 3D region of space, and tall objects that rise out of it are visible at a nearer distance than shorter ones or the surface of the ground.

Tests must be performed to identify these blind areas. Different procedures are outlined in [45] and [23]. The results of one of these tests is shown in Figure. 2.1 for a similar haul truck.

The regions in front of or behind the truck are each large enough to park a light vehicle in, meaning that a truck driver could start moving with a vehicle in its path and be unaware. This has been a factor in numerous accidents such as the ones in Fig. 1.1.

2.2   Using Technology To Eliminate Blind Areas

One method to overcome the driver's visibility problems is to remove the driver from the vehicle entirely and automate the vehicle. While this has already been done in some applications (such as driverless haul trucks), it is not technologically or economically viable to automate the entire mine operation, and in some mines automating haul trucks will not be feasible due to cost, specific environmental conditions or other limitations. Technology can be used to augment a driver's limited situation awareness in conditions with poor visibility, by sensing in blind spots and conveying information and warnings to the driver by visual and aural cues. Different solutions are varyingly called situational awareness systems, driver assistance systems, close proximity systems, hazard detection systems or visual aids.

2.2.1   System Requirements

Kloos [27] defines a **close proximity system** as "a system that provides an indication or warning to the operator of a dangerous machine (truck), that someone or something is in close proximity to the machine (truck) and will thus allow the operator (driver) to take appropriate action to avoid accidents".

Functional requirements for a close proximity system are given as [27] [59] [57]:

- Reliable detection of a light vehicle and human being within 6m;

- No false alarms and especially, false negatives;

- Full coverage of dangerous areas;

- Easy to read and understand operator interface;

- Ability to handle harsh environmental conditions;

- Reasonable cost (one quoted figure is under ~$15000US per vehicle in 2007 [59], which would be approximately ~$17000US in 2015).

Kloos stresses the need for reliable detection of objects in proximity, with low false alarms.

There exist international standards to help eliminate blind spots in large vehicles, specifically *ISO 5006 Earth-moving machinery - Operator's field of view - Test method*

and performance criteria, and *ISO 16001 Earth-moving machinery - Hazard detection systems and visual aids - Performance requirements and tests* [15] [24] [23]. There are no Australian Standard equivalents to these international standards.

ISO 5006 sets out a method for evaluating the field of view (and identifying any blind spots) for the driver of a given vehicle and recommends the use of mirrors or video cameras "where the direct visibility is considered inadequate." [15]. ISO 16001 focuses on evaluating situation awareness systems, or "hazard detection systems (HDS) and visual aids (VA) for detecting people" that might be used to sense objects within these blind spots.

An ideal situation awareness would provide the driver with total awareness of the environment around them, in any weather conditions, with infinite range, perfect accuracy, zero noise, infinitely fast update rate and in a form that was non-distracting and easy for the driver to interpret. Real systems can only approximate these goals; as ISO16001 states "It is essential to note that HDS and VA have both advantages and disadvantages. There is no device that works perfectly in all situations." [15]

This thesis argues that the sensor's reliability is paramount. It needs to sense the risks that are present, and return a minimum of **false alarms** (also called **false detections** or **false positives**). All real sensors will have limitations, whether it be in range, resolution, sensitivity, line of sight requirements or degraded performance in bad weather. Keeping the performance criteria for a close proximity system in mind, the available sensing modalities will now be examined in greater detail.

2.2.1.1 Cameras

Visible light cameras can possess a wide field of view allowing good volume coverage, a rich measurement modality that is familiar to drivers, and they are low-cost compared to other sensors that might be considered for the close proximity problem. A high measurement rate and excellent angular resolution have made them a popular sensor in many robotics and intelligent transport applications [4]. Stereo vision systems are under development for the consumer vehicle market, allowing cameras to provide range as well as bearing estimates, and segmenting and tracking features in the environment.

Visible light cameras require a source of illumination whose light is reflected off the objects in the environment and into the camera. Low or uneven light levels, shadowing,

excess light and image noise can produce low-quality images. Even when there is sufficient illumination of the scene, camera images can be degraded by the same weather effects that degrade driver visibility, as the images of bad weather conditions in Figures 1.5 show. Water droplets on a camera lens act as additional lenses, distorting the image. Once the lens is wet, dust will stick to it, obscuring that region of the image. In particularly dusty environments, the lens can become coated even while dry. In a mine, where unsealed roads lead to a dust trail behind every vehicle [76], this is almost inevitable.

Infrared cameras are becoming available that do not require external sources of light, as vehicles and human beings are typically hotter than the background. These cameras are capable of penetrating some amount of dust [4] but they will still struggle to perform when covered in it.

Existing situation awareness systems using visible light cameras display the camera feed directly to the driver, similar to the reversing cameras present in high-end consumer cars. Obstacle detection relies on the driver watching the video feed. The driver will need to inspect the image, detect a hazard and then make an action based on that. The driver is still responsible for detecting a hazard, which could come in a form they are not anticipating and so may not recognise in time. If the driver's attention is elsewhere, or the picture is severely degraded by rain, dust or mud, the system will provide little or no protection. Automating object detection is a challenge in this application due to the dynamic and unstructured environment, varying light and shadow conditions, inclement weather, and the fact that determining the range to an object requires multiple cameras and more complicated processing. Vision can be useful for confirming object presence once detected by another sensor [27] but as a primary detection sensor, it leaves a lot to be desired.

2.2.1.2   Laser Rangefinders

Laser rangefinders (or lidars, from light detection and ranging) mostly transmit a pulse of laser light into the environment where it can be reflected off an object back to a receiver in the sensor. The time of flight is measured and used to calculate the distance to the object. By rotating the laser beam, a 2D fan-shaped scan of range-bearing measurements of the environment can be rapidly built [4]. Some laser rangefinders return the strength of the reflection as well as the range, which can be useful for imaging purposes.

Lidars come in 2D and 3D variants, with 3D lidars being quite expensive.

Lidar does not require external lighting and so is equally reliable day or night, though it can be hampered by mirrored surfaces that reflect the beam without registering as a detected object, or dark surfaces that absorb the infrared beam without reflecting a sufficient amount of energy for detection. Lidar is affected by atmospheric particles such as dust, fog, steam and heavy rain, which cause multiple reflections along the path of the beam if they do not block it entirely. Newer lidars provide for multiple echo capabilities, allowing some of these effects to be handled more effectively with light rain or dust, but if the sensor becomes coated in dust or mud it will be unable to function [4] [21]. The narrow, planar nature of a 2D lidar scan means that it needs to be placed quite low on the large vehicle or it will overlook objects in the blind spots caused by the height of the cabin. But this will require it to be exposed to increased amounts of dust and mud kicked up by other vehicles.

3D lidar can be mounted higher to protect it from the elements, but these sensors are much more expensive. And if any dust does land on the sensor, its performance will degrade.

2.2.1.3 Radar

Radars transmit a pulse of radio signal into the environment, where it can be reflected off an object back to a receiver antenna. There are many different types of radars [4] but for this thesis we will focus on frequency modulated continuous waveform (FMCW) radars in the microwave/millimetre wave bands. These radars are comparatively small, inexpensive and possess good range and range resolution [7] [14].

Scanned FMCW radars are able to measure the range, bearing and radar cross-sections of objects in the environment. Using the Doppler effect, some are able to measure the relative speed of objects as well.

FMCW radar performance can be likened to other active range-bearing sensors like lidar - they transmit electromagnetic radiation into the environment and use the reflection time and power level to create a measurement. But radar presents unique challenges including higher cost, a lower measurement rate, higher susceptibility to clutter, lower bearing resolution (because of the larger beam size) and a more complex measurement to process [1] [43]. Despite these challenges, radars have the ability to perform well in much worse weather conditions than other sensing modalities. Unlike the

visible-wavelength radiation that visible-light cameras rely on, or the infrared radia-
tion used in lidars, the frequencies used in microwave or millimetre-wave FMCW radar
are less absorbed or reflected by rain, fog, snow or dust [7]. Under these conditions,
FMCW radar will suffer far less degradation in measurement quality than camera or
lidar [47] [76] [21].

An airborne cloud of dust can obscure objects from a camera or lidar, whereas
a FMCW radar is able to penetrate the dust and continue to detect objects within it.
The radome (or cover) of a FMCW radar can even be coated in a layer of dust or
mud and continue to function; these conditions would render nearly any other sensor
totally useless [4]. This is a major advantage in applications that require reliable all-
weather performance. This makes FMCW radar an attractive sensor for a situation
awareness system for all weather conditions and for that reason is already seeing use in
the field [4] [1]. For these reasons, FMCW radar is the sensor we will focus on for this
thesis.

2.2.2   Why Not Just Use A Camera?

The international standards for operator visibility in large vehicles and hazard detection
systems, ISO 5006 and ISO 16001, consider video cameras as the best substitute for
direct vision; they are inexpensive, low maintenance and do not require modification of
the environment. Other techniques such as RFID are considered inferior to video cam-
eras for overcoming blind spots, due to being themselves a "blind technology" capable
of detecting the presence of an object but not necessarily its exact location [24] [23].

This assertion of camera superiority is based on the assumption that

- a video camera will be able to sense an object reliably at all times, clearly enough
  that the driver will be able to see it onscreen

- the driver will be watching the video feed when an object is present.

- the driver will recognise the risk present and take appropriate action in time.

The onus is still on the driver to detect any risks present. This is not reliable enough
in complex, unstructured, unclear environments where there may already be multiple
tasks, screens and potential hazards vying for the driver's attention as they operate their
vehicle.

As discussed in Sec. 2.2.1.1, cameras are an intuitive but limited sensor that are difficult to automate in a cluttered, unstructured environment. Like human eyes, cameras are reliant on an external light source whose light is reflected off objects in the environment and into the sensor [16]. Even if the driver is paying attention, water or dust on the lens, or inadequate lighting will reduce the camera's ability to provide information.

2.2.3 How Bad Is Bad Weather?

Heavy vehicles must operate outdoors in what can be very challenging environments, and operations need to be able to continue even when weather conditions are poor. Rain, fog, snow, dust and low light conditions make it even more difficult for the driver of the vehicle to see in their limited field of view.

2.2.3.1 Mist, Fog, Rain and Snow

Precipitation will degrade driver visibility, either as water in the air reducing viewing distance or droplets or condensation on the windshield [76]. The presence of water in the air will also degrade the measurements of sensors using the non-visible electromagnetic spectrum. Fig. 2.2 shows the attenuation of different wavelengths of electromagnetic signals in different atmospheric conditions. The vertical axis is a logarithmic scale in decibels per kilometre; for the close proximity problem, attenuation is only an issue when it begins to approach 100 dB/km (since the close proximity problem is in the regions of tens of metres and any attenuation less than this will be very small over that distance).

A typical lidar operating at a wavelength of 900nm will not be too badly affected by drizzle or even heavy rain. Deluge may degrade performance, but fog will severely limit the usefulness of the sensors [4]. Even in lighter rain, water droplets on the lidar, and dust and mud that stick to them, obscure the laser beam and prevent the lidar from making measurements of the blocked regions.

2.2.3.2 Poor Lighting

Many mines operate 24 hours a day, and even those that do not will often be working at dawn and dusk [76]. The low sun can shine straight into the cabin of haul trucks, into the eyes of the driver. Lights from other vehicles can blind a camera, especially when
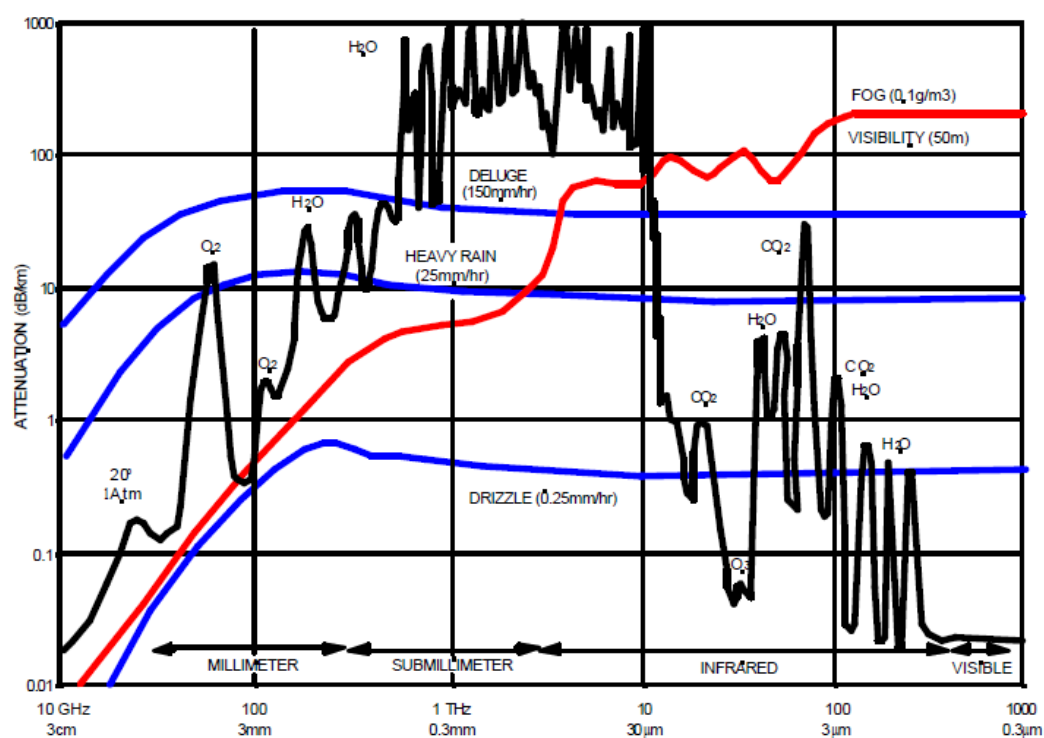
Figure 2.2: Atmospheric attenuation as a function of frequency with different precipitation rates [6]. The horizontal axis is frequency (or wavelength) of electromagnetic radiation. The vertical axis is attenuation in decibels per kilometre.

reversing, and glare can make it difficult for drivers to see and reduce the quality of any camera image [22]. When light levels are too low, a driver cannot detect potential risks and a camera image is dim. One survey of 508 mining incidents in Queensland found that visibility was a factor in 11% of accidents [53], either due to inadequate lighting or glare. Lidar and radar are not affected by poor lighting.

2.2.3.3   Dust and Mud

Dust is very common in many mine environments. It is thrown up by the tyres of vehicles travelling on unsealed haul roads [76], wind, machinery and explosive blasting. Most mines require water trucks to continually spray the haul roads to suppress dust, but this cannot completely eliminate it. Dust will accumulate on any exposed surface, especially in the presence of rain. This can be ruinous for sensor performance.

| | | | | | Attenuation (dB/km) | |
Source	Material	Density (g/cm^3)	Mass Concentration	Visibility (m)	$\lambda = 3mm$	$\lambda = 1 - 10\mu m$
Desert dust	Quartz	2.6	10	3.4	1.56	1300
Stack: stone mill	Quartz	2.6	80	0.5	12.5	10400
Stack: steel mill	Coal	1.5	10	3.4	2.2	1300
Light fog	Water	1	0.001	7000	0.004	0.2
Heavy fog	Water	1	1	55	4	200

Table 2.1: The attenuation of millimetre-wave radar ($\lambda = 3mm$) and infrared ($\lambda = 1 - 10\mu m$) through dust and fog. Infrared is severely degraded compared to radar [4].

As Table 2.1 shows, millimetre-wave radars can penetrate dust clouds far better than infrared. Lidar has very similar attenuation characteristics to visible light; dusty or foggy conditions that block vision will probably block lidar as well [4]. Radars can even continue to function with partial coverage of dust or mud on the antenna [6], and will degrade more gracefully with more extensive coatings [59] [21].

2.2.4   Sensor Selection

From the system requirements outlined in Section 2.2.1 and the effects of bad weather on drivers and sensors in Section 2.2.3, it is clear that

- reliability is paramount for any blind spot monitoring system.

- the bad weather conditions in which driver visibility is most degraded have nega-
  tive effects on most sensing modalities other than radar, reducing their reliability
  when they are most needed

Radar is the sensor modality providing the best all-weather performance without re-
quiring modification of the environment (such as with RFID tracker tags) [6] while still
being capable of at least partially automating the risk detection process. The remainder
of this chapter will focus on the characteristics, design and use of FMCW radars for
solving the close proximity problem.

2.3   FMCW Radar

Radar is originally an acronym for **radio detection and ranging** and involves transmit-
ting a pulse of electromagnetic energy into the environment as a directed beam. The
pulse is reflected off an object back to a receiver, and the distance calculated [4].

Different operating principles can be used for radars. **Frequency-modulated con-
tinuous waveform** (FMCW) is popular in conditions when good range resolution is
required, such as in robotics and intelligent transport systems. FMCW requires lower
peak power outputs than pulsed time-of-flight radars, a lower minimum range than pulse
compression radars, and shorter measurement time compared to stepped frequency
radars [6].

Automotive radar development began in the 1970s, when 35GHz antennas were first
able to be miniaturised for use in test vehicles [36]. The 77GHz frequency band was
allocated for automotive use, encouraging development of sensors at that frequency.
The first commercially available consumer vehicle with a radar was released in 1998
and had a 77GHz radar. Today, radars see widespread use in intelligent transport and
close proximity collision avoidance [3] [57]. Many high-end consumer vehicles are
fitted with radars for adaptive cruise control, to automatically maintain a safe distance
behind a vehicle in front without driver intervention [36]. A typical automotive radar
might have a beam $4°$ in yaw, $15°$ in elevation, operating at 77GHz and be mounted
behind the front grill of the vehicle approximately 1m off the ground [3]. Shorter range
radars at 24GHz are also becoming common [36]. While the radars are common in new
vehicles, they are not readily available for purchase separate to the car in most cases,
and can be very expensive if they are.
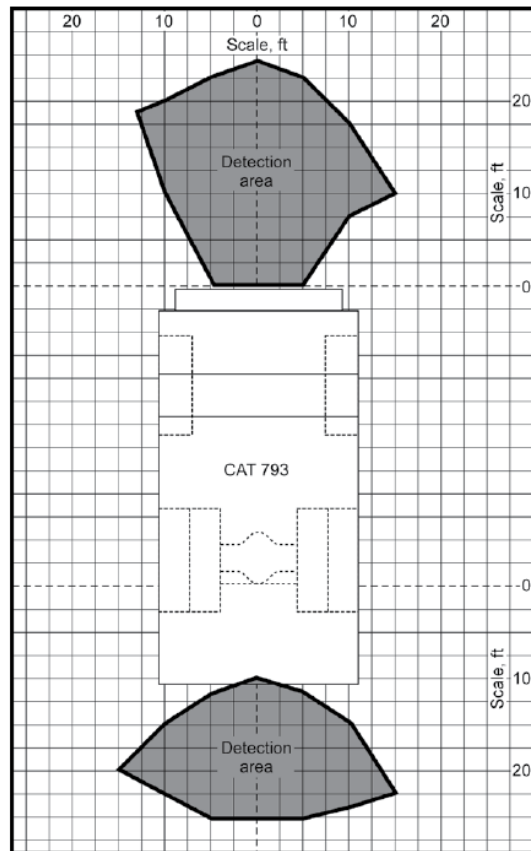
Figure 2.3: The areas in which a Preview Heavy Duty Radar System from Preco Electronics can detect a human being when mounted on the front or back of a haul truck [57].

A commercially-available radar for close proximity collision avoidance in large vehicles is evaluated by Ruff [59] [57]. It has a beam approximately 60° wide in azimuth and indicates the presence of an object via a row of LEDs. A diagram of the areas in which it is able to detect a human being when mounted on a haul truck is shown in Figure. 2.3.

The system displays the presence of objects within the beam via a row of lights that indicate the proximity of the nearest object. This system demonstrates some of the strengths and limitations of existing close proximity radars for large vehicles; the beam is wide and unscanned, giving poor cross-range resolution. **Nuisance alarms** are sometimes a problem; as shown, the detection zone extends beyond the width of the vehicle, and if the vehicle is being driven next to a strong stationary reflector (such as reversing parallel to a wall or beside another vehicle) it may detect the object beside it
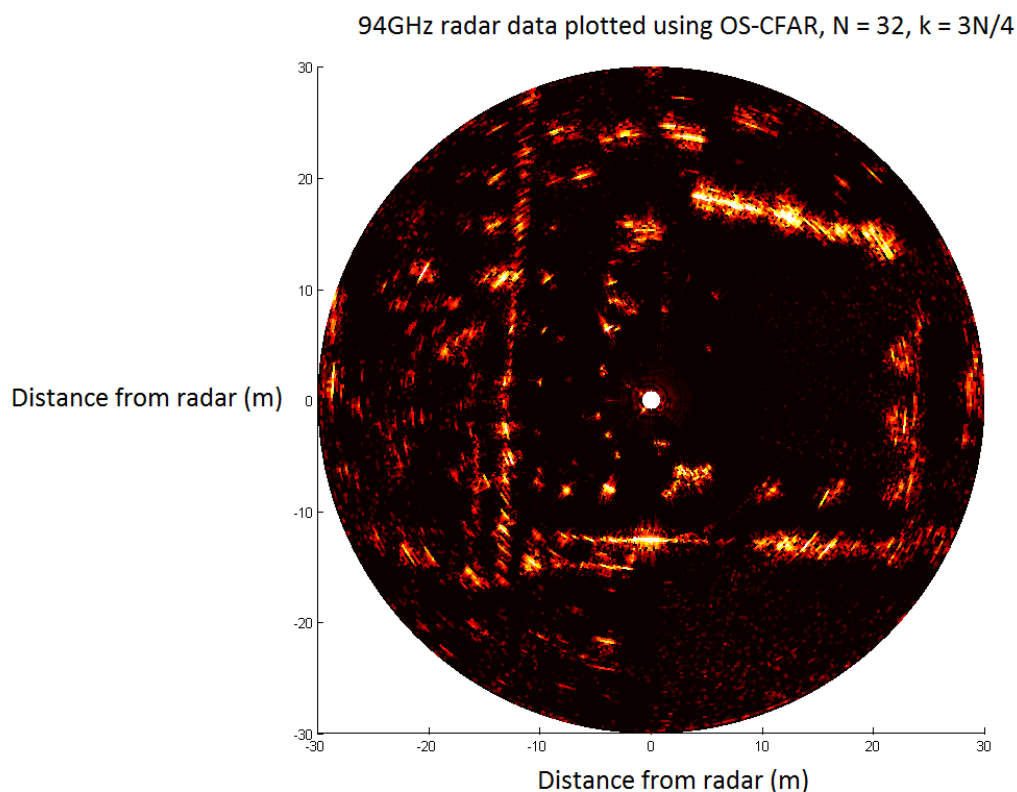
Figure 2.4: This plot shows data from a high-quality radar with a narrow beam in an outdoor environment, near buildings, walls, trees, poles and other structures, and some human beings. These are high quality measurements, beyond what is needed to solve the close proximity problem, but such radars are too expensive for widespread application.

and warn the driver, when in reality this object poses no threat.

More expensive radars are available, with narrower beams, longer range and better range resolution. These can produce detailed maps of the environment, but at much higher cost (on the order of tens or hundreds of thousands of dollars). This is not economical for a fleet of large vehicles. Fig. 2.4 shows a plot of a high-quality 94GHz radar in an outdoor environment, plotted with the bright areas depicting areas containing strong reflectors while the black areas are empty space. A satellite photograph of this scene is shown in Fig. 2.5. In Fig. 2.4, walls are visible as long bright lines, while trees and poles form smaller bright points. This radar has excellent range and bearing resolution, and is more than precise enough to solve the close proximity problem were it mounted on a large vehicle; this is not viable due to the sensor's high cost.

Figure 2.5: This satellite photo shows the scene scanned by the radar in Fig. 2.4. The buildings are clearly visible, as are some of the trees and other structures[17].

2.3.1 FMCW Radar Measurement Structure

FMCW radar is unlike many other rangefinders such as lidar in that each measurement does not measure the range to a single target "point". Instead, each measurement returns a vector of a fixed number of power bins. For the purpose of this thesis, one **measurement** is a vector of **bins** containing **power values** in decibels, produced from a single transmit-receive cycle. The radar beam can be swept across the environment either mechanically or electronically, or kept static. A series of measurements across the full extent of its movement will be referred to as one **scan**.

A single measurement can detect multiple objects if the objects closest to the radar do not totally occlude the ones further away, or if the nearer object allows the radar to penetrate through it. This will depend on the nearer object's material type and thickness, and the frequency of the radar [44]. Lower frequencies have better penetrative properties than higher frequencies.

Bins whose corresponding beam-volumes contain reflectors will contain a larger power value than those that do not, but every power value of every measurement will
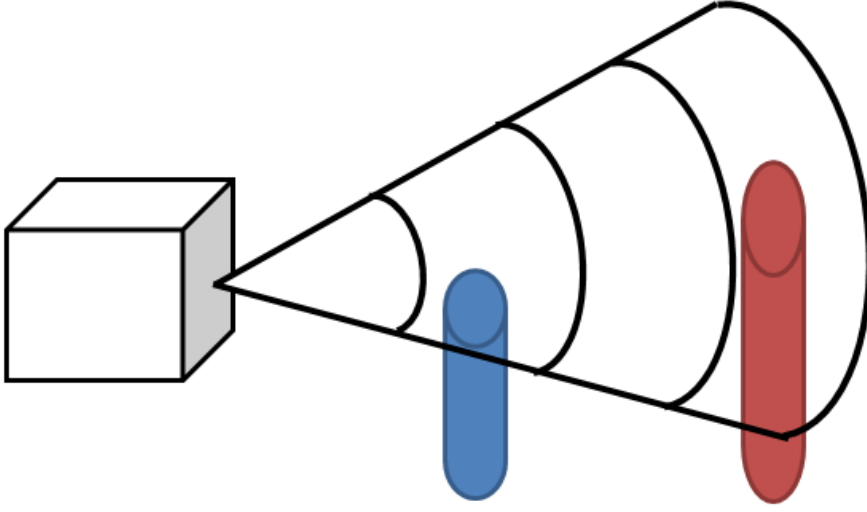
Figure 2.6: Each radar measurement consists of a vector of bins. In this simplified
            diagram, objects are present within the beam volume corresponding to the
            second and fourth bins. As the nearer object does not totally occlude the
            farther one, both should be detectable.

also be corrupted to some extent by noise within the radar and clutter within the envi-
ronment. These are described in greater detail in Sec. 2.4.1.

2.3.2   Radar Cross-Section (RCS)

Rather than returning an image similar to how a human being would see the world the
way a camera does, or the range to objects in the environment the way a lidar does,
radars sense the **radar cross-section** (RCS) of objects within the beam's volume.

Radar cross-section "qualitatively relates the amount of power that strikes the target
to the amount of power that is reflected into the receiver" [4]; it is a measure of the
"size" of the object as viewed by a particular wavelength radar. Radar cross-section is
typically measured in square metres (m^2) but due to the large range in RCS this is often
given in dBm^2.

The power P_r (W) reflected from an object isotropically (uniformly in all directions)
when it is struck by a radar wave of power density S_i (W/m^2) is directly proportional to
the object's radar cross-section σ (m^2) [4].

$$P_r = \sigma S_i \qquad\qquad (2.1)$$

An object's RCS is a function of its size, shape, material and surface finish, and the radar's frequency, polarisation and aspect angle to the object. Some of these factors will be discussed in greater detail.

2.3.2.1 The Effect of Object Size on RCS

The relationship between an object's size and its RCS is complex. Large objects will be composed of multiple reflectors. For example, a car's engine, chassis, headlights and mirrors might all reflect some of the transmitted energy back to the receiver. The reflections from each component of a complex reflector (or of multiple nearby objects) can interfere with one another, both constructively and destructively. A small change in the aspect angle between the radar and the objects can shift the interference substantially, causing a large change in the effective RCS. This phenomenon is called **scintillation**[62]. This makes it difficult to predict the RCS of complex targets; sophisticated software tools are available but empirical testing is the best approach [4]. It also contributes to **speckle noise** where multiple weak clutter reflectors can combine to appear as a single strong reflector, or cancel completely, creating spurious measurements of small targets on empty ground.

An object that is physically very large can have a very small radar cross-section if it is made in a radar-scattering shape and/or out of radar-absorbing or scattering materials. This is used in the design of "stealth" aircraft and naval vessels. Even modern cars with their blended curves and aerodynamic shaping have reasonably low RCS for their size, compared to angular industrial or utility vehicles.

An object that is physically very small can have a very large radar cross-section if it has the right shape. This is used in the design of radar targets such as trihedral corner reflectors. A corner reflector reflects a significant proportion of received energy across a fairly wide range of aspect angles [14]. For a triangular trihedral corner reflector of side length a m, its radar cross-section when measured by a radar with wavelength λ is given by:

$$\sigma_{ref} = \frac{4\pi a^4}{3\lambda^2} \tag{2.2}$$

2.3.2.2 The Effect of Aspect Angle on Object RCS

Aspect angle is the angle between the beam and the flat reflecting surface - a $0°$ aspect angle means the object's bearing is directly towards the beam, giving best reflection.
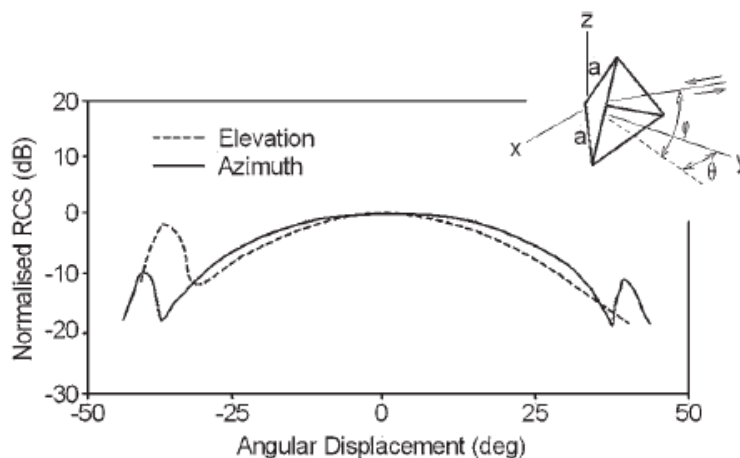
Figure 2.7: Trihedral corner reflectors make good radar targets as they can be mis-
    aligned relative to the radar by a few degrees with minimal decrease in
    RCS [4].

Larger aspect angles can reflect more power away from the radar and less power back.

Only spheres and cylinders have a fairly constant cross-section, with the latter being specular around one axis. Other objects have a statistical RCS that varies as a function of aspect angle [4]. To describe the RCS of a target it is necessary to measure it over a range of aspect angles, by rotating the target relative to the radar, or orbiting the radar around the target.

Figure 2.7 shows the variation in normalised RCS of a trihedral corner reflector with varying aspect angles in azimuth and elevation. The RCS changes smoothly and gradually, making this shape a good choice for synthetic radar targets. Figure 2.8 shows the RCS of a Toyota utility vehicle rotated on a turntable and scanned with a 35GHz radar at $0°$ grazing (pitch) angle [5]. The RCS fluctuates over a wide range, even for small changes in aspect angle, due to interference between component reflectors.

Tests like this are important for characterising radars to be used for close proximity systems; it is of little use if objects as large as a light vehicle cannot be detected at certain angles.

2.3.2.3   The Effect of Object Material on RCS

The object's material plays a large role in determine RCS. The radar reflectivity characteristics of a material are inversely proportional to the material's relative dielectric constant [4]. Metals and water reflect well; plastics and wood, less so. Human beings
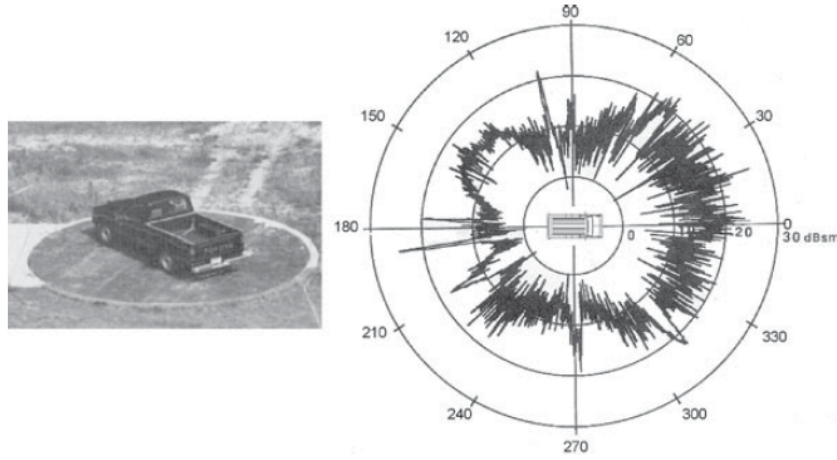
Figure 2.8: This light vehicle was rotated on a turntable and scanned with a 35GHz radar. The spider plot on the right shows the RCS in decibel square metres as a function of aspect angle (further from the centre is higher RCS, while the numbers around the edge show the aspect angle). Its RCS fluctuated considerably with even small changes in aspect angle. [4].

can be approximated as ellipsoids of water, with an RCS of one quarter their surface area; assuming a surface area of $\approx 2\text{m}^2$, this should be an RCS around 0.5m^2 (higher at higher frequencies).

2.3.3   Swerling Models of RCS

Since the fluctuation of objects' RCS is so difficult to predict, it is typically modelled mathematically. Swerling [65] describes five different models of the fluctuations in targets' RCS that are widely used.

- Swerling 1: Scan-to-scan fluctuation. The target's RCS is constant for pulses in a single scan, leading to reflected power of constant amplitude in the scan, but the RCS fluctuates independently from scan to scan. The target's RCS follows a Rayleigh probability distribution.

$$p(\sigma) = \frac{1}{\sigma_{av}} e^{\frac{-\sigma}{\sigma_{av}}} \tag{2.3}$$

where σ is the measured target RCS, and σ_{av} is the mean of σ.

- Swerling 2: Pulse-to-pulse fluctuation. The target's RCS fluctuates independently

from pulse to pulse, even within a single scan. The target's RCS is still Rayleigh distributed.

- Swerling 3: Scan-to-scan fluctuation, and the target's RCS has a chi-squared probability distribution function.

$$p(\sigma) = \frac{4\sigma}{\sigma_{av}^2} e^{\frac{-2\sigma}{\sigma_{av}}}) \tag{2.4}$$

- Swerling 4: Pulse-to-pulse fluctuation with a chi-squared RCS distribution.

- Swerling 5: Nonfluctuating RCS.

The Swerling 1 and 2 models describe a target that has multiple reflectors (at least 5) of approximately equal RCS, such as aeroplanes. The Swerling 3 and 4 models describe a target which consists of one large reflector and several smaller reflectors, such as missiles [1]. The Swerling 5 model describes a spherical or almost spherical object.

For the purposes of solving the close proximity problem, the Swerling 1 model is assumed. Objects such as vehicles and buildings are large and made up of many reflectors, but they are not typically moving so rapidly that the RCS will fluctuate much within a single scan. The net effect of this is that the power values corresponding to this object, being linearly proportional to the RCS, will follow a Rayleigh distribution.

2.3.4   The Radar Range Equation

Like all radars, FMCW radar performance is described by the radar range equation [4].

For a target of RCS σ (m^2) at a range of R (m) from a radar with a wavelength of λ (m) transmitting power P_t (W) through an antenna with gain G, with system losses L, the power density of the reflected echo at the receiver S (W/m^2) is

$$S = \frac{P_t G^2 \lambda^2 \sigma L}{(4\pi)^3 R^4} \tag{2.5}$$

As Equation 2.5 shows, the received power is inversely proportional to range raised to the power of 4. This is due to the inverse-square power radiation law being applied twice: once when the beam is transmitted into the environment, and once when it is reflected back from an object in the environment. This leads to a very rapid dropoff in

received power with increasing distance to the target, and so a very large range in the power measurements that will be received.

To express this range more compactly without high-power measurements dominating so severely, radar power measurements are commonly converted to decibels (dB) [4]. Decibels are a logarithmic scale relative to some reference value.

$$P_{dB} = 10log_{10}\frac{P}{P_{ref}} \tag{2.6}$$

Whether the reference power is known or not is irrelevant; the decibel measurement is a ratio and measurements made from the same radar can be compared relative to one another regardless of the true value of the reference.

Converted to decibels, the radar range power equation is

$$S_{dB} = P_{tdB} + 2G_{dB} + 10log_{10}\left(\frac{\lambda^2}{(4\pi)^3}\right) + \sigma_{dB} + L_{dB} - 40log_{10}R \tag{2.7}$$

It is useful to rearrange this to calculate the maximum detection range. For a target of RCS σ (m^2) measured with a radar of wavelength λ (m) transmitting power P_t (W) through an antenna with gain G with system losses L, and minimum detection power S_{min} (W/m^2), the maximum detection range R_{max} is:

$$R_{max} = \left[\frac{P_t G^2 \lambda^2 \sigma L}{(4\pi)^3 S_{min}}\right]^{\frac{1}{4}} \tag{2.8}$$

Derivation of this equation is given in [4].

We can see that the maximum range is proportional to the transmitted power and the target's RCS raised to the one-fourth; a more powerful radar can detect a smaller object at longer range, and larger objects can be detected more easily by less powerful radars.

This has limited use in practice; the maximum detection range is typically much smaller than this equation would suggest, due to additional losses and noise [4].

2.3.5 Radar Range Resolution

The **range resolution** of a radar is the minimum distance required between two objects for them to be identified as distinct objects rather than one extended object.

A fine range resolution is desirable for precise mapping. This can be achieved using pulsed amplitude modulation with a very short rectangular pulse. As equation 2.8

shows, the maximum range of the radar is proportional to the transmitted power; to achieve both good resolution and good range, a very large amount of power will need to be transmitted in a very short time [4].

Alternatively, **frequency modulation** (the FM in FMCW) involves transmitting a **chirp** or sweep of frequencies. The range resolution δR (m) in terms of the signal bandwidth Δf (Hz) is

$$\delta R = \frac{c}{2\Delta f} \tag{2.9}$$

where c is the speed of light [4]. To minimise δR, a large frequency sweep is required. Frequency modulation allows for better range resolution and spreads the transmitted energy out over a longer period, allowing a lower peak power [4]. More details on the operating principles of FMCW are given in Sec. 2.3.9.

2.3.6   Radar Bearing Resolution

2.3.6.1   Beam Width

The radar's antenna has a **radiation pattern** that describes how strongly it transmits and how sensitively it receives as a function of the angle around it (assuming it is a reciprocal antenna, that has the same pattern for transmission and reception [14]). The "strength" or "sensitivity" of the antenna in a certain direction is called the **gain**. It is expressed as a ratio relative to an omnidirectional antenna (one that transmits uniformly in all directions i.e. a spherical radiation pattern), in units of decibels of transmission power intensity (dBi) [14].

An example radiation pattern is shown in Fig. 2.9. The **mainlobe** has the peak of the gain in azimuth, which then decreases away from that axis. The gain does not continuously drop, but takes the form of periodic peaks. This indicates the presence of secondary lobes, or **sidelobes**, either side of the mainlobe [4]. By concentrating the power/sensitivity in one direction, the maximum range of the radar is increased (see Eq. 2.5).

The radar beam is often considered as consisting solely of the highest-gain section of the mainlobe. The beamwidth is defined as the angle range over which the mainlobe's gain drops by 3dB (a 50% reduction). This is a function of the radar's wavelength λ (m) and the aperture diameter D (m). In a conventional radar configuration, the signal level of the received echo would be reduced by 6dB (3dB on transmit and 3dB on receive).
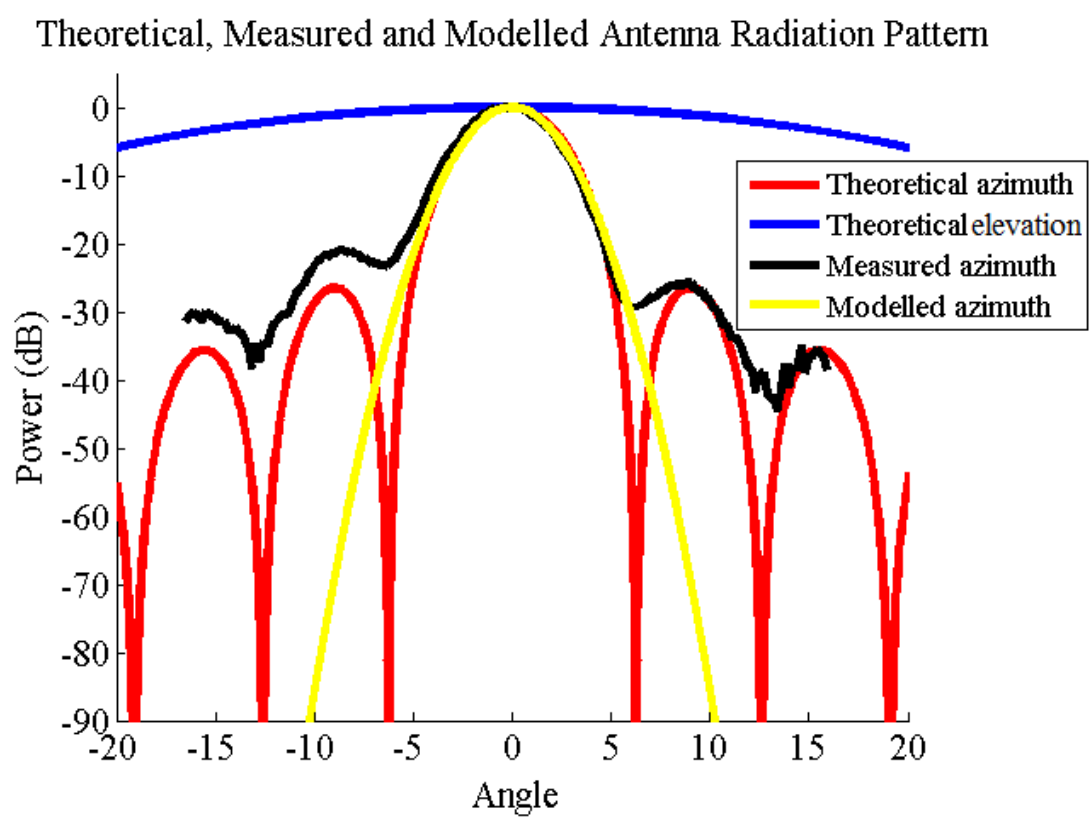
Figure 2.9: This plot shows theoretical and measured radiation patterns for a radar antenna.

In degrees, the beamwidth θ_{3dB} (°) is [4]:

$$\theta_{3dB} = \left(\frac{70\lambda}{D}\right)^{\circ} \tag{2.10}$$

The maximum antenna diameter allowable is typically a function of the application. For the close proximity problem, the size is fairly constrained; the radar needs to be small enough to mount on a vehicle; this requires an antenna diameter of ≈30cm or less.

A narrower beam is useful for precise imaging [14], as it has better **angular resolution**, which allows the radar to detect whether two nearby objects are separate or are components of the one large object [4]. A narrow beam is worse for volume monitoring as it will need to be scanned to monitor a wider scene and risks overlooking objects if the elevation angle is such that the beam does not illuminate the target.

As described in Sec. 2.3.5 and Sec. 2.3.6, a radar bin covers a volume of space rather than a point or a line. The volume is determined by the range and the beam width, the latter of which is determined by the antenna diameter and the operating frequency.

The fact that the radar beam covers a volume of space rather than just a point or a line can make it difficult to determine exactly where a detected object is. A widely used approach is to project from the polar coordinates used in the radar measurement into a discrete Cartesian grid [5]. The chosen grid resolution will determine the precision of the plotted data as well as the ease of calculating which bins overlap which grid cells; small cells will make a finer map but the projection into the grid will be more complicated due to the larger number of intersections between the beam footprint and the grid.

2.3.6.2 Sidelobes

Outside of the 3dB width, the mainlobe and any sidelobes may still be capable of detecting some targets, either those at very close range or with very large radar cross-sections. For this reason, strong sidelobes are not desirable; they degrade bearing resolution by introducing uncertainty about which lobe detected the object.

A FMCW antenna's gain pattern approximates the sync function $\left|\frac{sin(x)}{x}\right|$ [4]. The first sidelobes can be seen to be 13dB lower than the mainlobe (effectively 26dB lower after both transmitting and receiving). This can be unacceptably high for some applications, so the received signal is subjected to a process called **windowing** where a varying

weighting function is applied to the received signal to weight it over the angular domain, which has the effect of drastically reducing the sidelobe strength (to -31dB or less compared to the mainlobe) but slightly weaken and substantially widen the mainlobe [4].

2.3.6.3 Scanning Speed

If the radar is scanned across the environment, the measurement rate and rotation rate will also impact the bearing resolution; the rotation resolution should be better than the beam width, and the speed should be matched to the measurement rate to achieve the desired overlap between successive measurements. A 50% overlap between consecutive measurements is a good compromise between scan speed and repeated measurements of the same point and by integrating the successive measurements, the cross-range resolution improves by a factor of two [5].

2.3.7 Near and Far Field

The antenna radiation pattern is not uniform along its entire range. The beam is divided into two regions, the **near-field** and **far-field** regions. It needs to be noted that Eq. 2.10 only holds in the antenna's far field. In the **far field**, the beam has constant divergence and its radiated power density is inversely proportional to the square of range [14]. In the **near field**, the antenna's radiation pattern is less predictable, but is often modelled as a cylinder with minimal divergence. The junction between near and far field is not clearly delineated; for convenience, the far field will be defined as beginning at a distance R_{min} (m) given by [14]

$$R_{min} = \frac{D^2}{\lambda} \tag{2.11}$$

Like the beam width, the length of the near field is dependent on the antenna diameter D (m) and the wavelength λ (m), which is directly coupled to the frequency. A higher frequency radar will have a longer near field. This is not crippling but the near-field beam pattern shape may need to be considered.

2.3.8 Frequency Selection

The frequency used for an FMCW radar with a given aperture will have an impact on:

- The beam width;

- The attenuation of radar measurements in dust and fog;

- The detection characteristics of different objects and maximum range of the radar;

- The start of the far-field antenna region (see Sec. 2.3.7);

- The total cost of the system.

A higher frequency will produce a narrower beam from a fixed aperture with a longer near-field region, higher propagation losses (decreasing maximum range) and higher component costs [4].

   This means that small, low-cost, lower-frequency radar systems can have wide beams which offer poor bearing resolution to targets. Nonetheless they can have good detection characteristics and low attenuation.

   Microwave and millimetre-wave frequencies can be used to produce a narrower beam; these frequencies are in the gigahertz band. Radars are not generally built in the terahertz range; atmospheric attenuation becomes extremely high, with attenuations of thousands of dB/km being common [14].

2.3.9   FMCW Operating Principles

Frequency modulated continuous waveform radar operates in the following way.

1. A radio "chirp" of linearly increasing frequency is generated and transmitted.

2. The radio signal is reflected from an object in the environment.

3. The reflected signal is received by the receiver.

4. The received signal is mixed (multiplied) with a stored copy of the transmitted signal.

5. A fast Fourier transform (FFT) is performed on the mixed signal to extract the frequency components.

   The received signal is a time-delayed copy of the transmitted signal, with the delay proportional to the range to the target that it is reflected off. Mixing this with the transmitted signal produces a beat signal, with frequency proportional to the range to the target [5] [26]. The FFT transforms this from a mix of frequencies to a vector of
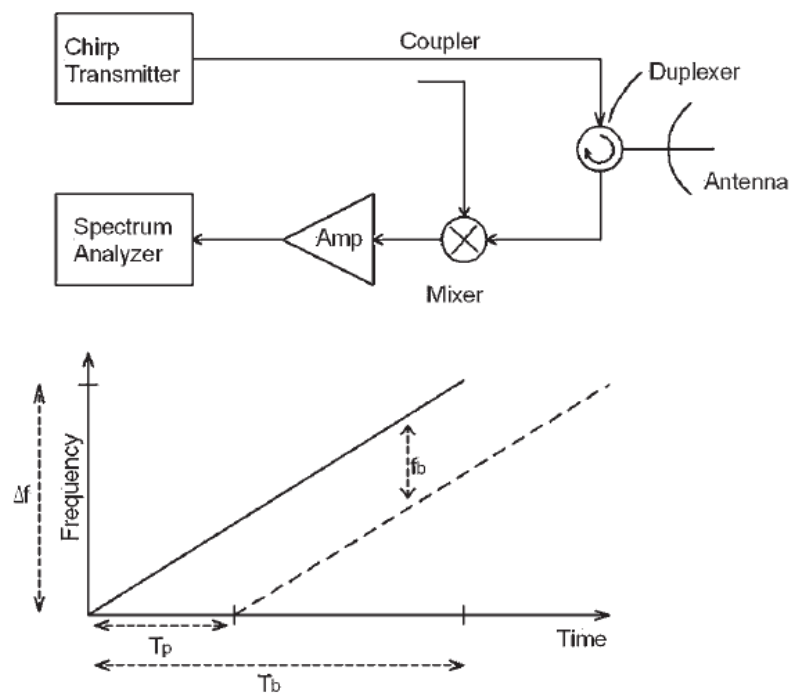
Figure 2.10: Schematic diagram of FMCW radar hardware and operating principle [4]. The signal transmitted and received by the antenna is mixed with a copy of the original transmitted signal, mixed to produce a beat signal, amplified and processed via FFT.

range bins. Each bin contains a power value proportional to the net radar cross-section of the objects in the volume of space at a corresponding distance from the radar within the region defined by the beamwidth.

2.3.9.1   Chirp Generation

An FM radio chirp can be described as

$$\omega_b = A_b t \tag{2.12}$$

where ω_b (Hz) is the change in frequency with time, at a rate of A_b Hz/s.

When combined with a carrier wave signal of ω_c (Hz), the change in frequency is integrated to produce a quadratically increasing phase. The voltage output for a time-varying sinusoidal frequency modulation driven by a voltage-controlled oscillator (VCO) $v_{fm}(t)$ (V) is given by

$$v_{fm}(t) = A_c cos \left[\omega_c t + \frac{A_b}{2} t^2 \right] \tag{2.13}$$

with carrier wave frequency ω_c (Hz), a constant of proportionality or chirp gradient A_b (Hz/s) and time t (s) [4].

This is under ideal circumstances. In reality, the chirp is not linear due to imperfections in the voltage-controlled oscillator that generates it; the chirp gradient A_b is not a constant but a noisy function of time, frequency and voltage. This has the effect of degrading the range resolution after mixing.

2.3.9.2   Propagation and Reflection

The antenna transmits the chirp into the environment. It travels through empty space until it is reflected off an object at range R_{obj} (m). The echo returns to the receiver after a total propagation time T_p (s) from transmitter to object to receiver:

$$T_p = \frac{2R_{obj}}{c} \tag{2.14}$$

This assumes a single reflection off the target. This is not always the case; multipath measurements involve reflections off multiple objects, leading to objects being detected at incorrect ranges. Sec. 2.4.1.5 describes this in greater detail.

2.3.9.3 Mixing

The received signal at time t is mixed (multiplied) with a time-delayed copy of the transmitted signal. With total round-trip propagation time T_p (s), the time-delayed signal is of the form

$$v_{fm}(t - T_p) = A_c cos \left[\omega_c(t - T_p) + \frac{A_b}{2}(t - T_p)^2 \right] \tag{2.15}$$

The mixed signal can be described by

$$v_{fm}(t - T_p)v_{fm}(t) = A_c^2 cos \left[\omega_c t + \frac{A_b}{2}t^2 \right] cos \left[\omega_c(t - T_p) + \frac{A_b}{2}(t - T_p)^2 \right] \tag{2.16}$$

Using the cos-product rule

$$v_{out}(t) = \frac{A_c^2}{2} \left[cos \left(\left(2\omega_c - A_b T_p \right)t + A_b t^2 + \left(\frac{A_b}{2}T_p^2 - \omega_c T_p \right) \right) + cos\{ A_b T_p t + \left(\omega_c T_p - \frac{A_b}{2}T_p^2 \right) \} \right] \tag{2.17}$$

The term $cos\{ \left(2\omega_c - A_b T_p \right)t + A_b t^2 + \left(\frac{A_b}{2}T_p^2 - \omega_c T_p \right) \}$ is a phase-shifted linearly increasing FM chirp at about twice the transmitter frequency that is filtered out.

The term $cos\{ A_b T_p t + \left(\omega_c T_p - \frac{A_b}{2}T_p^2 \right) \}$ describes a beat signal at a fixed frequency.

$$f_b = \frac{A_b}{2\pi}T_p \tag{2.18}$$

That is, the beat frequency f_b (Hz) is directly proportional to the propagation time T_p. Substituting in Eq. 2.14 gives the frequency in terms of the range to the target

$$f_b = \frac{A_b}{2\pi}\frac{2R_{obj}}{c} \tag{2.19}$$

Chirp gradient $\frac{A_b}{2\pi}$ is equal to the chirp bandwidth Δf (Hz) divided by the chirp duration T_b (s), giving

$$f_b = \frac{\Delta f}{T_b}\frac{2R_{obj}}{c} \tag{2.20}$$

Eq. 2.20 is the FMCW range equation.

2.3.9.4 Fast Fourier Transform Processing

A fast Fourier transform (FFT) is used to convert from a mix of beat frequencies to a set of bins containing power values. This takes as input a series of samples of the

beat signal (with a sample frequency at least twice that of the highest expected beat frequency, to comply with the Nyquist sampling theorem).

The number of bins is dependent on the sample rate and the observation time, which is generally equal to the chirp duration. The more bins there are in the FFT, the slower the FFT will be to calculate, but combined with a higher chirp bandwidth, the finer the resolution will be.

The bin length should be no longer than the range resolution, but shorter is better. Longer observations with more samples will result in shorter bins but be more costly computationally.

2.4   Characteristics of Radar Measurements

Meaningful interpretation of radar measurements requires an understanding of the characteristics of radar as a sensor. One reason that radar is less widely used than sensors such as cameras and lidars in intelligent transport applications is that the measurement is more difficult to interpret. Due to the characteristics of the radar beam and the FMCW process, additional processing is required to extract features from the measurement.

A single radar measurement can tell us several things about the observed region, with varying reliability:

- Whether an object is present or not (object existence)

- The location of any detected objects (object position)

- The radar cross-section of any detected objects (object RCS)

- The size of detected objects (object extent)

This information is stochastic in nature; due to finite resolution and measurement noise, there is uncertainty in all aspects of the information that a radar provides. The level of uncertainty is dependent on the radar itself and the object. It is important to understand all the different sources of uncertainty to account for them in radar mapping.

Unfortunately, there are numerous sources of noise and clutter that contaminate the measurement. Some originate within the radar, and some from in the environment. The bins corresponding to empty space will not contain power values of 0 dB, and the bins corresponding to objects will have their values affected by noise. Sources of

false positives are discussed in Section 2.4.1. Detecting an object is a signal processing problem [43] [55]; some technique must be used to distinguish between power values corresponding to objects and power values corresponding to noise or clutter. Different methods have been developed for performing this feature extraction. These are discussed in Section 2.5.

Dynamic range is the ratio between the maximum power and minimum power that the radar is able to detect. This is typically expressed in decibels. A large dynamic range is desirable, as it will allow the radar to be more sensitive to weaker reflectors without noise or clutter interfering with detection as badly, and to detect objects with larger RCS without the power value saturating. A small dynamic range increases the risk of the measurement saturating in the presence of heavy clutter.

2.4.1 Sources of false positives

Every FMCW radar measurement will contain noise from one or more sources. "Noise can be defined as unwanted signals, usually of a random nature, that interfere with the detection or analysis of a signal carrying information." [4]. Clutter is another source of unwanted measurements and can be defined as "background reflectors, undesired...from the standpoint of detection and tracking" [58].

It is important to differentiate between clutter and noise. Noise is typically random and varying in time (phase noise could be considered an exception, as it is related to the presence of strong reflectors and the radar transmitter's characteristics). There are many sources of noise, including thermal noise, speckle noise and phase noise. Noise power levels are largely dependent on the hardware of the radar as well as the strength of targets in the case of phase noise.

Clutter is caused by the radar detecting targets that exist but that are of no interest; typically these are things like the ground and small objects on it. Clutter power levels are dependent on the environment and the grazing angle of the radar; a steeper grazing angle and rougher terrain will produce higher clutter levels.

There are other sources of false detections that affect FMCW radar. The ones considered in this section are multipath, poor chirp linearisation, and the wide radar beam with sidelobes.

- The different types of noise encountered include:

Thermal noise is primarily due to the internals of the radar and is largely a function of temperature and components.

Speckle noise is due to constructive and destructive interference between weak reflectors on the ground. Unlike regular clutter measurements, it fluctuates rapidly. It appears as random fluctuations in the measurements of the ground [4].

Phase noise is caused by the VCO's generation of a spectrum of frequencies rather than a single frequency at any point in the chirp, and appears most visibly in the presence of very strong targets [14].

- **Clutter** is caused by the radar detecting unwanted background reflectors, typically the ground.

- **Multipath measurements** are reflections of the radar beam off two or more objects in the environment, rather than just one.

- **Poor linearisation** of the FMCW chirp can cause targets to smear over multiple bins rather than just one, or even split into two targets [26] [5].

- The **wide radar beam** extends beyond the estimated 3dB width, and has sidelobes either side of the mainlobe. These can inflate the width of particularly strong or nearby targets, causing them to appear wider or even as one strong target with two smaller targets on either side [4].

As well as false positives, false negatives (missed detections) will also occur, due to objects occluding one another or because of RCS interference effects (as described in Section 2.3.2). These are difficult to model.

2.4.1.1   Thermal Noise

Thermal noise is the dominant source of noise. It appears as nonzero power measurements even in the absence of targets (even when the radar is pointed at the sky). This forms a **noise floor**, the minimum power that the radar returns. It is a function of the radar's hardware and temperature.

Thermal noise power P_N (W) is given by [4]

$$P_N = kT_0\beta \tag{2.21}$$

where k is Boltzmann's constant ($1.38 * 10^{-23}$ J/K), T_0 (K) is the system temperature and β (Hz) is the receiver's noise bandwidth. Thermal noise is Gaussian in nature [4]. After being passed through a narrow band filter and detected, a Gaussian probability distribution is transformed to a Rayleigh function [4]. As measurements are converted to decibels, this transforms the Rayleigh function to log-Rayleigh.

2.4.1.2 Speckle Noise

Speckle noise is produced by constructive and destructive interference from a large number of small reflectors in the beam volume. Due to the fluctuating nature of RCS as described in Sec. 2.3.3, some reflections from the environment will be more powerful and some will be less. Even when only very weak reflectors are present in the environment, these can interfere constructively or destructively with one another to cause random targets to be detected [4].

Speckle is distinct from clutter in that it is random and time-varying. Different filtering approaches are used to filter out speckle while retaining true targets. A common and simple approach is multilook filtering, where multiple observations of the same positions are integrated and averaged [1].

2.4.1.3 Phase Noise and Frequency Spurs

Phase noise appears as wide skirts along the range axis around real targets. This degrades range resolution and makes objects appear larger than they are along the radial axis, and reduces sensitivity by lifting the noise floor [1] [5].

Phase noise is generated by the voltage-controlled oscillator that the radar uses to generate a varying-frequency signal. Ideally it would output only one frequency, and by varying the input linearly voltage with time, a linear chirp of frequencies could be generated. Realistically, for any given voltage, the VCO generates a spectrum of frequencies with finite bandwidth. This causes transmitted energy to leak into the mixer and be reflected into the receiver channel due to antenna impedance mismatch [14], manifesting itself as the target spectral widening discussed earlier. Additionally, spurious frequency components are often generated by clock signals that couple onto the VCO control voltage.

Phase noise is very difficult to model and filter out. It can be very noticeable around
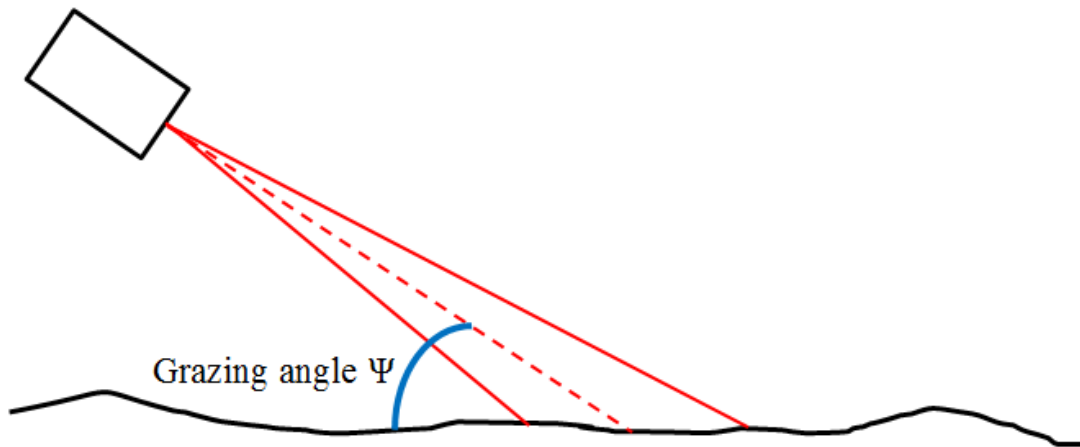
Figure 2.11: Clutter increases at steeper values of grazing angle ψ.

targets with a high RCS, and is often stronger than the measurements of weaker reflectors. It is difficult to come up with a reliable way of discriminating between targets that are "ghosts" caused by phase noise, and targets that appear behind one another.

2.4.1.4   Clutter

Clutter differs from noise in that it is not randomly distributed from nonexistent targets. The targets exist, but they are of no interest to the driver of a large vehicle. For the purpose of this thesis, clutter is typically measurements of the ground or grass. Clutter is strongest in areas where the ground is uneven or hilly, or when the radar is pitched down more steeply towards the ground. These conditions can be common in the environments large vehicles are used in, so filtering clutter is an important issue.

The reflectivity of ground clutter fluctuates, so the mean reflectivity is used. It is modelled using the constant γ model:

$$\sigma^0 = \gamma \, sin\psi \tag{2.22}$$

σ^0 is the reflectivity (cross section per unit area, m^2/m^2), ψ (rad) is the radar grazing angle, and γ describes the scattering effectiveness, which is a function of the roughness of the terrain [4].

The probability distribution function for the combined noise and clutter or noise and target signal is expected to follow the Rice distribution [4]; the expected PDF for the

clutter and the targets is log-Ricean.

2.4.1.5 Multipath

Multipath measurements are reflections off two or more objects in the environment, rather than just one; the beam reflects off the first object, onto the second and then returns to the receiver. This is particularly problematic in environments containing large, flat objects such as walls and smooth ground [1].

Multipath displaces objects from their correct range and bearing, and can cause an object to appear duplicated if some of its reflected power suffers multipath and some does not [5].

2.4.1.6 Poor Chirp Linearisation

FMCW radar frontends will ideally produce a perfectly linear chirp; i.e. the radio frequency will increase at a uniform rate. In reality the produced chirp may be nonlinear. When a nonlinear chirp is mixed with a nonlinear echo signal, the resulting beat signal's frequency will not be proportional to the range to the reflecting object; there will be multiple frequencies instead. Targets appear spread over several bins rather than as narrow peaks in just one. The worse the nonlinearity in frequency, the wider the spread will be. If the linearisation is particularly bad, a single target may split into two peaks [26] [5]. The peaks will be lower due to the power being spread over multiple bins rather than in one. Chirp linearisation is dependent solely on the characteristics of the radar, rather than the environment. This can be adjusted for either pre-transmission, by determining the nonlinearity and either compensating for it by adjusting the voltage supplied to the VCO over the chirp duration [4], or post-reception by processing the echo signal to remove the effects of the nonlinearity [37].

2.4.1.7 Wide Beam and Sidelobes

A radar beam has finite width in both azimuth and elevation, proportional to the aperture diameter and frequency (see Sec. 2.3.6.1). When the azimuth width is projected onto the ground in 2D, the beam is said to have a **footprint** which is approximately trapezoidal in shape, defined by the bin length and beam width. The cross-range width of the footprint increases proportionally to the range.

The beam width limits the maximum detection range of the radar; at longer range, the wider beam is spread over a larger area of potential clutter, increasing potential clutter returns, while a point target will not reflect any more. At a long enough range, the clutter floor will raise above the reflected power of the target.

The antenna's gain drops off the further off-axis the target is. Strongly reflective targets further to the side of the mainlobe than the 3dB limit may still be detected (as is the case with sidelobes) but will be assumed to lie within the mainlobe's beam footprint. Increased range increases the uncertainty in the cross-range position of any detected target.
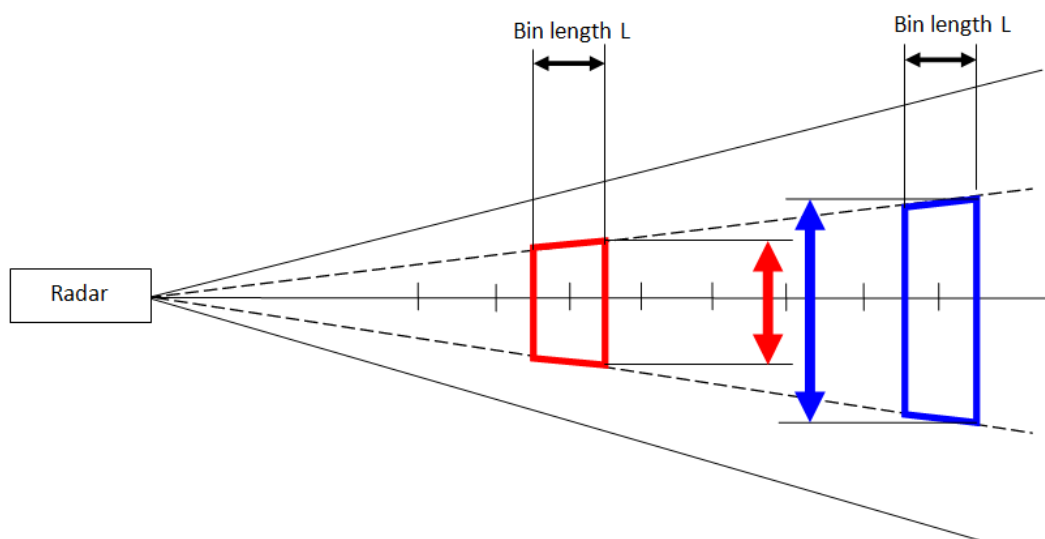
Figure 2.12: The footprints of individual bins are equal in length but have widths proportional to their distance from the radar, when in the far field of the beam.

Sidelobe-detected targets appear as smaller, weaker targets either side of the true target along the normal to the beam direction, or even as the three targets blurring together into one wide target. Combined with phase noise or poor linearisation, a target's extent can appear greatly inflated.
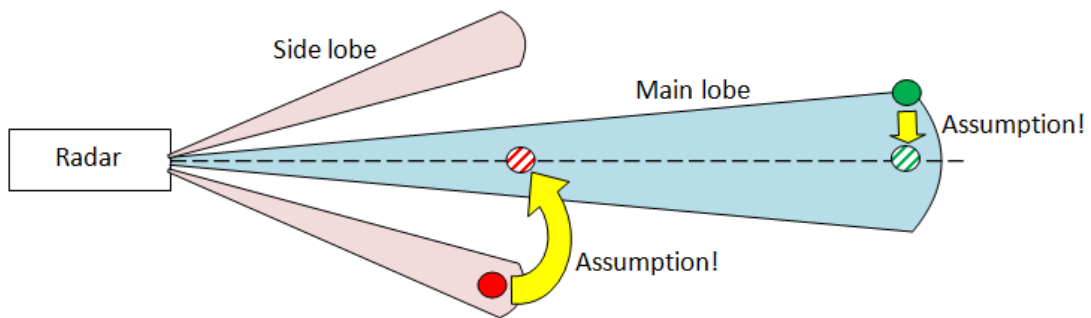
Figure 2.13: A FMCW radar has a main lobe (blue) with non-zero width as well as sidelobes (pink). Measurements are assumed to be of objects lying on the centre axis of the mainlobe.

2.5 Static and Adaptive Detection Filters In Radar

As a FMCW radar only returns a noise-contaminated vector of power values (rather than, for example, a list of the ranges to objects within the beam), further processing is required to actually identify the presence of any objects within the volume of space observed for the measurement. Detecting a target within a measurement is a signal processing problem on the measurement [43] [58]. Some sort of **detection filter** is needed to identify which bins' power values correspond to a detection of an actual object within the beam footprint, and which bins correspond to empty space, noise or clutter. The most common method is to calculate a detection threshold for each bin. Any bin containing a power value above the threshold is considered a detection of a real object; all other bins are treated as containing noise or clutter.

Detection thresholds can be adaptive or static. An **adaptive detection threshold** is calculated as a function of each individual measurement. A **static detection threshold** does not change and is independent of the individual measurements; it is constructed using a priori knowledge of environment and target characteristics. It is fixed and unchanging, and can only work well when the environment and target characteristics are well known and the radar's characteristics are not expected to change. Test data would be gathered using the sensor in typical operating environments with typical targets and this would inform threshold design.

Adaptive detection thresholds are able to respond to varying environments and targets. This reduces the amount of testing and calibration required before deployment

and allows them to respond to changing environments. The most commonly used adaptive threshold method is constant false-alarm rate (CFAR) [58]. As the name suggests, CFAR is designed to give a fixed false-alarm rate, typically 10^{-6}.

There are different CFAR methods of calculating this detection threshold. They typically utilise a sliding window to examine different sections of the measurement in isolation; this reduces the effect of distant targets on one another and the effect of the received power drop-off at longer range. The most popular approach is OS-CFAR, which involves sorting the N power values within the window in ascending order and then selecting one position (e.g.: second or third or fifth) as the threshold and applying a calculated scaler. The window size and exactly which position is chosen as the threshold will influence the resulting false alarm rate; Rohling recommends using a window N between 24 and 32, and a threshold index k between $\frac{N}{2}$ and $\frac{3N}{4}$ [55]. $N - k$ should be at least double the maximum target length to prevent two nearby targets from masking one another's presence [55].
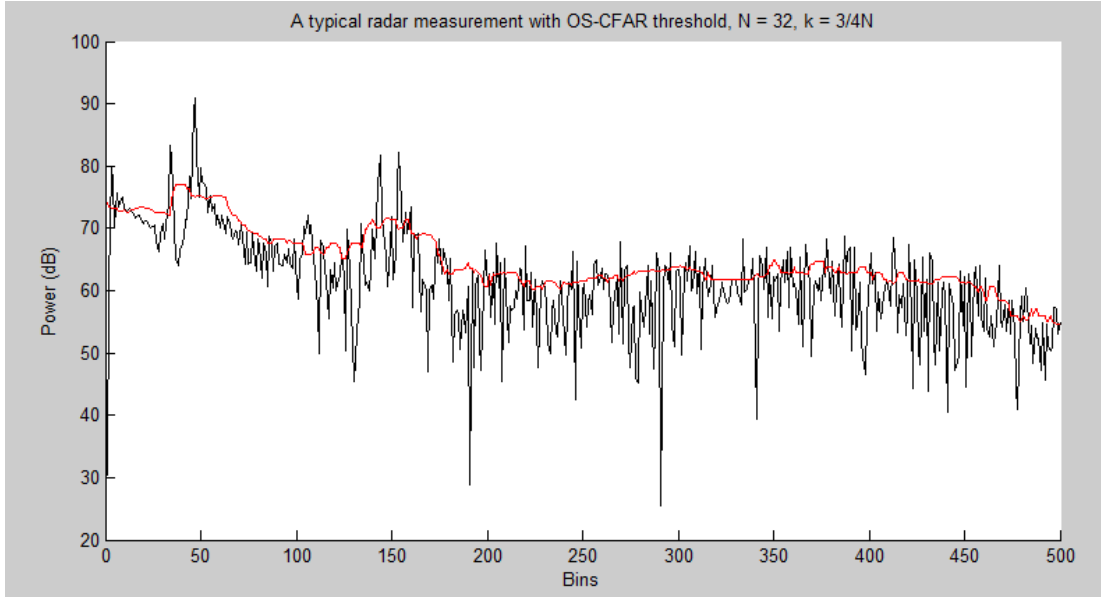
Figure 2.14: A typical radar measurement with an OS-CFAR threshold calculated with a window width N = 32 and threshold index k = 3N/4. The dark line is the measurement power, the pale line is the calculated threshold for each bin.

2.6 Mapping With FMCW Radar

2.6.1 The Mapping Problem

Assuming we have a narrow-beam radar that is scanning across the environment, a single measurement is not enough to reliably warn the driver of all nearby hazards. As well as covering only a fraction of the environment, there is a risk of false alarms or missed detections. By combining multiple measurements, a representation of the wider scene can be obtained, and the risk of false alarms or missed detections drastically reduced.

Mapping is the process of combining sensor measurements together into a spatial model of the environment [1] [67]. It is an important problem in robotics and intelligent transport, being a key component of any navigation, localisation or collision avoidance system [67] [68].

Traditionally, mapping assumes that the sensor's pose is known. **Localisation** is the related problem of determining the pose of a sensor (or some other agent) based on observations and a known map [68]. The **simultaneous mapping and localisation** (SLAM) problem addresses concurrently creating a map and using it to estimate the position of the sensor. Localisation and SLAM are very active fields of research but are beyond the scope of this thesis. For solving the close proximity problem, a global map of the wider environment is not necessary; a map of the local environment relative to the sensor mounted on the vehicle is sufficient.

Probabilistic mapping approaches have become the most popular approaches in recent years. There is noise and uncertainty inherent in all sensor measurements and estimated sensor positions [67]; it is desirable to be able to directly model this uncertainty using statistical and probabilistic techniques [1]. Virtually all these mapping approaches are based on Bayes' rule [67]. This states that the probability of state A given the information B is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2.23}$$

$P(A|B)$ is the probability of state A given the state of information B. $P(B|A)$ is the probability of receiving information B assuming state A. $P(A)$ is the existing probability of A, called the **prior** [67].

This can be written as

$$P(A|B) = \eta P(B|A)P(A) \tag{2.24}$$

where η is a normaliser.

Bayes' rule can be used to recursively fuse consecutive measurements B about some state of the environment to update the estimated state A [67]. Additional assumptions and compromises are made by different mapping approaches for the prior distribution $P(A)$ and sensor model $P(B|A)$. These can generally be divided into grid-based representations and feature-based representations. These approaches are discussed in greater detail in Section 2.6.2 and Section 2.6.3.

2.6.2   Grid Maps

2.6.2.1   Definition

Occupancy grids were first introduced by Moravec and Elfes [41] and while this work has been expanded on (such as in [34] [30]) the basic concept remains largely the same: the environment is divided into a tessellating pattern of identical cells that contain variables that estimate one or more states of the environment (usually occupancy or emptiness) [1, 14, 30, 34, 41, 43, 68].

These variables are updated with sensor measurements, typically using a Bayesian update [34, 41] although other approaches have also been used [43].

There are different implementations of occupancy grid updates. A simple and effective one is the Bayesian method used by Matthies & Elfes [34] although approaches include Dempster-Shafer evidence [50], heuristics [8] and neural networks [66]. Other work focuses on extending the occupancy grid, traditionally used for mapping static environments, to dynamic environments [9] [8]. Foessel [14] creates 3D occupancy grids using millimetre-wave radar, representing probability as the logarithm of odds for "better dynamic range and computational efficiency" [14].

2.6.2.2   Advantages of Grid Mapping for FMCW radar

Occupancy grids are intuitive, efficient and widely used for radar mapping [1, 14, 43].

They explicitly model empty space rather than just extracting small targets from the environment, which allows all the information in the radar measurement to be utilised in performing the update. The beam footprint can be projected onto the grid and every cell it covers (fully or partially) can be updated [14].

There is no need to perform data association between measurements and targets because individual targets are not explicitly modelled, just the location in space [1]. This removes the difficult data association stage from the mapping problem, where measurements must be matched against known features or added as new ones [67]. The occupancy value provides an estimate of whether a target exists or not, modelling the uncertainty in sensor measurements due to noise, clutter or missed measurements [1]. The measurement likelihood model used in the Bayes update can be designed to reflect the expected target RCS distribution, expected levels of clutter or missed detections [14].

2.6.2.3 Disadvantages of Grid Mapping for FMCW radar

The grid size and resolution of an occupancy grid are fixed at initialisation, limiting the extent that they can be used in large environments. Memory requirements are quite large; for an increase in map size or map resolution by a factor of x, the memory required increases proportional to x^2. For a fine grid or a long-range or wide beam, the computation involved to project the beam onto the grid quickly becomes expensive [1].

For a Bayesian update, a measurement likelihood model is required, which must be based on the sensor's measurement properties and will have a large impact on mapping performance [14, 33]. If this measurement likelihood model does not accurately represent the sensor's true characteristics this will impact the accuracy of the map. Creating an accurate measurement likelihood requires calibration or learning in the target environment [1].

Although the occupancy grid effectively represents the uncertainty about a target's existence, it is not effective at representing uncertainty about a target's position. When a target is detected in a radar bin, an update must be performed over all the occupancy grid cells covered by the beam footprint. At longer range, where the beam fans out wider, this can be a fairly large number of cells. This has the effect of "smearing" the probability of change over a number of cells, so an update that would cause a larger probability change in a few cells at close range produces a smaller probability change in many cells at long range [1, 14].

Occupancy grids can develop significant inertia or persistence in the occupancy state of the cells. The standard Bayes update is commutative; that is, the result will be the same regardless of the order of the measurements [8]. Therefore a sequence of 100 measurements of emptiness followed by 100 measurements of occupancy will produce the

same final occupancy estimate as 100 measurements of occupancy and then 100 measurements of emptiness (assuming neither occupancy or emptiness reach the saturation threshold), even though the final state of the cell is quite clearly different. Coupled with the diluting effect of multiple cells being covered at longer range, this can lead to an undesirably slow response to changes in the environment [8]. The sensor model can be tuned to avoid this, or the maximum/minimum probabilities of occupancy can be bounded to speed the response time, but this is an ad-hoc approach.

Transmitting the map for the purpose of cooperative mapping is difficult and requires either iterating through the map to segment and cluster cells together, or transmitting the entire grid. Both of these options can be computationally expensive.

2.6.3   Feature-based Maps

2.6.3.1   Definition

Feature-based maps model the environment as a list of features, "interesting" and recognisable landmarks that can be defined by a simplified model such as a point, line, corner or circle. The map consists of a list of the estimated positions of these features and their defining properties, and often a representation of the uncertainty of the positions (represented as a covariance matrix in an extended Kalman filter). Map management routines are needed to add new features to the map, or identify what measurements correspond to what known targets [1]. Feature-based maps can be implemented in different ways, such as using Gaussian noise-approximations (extended Kalman filters [63] or unscented Kalman filters [68]) or particle filters [67].

2.6.3.2   Advantages

Feature-based maps are much more sparse than occupancy grid maps; they only track detected targets and assume all other space is empty [1] [67]. Features are compressed into simple locations and shapes, which can be processed more efficiently as well as transmitted to other agents for the purpose of cooperative mapping. This requires less memory compared to an occupancy grid of similar resolution, which is why this approach can be more suitable to large, outdoor areas.

The uncertainty in a feature's position can be explicitly modelled by tracking the covariance of the estimate of the feature's position [68]. This is a superior approach to

that used by occupancy grids, as it reflects the uncertainty without diluting the certainty of the feature's existence and allows a faster response if the environment changes.

A prediction step can be implemented efficiently to reflect motion of the sensor or the features, updating position or uncertainty as time passes in the absence of measurements.

2.6.3.3 Disadvantages

A feature extraction algorithm is required to extract targets from the measurement. This is usually heuristic-based and for more complicated models, can be susceptible to clutter or noise [1, 75].

Map management techniques are required to add or remove features, and identify features on the map. Data association can be slow and introduces the potential for significant inaccuracies if it makes a mistake. Data association and other heuristic-based methods used for map management do not guarantee Bayes optimality in the updates. The presence of clutter and the potential for missed observations increases the likelihood of data association errors [1].

In the popular Kalman Filter approach, uncertainty is modelled as being in range and bearing on the measurement, but uncertainty in detection (due to missed detections, clutter, noise or bad data association) is not represented by these [1]. A separate filter can be used to track probability of existence, but this introduces mathematical inconsistencies [1].

2.7 Occupancy Grid Maps with FMCW Radar

Each cell of an occupancy grid contains a value of the estimated probability of occupancy or emptiness. For simplicity, occupancy and emptiness are estimated with a single value. 0 is absolute confidence of emptiness, 1 is absolute confidence of occupancy, and 0.5 is maximum uncertainty. Therefore, with probability of occupancy $P(occ)$ and probability of emptiness $P(emp)$:

$$P(occ) + P(emp) = 1 \tag{2.25}$$

$$P(emp) = 1 - P(occ) \tag{2.26}$$

Using an occupancy grid with occupancy measurement R, Bayes' theorem (Eq. 2.23) can be written as:

$$P(occ|R) = \frac{p(R|occ)P(occ)}{p(R|occ)P(occ) + p(R|emp)P(emp)} \tag{2.27}$$

For numerical stability reasons (as well as to represent the limited capabilities of sensors) the occupancy/emptiness value is bounded to some range within 0 and 1. It is possible to represent occupancy in odds form for improved processing efficiency and numerical stability [14], but not essential.

There is no prior information about the state of the environment, therefore all cells start with the uncertainty occupancy value $P(occ) = 0.5$. An occupancy measurement R is used to update each cell within the measurement footprint. Substituting Eq. 2.26 into Eq. 2.27 allows a posterior over the cell occupancy to then be calculated from the previous occupancy estimate and the newest measurement [67].

$$P(occ|R) = \frac{p(R|occ)P(occ)}{p(R|occ)P(occ) + (1 - p(R|occ))(1 - P(occ))} \tag{2.28}$$

$p(R|occ)$ is determined by the **sensor model** or **perceptual model** [67]. It is a function that converts from a sensor measurement to a probability of a certain map state (in this case, the probability of occupancy). Different sensors in different environments will have different sensor models. They can be learned or generated manually.

Occupancy grid mapping with radar is typically reliant on several assumptions:

- Radar cross-section distribution is independently distributed among all occupancy grid cells [34] [30] [14]. If they were assumed to be dependent on one another, the update would quickly become non-tractable for any but the smallest grid [14].

- One or no target is assumed to be detected per bin per measurement [50]. At close range this is a reasonable assumption, but at longer range the beam spreads out to cover a larger number of cells over several square metres, making this assumption less reasonable.

- A single object is present per grid cell [30].

- The environment is mostly static. This removes the need to model grid cell state transitions over time other than through measurement updates.

2.7.1 Sensor Modelling with Radar

In Eq. 2.27 the sensor model is given as $p(R|occ)$, the probability of receiving a measurement R given that the cell is occupied. A more generalised sensor model is a probability function of the form $p(R = R_i|s)$, the probability of the sensor returning measurement $R = R_i$ given that the environment has property s [30].

For a typical FMCW radar, the sensor model will convert a power value in decibels to a probability somewhere between 0 and 1. It will need to be applied for each bin of each measurement. Radar is a difficult sensor to model compared to many range-bearing sensors. Factors to consider include the beamwidth, beam sidelobes and range resolution; the type, distribution and radar cross-section of targets and clutter; the noise level; and more characteristics described in Section 2.4 [14, 43, 42]. Some of these factors are known or easily measurable, such as the beamwidth and range resolution. Others, such as the environmental RCS distribution, need to be estimated. Foessel[14] examines different distributions for the radar cross-section of objects in the scene, considers the effect of occlusions and wide beams with multiple sidelobes, and produces a geometric representation of the beam that is adopted in this thesis. Although Foessel discusses "the importance of a good characterization of the RCS range for the objects present in a scene", the characterisation process is not described and detection probabilities are not considered [14].

Mullane et al. emphasise the signal processing and target detection problem [43, 42] and use a measurement likelihood approach involving a particle-based occupancy filter to create 2D maps, but do not use a sensor model. Detection likelihoods are unknown and a method based on a particle filter is used to estimate occupancy [42].

Brooker [5] performs a thorough analysis of the characteristics of environmental clutter and noise characteristics, as well as considerations for building images, such as speckle and polar to Cartesian transformation. However, the main focus of the work is on imaging rather than occupancy mapping.

2.8   Feature-Based Mapping with the GM-EKF-PHD filter for FMCW Radar

The probability hypothesis density filter (PHD filter) has been growing in popularity for use in radar tracking and mapping in recent years [1]. It is a feature-based approach able to represent both the uncertainty of object existence, and the uncertainty of object position, for an unknown number of objects being tracked in a dynamic environment.

2.8.1   Finite Set Statistics

Occupancy grids, Kalman filters and particle filters are all implementations of Bayes filters. The Bayes filter is an effective means of tracking a known number of targets but is generally computationally intractable without specific assumptions (such as a limited state-space in a grid, or linear Gaussian systems in a Kalman filter, or approximating the distributions using particle filters [75]). They all rely on having a known number of targets, either through a fixed number of states in a grid, or data association in the feature-based methods [1, 44]. The map and the measurement are represented as vectors and both the size of the vector and the sequence of the values are important (so the first observation corresponds to the first map feature, the second observation corresponds to the the second map feature, and so on). That is, for a map M_v and observations Z_v represented as vectors, with four objects each in the map and observation:

$$M_v = [m_1, m_2, m_3, m_4]^T \tag{2.29}$$

$$Z_v = [z_1, z_2, z_3, z_4]^T \tag{2.30}$$

If the size or sequence of these vectors are not the same, data association is necessary before the update can take place [1, 44].

An alternative representation is to use random finite sets, where the size and sequence of the sets are treated as random variables in a multitarget Bayes filter. This approach uses the finite set statistics (FISST) developed by Mahler [33]. In a random finite set (RFS), the length and sequence of the map and observation sets do not need to be the same [1, 44]. That is, for a map M_s and observations Z_s represented as random

finite sets with four objects each in the map and observation sets:

$$M_s = \{m_1, m_2, m_3, m_4\} = \{m_2, m_1, m_3, m_4\} = ... \tag{2.31}$$

$$Z_s = \{z_1, z_2, z_3, z_4\} = \{z_4, z_2, z_1, z_3\} = ... \tag{2.32}$$

In a typical Bayes implementation, the order of map features and observations matters, requiring an expensive data association step and steps such as χ^2 tests to match observations to known map features or exclude poor-quality observations.

FISST processing is unaffected when the two sets are of different lengths, as would be expected from missed detections (more map features than observations) or clutter (more observations than map features). All sensor measurements are bundled into one meta-sensor (but still retain their individual characteristics), and all targets into one meta-target so that they may be solved through techniques analogous to those used in single-sensor single-target problems [33].

Like the single target Bayes filter, the multitarget Bayes filter is computationally intractable [75] but can be approximated using the probability hypothesis density (PHD) filter. The probability hypothesis density is the first moment of the multitarget Bayes filter [1, 75] (in the same way that the mean is the first moment of the single-target Bayes function [75]). The integral of the PHD function over a region of space is the expected number of features within that region, and the relative peaks are the locations with the highest likelihood of feature existence.

Due to its basis in FISST, the PHD filter can offer several advantages over both the grid-based and feature-based mapping approaches for radar mapping in dynamic environments:

- Explicit representation of detection probabilities, clutter and noise [75].

- Representation of uncertainty in both target position and target existence [1].

- Minimal use of additional heuristics; there is no need for data association [74].

- Compact map representation to minimise memory use and allow it to be used in large areas and shared between agents.

There is still a need for some map management (initialisation, merging and pruning of targets) but simple methods are effective for this [74].
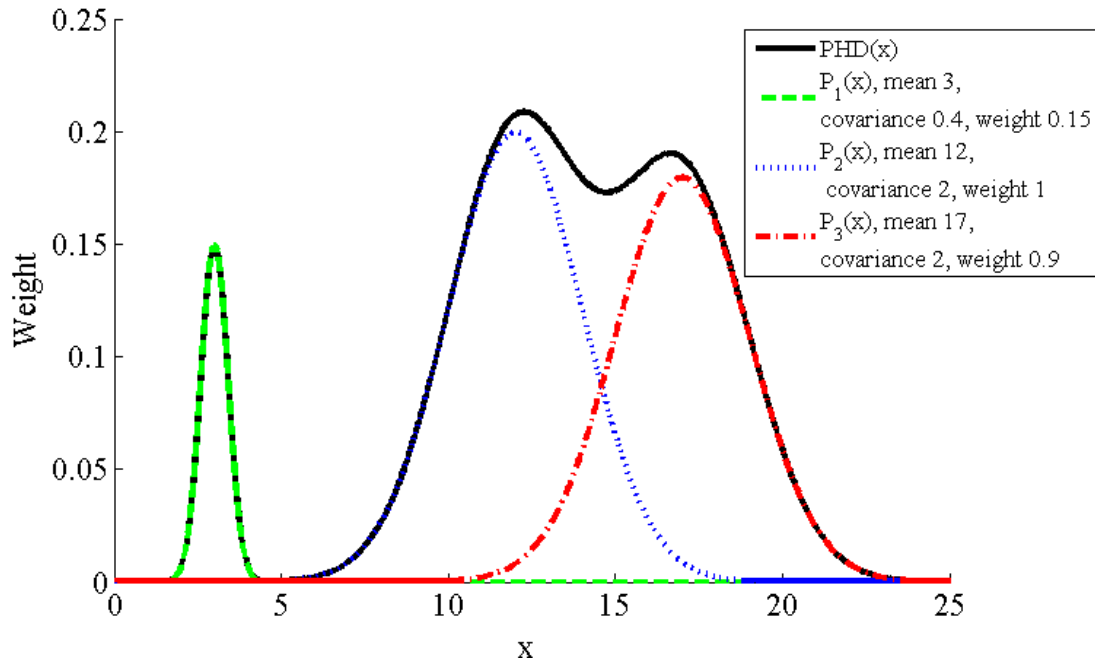
Figure 2.15: A plot of probability hypothesis density versus x for simulated one-dimensional measurements, and its Gaussian mixture components. The mixture is composed of three different Gaussians, with different means, covariances and applied weighting. The integral of the PHD is the expected number of targets, calculated from the area of the integral of the PHD (or simply summing the weights) - in this case, 2.05. The locations of the peaks are the most likely locations of the targets. In this plot, the targets at x-values of 12 and 17 are most likely to exist.

A simulated probability hypothesis density is plotted in Fig. 2.15 and illustrates some of the characteristics of the PHD filter. Each Gaussian in the mixture has three components :

- A **mean**, the centre of the distribution, where an object is most likely to be located.

- A **covariance**, which controls the width or spread of the distribution, showing the distribution of possible locations of an object. A low covariance can be said to represent a high confidence in the object's position as being near the mean; a high covariance, which spreads the distribution out, shows higher uncertainty.

- A **weight**, indicating the "strength" of the estimate of the object's existence.

Existential uncertainty is represented by the weight, which strongly influences the height

of the peaks. Position uncertainty is represented by the width of the covariance, which influences the width and height of the peaks. The integral of this PHD is equal to 2.05, but fractional targets cannot exist, so an assumption must be made about the true number of targets present. There are three relative peaks; the simplest approach is to take the two highest peaks. However, these have a relatively high covariance and from their shapes and positions, it is possible that two of them (red and blue) form one large target. More measurements are needed to clarify this. Additionally, identifying the highest peaks is a non-trivial task; the effects of weights, distribution covariances and the summation of multiple components must be considered. In a more realistic map, the distributions would be at least two dimensional, and it would be possible to add a third dimension in space or velocity component.

The cardinalised probability hypothesis density (CPHD) filter estimates the distribution on the number of features in addition to the probability hypothesis density [1]. In this thesis, the simpler PHD filter will be examined.

2.8.2 The Gaussian Mixture Probability Hypothesis Density Filter (GM-PHD)

The PHD filter itself is computationally intractable but approximations exist, using sequential Monte Carlo methods (SMC-PHD) [72] and Gaussian mixtures (GM-PHD) [74]. Both have been shown to converge as the number of particles or Gaussian components trends to infinity [1]. The SMC-PHD can handle more nonlinear problems, but the GM-PHD is more efficient and is the one this thesis will focus on. GM-PHD uses lists of weights, target state estimates and covariance matrices to represent the intensity function as a sum of Gaussians, as shown in Fig. 2.15. As Wood [75] points out, if we neglect the underlying mathematical machinery derived by Mahler, the implementation is very similar to an array of independent Kalman filters.

There are a six key assumptions made in the derivation of GM-PHD filter[74]:

- Each target moves independently of one another and generates observations independently of one another.

- Clutter is Poisson distributed and independent of target-originated measurements

- The predicted multi-target random finite set governed by density $p_{k|k-1}$ is Poisson distributed.

- Targets follow a linear Gaussian dynamical model such that, given a previous target state of ζ, the probability density of transitioning to a new target state x is given by

$$f_{k|k-1}(x|\zeta) = \mathcal{N}(x; F_{k-1}\zeta, Q_{k-1}) \tag{2.33}$$

  For some linear motion matrix F and covariance matrix Q; and the sensor has a linear Gaussian measurement model such that, given a target state of x, the probability density of receiving observation $z \in Z_k$ is given by

$$g_k(z|x) = \mathcal{N}(z; H_k x, R_k) \tag{2.34}$$

  for some linear observation model H_k and measurement covariance R_k.

- The target survival probability is state-independent, as is the target detection probability.

$$p_S(x) = p_S \tag{2.35}$$

$$p_D(x) = p_D \tag{2.36}$$

- The intensities of the birth RFS and spawn RFS are Gaussian mixtures.

Some of these assumptions can be relaxed. The state-independent survival and detection probabilities can be extended to be state-dependent, which can be useful in environments where objects can occlude one another; an object that temporarily moves behind another should probably not be treated with the same detection probability as one in plain sight. A way of calculating a state-dependent probability of occlusion is shown in [20]. The linear observation and motion assumptions can be relaxed to non-linear models:

$$x_k = \phi(x_{k-1}, u_{k-1}) \tag{2.37}$$

$$z_k = h_k(x_k, \epsilon_k) \tag{2.38}$$

Where u_{k-1} is the dynamic model noise and is a zero-mean Gaussian with covariance Q_{k-1}. ϵ_k is the measurement noise, a zero-mean Gaussian with covariance R_k. These allow the use of the GM-PHD algorithm in a much wider range of applications, but require the use of an extended Kalman filter implementation.

Full algorithms for GM-PHD filters are given in [1, 44, 74]. The algorithms described below are the GM-PHD (Kalman filter) and EK-GM-PHD (extended Kalman filter) implementations, for linear and nonlinear target models, from [74].

2.8.2.1   Prediction Step

GM-PHD uses a linear dynamic prediction model with the potential for additional terms for target survival and new target birth or spawning [75]. New targets can appear by spontaneous birth, by spawning from existing targets, or due to clutter measurements. The multi-target state $X_{k|k-1}$ at time k is given by:

$$X_{k|k-1} = \left(\bigcup_{\zeta \in X_{k-1}} S_{k|k-1}(\zeta) \right) \cup \left(\bigcup_{\zeta \in X_{k-1}} B_{k|k-1}(\zeta) \right) \cup \Gamma_k \qquad (2.39)$$

Where $S_{k|k-1}(\zeta)$ is the single target dynamical model and includes target survival probability. $B_{k|k-1}(\zeta)$ is the set of new targets spawned from existing targets with previous states ζ (**spawned** targets are those that appear by splitting off another target that existed in the previous timestep), including a dynamical model. Γ_k is the set of newly birthed targets (**birthed** targets are those that appear spontaneously). $B_{k|k-1}(\zeta)$ can be neglected entirely in maps where objects are assumed to be static but the sensor is moved; all new targets can be treated as birthed. The forms of $B_{k|k-1}(\zeta)$ and Γ_k are problem dependent.

If the posterior intensity v at time $k-1$ is a Gaussian mixture

$$v_{k-1}(x) = \sum_{i=1}^{J_{k-1}} w_{k-1}^i \mathcal{N}(x; m_{k-1}^i, P_{k-1}^i) \qquad (2.40)$$

Then the predicted intensity for time k is a Gaussian mixture

$$v_{k|k-1} = v_{S,k|k-1}(x) + v_{\beta,k|k-1}(x) + \gamma_k(x) \qquad (2.41)$$

The predicted intensity is the union of the set of previously tracked targets that have survived, and the targets that have just been spawned or birthed.

$v_{S,k|k-1}$ in Eq. 2.41 is the intensity of $S_{k|k-1}$, the RFS of targets that survived from

the previous timestep.

$$v_{S,k|k-1} = p_S \sum_{j=1}^{J_{k-1}} w_{k-1}^j \mathcal{N}(x; m_{S,k|k-1}^j, P_{S,k|k-1}^j) \tag{2.42}$$

For the targets surviving from time $k-1$, J_{k-1} is the number of targets that survived to timestep k, p_S is the probability of target survival, w_{k-1} is the weights of the targets, $m_{S,k|k-1}$ is the target state estimates (and location of peaks of the intensity), and $P_{S,k|k-1}$ is the target state covariances.

$$m_{S,k|k-1}^j = F_{k-1} m_{k-1}^j + u_k \tag{2.43}$$

where F is the target motion matrix and u_k is zero-mean Gaussian noise.

$$P_{S,k|k-1}^j = Q_{k-1} + F_{k-1} P_{k-1}^j F_{k-1}^T \tag{2.44}$$

Where Q is the surviving target motion covariance.

$v_{\beta,k|k-1}$ in Eq. 2.41 is the intensity of the RFS of newly spawned targets $B_{k|k-1}(\zeta)$.

$$v_{\beta,k|k-1}(x) = \sum_{j=1}^{J_{k-1}} \sum_{l=1}^{J_{\beta,k}} w_{k-1}^j w_{\beta,k}^l \mathcal{N}(x; m_{\beta,k|k-1}^{j,l}, P_{\beta,k|k-1}^{j,l}) \tag{2.45}$$

For the targets spawned at time k, $J_{\beta,k}$ is the number of targets spawned, $w_{\beta,k}$ is the weights of the targets, $m_{\beta,k|k-1}$ is the target state estimates (and locations of the peaks of the intensity), and $P_{\beta,k|k-1}$ is the target state covariances.

$$m_{\beta,k|k-1}^{j,l} = F_{\beta,k-1}^l m_{k-1}^j + d_{\beta,k-1}^l \tag{2.46}$$

Where F_β is the spawned target motion matrix, which has offset d_β.

$$P_{\beta,k|k-1}^j = Q_{\beta,k-1}^l + F_{\beta,k-1} P_{\beta,k-1}^{j,l} (F_{\beta,k-1}^{(l)})^T \tag{2.47}$$

Where Q_β is the spawned target motion covariance.

EK-GM-PHD uses the same initialisation of birthed and spawned targets, but a non-linear prediction model for existing targets.

$$m^{(i)}_{k|k-1} = \phi_k(m^{j}_{k-1}, 0) \tag{2.48}$$

Weights are updated in the same way, but the covariance update relies on the calculation of Jacobians of the motion model with respect to the state and noise, to linearise the prediction model about that point.

$$P^{(i)}_{k|k-1} = G^{(j)}_{k-1} Q_{k-1} \left[G^{(j)}_{k-1} \right]^T + F^{(j)}_{k-1} P_{k-1} \left[F^{(j)}_{k-1} \right]^T \tag{2.49}$$

Where the Jacobians $F^{(j)}_{k-1}$ and $G^{(j)}_{k-1}$ are:

$$F^{(j)}_{k-1} = \left. \frac{\partial \phi_k(x_{k-1}, 0)}{\partial x_{k-1}} \right|_{x_{k-1}=m^{(j)}_{k-1}} \tag{2.50}$$

$$G^{(j)}_{k-1} = \left. \frac{\partial \phi_k(m^{(j)}_{k-1}, u_{k-1})}{\partial u_{k-1}} \right|_{u_{k-1}=0} \tag{2.51}$$

As Eq. 2.45 shows, the $J_{\beta,k}$ spawned targets are treated as all being spawned from all J_{k-1} survived targets. This creates a total of $J_{\beta,k} * J_{k-1}$ targets in $v_{\beta,k|k-1}$. These duplicated spawned targets will need to be removed in the merging/pruning step.

$\gamma_k(x)$ in Eq. 2.41 is the intensity of Γ_k, the random finite set of birthed targets at time k.

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} = w^{(i)}_{\gamma,k} \mathcal{N}(x; m^{(i)}_{\gamma,k}, P^{(i)}_{\gamma,k}) \tag{2.52}$$

For the targets birthed at time k, $J_{\gamma,k}$ is the number of targets spawned, $w_{\gamma,k}$ are the weights of the birthed targets, $m_{\gamma,k}$ are the means of the birthed targets, and $P_{\gamma,k}$ are the birth covariances.

The mean predicted number of targets is the sum of all the weights, and will be given by

$$\hat{N}_{k|k-1} = \hat{N}_{k-1} \left(p_{S,k} + \sum_{j=1}^{J_{\beta,k}} w^{(j)}_{\beta,k} \right) + \sum_{j=1}^{J_{\gamma,k}} w^{(j)}_{\gamma,k} \tag{2.53}$$

All the targets surviving from timestep k have their weights reduced by the survival probability $p_{S,k}$, and are treated as potentially spawning $J_{\beta,k}$ targets, in addition to $J_{\gamma,k}$ targets that are birthed independently.

There will be $\hat{N}_{k-1} + \hat{N}_{k-1} * J_{\beta,k} + J_{\gamma,k}$ Gaussian components in the mixture after prediction.

2.8.2.2 Update Step

The sensor generates a random finite set of measurements Z_k at time k. These consist of measurements generated from target set X_k in the environment and noise or clutter-generated measurements. Targets have a probability of being detected $p_{D,k}$ or not detected $1 - p_{D,k}$. If detected, the target will generate a random finite set Θ_k where

$$\Theta_k(x_k) = \begin{cases} z_k & : x_k \text{ is detected} \\ \emptyset & : x_k \text{ is not detected} \end{cases} \tag{2.54}$$

Measurements due to noise and clutter form the RFS K_k, which has intensity κ_k. The form of K_k is problem-dependent. The measurement set Z_k is therefore the union of these sets:

$$Z_k = \left[\bigcup_{x \in X_k} \Theta_k(x_k) \right] \cup K_k \tag{2.55}$$

The measurements $z_k \in Z_k$ can then be used to update the predictions of target states. Given that $v_{k|k-1}$, the predicted intensity at time k, is of the form:

$$v_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^i) \tag{2.56}$$

Then $v_{k|k}$, the posterior intensity at k after updating, is:

$$v_k(x) = (1 - p_{D,k})v_{k|k-1}(x) + \sum_{z \in Z_k} v_{D,k}(x; z) \tag{2.57}$$

The posterior is composed of the set of targets that were predicted but unobserved, and the set of predictions updated with sensor measurements. The unobserved targets are a duplicate of the predicted targets with their weights reduced; if a target goes unobserved enough times, it will be pruned out in the merging and pruning step. The mean and the covariance of the unobserved targets are unchanged from prediction.

If a target is observed, the merging-and-pruning step should merge the duplicate created in the non-observed set into the version that is updated with sensor observations.

The detected targets' intensity is represented in $v_{D,k}$, which is given by

$$v_{D,k}(x;z) = \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^j \mathcal{N}(x;m_{k|k-1}^j,P_{k|k-1}^j) \tag{2.58}$$

The update of the mean and covariance follows the standard Kalman filter update.

$$m_{k|k}^j(z) = m_{k|k-1}^j + K_k^j(z - H_k m_{k|k-1}^j) \tag{2.59}$$

$$P_{k|k}^j = [I - K_k^j H_k]P_{k|k-1}^j \tag{2.60}$$

$$K_k^j = P_{k|k-1}^j H_k^T (H_k P_{k|k-1}^j H_k^T + R_k)^{-1} \tag{2.61}$$

For EK-GM-PHD, the observation model is nonlinear and once more Jacobians are required.

$$m_{k|k}^j(z) = m_{k|k-1}^j + K_k^j(z - h_k(m_{k|k-1}^j,0)) \tag{2.62}$$

$$S_k^j = U_k^j R_k \left[U_k^j\right]^T + H_k^j P_{k|k-1}^j \left[H_k^j\right]^T \tag{2.63}$$

$$K_k^j = P_{k|k-1}^j \left[H_k^j\right]^T \left[S_k^j\right]^{-1} \tag{2.64}$$

$$P_{k|k}^j = [I - K_k^j H_k^j]P_{k|k-1}^j \tag{2.65}$$

A standard Kalman filter has a data association step to match each measurement to the target most likely to have generated it. The PHD filter updates every predicted target with every measurement, generating $|Z_k| * J_{k|k-1}$ new updated targets. The updated targets have their weights recalculated based on how well the measurement matches the prediction, as well as the clutter distribution. This is illustrated in Fig. 2.16, Fig. 2.17 and Fig. 2.18.

(a) Track initialisation   (b) Objects being re-observed   (c) Data association
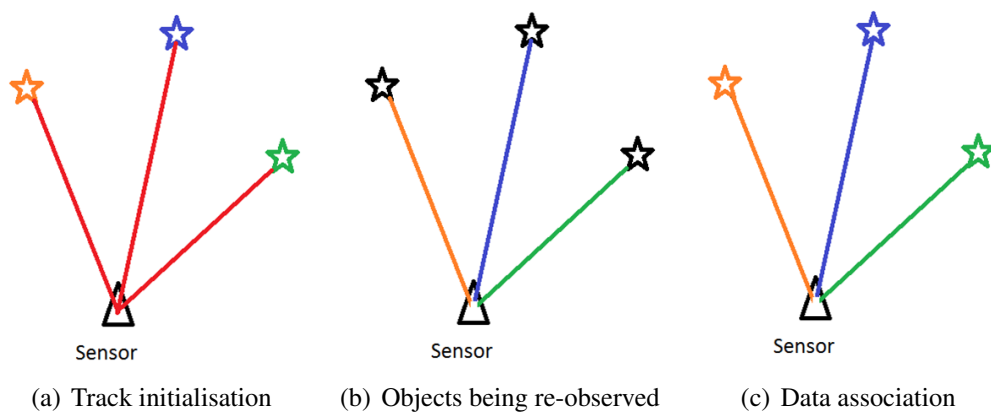
Figure 2.16: This diagram shows a sensor detecting three objects in the environment. These are point objects, returning a single measurement each. In (a), these objects are first observed and the measurements are used to initialise tracks of these objects. (b) shows these objects being re-observed. (c) illustrates how an update would be performed in a filter that relied on data association: each measurement would be matched to the tracked object best matching it (matching measurement colour to object colour), and would be used to update it.
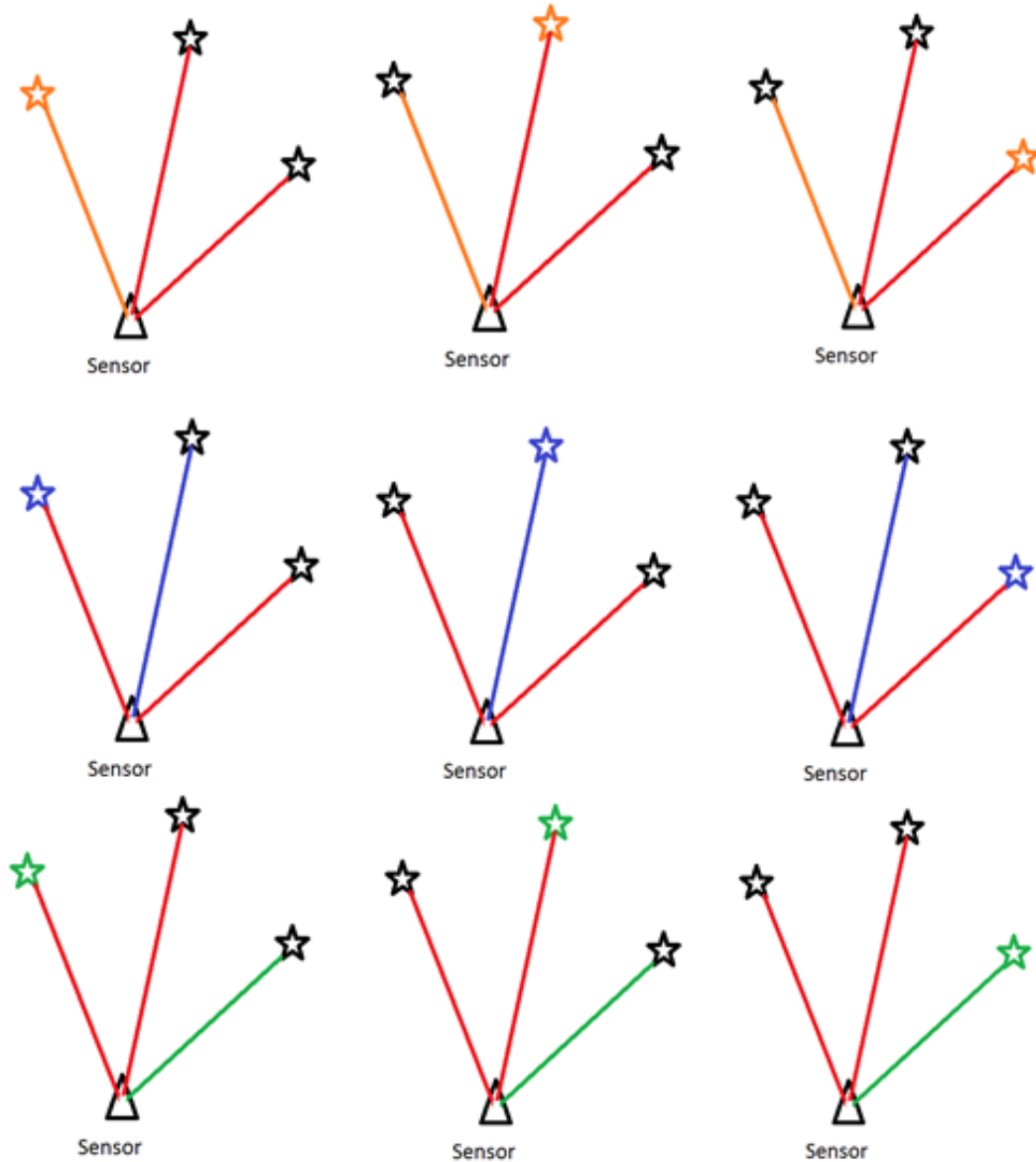
Figure 2.17: In the PHD filter update, every measurement is used to update the esti-
mated position of every object being tracked. In this diagram, the uniquely
coloured measurement in each row (orange, blue or green) is used to up-
date the estimate of each tracked object (with one object in each column
coloured with the unique colour of that row as it is updated). For $|Z_k|$ mea-
surements updating N_k objects, this generates a total of $|Z_k| * N_k$ updated
objects. Most of these will be the result of incorrect measurement-object
matchings.
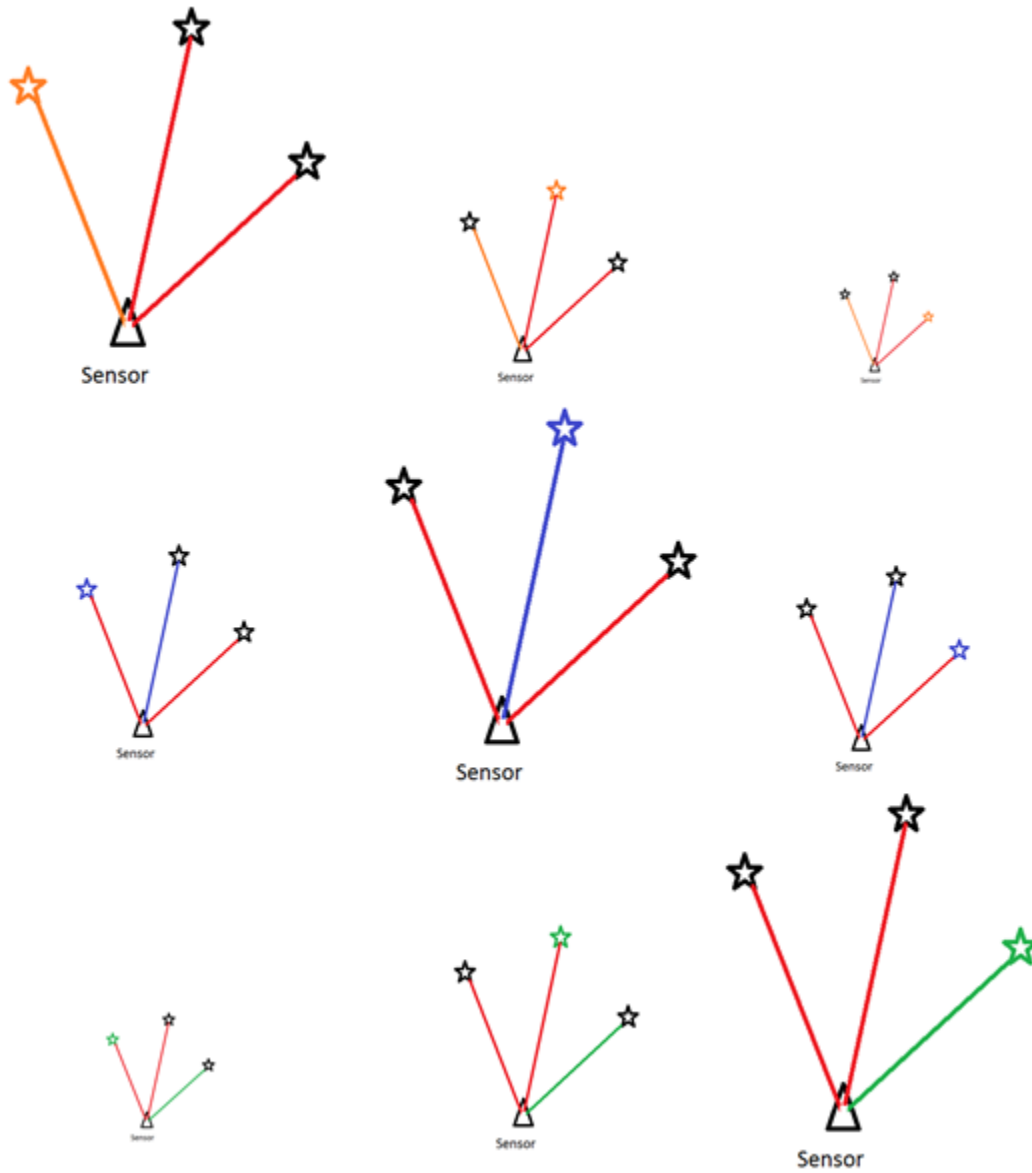
Figure 2.18: After the update of the object state estimates, the weights of the objects are recalculated based on the linear Gaussian measurement model. Objects and measurements that match well will have higher weights than objects with a low likelihood of generating the measurement. The weights are symbolised by the size of the diagram - large-weight objects have larger diagrams.

The new weights w_k of the updated targets, given by

$$w_k^j(z) = \frac{p_{D,k} w_{k|k-1}^j q_k^j(z)}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^l q_k^l(z)} \tag{2.66}$$

Where q_k is the term that uses the sensor model to update the weight.

$$q_k^j(z) = N(z; H_k m_{k|k-1}^j, R_k + H_k P_{k|k-1}^{(j)} H_k^T) \tag{2.67}$$

Where H_k is the sensor measurement model, and R_k is the measurement model co-variance (noise covariance model). The adjusted weight is simply the probability of obtaining sample z from the normal distribution centred on the sensor model's expected measurement of target $m_{k|k-1}^j$ and with a covariance based on the sensor model and noise covariances. This is illustrated in Fig. 2.18, where the targets on the diagonal are being updated with the correct measurements and maintain high weights. The off-diagonal targets, with mismatches between known targets and new measurements, are weighted down.

κ_k is the intensity of the clutter random finite set K_k at time k. Its form is depen-dent on the implementation; commonly it is treated as uniform across the environment. Its effect is to adjust the newly calculated weight depending on how prevalent clutter measurements are; in a high-clutter environment, it acts as a brake on the increase of target weight as there is less confidence in the authenticity of the measurement. In a low-clutter environment, this effect is reduced and weights increase faster. Therefore κ_k can be said to have a similar sort of effect on weight w_k to what large R_k has on $m_{k|k}$.

The mean updated number of targets is once more the sum of all the weights, and will be given by

$$\hat{N}_k = \hat{N}_{k|k-1}(1 - p_{D,k}) + \sum_{z \in Z_k} \sum_{j=1}^{J_{k|k-1}} w_k^j(z) \tag{2.68}$$

This is the sum of the weights of all the predicted targets treated as unobserved, and the weights of all the updated targets formed by matching all observations to all predicted targets. There will be $J_{k|k-1} + |Z_k| J_{k|k-1}$ Gaussian components to the mixture.

2.8.2.3   Merging and Pruning, Target Extraction and Track Labelling

There are combinatorial explosions of the number of Gaussian components in both the prediction and the update steps of the GM-PHD filter. This is illustrated in Fig. 2.17, where 3 targets are updated with 3 new measurements, creating 9 updated targets. There are also 3 non-updated targets (not shown in Fig. 2.17) to account for the possibility of missed detections. This is an increase from 3 known targets to a potential 12. The results after reweighting are shown in Fig. 2.18.

As [74] points out, at time k the required number of Gaussian components will be

$$(J_{k-1}(1 + J_{\beta,k}) + J_{\gamma,k})(1 + |Z_k|) \tag{2.69}$$

Therefore some sort of pruning is necessary for this to be computationally feasible. [74] gives a simple merging and pruning algorithm where the weights below a certain minimum threshold T are discarded, and the remainder are merged where they fall within a certain Mahalanobis distance threshold. Heuristic values for the pruning and merging thresholds are required, and these will vary between implementations.

The merging and pruning algorithm requires as inputs:

- weights w_k^i

- mean states m_k^i

- covariance matrices P_k^i

- a minimum weight threshold T

- a merging threshold U

for $i = 1..J_k$, where J_k is the number of Gaussian mixture components.

The algorithm produces outputs:

- merged weights $\tilde{w}_k^i$

- merged mean states $\tilde{m}_k^i$

- merged covariances covariances $\tilde{P}_k^i$

for $i = 1..l$, where l is the number of merged and prune mixture components (so $l \leq J_k$).

The algorithm is as follows [74]:

$l \leftarrow 0$

$I \leftarrow \left\{ i = 1, \ldots J_k | w_k^{(j)} > T \right\}$

repeat

    $l \leftarrow l + 1$

    $j \leftarrow argmax w_k^{(i)} \text{for } i \in I$

    $L \leftarrow \left\{ i \in I \left(m_k^{(i)} - m_k^{(j)} \right)^T \left(P_k^{(i)} \right)^{-1} \left(m_k^{(i)} - m_k^{(j)} \right) \leq U \right\}$

    $\tilde{w}_k^{(l)} \leftarrow \sum_{i \in L} w_k^{(i)}$

    $\tilde{m}_k^{(l)} \leftarrow \frac{1}{\tilde{w}_k^{(l)}} \sum_{i \in L} w_k^{(i)} m_k^{(i)}$

    $\tilde{P}_k^{(l)} \leftarrow \frac{1}{\tilde{w}_k^{(l)}} \sum_{i \in L} w_k^{(i)} \left(P_k^{(i)} + \left(\tilde{m}_k^{(l)} - m)k^{(}i \right) \right) \left(\tilde{m}_k^{(l)} - m)k^{(}i) \right)^T \right)$

    $I \leftarrow I \backslash L$

until $I = \emptyset$

return $\left\{ \tilde{w}_k^{(i)}, \tilde{m}_k^{(i)}, \tilde{P}_k^{(i)} \right\}_{i=1}^{l}$

Assuming that the minimum weight threshold T were well chosen, the off-diagonal estimates in Fig. 2.18 would be pruned out, leaving only the targets that had been updated with the correct measurements. Pruning is not illustrated here, but would be performed if there were two very close targets.

After updating, pruning and merging, the states of the targets must then be extracted from the Gaussian mixture to be output to the user. The means of the Gaussian components provide the positions of the local peaks, but the height of each peak is a function of the weights, covariances and position relative to other peaks. The simplest method is to assume that covariances are small, and that there is a reasonable separation between targets after the nearby ones have been merged together (so superposition effects from other Gaussians will be low). Simply selecting the targets with highest weights and extracting these should produce satisfactory results. [74] uses a weight threshold of 0.5 as the extraction threshold; the target is more likely to exist than not.

The algorithm to extract presumed-existant targets $\tilde{X}_k$ is as follows[74]:

$\tilde{X}_k \leftarrow \emptyset$

$i \leftarrow 1$

for $i = 1$ to J_k **do**

    if $w_k^{(i)} > 0.5$ **then**

        for $j = 1$ to $round w_k^{(i)}$ **do**

            $\tilde{X}_k \leftarrow \left[\tilde{X}_k, m_k^i \right]$

        end for

     end if
   end for
   return $\tilde{X}_k$

2.9 Summary

The close proximity problem requires an understanding of several diverse fields: situation awareness, sensing technology, and probabilistic mapping. This chapter has discussed the most relevant ideas from these fields.

Radar has been identified as the most appropriate sensor for the close proximity problem in all-weather conditions. But it presents challenges of measurement interpretation and susceptibility to noise and clutter. For a radar to be useful for this application, its performance must be well understood and techniques implemented to discriminate between true detections of objects in the environment, and spurious measurements caused by nonexistent or uninteresting objects.

The reliability of radar information can be improved by fusing successive measurements together in a map, but different mapping approaches have different strengths and require different models and parameters based on the radar's performance. The approaches considered are occupancy grid mapping and probability hypothesis density filtering.

Occupancy grid maps have been used with radar with good results[14] but require a measurement likelihood (sensor model) for to that sensor. Without additional heuristics they do not track individual targets, just the state of the total environment.

Probability hypothesis density filters, based in finite set statistics, offer a feature-based map that can represent both target existence uncertainty and position uncertainty. They require a feature extraction module to identify the targets in raw measurements, and require explicit representation of clutter, noise, and detection probabilities, as well as representations of the noise covariances in sensor measurements and target prediction.

In the next chapters, this information will be used in the development of a new sensing system to solve the close proximity problem. The requirements for a new FMCW radar sensor for driver assistance will be analysed; different detection filters will be trialled and compared; a sensor model for occupancy grid mapping with the radar will

be developed; and the EK-GM-PHD algorithm will be implemented, with the control parameters being determined based on the characteristics of the radar sensor.

Chapter 3

Sensor Characterisation and Modelling

In this chapter the characterisation and modelling process for a sensor to solve the close proximity problem for large vehicles is presented. The sensor will be mounted on the outside of the vehicle where it will monitor the vehicle's blind spot, and warn the driver of the presence of any obstacles in the path of the vehicle.

FMCW radar is the modality of choice for its all-weather performance, but there are no existing radar system that meet the requirements for our desired situation awareness system. It is therefore necessary to develop a new radar that does, called the close proximity detector (CPD) radar.

In Sec. 3.1, the performance requirements of a close proximity sensor are analysed. These are used in Sec. 3.2 where the characteristics and expected performance of a close proximity sensor are selected and the sensor designed. In Sec. 3.3 the sensor's true performance is tested. In Sec. 3.4, a static detection filter is developed using detection statistics for a known environment.

The contributions of this chapter are the performance requirement analysis for a sensor to solve the close proximity problem, the description of a characterisation procedure for a FMCW radar, and a method for designing a static detection threshold for a close-range FMCW radar with a short measurement vector.

3.1 Sensor Requirement Analysis

To summarise the requirements listed in Sec. 2.2.1, for a system to solve the close proximity problem it needs to reliably and accurately inform the driver of the presence

and position of objects that pose a collision threat to the vehicle in all weather, at a reasonable cost and in an easy-to-understand format.

For the sensor to be trusted by drivers, the most important performance criteria are:

- Reliability - how trustworthy a sensor is in terms of its probability of false alarm and probability of missed detection;

- Accuracy - how close the sensor's estimate of an object's position is to its true position;

- Coverage - how much of the dangerous area around the vehicle the sensor can observe;

- Ease of use - the operator interface needs to be easy to read and understand.

Constraints limiting these are:

- Size - the sensor must fit easily onto a large vehicle in a position where it will be able to monitor the required area;

- Power - the sensor must be safe for humans to work around, which limits the amount of power that it can output;

- Cost - the sensor must be priced economically.

For an FMCW radar, there are many design factors that will influence and be influenced by these criteria.

- Frequency affects beam width, maximum range and attenuation effects due to weather

- Bandwidth affects range resolution

- Antenna diameter affects beam width

- Power affects range

3.1.1 Reliability

A sensor system's reliability can be described by its probability of false detection P_{fa} and probability of detection P_d. These are functions of the sensor's sensitivity, noise levels, target type, target range and target aspect, sensor noise and environmental conditions such as weather and clutter distribution. A low false alarm probability and a high probability of detection are important for any sensor used in a situation awareness context. Too many false alarms and the system will be an annoyance to vehicle operators; missed detections can render the system less useful and potentially dangerous if the driver trusts it to detect all hazards.

For the close proximity problem, the sensor should be able to detect any object that poses a collision threat within a danger zone around the large vehicle. Objects of interest include other large and light vehicles, pedestrians, buildings and other structures such as walls, fences, trees, light poles and machinery. Some of these are much easier to detect than others; it is desirable to be able to detect a human being anywhere in this region, but this may be more difficult at range as humans have a fairly small RCS. It will be easier to detect a human being at closer range, within 10m, but a high probability of detection at longer range is always better.

In bad weather conditions, the sensor's performance should be expected to degrade slightly, but not substantially. A graceful degradation [4] of performance may be tolerable, but total failure would not be. Graceful degradation is when a sensor does not completely cease making measurements when hampered by bad weather, but the quality of the measurement gradually decreases, such as a lower sensitiivity, range or resolution.

False or missed detections, and bad weather reliability, can be influenced by processing of the sensor data as well as in designing the sensor hardware itself. However the same filter that excludes noise and clutter to reduce the P_{fa} may exclude real objects, decreasing P_d.

3.1.2 Accuracy

It is not enough to know that an object is present *somewhere* within a broad region near a large vehicle. Knowing the relative position of an object with greater precision will allow the driver to determine whether it can be safely avoided when the vehicle moves, or if it is not safe to move the vehicle at all. A sensor's accuracy determines how close

its measurements of object position are to the true position.

This information needs to be provided in a timely fashion; long delays in sensor measurements and processing are not tolerable. If the environment changes, the information should be updated quickly. The driver will not want to wait for a slow update; more than a second or two will become intolerable.

The majority of close proximity accidents occur at low speed, such as when the vehicle is being started from rest [40], so the vehicle will not be moving at speed or turning in most cases when the system is in use. It should be enough to determine whether an object is within a danger region close to the path of the vehicle, in a short enough timeframe that the driver does not become annoyed. We will say that it would be enough to know an object's position to within ~1m within ~1 second to allow the driver to determine if it can be safely avoided.

Sensor accuracy is influenced by the sensor's range and bearing resolution and any influences that may distort the measurement, such as noise. The range resolution of a FMCW radar is determined by its bandwidth, and by the bin size of the FFT performed on the raw measurements (which depend on sampling period and rate). Phase noise and chirp nonlinearity will have an effect, stretching out an object into several range bins rather than just one. The bearing resolution of the radar is determined by the beam width, which is a function of frequency and antenna diameter, as well as the measurement rate and scan speed.

The requirement to keep costs low prevents the use of a high-frequency high-performance FMCW radar with a narrow beam (~1°). These capabilities would be more precise than are needed for collision avoidance and are not justified by the sharp increase in cost. However, existing lower cost approaches use wide, fan-shaped beams that provide poor bearing resolution [59]. Multiple static, wide-beam radars can be used in an overlapping fashion, and an object's position computed more accurately using knowledge of their intersection areas as Fig. 3.1 illustrates. Having additional sensors would provide redundancy should one fail or produce noise measurements. However, using a single sensor is greatly preferred; each additional radar would increase the system cost, complexity, and the maintenance and calibration requirements. A narrower, scanned beam is preferable.
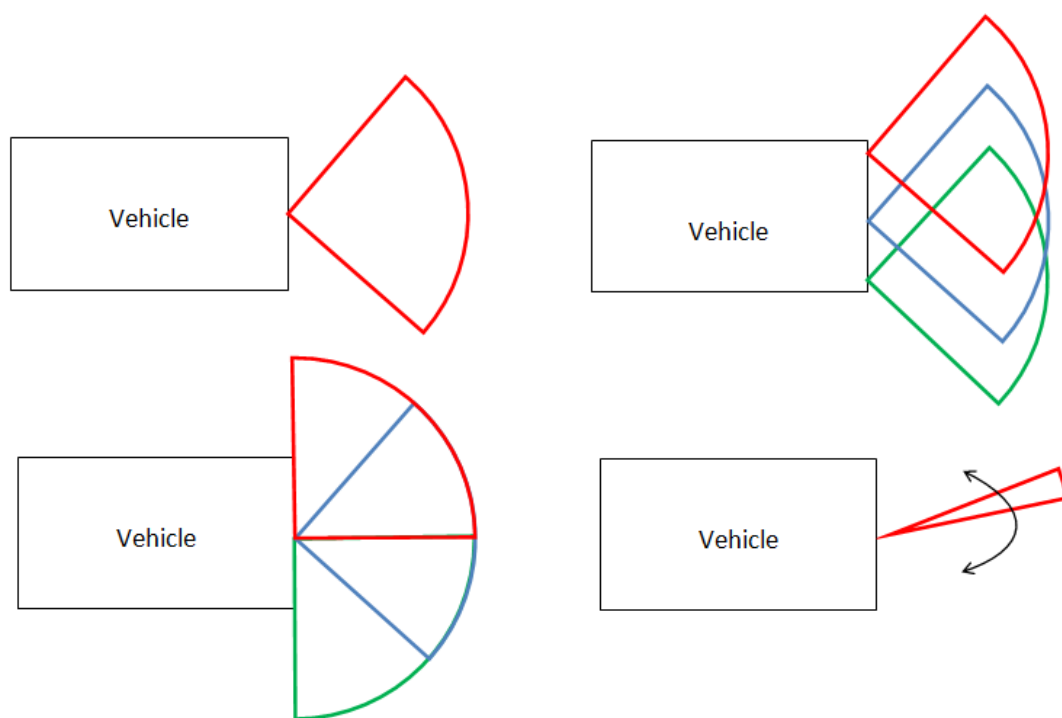
Figure 3.1: Different layouts of radars are considered for the close-proximity prob-
lem such as single or overlapping wide-beam radars, but a single, narrow,
scanned beam was decided on.

3.1.3 Coverage

The **surveillance volume** is the region of space that the sensor is able to monitor, whether by encompassing a large region in a single measurement or by being scanned across a larger region in multiple successive measurements. At a minimum, the surveillance volume needs to encompass the "danger zone" of blind spots around the vehicle. It is desirable for the sensor to function reliably at longer range, to provide earlier warning of possible hazards and provide greater aid to the driver. Throughout this thesis, "surveillance volume" will be used interchangeably with "observation volume" and "measurement volume".

The danger zone primarily consists of the blind volume within the path of motion of the large vehicle in front and behind it. The size and shape of this blind volume will vary with the type of large vehicle. For the haul truck shown in Fig. 1.2, the most dangerous blind spot in front of the truck consists of a rectangular region. There is also a dangerous blind spot directly behind the truck, spanning its width.

It is not possible to have one radar covering all blind spots; the bulk of the truck prevents it from viewing both front and back. For now we will focus on the front half of the truck as it has the wider blind spots, and assume that any solution that works in the front can be adapted to work in the back. One thing that needs to be noted is that a sensor at the front can be mounted on the front edge of the vehicle, but at the back will need to be mounted between the wheels, where it will need to *not* trigger nuisance alarms caused by detecting the wheels.

3.1.3.1 Surveillance volume in the horizontal (azimuth) plane

As Figure 1.2 shows, in the 2D horizontal or azimuth plane, the desired scan volume can be approximated as a 180° field of view centred on the front bumper of the large vehicle. Different ways to achieve coverage of such an area are illustrated in Fig. 3.1. Some sensors such as lidar can easily achieve a 180° or higher azimuth field of view. A camera would typically have a narrower field of view, though wide-angle lenses can expand this. Observing exactly 180° with radar can be difficult; the front of the vehicle is a very strong reflector at very close range, and could easily be detected by the edge of the beam, causing false alarms with every scan.

Therefore, it is necessary to reduce the scan region. This leaves a sector of the blind spot unmonitored. The size of this region should be considered; if it is too small to
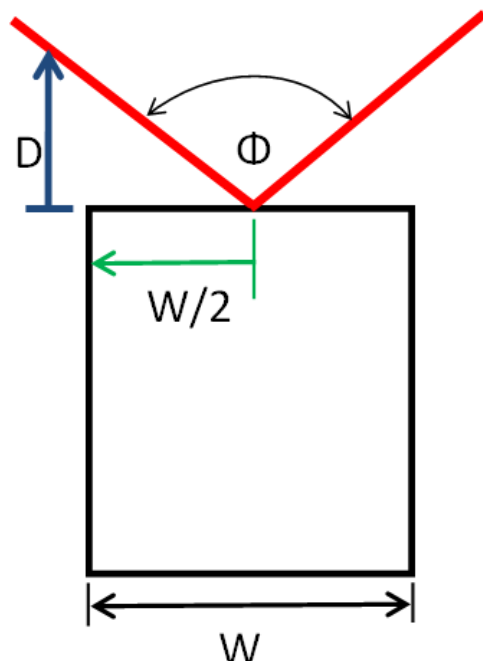
Figure 3.2: The maximum distance *D* between the front of a large vehicle and the edge of the surveillance volume is a function of the total scan range Φ and the distance of the radar from the edge of the vehicle. When the radar is centred on the vehicle, as shown here, the distance from the edge is *W/2*.

contain an object of interest, it is acceptable to ignore it.

The maximum distance between the edge of the radar beam and the front of the vehicle should be considered. This is the maximum size of an object that could be undetected at the extremities of the vehicle's front edge, and is a function of the radar's total swept angle and the distance in that the radar is mounted from the edge of the vehicle.

Returning to the haul truck in Fig. **??**, the Komatsu 730E is 6.68m wide. Detecting an object that is 1m wide or larger just in front of the corner of the vehicle will require a scan range of approximately 150°. Modelling a human being as 0.5m in diameter, to detect them reliably would require a scan range of approximately 165°.

Sweep range is the sum of the beam width and the scan range of the radar. Existing automotive and collision avoidance radars are unswept in azimuth; they point in a fixed direction and rely on either a wide beam (in the case of collision avoidance radars) or knowledge about the relative position of objects (in the case of automotive radars) to detect possible hazards. Beams can be mechanically or electronically scanned. Electronic
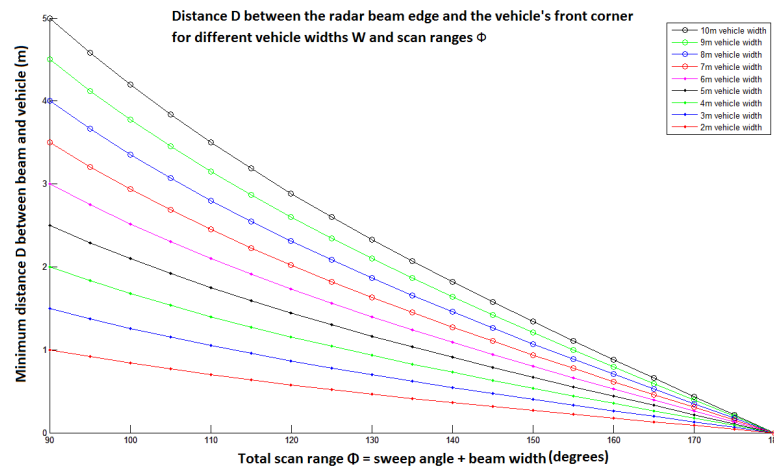
Figure 3.3: When scanning less than 180°, there is a gap between the edge of the radar beam and the front edge of the vehicle. This distance varies with total scan range and the distance of the radar from the edge of the vehicle. The radar is assumed to be centred on the vehicle and point straight ahead.

scanning is less maintenance intensive and allows for a simpler mechanical design, but requires a more complex electronic design.

3.1.3.2   Surveillance volume in the vertical (elevation) plane

If the radar beam is too narrow in the elevation axis, there is a risk of it overlooking objects in the environment with slight changes in the vehicle's pitch angle due to uneven terrain. Even on flat ground, it may miss some low-lying objects if it is mounted too high. The radar can be scanned in elevation as well as azimuth, but this will complicate mechanical design substantially.

There are several competing design factors present in determining the surveillance volume in the elevation plane: the beam's grazing angle, the beam's width in the vertical plane, and the sensor mounting height.

A low grazing angle gives a longer range before the beam is buried in the ground and gives lower clutter measurements (as described in Sec. 2.4.1.4). A high grazing angle will better detect low-lying objects nearby, but with higher clutter measurements and shorter maximum range before the beam hits the ground.

A "tall" beam that is wide in elevation will more reliably monitor the blind volume as the vehicle pitches up and down, but will be more susceptible to detecting clutter from small objects on the ground, or hilly terrain. A "flat" beam that is narrow in
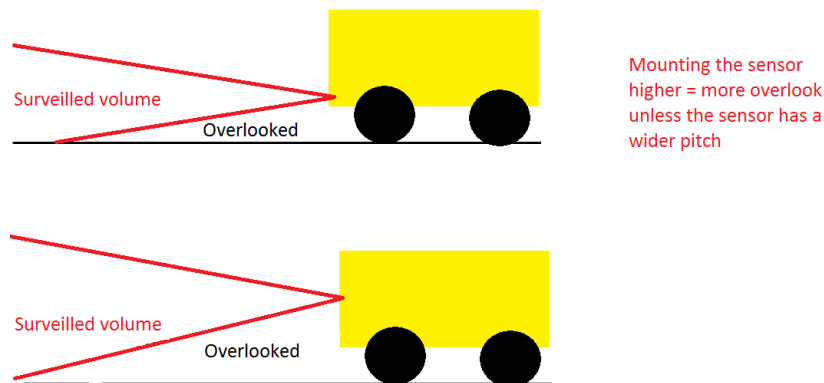
Figure 3.4: The height at which a sensor is mounted has a large impact on its surveillance volume.

elevation will have fewer clutter measurements but higher risk of overlooking an object closer to the vehicle.

A sensor mounted higher on will be better protected from dust and mud spatter, and will be less likely to be damaged by falling rocks if it is mounted on a more protected region of the vehicle. But it may have blind spots due to overlooking nearby objects, in the same way as the driver, unless it has a tall beam or high grazing angle. A sensor mounted lower on the vehicle will be more exposed to the elements but will be able to detect nearby objects reliably with a flatter beam and a lower grazing angle.

Unless the radar is mounted at ground level, or pointed down at a very steep grazing angle, it will overlook some volume of space for the same reasons that the driver does. Neither of these options is practical, so some compromise will need to be reached between grazing angle and sensor mounting height that is overlooked will need to be calculated. Fig. 3.5 shows the relationship between grazing angle and distance for the lower beam edge to reach the ground, for a range of different sensor mounting heights. This can be used for either true grazing angle, to find when the beam axis hits the ground, or the lower edge of the beam (using grazing angle plus half the beam width). For example, a beam with an effective grazing angle of 25° and a mounting height of 2m will strike the ground slightly less than 5m in front of the vehicle.

From these figures, it is clear that the mounting height and grazing angle are key factors in determining the surveillance volume.
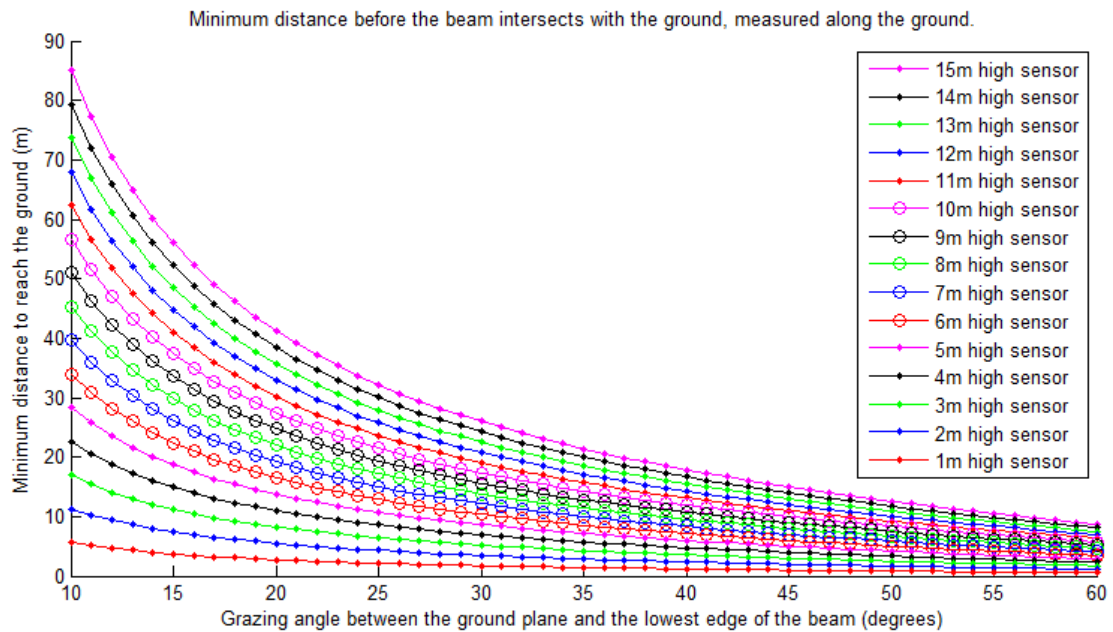
(a) The effect of grazing angle and sensor mounting height on minimum distance to reach the ground, over long ranges.
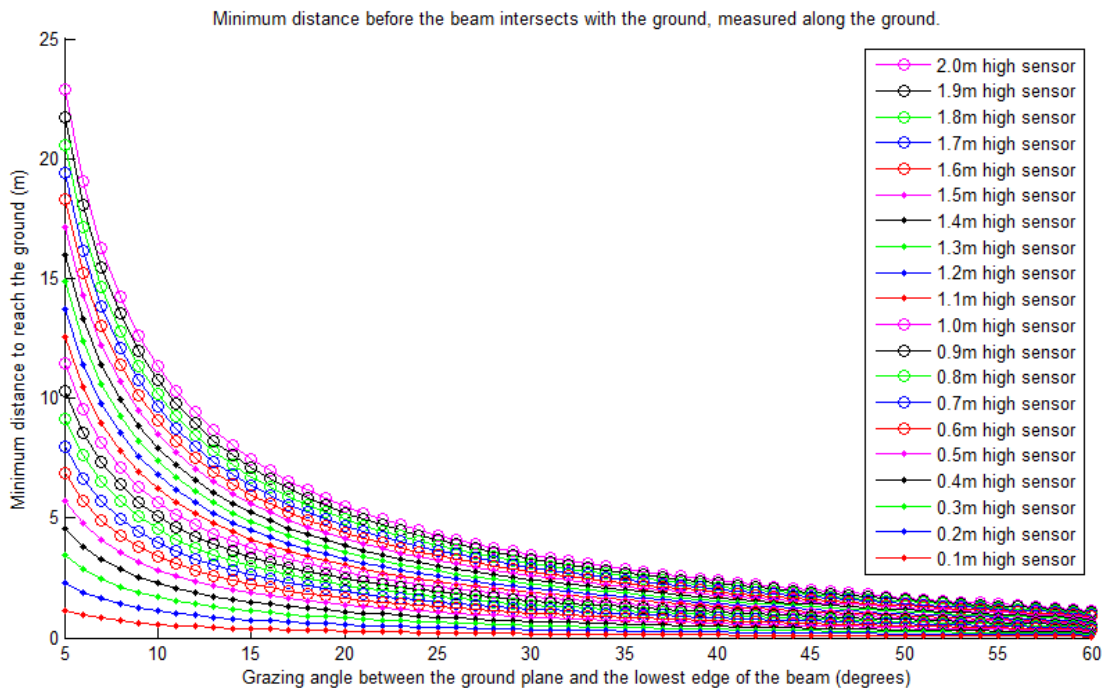
(b) The effect of grazing angle and sensor mounting height on minimum distance to reach the ground, over short ranges.

Figure 3.5: Depending on the height of the sensor mounting and the grazing angle of the beam, the beam strikes the ground at a different distance in front of the vehicle. A low distance for the beam to intersect with the ground is desirable to prevent overlooking obstacles. A shallow grazing angle is desirable to reduce clutter measurements.

3.1.4   Ease Of Use

The sensor needs to present its information to the driver in an easy to understand fashion. Drivers of large vehicles have many stimuli competing for their attention. A complicated display of sensor data that requires effort to interpret will be more of a distraction than a useful safety tool. Ideally a system would be automated and not require the driver to pay much attention until an alarm warning is given visually or aurally.

FMCW radar produces rich measurements that are not easily interpreted by humans in their raw numerical form. To make them more human-readable, they are typically plotted as a heat map of power values, with higher values being drawn brighter and weaker values being drawn more dimly. This can be drawn in a Cartesian grid or directly from the polar coordinates of the radar measurements. Audible alarms can be generated to warn the driver of a nearby hazard, prompting them to check the radar plot, rather than requiring them to continuously monitor it.

3.2   Sensor Design

At the time this thesis was commenced, there was no existing radar capable of meeting these requirements to a satisfactory level; there were radars that are too expensive and more precise than is necessary, and radars that were affordable but not precise enough or did not provide sufficient coverage.

Rather than compromising on cost or angular resolution, a new radar was designed. This would be called the **close proximity detector (CPD) radar**. The most important design considerations for the CPD radar were that the it needed to have low cost but maintain good detection capabilities, good bearing resolution, reasonably good range resolution and provide adequate coverage of the blind spot in front of a large vehicle. Some photos of the CPD radar are shown in Fig. 3.6.

Since this research commenced, similar radars have been developed independently that offer performance closer to what is needed for this problem, such as the one used in [31] and the IMST sR-1200 [25].

(a) The CPD radar, mounted on a small mining haul truck.
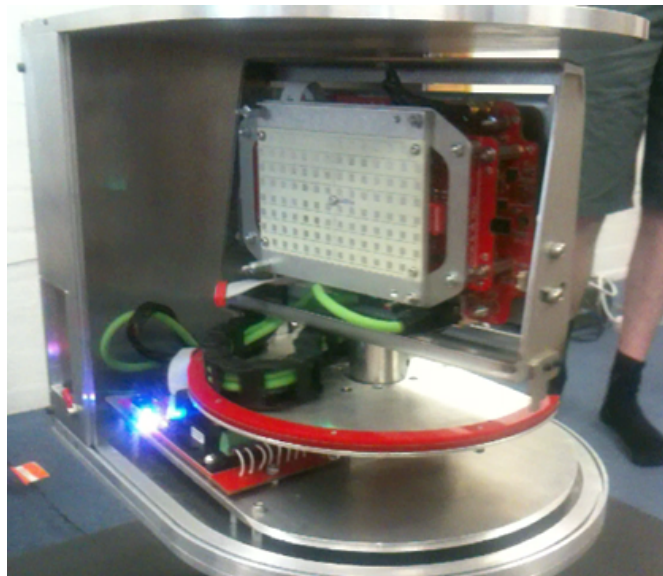
(b) The CPD radar with its cover removed to show the antenna

Figure 3.6: The close proximity detector (CPD) radar.

A single radar was to be used, with a beam that was fairly narrow in azimuth but was mechanically scanned across a variable range, up to 160° . The beam decided on was designed to be ≈8° wide in azimuth and ≈17 degree in elevation. This provides good cross-range resolution and a taller survived volume, without being so tall that energy will dissipate too rapidly or ground clutter will be detected too strongly. The measured beam width was closer to 6° in azimuth. The beam and scan geometry are shown from above in Fig. 3.7.

CPD radar beam and scan geometry, viewed from above

Beam width: ~6°

Total scanned width: 160°

One range bin = 1m
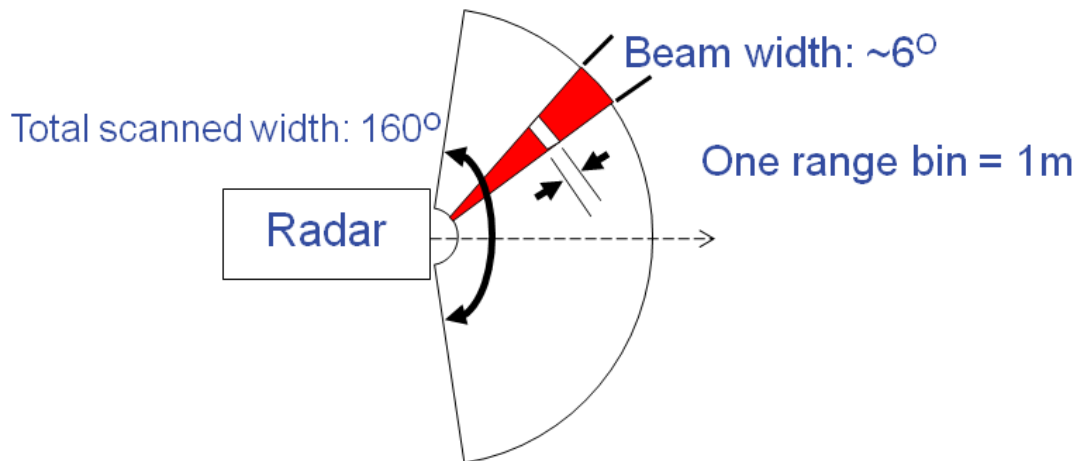
Radar

Figure 3.7: This diagram illustrates the geometry of the CPD radar beam and scan range. The radar scans mechanically across a maximum range of 160° . The beam was ≈6° wide, and range bins were 1m long.

The measurement rate is limited by the rate at which FFTs can be performed by the processor. The measurement rate limits the scan speed; scanning too quickly will leave too large a gap between measurements and objects could be missed.

The radar needed to be mountable on a vehicle, which limits the antenna size, which affects the width of the beam.

The nominal characteristics of the close proximity radar as designed are listed in Table 3.1.

Table 3.1: CPD radar properties

Parameter	Value
Frequency	24GHz
Bandwidth	146MHz
Range resolution	1.03m
Measurement rate	100Hz
Maximum rotation range	160° (azimuth)
Azimuth beamwidth	8.83° (nominal) 6.02° , standard deviation 0.72° (measured)
Sidelobe angle	±9° (nominal) ±9.45° , standard deviation 1.16° (measured)
Vertical beamwidth	16.81° (nominal)
Maximum scan speed	1200°/s (nominal)
FFT vector length	512 bins (256 bins in measurement, only first 16-64 used)
Bin length	1m
Received sidelobe power	-26dB (nominal) -19.44dB (left, tested) -17.72dB (right, tested)

3.2.1   Plotting the radar measurements

The radar data can be plotted as a heat map for easy visualisation of the data. Each bin is projected into the horizontal plane as a trapezium of length equal to the estimated bin length, and with its width calculated from the range and the estimated beam width. Overlapping cells are drawn on top of one another. The brightness of each cell is proportional to the power measurement recorded for that bin.

Cells can be upsampled along the range axis into two or more subcells and drawn with an intensity corresponding to a weighted average of adjacent bins, for a smoother plot. This plotting method is fairly resource-intensive at high angular resolution or with more upsampling of the cells, but produces smoother, more attractive plots.

The presence of noise and clutter measurements in otherwise empty space can make it more difficult to interpret a plot produced in this way; areas that are equally empty in reality will have different power values and so will not necessarily be drawn with the same colour, making it appear as if there are small, weak targets present. A wider power range of noise values stretches out the colour range, reducing the colour resolution. As

Sec. 2.5 outlines, a detection threshold can be used when plotting radar measurements when the structure of the environment is known. All measurements below the detection threshold can be treated as zero and drawn as black; this filters out a lot of bad detections and produces much clearer plots. A saturation threshold can be defined as some ratio of the detection threshold (for example, N dB higher than the threshold for each bin) and the colour of each cell scaled between a minimum (black) at the detection threshold and a maximum (white) at the saturation threshold. Since the structure of the environment is known, the detection threshold can be adjusted manually until the plot looks acceptable.

Fig. 3.8 shows a photograph of a carpark environment, with certain features labelled, and below it a plot of a radar scan of that environment with the same features marked. The brighter regions of the plot are regions that return more power to the radar than the darker ones. The features that the radar detects are the ones that present large, flat reflective surfaces such as the columns marked (1) and (4), the highly reflective shape of the trihedral corner reflector (2), or the white station wagon (3) which contains elements of both flat surfaces and tail lights with similar geometry to the corner reflector.
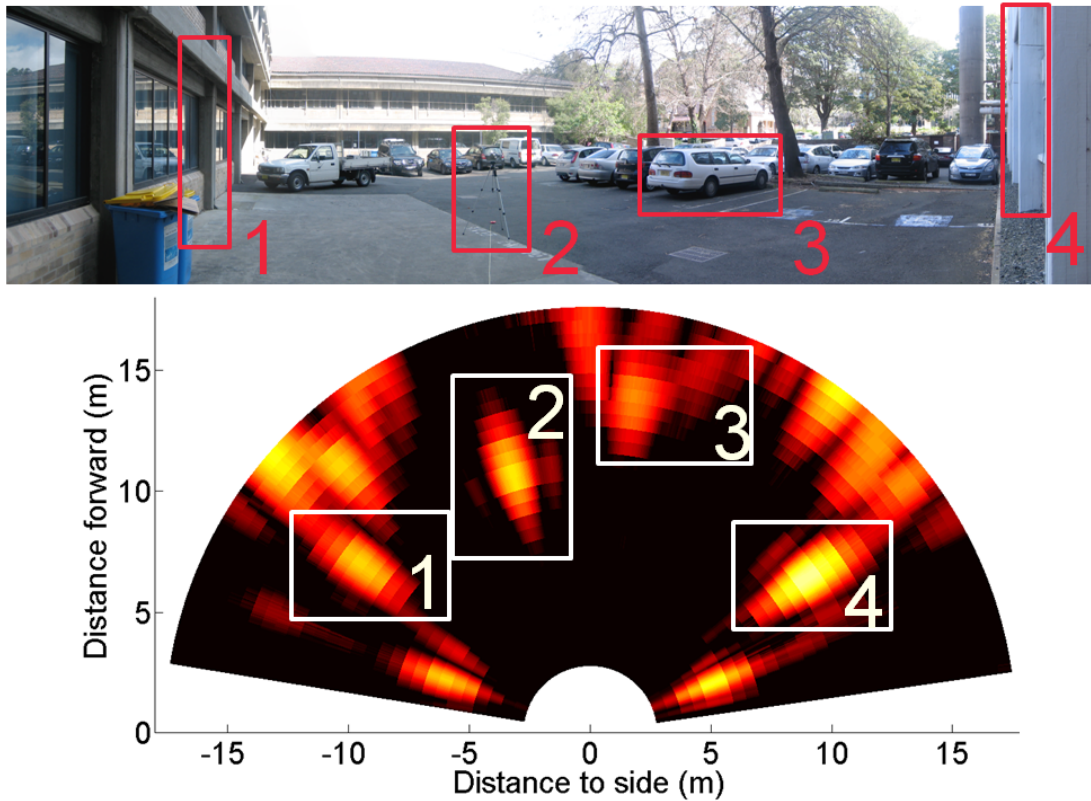
Figure 3.8: A photograph and radar scan of the environment are shown here. Some of the same key features are labelled in both images for ease of comprehension. Feature 1 is a column of the building, feature 2 is a corner reflector on a tripod, feature 3 is a parked car, and feature 4 is a column on another building. Additional features such as columns and vehicles are also present in both images. This is plotted using a uniform threshold.

In Fig. 3.9 different uniform thresholds are used to remove noise and clutter when plotting the same car park scene. In each scan, all bins have the same threshold value; this works because only the closer bins are considered. Increasing the threshold value between each scan produces substantially different plots of the same scene. A higher uniform threshold allows clearer discrimination between empty space and objects, but if it is too high, real objects may be rejected. This is discussed in Sec. 3.4.
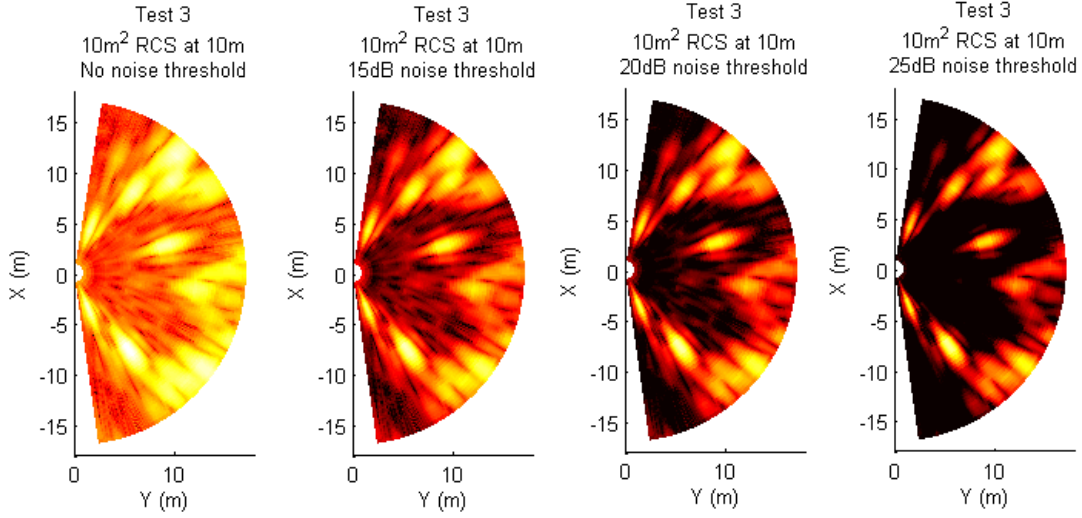
Figure 3.9: The CPD radar measurements of the same car park scene shown in Fig. 3.8 are drawn here, rotated 90° clockwise (for space) and with different noise thresholds, to illustrate the impact of varying the threshold. The brighter the area of space, the higher the measurement is above the noise threshold. The uniform threshold is increased in each plot from left to right. By increasing the threshold, empty space becomes clearer and targets are resolved more distinctly.

A uniform threshold needs to be manually adjusted to suit the environment to produce an attractive radar image. This approach is useful when the structure of the environment is well known but the characteristics of the radar are not. Once the radar has been characterised and needs to be used in unknown environments, a better method of designing a detection threshold will be needed.

3.3 Sensor Linearisation and Characterisation

As described in Sec. 2.3.9.1 and illustrated in Fig. 3.10, the FMCW radio chirp is generated by a voltage-controlled oscillator (VCO) which outputs a frequency proportional to the input voltage. Due to hardware limitations, especially with less expensive frontends, the VCO will not have a perfectly linear frequency response. This degrades the range resolution and raises the noise floor, as described in Sec. 2.4.1.6. The frontend manufacturer provides a graph of the expected frequency response nonlinearity; one such graph is plotted in Fig. 3.11. Before we characterise the radar, this nonlinearity should be compensated for as much as possible.
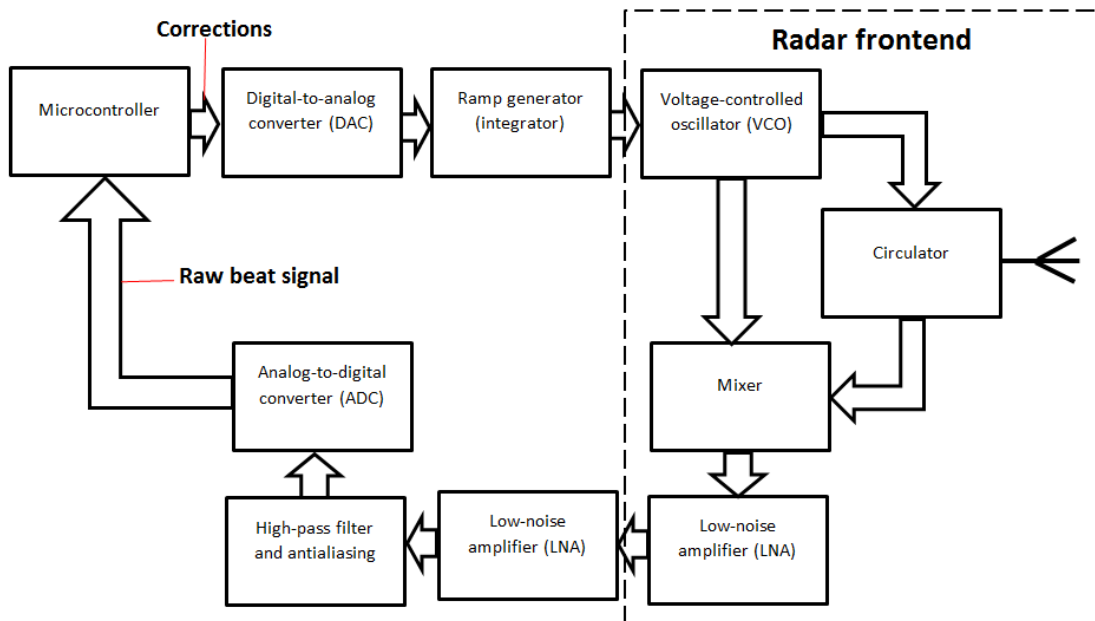
Figure 3.10: The hardware block diagram for the CPD radar. To linearise the chirp, the corrections passed from the microcontroller to the digital-to-analog converter must be calculated. This involves testing under controlled conditions and analysing the raw beat signal.
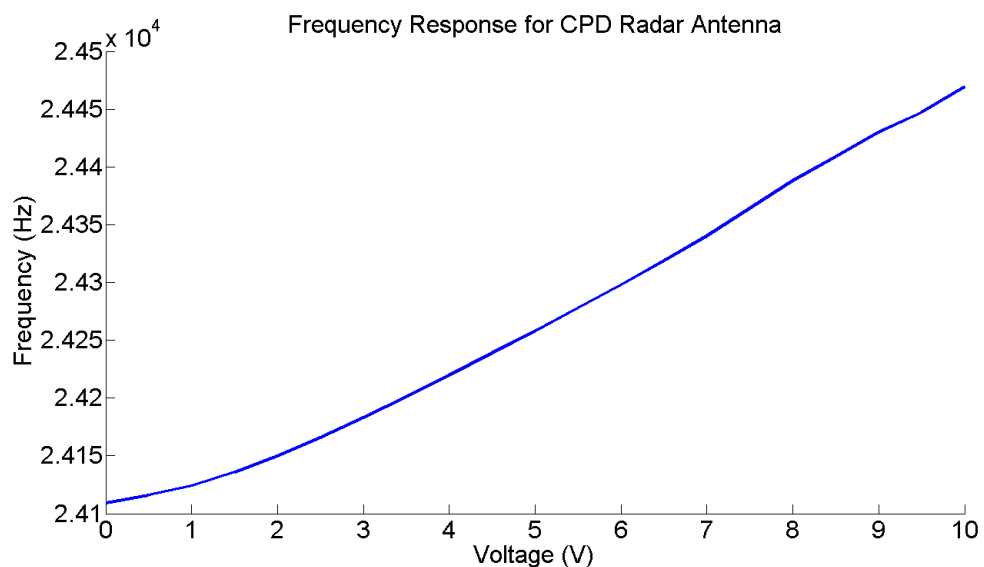
Figure 3.11: The CPD radar frontend has a nonlinear response, as visible in this graph of frequency response versus voltage provided by the manufacturer.

Once the chirp has been linearised, further tests are needed to measure the radar's properties. The radar's detection characteristics are difficult to predict; the radar cross-section of complex objects is difficult to model and is a function of size, material, surface texture and viewing angle, and the radar range equation is only an approximate guide for actual performance. The only way to determine the true performance of the radar is to perform characterisation tests, using targets of known RCS at known ranges in a controlled environment.

3.3.1   Chirp Linearisation

Linearisation refers to the process of modifying software parameters that are used in FMCW chirp generation to create a more linear frequency chirp. A more linear chirp produces better range resolution and signal to noise ratio. Once linearisation has been completed, all the characterisation tests can be performed.

The hardware block diagram for the CPD radar is shown in Fig. 3.10. Linearisation is performed by generating a series of corrections in the voltage input to the VCO. A look-up table is used to feed values into a digital-to-analog converter that generates a voltage in opposition to the VCO integrator. Selecting the correct values for the look-up table will produce a nonlinear signal that counteracts the nonlinearity inherent to the frontend. The linearisation process is how these values are calculated.

The linearisation process employed here used a single target with a fairly low RCS, situated closely enough to be a strong, clear target but not so close to saturate the measurement. The target is set up in such a way that there are no objects in the background. Typically this involves mounting the target on a tall stand in an open area and placing the radar low to the ground, angled up, so that there is only sky behind the target. The radar is not scanned and is pointed directly at the target, and data is logged.

Rather than perform a FFT on the raw beat signal, the phase of the sinusoidal signal is obtained and "unwrapped"; rather than wrapping around from 2π to 0, it continues up to 3π, 4π, etc. Ideally, the presence of a single target would result in a single beat frequency, which would be represented by a constant **phase gradient**. The presence of noise (from nonlinearity, additional targets or thermal noise) would appear as nonlinearities in the phase gradient. Some samples of raw beat signal for different uniform corrections are shown in Fig. 3.12. This unwrapped phase data of these signals are shown in Fig. 3.13.
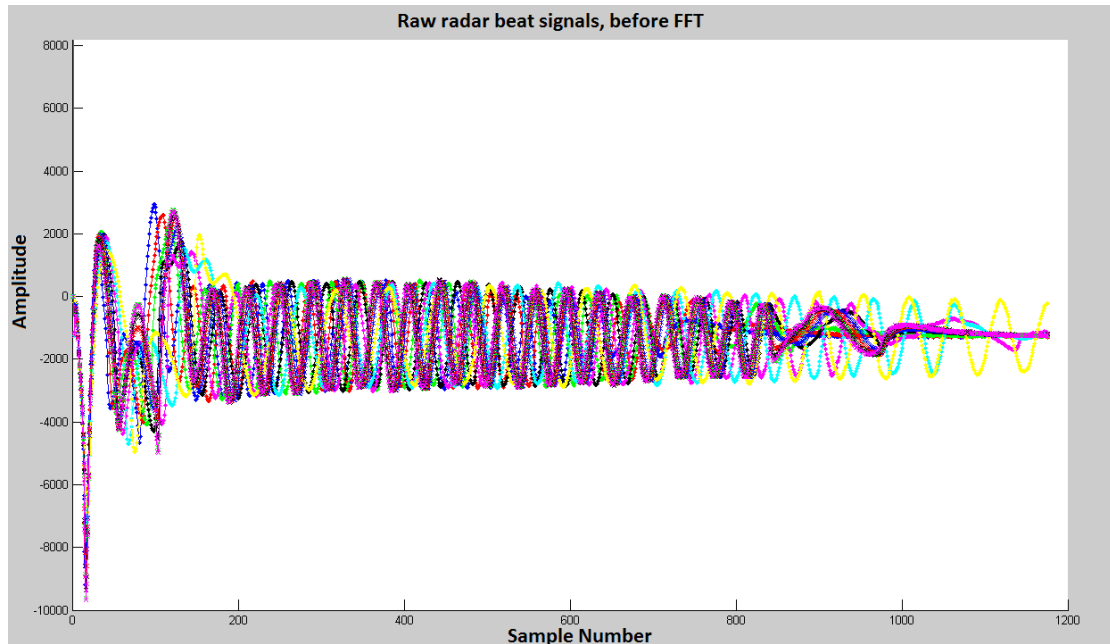
Figure 3.12: The raw beat signals for a number of different correction values, both uniform and linearised. The horizontal axis is sample number, the vertical axis is voltage. A perfectly linear chirp would have constant frequency. The worst nonlinearity is present at the start and end of the chirp.

Since a constant time between samples is assumed, phase gradient can be calculated by just taking the difference between every consecutive pair of measurements. Ideally this would form a horizontal line when plotted, with some variation at the start and end of the chirp as the frequency begins to ramp up and levels off. In practice, it is more of a sloped or wavy line. The gradients of some unwrapped phase signals, both uniform corrections and linearised corrections, are shown in Fig. 3.14.

Linearisation is performed by forcing the line to be more horizontal, by either increasing or decreasing the gradient at the points where it is lower or higher than it needs to be. Since the chirp duration is fixed, linearisation can reduce the total bandwidth of the chirp if it reduces the average phase gradient over the duration of the chirp. A high bandwidth is desirable for good range resolution. Therefore, the goal of linearisation is to produce a linear chirp with minimum correction downwards to allow as high a gradient as possible which will produce a chirp with higher bandwidth.

Corrections were integers between 0 and 65532. The CPD radar takes 1174 samples over the duration of the chirp, and a subset of 512 of these are used in the FFT. There are inherent nonlinearities at the start and end of the chirp, so the 512 samples chosen
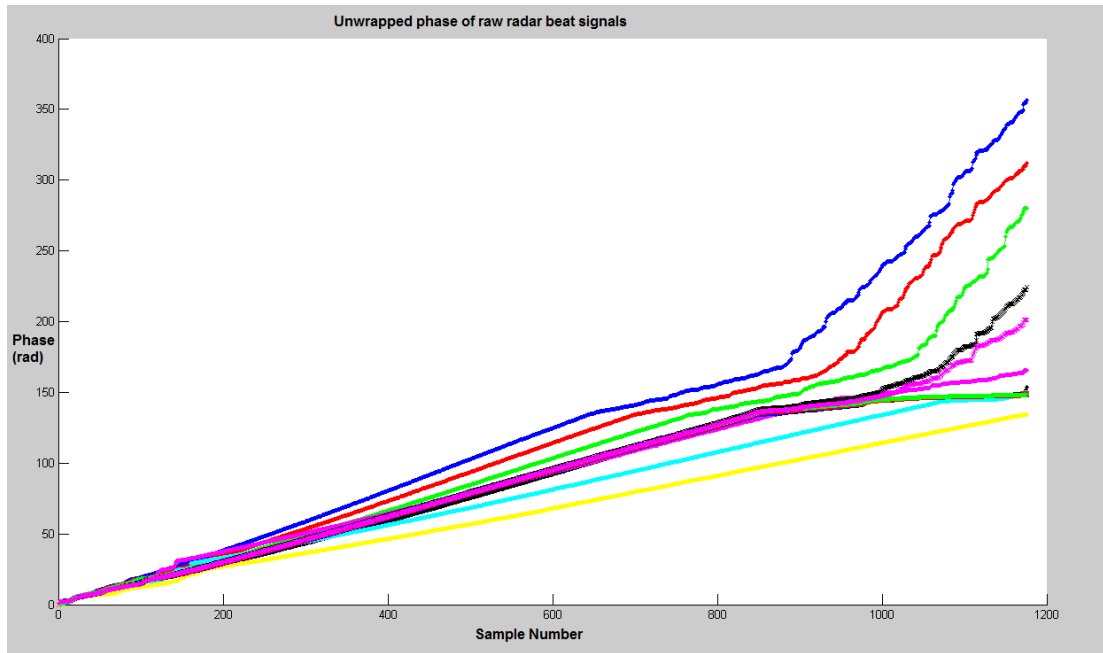
Figure 3.13: The unwrapped phase signals for a number of different correction values, both uniform and linearised. The different gradients show the effect of the correction on the chirp gradient: higher corrections produce lower gradients, as the corrections reduce the voltage being input to the VCO. The blue and red curves have lower uniform corrections; the yellow and cyan have higher uniform corrections. The linearised curves include the black and magenta curves in the middle. The beginning and end of the datasets, where the chirp begins and levels off, are most nonlinear; therefore the FFT samples are drawn from the more linear middle section. The steeper lines correspond to lower corrections, while larger corrections drive down the rate of phase increase and cause a gentler slope.

are taken from the most linear region in the middle of the chirp.

Reference data is first obtained by performing identical tests with different uniform corrections with a range of values (all 0, 1000, 2000, 5000, 10000, 15000, 20000). The results from these can be used to calculate a correction vector by interpolating between the correction values to obtain a desired gradient at a certain phase and timestep.

Chirp gradients: Uniform and linearised
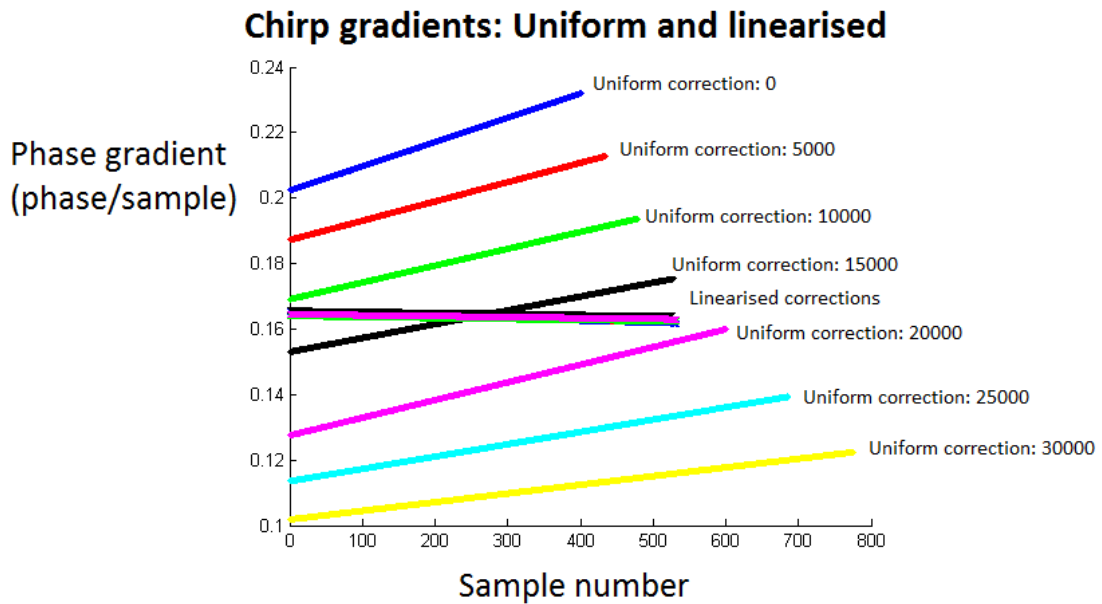
Figure 3.14: The phase gradients of the uniform corrections and linearised corrections. The uniform corrections all have varying gradients, which shows their non-linearity. The iterated linearised corrections have a much more constant, almost horizontal gradient, showing that they produce a much more linear signal.

The results of linearisation can be seen by taking a the FFT of the linearised beat signals, shown in Fig. 3.15. For the sake of clarity, only a subset of all the linearised and unlinearised signals are plotted. The unlinearised signals (the red 15000 and blue 20000 constant corrections) have lower, broader peaks than the linearised ones. This is more noticeable in the close-up shown in Fig. 3.16; the linearised peaks are around 1dB higher than the non-linearised ones for the peak, and 6dB and 10dB lower in the bins either side of the peak. This is a substantial narrowing of the detected target, due to a more linear chirp. The noise floors in the linearised signals are lower. This all contributes to targets being easier to detect.
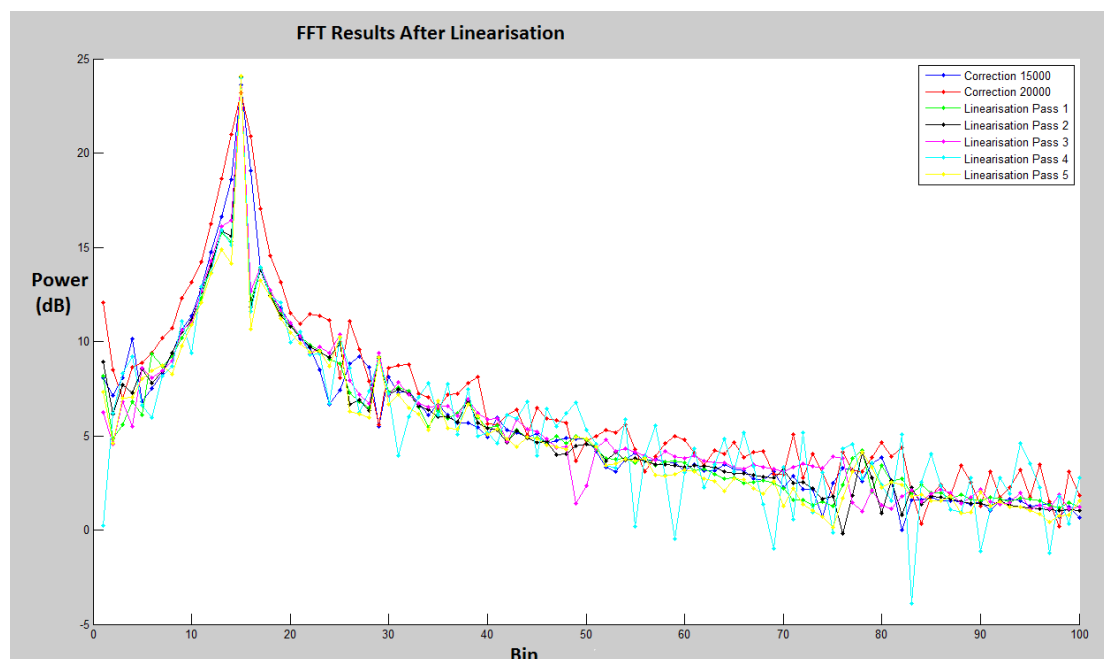
Figure 3.15: A FFT is taken of some of the linearised chirp and non-linearised chirp signals to show the effect of linearisation. The linearised signals have a higher peak and lower noise floor.
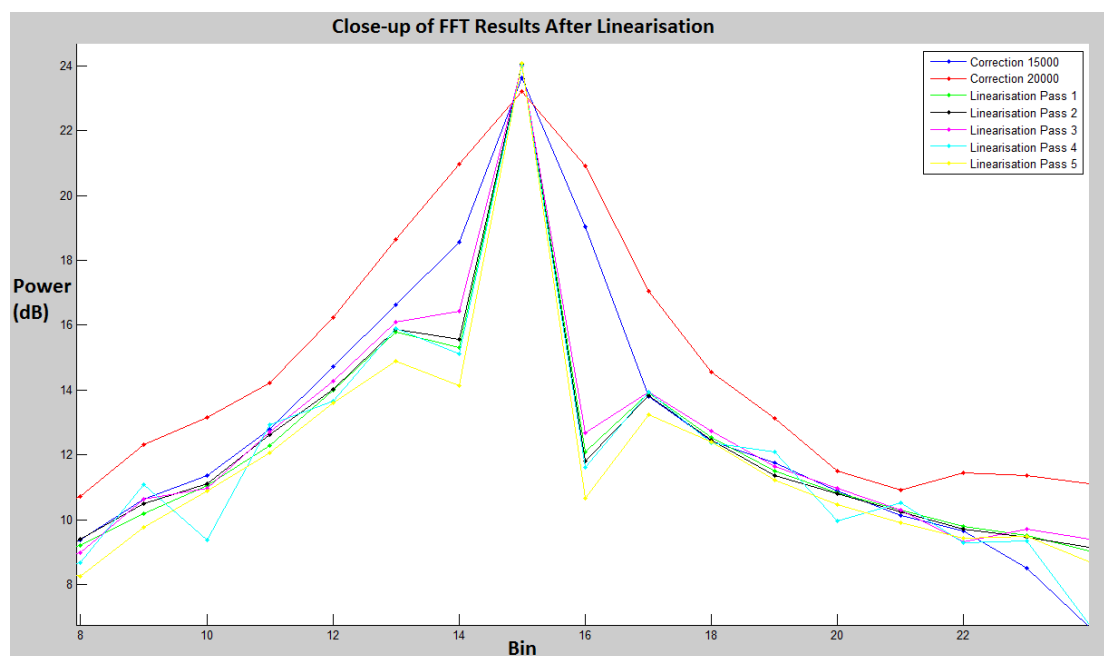
Figure 3.16: In this close-up of the FFT of the linearised and non-linearised signals shown zoomed-out in Fig. 3.16. The linearised chirps produce a taller, narrower peak than the unlinearised ones.

It is not possible generate a "perfect" linear chirp, so linearisation is only done to some level of satisfaction; if the quality of the linearisation starts to degrade rather than improve (due to errors in the sampling, excessive or inaccurate filtering or other flaws in the process) it is terminated. The most linear curve is identified and its corrections are used. In this case, we see that Linearisation Pass 1 and Linearisation Pass 2 (the green and black lines in Fig. 3.15 and Fig. 3.16) are the best candidates; further iterations of the linearisation produce increased oscillations in the noise floor without substantial improvement in the height or narrowness of the target peak. We use the corrections for Linearisation Pass 2 on the CPD radar.

3.3.2 Sensor Characterisation

Once the chirp has been linearised, further tests are needed to measure the radar's performance characteristics. The only way to determine the true performance of the radar is to perform characterisation tests, using targets of known RCS at known ranges in a controlled environment. The radar range equation is only an approximate guide for actual performance, and the radar's detection characteristics are difficult to predict.

Most of these tests use artificial targets (such as corner reflectors), as they provide consistent and reliable detection properties. Natural targets such as a light vehicle, buildings and human being are used in other tests, to verify that the CPD radar is capable of detecting the objects that it will encounter naturally.

The standard test setup is shown in Fig. 3.17. A corner reflector target of known radar cross-section is mounted on a tripod in a wide, open space (such as a lawn or car park) and the radar scans across it. It helps to scan very slowly to give good angular resolution but this is not essential. The ground truth distance to the target is measured using either a SICK LMS-200 lidar (with accuracy within $\approx$10cm) or a tape measure (accurate to within <1cm).

The radar beam is scanned across the target

Radar

The environment is flat and open with no other objects the radar should detect

A corner reflector target is mounted on a tripod
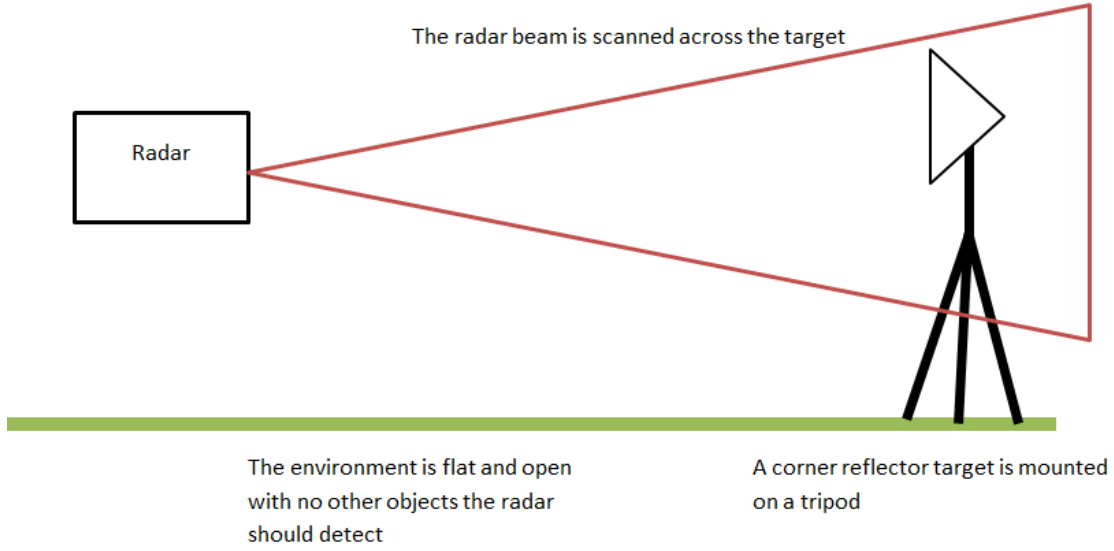
Figure 3.17: Most characterisation tests followed this basic setup.

3.3.2.1 Target Properties

The corner reflector targets used were $100m^2$ and $10m^2$ RCS at 77GHz. Remembering the RCS of a corner reflector of vertex length a is:

$$\sigma = \frac{4\pi a^4}{3\lambda^2} \tag{3.1}$$

To convert RCS@77GHz to RCS@24GHz for the CPD radar:

$$\sigma_{@77} = \frac{4\pi a^4}{3*(c/(77*10^9))^2} \tag{3.2}$$

$$\sigma_{@24} = \frac{4\pi a^4}{3*(c/(24*10^9))^2} \tag{3.3}$$

$$\sigma_{@24} = \left(\frac{24}{77}\right)^2 * \sigma_{@77} \tag{3.4}$$

$$\sigma_{@24} = 0.0971 * \sigma_{@77} \tag{3.5}$$

$$\sigma_{@24} \approx 0.1 * \sigma_{@77} \tag{3.6}$$

Therefore the 100m^2 and 10m^2 at 77GHz RCS targets would be 9.71m^2 and 0.971m^2 at 24GHz. For the sake of simplicity, these will be referred to as 10m^2 and 1m^2 RCS reflectors from this point, but the true values will be used in calculations.

3.3.2.2 Bin Length and Range Resolution

The bin length is controlled by the sample frequency and the number of bins in the FFT. Both of these are limited by the processing speed within the radar; high sample frequencies and larger FFT sizes will require faster hardware.

The maximum beat frequency is half the sampling frequency (due to the Nyquist criterion).

$$f_{b,max} = \frac{f_{sample}}{2} \tag{3.7}$$

This frequency range is divided evenly between the bins of the FFT.

$$f_{resolution} = \frac{f_{max}}{N_{bins}} \tag{3.8}$$

In the CPD radar, each measurement consists of 1174 samples taken at $f_{sample} = 850\text{MHz}$. A 512-point FFT is performed on the most linear region of this data, which means there will be 256 bins (since the number of bins is half the FFT size). The frequency resolution per bin is therefore $f_{resolution} = 1.66\text{MHz}$.

Rearranging and substituting into the FMCW range equation, Eq. 2.20 gives

$$f_{resolution} = \frac{\Delta f}{T_b} \frac{2R_{bin}}{c} \tag{3.9}$$

Chirp duration T_b is the FFT size (512) divided by the sample frequency (850MHz). The bandwidth Δf (Hz) was designed to be 160MHz, but linearisation can alter this (as the process of linearisation can involve driving down the chirp gradient, reducing total bandwidth over the span of the chirp). The true bandwidth is calculated by making measurements of a target at a known range, measuring the beat frequency, and rearranging and substituting into Eq. 2.20 to solve for Δf. This is found to be $\Delta f = 146\text{MHz}$.

Substituting these values into Eq. 3.9 and rearranging gives $R_{bin} = 0.999\text{m}$, which is simpler to approximate as 1m. Tests using targets at known ranges confirm this value.

The bin length should be at most equal to the range resolution of the radar. Ideally it would be about half, to allow for higher precision without excessively increasing processing load. The theoretical range resolution of the radar is stated in Eq. 2.9 as a function of the bandwidth: $\delta R = \frac{c}{2\Delta f}$. For the 146MHz bandwidth of the CPD radar, this gives a range resolution of 1.0267m, which is very closely matched to the bin length. A smaller bin would be preferable, but it is smaller than the theoretical range resolution so there is less loss of resolution.

Range resolution of 1m is fairly poor. Two separate objects any closer together than 2m will be indistinguishable from one object sitting on the boundary of two bins, or one large object extending into two bins. But for the close proximity, poor range resolution is not totally crippling flaw. The driver needs to be aware of the presence of an object; whether it is one extended object or two distinct ones is less critical.

3.3.2.3   Beam Width Measurement

Unless the radar scans very fast, a target will be detected by multiple measurements within a single scan. The beam width can be measured from the range of angles within which a small target is detected. The power value reaches a peak within the measurement that is made when the beam is pointed directly at the target, but measurements made slightly off-axis will still detect the target, albeit at a diminished gain.

Eq. 2.10 gives the beam width as a function of wavelength, $\theta_{3dB} = \frac{70\lambda}{D}^{\circ}$. At 24GHz, and with an aperture width of 99mm wide in azimuth and 52mm in pitch, the calculated beam width is 8.83° in azimuth and 16.81° in pitch.

The beam width is measured by plotting the power values for a single bin in multiple measurements as it sweeps across a target. Fig. 3.18 shows this data. There is an increase as the leading edge of the mainlobe passes over the object, a peak as the mainlobe axis points directly at the object, and then the power will drops as the mainlobe moves away. The beam width is defined as the angle across which there is a 3dB drop either side of the peak. Since these measurements are made after both transmission and reception, the power will have dropped by 6dB (since it loses 3dB each way).

From measurements made using the $1m^2$ and $10m^2$ corner reflectors, the width of the beam was found to be somewhere between 5° and 7° ; it is assigned a nominal width of 6°. This is narrower than the expected width, but not excessively so.

The same data is plotted in 3D in Fig. 3.19 to give an idea of the shape of the beam

volume.
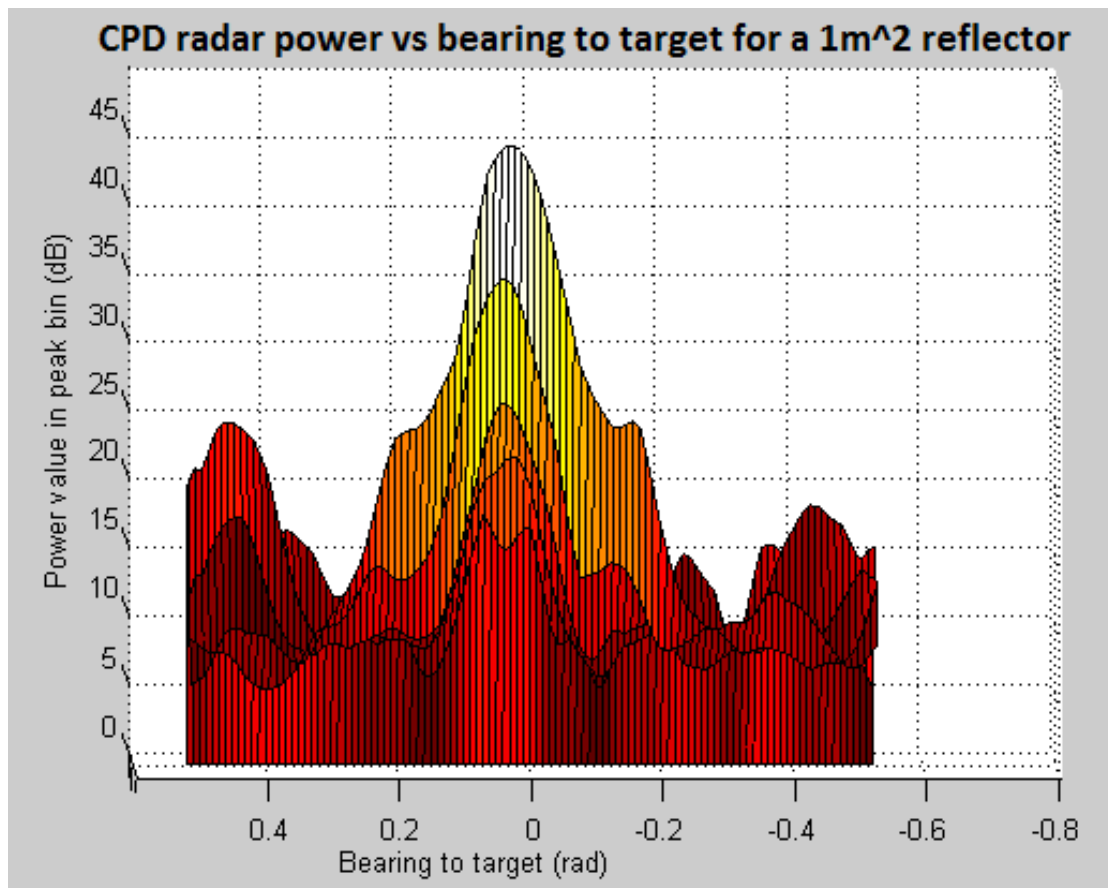
Figure 3.18: The CPD radar beam profile as measured at different ranges using a ≈1m2 RCS target on bearing zero. The uppermost curve is the mean profile between 0m and 5m, and the ones below are measured at 5m increments to 25m.
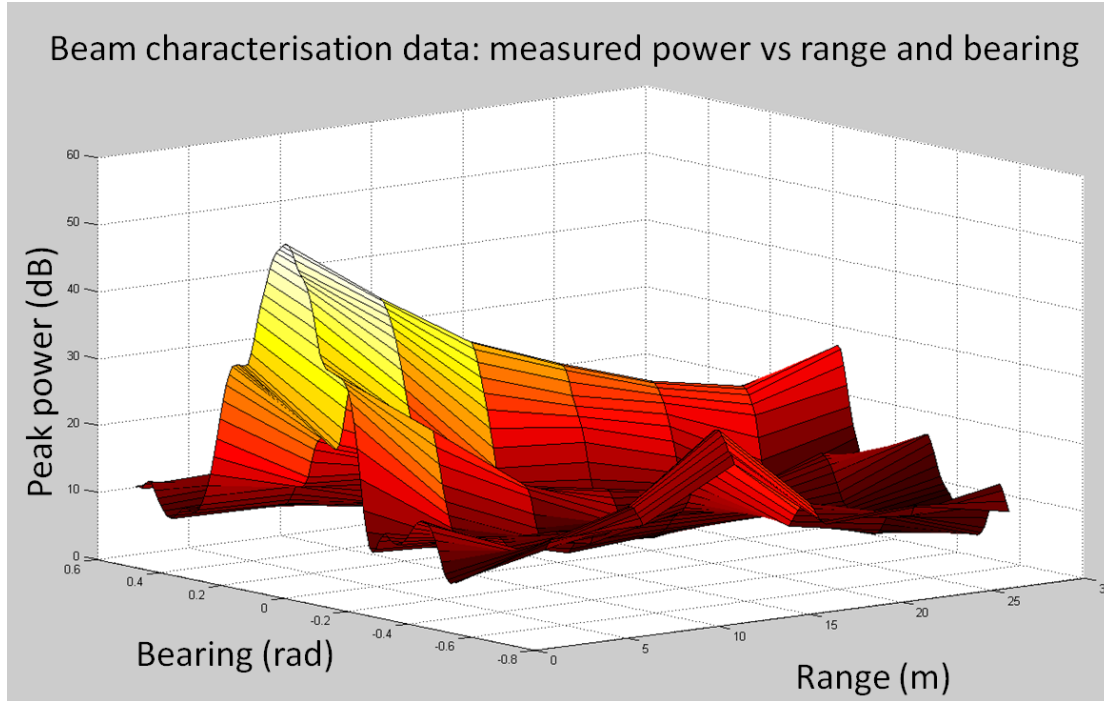
Figure 3.19: The CPD radar beam profile as measured at different ranges using a ≈1m2 RCS target on the zero-bearing line, plotted in three dimensions. The mainlobe and sidelobes can be seen clearly, with sidelobes dying off around 10m. The mainlobe appears to pick up again at 30m, but this is due to a wall in the environment at that range.

Sidelobes are clearly visible in Fig. 3.18, approximately 0.2 radians either side of the mainlobe. No windowing is used, so the sidelobes are expected to be 26dB lower than the mainlobe [4]; here they are approximately 22dB, which is within tolerance, though this varies with the target and test scenario; the difference may be due to noise, clutter or experimental error. Fig. 3.20 tracks the power values of the mainlobe and the sidelobes for a 10m^2 RCS@24GHz corner reflector over ranges from 5m to 30m. The difference is fairly consistently around 20dB.
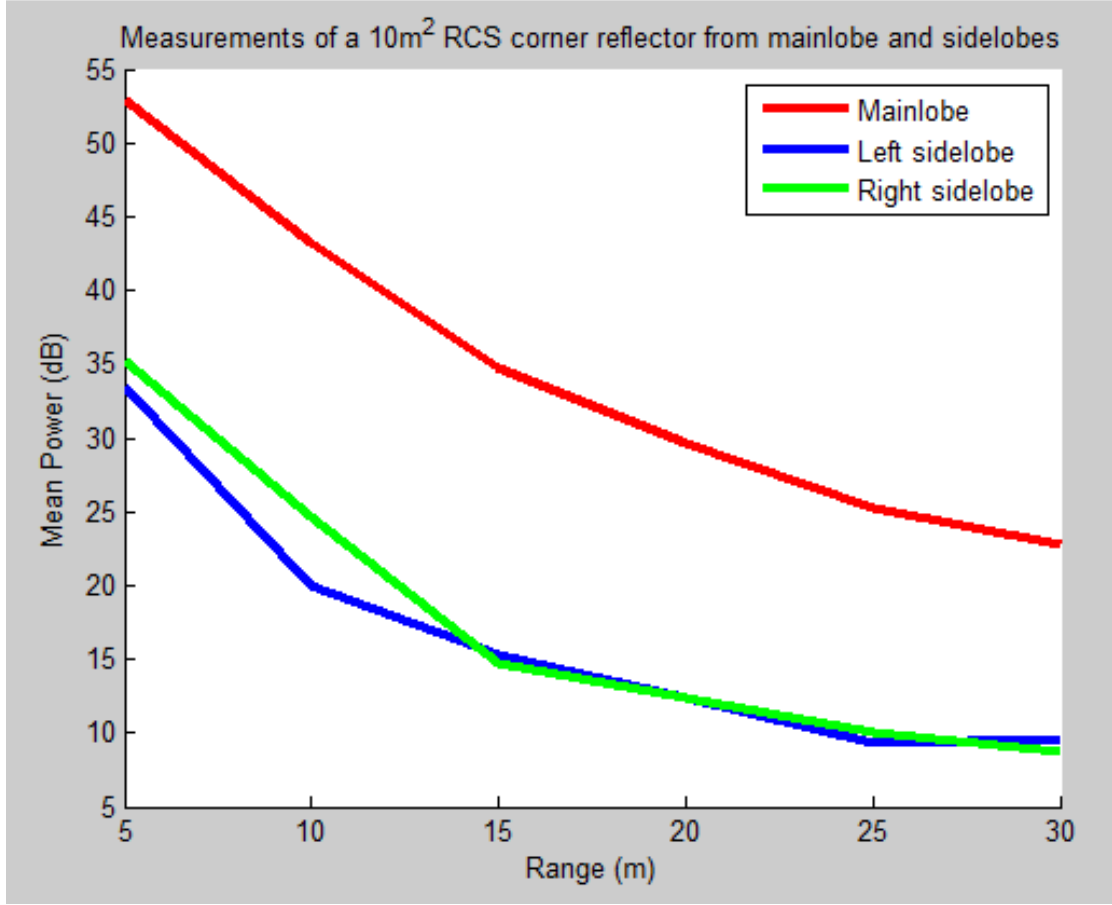
Figure 3.20: The CPD radar sidelobes are approximately 20dB lower than the mainlobe.

3.3.2.4 Radar Sensitivity and the RCS of Natural Targets

The CPD radar needs to be able to detect light vehicles, human beings and a wide range of other objects such as buildings, walls, trees, posts and poles.

By varying the RCS and range of the known targets and comparing measurements to unknown targets, the RCS of the unknown targets at corresponding ranges can be calculated. The RCS of a human being and a light vehicle are calculated here.

At a constant range, the radar's power density S_i (W/m^2) should be constant; therefore, using Eq. 2.1 $P_r = \sigma S_i$ for two different targets,

$$\frac{P_{r,1}}{\sigma_1} = \frac{P_{r,2}}{\sigma_2} \tag{3.10}$$

$$\sigma_2 = \sigma_1 \frac{P_{r,2}}{P_{r,1}} \tag{3.11}$$

The unknown RCS σ_2 can be calculated using Eq. 3.11, using measurements made of it and a known target at equal ranges.
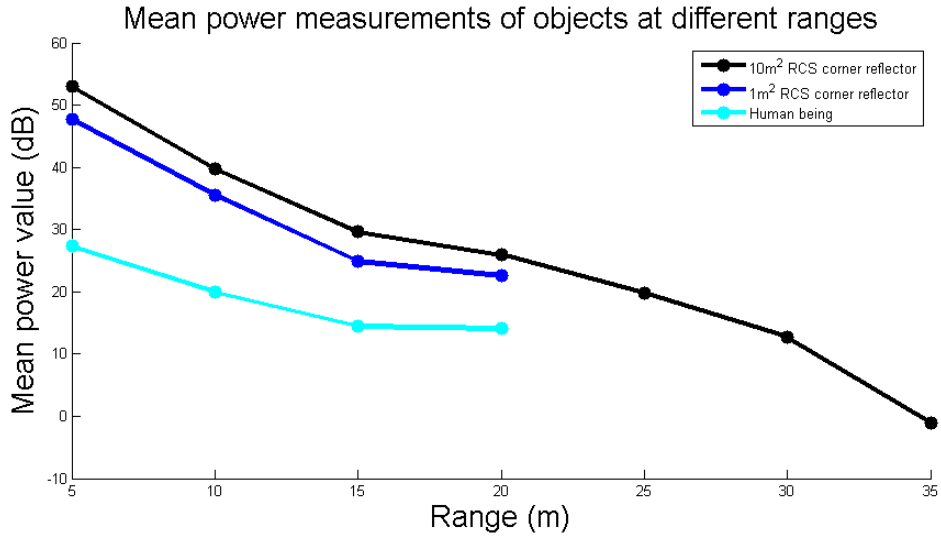
Figure 3.21: Received power from measurements of a $10m^2$ and $1m^2$ RCS targets, and a standing human being, at 5m distance increment.

Fig. 3.21 shows the different power values obtained from tests using different sized corner reflectors and a human being. From this data, the RCS of the human being can be calculated.

Due to the statistical nature of radar measurements, the RCS of the human covers a distribution, between 0.05 and 1.0 m^2 with a mean of around 0.22m^2. Fig. 3.22 shows a histogram of human RCS, calculated over ranges between 3m and 25m. Measurements are clustered around 0.20m^2 (or less) but there is a long tail.
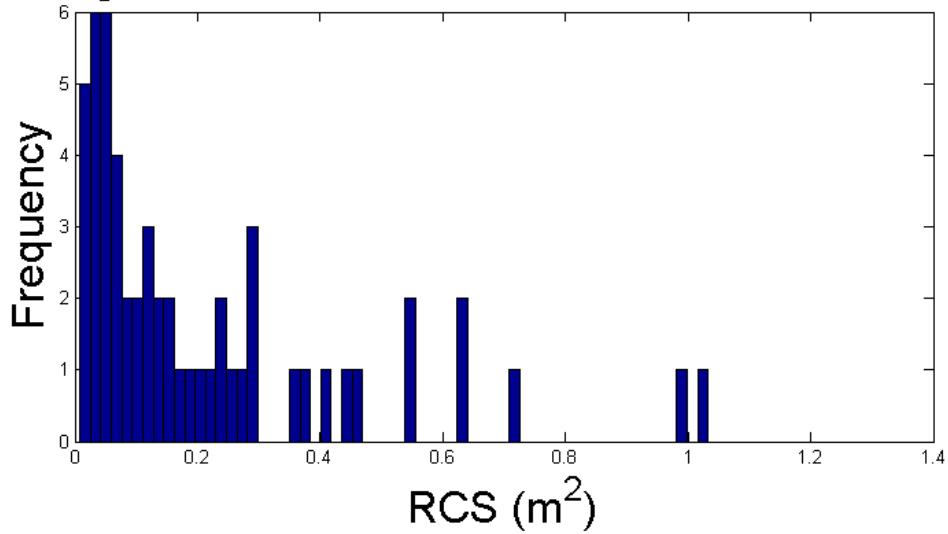
Figure 3.22: As this histogram shows, the RCS of a human being at 24GHz has a long-tailed distribution, with a peak around $0.1m^2$ and a mean around $0.2m^2$.

3.3.2.5 Vehicle Detection

It is important to be confident that the CPD radar can detect a light vehicle regardless of the angle of orientation of the vehicle relative to the radar. If a vehicle is invisible to the radar at a certain range and orientation, that presents a potential collision risk. Tests were performed across a range of aspect angles. Assuming that the light vehicle is roughly symmetrical, it is enough to rotate the radar 180° around the vehicle's centroid from front to back, maintaining a distance of 10m from the centroid. Fig. 3.23 shows the vehicle set up for testing.

Figure 3.23: The target vehicle set up for radar testing.

The literature shows that the RCS of a vehicle does fluctuate significantly with aspect angle. Typically it is strongest when viewing the flat surfaces of the vehicle straight-on, that is, viewing the vehicle from the front, rear or side. The RCS is lowest when viewing the vehicle at an oblique angle where most of the radio energy is reflected away from the receiver. Table 3.2 shows some approximate RCS values [4]. These are for illustrative purposes only; RCS changes as a function of bearing to the vehicle, so it cannot be represented by just one number. Large fluctuations in RCS are present due to phase interference effects discussed in Sec. 2.3.2.1.

Table 3.2: Approximate radar cross-section of vehicles

Target	Direction	RCS range (m^2)	Approx. RCS average (m^2)
	Front	1.5-150	15
Ute	Sides	0.5-140	1.5
	Back	0.5-140	5
	Front	0.5-10	2
Truck	Sides	0.5-10	2
	Back	0.5-130	4

Fig. 3.24 shows the calculated RCS of a light vehicle (specifically, a Toyota Land Cruiser) that had the CPD radar rotated around it at resolution of approximately $10°$.

An aspect angle of $0°$ corresponds to the vehicle facing towards the radar head on, $90°$ is side-on and $180°$ is the rear of the vehicle facing the radar. The green line is RCS calculated from unscanned radar measurements, made with the radar's beam fixed and pointing directly at the vehicle. The red line is calculated from the highest power measurement when the beam is scanned across the vehicle while it sits at different orientations relative to the radar. Both are approximately the same shape, with peaks around $0°$, $90°$ and $180°$ as would be expected when there are large, flat metal panels presented towards the radar.
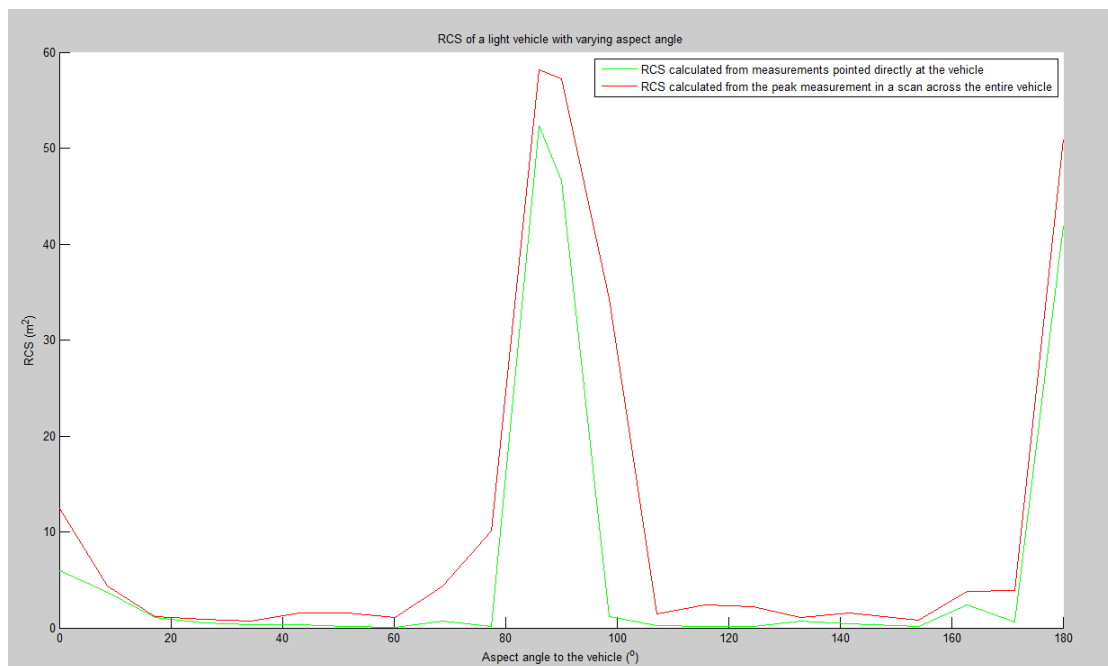
Figure 3.24: The radar cross section of a light vehicle varies greatly with varying aspect angle.

Fig. 3.25 shows the same data plotted in decibel square metres (dBm^2) at 24GHz as opposed to square metres. Compressing the vertical scale this way prevents the large RCS values from dominating so much. It becomes clear that the RCS gets quite small at particular angles when using only the measurements made while the beam is pointed directly at the radar, down to a minimum of around $-10dBm^2$ or $0.1m^2$, which is smaller than some measurements of a human being! When data from the entire scan of the vehicle is considered, the minimum RCS of the light vehicle is around $-1dB,^2$, or $0.794m^2$, substantially larger than the unscanned case. Therefore, at particular aspect angles, any single CPD radar measurement has the chance of receiving a very weak

reflection from a light vehicle, but through an entire scan there will be some stronger measurements. As long as the detection threshold is designed with this consideration in mind, the CPD radar will be able to detect the light vehicle regardless of aspect angle.
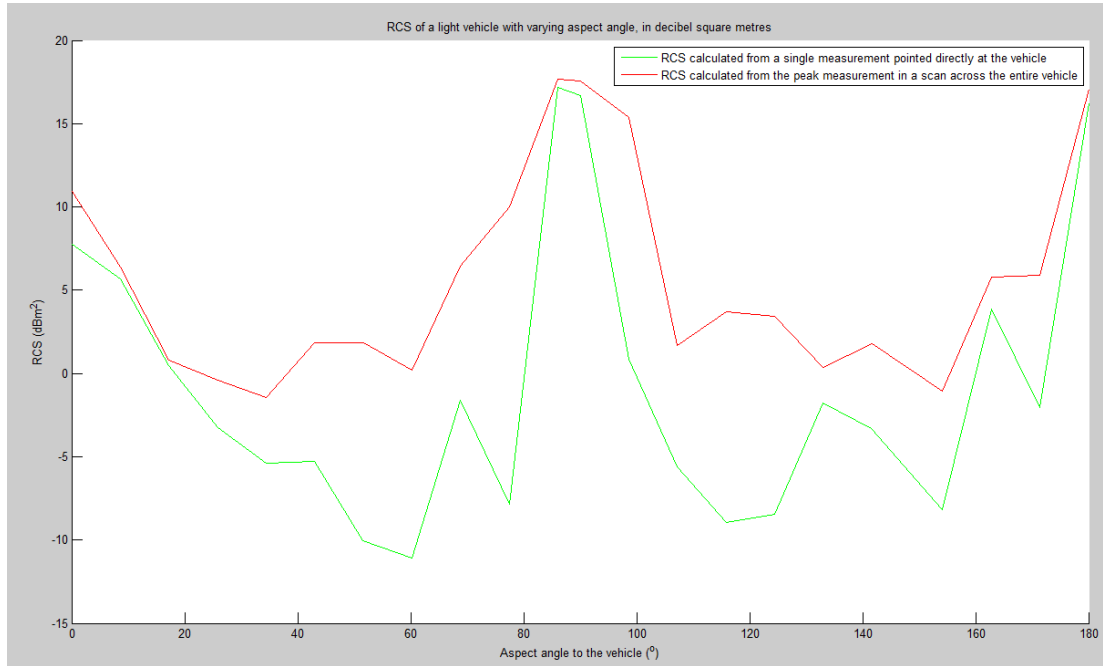
Figure 3.25: Plotting the RCS of the light vehicle with varying aspect in decimal square metres reduces the dominance of the larger RCS values. For reference, $0\text{dBm}^2 = 1\text{m}^2$ and $-10\text{dBm}^2 = .1 \text{ m}^2$.

3.3.2.6 Pedestrian Detection

Rather than a rigorous analysis of a human target's radar cross-section as a function of aspect angle, tests were performed to investigate the qualities of measurements from human beings. These included whether a human being could be detected equally well from the front, back and side, or if a human being lying on the ground could be detected at all.

A human being was scanned facing towards the radar at the same ranges as the corner reflector(s), as well as at a range of 10m facing towards the radar, away from the radar, and lying on the ground. Fig. 3.26 shows the poses that the human was scanned in and the plots of these scans.

Human target tests: poses and plots
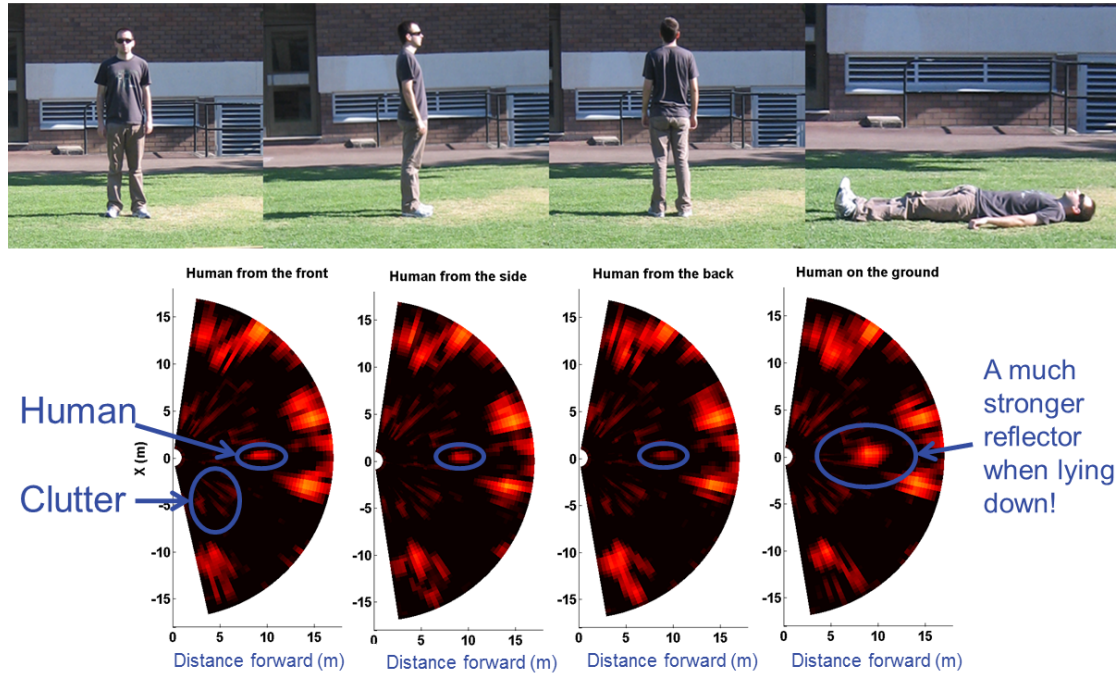
Figure 3.26: The poses that the human was scanned in, and the CPD radar scans of the human. Some clutter is present in the measurements; the human being is the largest blob at a range of 10m directly in front of the radar.

As expected a human being was a much weaker target than a vehicle. Unexpectedly, a human being has a much larger RCS when lying on the ground perpendicular to the beam axis. This could be because the corner between the side of their body and the ground forms a strong reflector. Importantly, the tests confirmed that a human being could be detected by the radar above the surrounding clutter levels.

3.3.2.7 All-Weather Performance

The ability of microwave and millimetre-wave radar to penetrate rain and fog is well established, and testing for these conditions is done in [21].

It is assumed the CPD will suffer little power attenuation in rain and fog according to the values given in Fig. 2.2; at 24GHz, the worst attenuation in 150mm/hr of rain is on the order of 10dB/km.

To test the radar's performance under the effect of mud on the sensor housing, tests

are performed with mud coating the sensor. The radar is scanned across a controlled environment containing a corner reflector on a tripod, trees, lights on poles, and a building in the background. The test is repeated with a thin layer of mud applied to the front of the radar housing. At its thickest, the mud was approximately 1cm thick. The mud used here was produced by mixing soil with water and applied by hand. Different soil types may have different effects. It should be noted that this mud was quite wet; drier mud may stick more effectively, but its lower water content would prevent it from blocking as much radar power.

For comparison purposes a SICK LMS-200 scanning laser rangefinder undergoes the same tests, having approximately the same density of mud applied to the front of its housing. The sensors are shown with their mud coating in Fig. 3.27. A similar test could have been performed with a camera, but the result would have been a foregone conclusion; there is no way for a visible light camera to see through mud on its lens.

Figure 3.27: The sensors with a light mud coating.

The test scene consisted of an open field with a corner reflector on a tripod, with a building and several trees and light poles in the background. The results are shown in Fig. 3.28. Before the mud is applied, the radar is able to detect the corner reflector at 7m, as well as the trees, light pole and building behind it. The laser rangefinder detects all these also, as well as objects in the wider field of view such as a hedge to the left and more trees scattered around the lawn, but is unable to detect the corner reflector on the
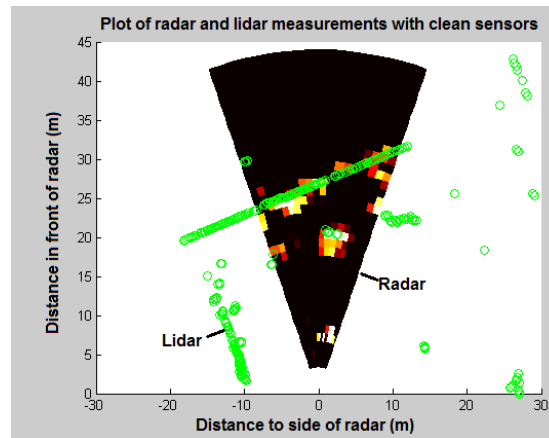
(a) Plot of the environment with no mud on the sensors. The long, straight building in front of the sensors is clearly visible, as are several trees and other objects in front of it.
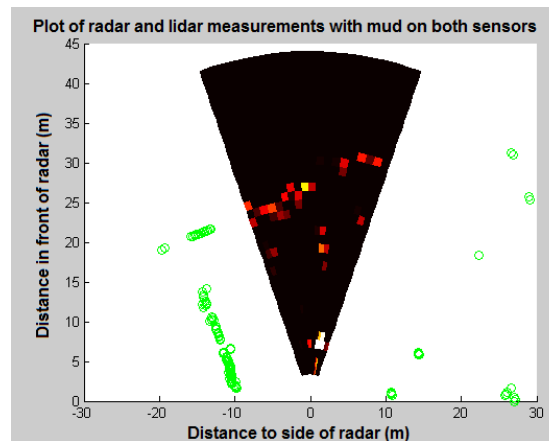
(b) Plot of the environment with mud on the sensors. In the mud-affected region, the lidar cannot make measurements. The CPD radar's measurements are of lower power (as shown by the dimmer colours in the cells of the plot) but the objects are still detected.

Figure 3.28: Radar and lidar plots of the environment with clean and muddy sensors. All dimensions are in metres. The lidar measurements are green circles across a region close to 180°. The radar measurements cover a much narrower region for these tests; the brighter squares correspond to high power values in the radar measurements (detected objects), while dimmer or black regions are low power (empty space). The lidar has better angular and range resolution than the radar and performs well in the clean test, but is unable to detect anything directly in front of it when muddy. When muddy the radar was slightly degraded, but still able to perform.

tripod at close range (it is mounted too low to detect the corner reflector itself, and the legs of the tripod are evidently too narrow or do not reflect the laser back to the sensor).

Once the mud is applied, the laser rangefinder becomes completely nonfunctional in the mud-affected sector straight in front of it. The CPD radar's measurements are weaker, particularly at longer range, but are still more than strong enough to detect all the same key features as earlier including the target, trees and building behind it.

This demonstrates the more graceful degradation of radar measurements in bad weather. It is unlikely that a mud layer quite this thick would be applied to the sensor (though it is still possible). A thinner layer of mud would absorb proportionately less energy and allow improved performance, but a laser rangefinder would have little improvement unless it was almost totally clear of mud.

3.4   Detection Threshold Design

The previous section outlines tests done in a simple, controlled environment where target positions and characteristics are known. Processing of radar data is less complicated under these circumstances than in complex, cluttered environments. A simple uniform detection threshold can be tuned by hand until it produces a plot where the known targets are visible, and the empty spaces appear dark. This processing technique cannot be applied generally to larger environments that are not known and controlled.

A single radar measurement consists of a vector of bins containing power values proportionate to the RCS of the objects within the beam volume, as well as noise contributions [14]. Any, all or none of these may be due to objects in the environment. As Sec. 2.5 outlined, a **detection filter** is needed to identify the measurements that correspond to valid targets as opposed to noise or clutter. A detection filter is a way of calculating the minimum required power value required to be classified as an object being detected, rather than noise or clutter. In plotting raw radar measurements, it acts as a detection threshold; above the threshold are power values due to objects in space, and below the threshold are noise and clutter values. In the following section, different types of detection filters are examined for use with the CPD radar.

3.4.1 Different Types of Detection Filter

3.4.1.1 Adaptive Constant False Alarm Rate (CFAR) Threshold

An adaptive constant false alarm rate (CFAR) threshold would be desirable but will not perform well with the short measurement vector of the CPD radar. As Sec. 2.5 outlined, a CFAR threshold is recalculated for each measurement using the values in the measurement itself. CFAR is widely used in radar processing, but unfortunately, CFAR does not perform as well for the CPD radar. CFAR relies on a number of assumptions:

- There are a large number of bins per measurement;

- There are a small number of targets detected per measurement (ideally only one in the sampling window, though some implementations can handle more [35]);

- Each target is small and occupies a low number of consecutive bins (ideally just one bin per target).

The CPD radar fails on these three counts:

- The measurement vector only has approximately 30 reliable bins per measurement;

- In a crowded environment, it may detect a large number of targets in a single measurement;

- Typical targets such as vehicles and buildings are longer than the bin and so take up several bins.

The net effect of these properties of the CPD radar is that strong targets can mask weaker ones nearby, and extended targets can appear "hollowed out" if the CFAR sampling window is too narrow.

A fixed detection threshold is instead calculated separately for each bin. This has the disadvantage of requiring the environment and target characteristics to be known beforehand, but this data is fairly quick to acquire, and once obtained the fixed threshold will allow the radar to be used reliably.

3.4.1.2   Uniform Static Threshold

The simplest static threshold is a uniform threshold. This threshold is equal for all bins i in the measurement:

$$UniformThreshold[i] = k \qquad (3.12)$$

for some constant k. This is fairly ad-hoc; it needs to be tuned by hand based on the clutter levels in the environment and the typical power value of detected objects.

Uniform static thresholds are simple to implement and produce good results for short ranges; the radar plots shown in this and previous chapters were mostly generated using uniform static thresholds. However they are an ad-hoc solution and their performance is limited by the fact that radar measurement power drops inversely proportional to range raised to the fourth power (as described in Eq. 2.5). In decibel form, this is a dropoff rate of $-40log_{10}R$ (dB/m). A uniform threshold cannot suit both the close and long range; it will either be too low to filter out all the noise at close range, or too high to detect weaker targets at long range. As the close proximity problem is most concerned with short-range hazards, using a uniform threshold in this application would require it be selected for short-range performance. This would hamper its performance at longer range.

Since close-proximity accidents occur at close range, greater consideration needs to be given to the short-range performance, which means the sacrifice of long range detection capability, if a uniform threshold were used.

3.4.1.3   Decreasing Static Thresholds

Since the reflected power decreases at longer range, a detection threshold that similarly decreases with range could be used. Eq. 2.7 shows a dropoff in power at range R of $-40log_{10}R$ (dB), but some sort of offset would be needed to account for the other terms in the equation. Fig. 3.21 shows mean power measurements of targets with different RCS following the logarithmic decrease that is expected, with different vertical translations because of their different reflectivity.

The threshold for bin i can be defined as:

$$DecreasingThreshold[i] = k_1 - k_2 * log_{10}(i * binLength) \qquad (3.13)$$

This approach is beginning to use the detection properties of the radar, but still relies

on manual adjustments to determine offset k_1 and scaler k_2. A higher k_1 will produce a threshold that filters out clutter better but will lower the probability of detection for weaker targets. A higher k_2 means a faster decrease. An intuitive value for k_2 is 40, since the radar range equation suggests a power drop-off of $-40log_{10}R$. Testing with empirical data will be needed to investigate this.
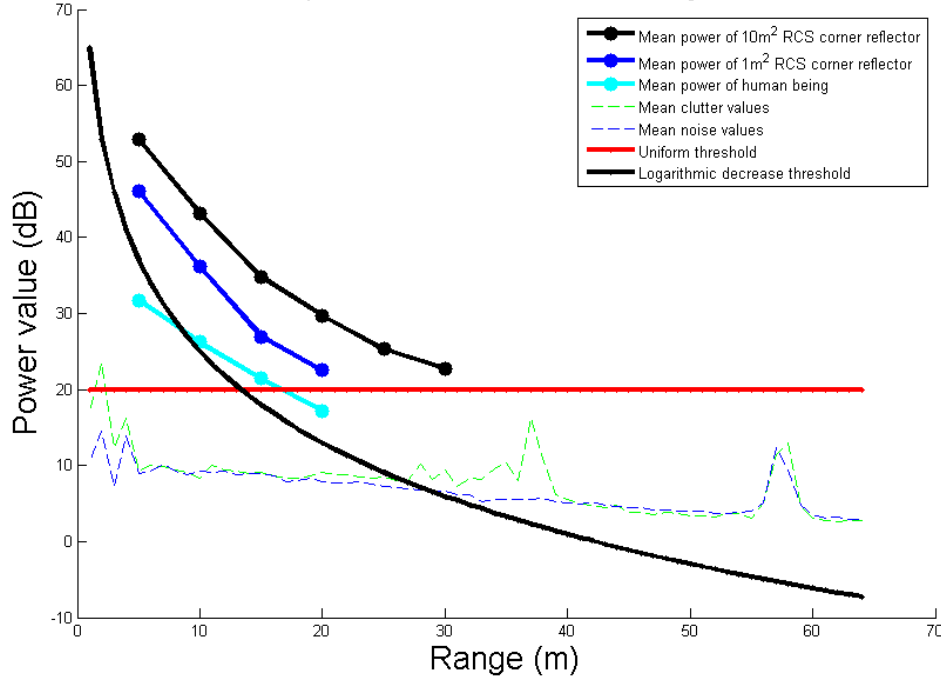
Figure 3.29: The mean power measurements of different-sized targets and the uniform and decreasing static thresholds are plotted here. The uniform threshold has a value of $k = 20$dB. The decreasing threshold has $k_1 = 65$dB and $k_2 = 40$.

3.4.1.4 Translated mean noise power threshold

Another approach to threshold design is to analyse the typical power levels for noise, clutter and/or targets in each bin, and use these as a starting point for the threshold. The mean noise power for each bin can be used as a baseline for shape and some offset added to translate this above the mean. A uniform offset would preserve the shape of the mean noise threshold, requiring a constant ratio of power above the mean to exceed the threshold. For example, an offset of 3dB would be a threshold that is twice the average power of the noise. Alternatively, the offset could vary with each bin - but this

raises the question of exactly how to vary the offset. The exact value of the required offset will depend on the sensor's dynamic range and the typical target power value.

The threshold for bin *i* can be defined as:

$$MeanNoiseThreshold[i] = mean(noise[i]) + k[i] \qquad (3.14)$$

where *noise*[i] is the set of noise measurements logged for bin *i*, and *k*[i] is some offset for bin *i*.

Pure noise measurements can be obtained by pointing the radar at the sky and logging data for several minutes. If the mean clutter measurement were to be used instead, this would involve measuring a typical environment that is devoid of targets, preferably a large area viewed from several different positions to mitigate any local variation and prevent overfitting to that region.

This approach is more closely grounded in the radar's noise and detection characteristics than the previous approaches. It is still fairly easy to implement and provides decent results. The amount of testing required is fairly minimal. But the value of the offset still needs to be tuned by hand. Using a constant offset is assuming a constant ratio between real object measurement power and noise or clutter power; this is an ad-hoc method. And using the mean of the noise distribution as the start point is also ad-hoc; it is simply an easy quantity to calculate. The Rayleigh probability density function is not symmetrical and the mean of a distribution is not especially representative of its shape.

3.4.1.5   Noise/Clutter Power Distribution-based threshold

It makes more sense to use the shape of the power value distribution to calculate the threshold, rather than taking a property of it and adding an offset. Numerical integration of the noise distribution can be used to determine a threshold that a given proportion of the noise values fall below.

The probability distribution function (PDF) of each bin's noise measurements can be calculated by taking a histogram of the measurements and dividing by its integral.

In keeping with the standard used in CFAR, we chose a desired $P_{fa} = 10^{-6}$. The detection threshold will be a point on the horizontal (power) axis that divides the area under the noise PDF in two - the noise measurements that will be (correctly) classified as noise to the left, and the one-in-one-million noise measurements that will be misclassified as targets (false alarms) to the right. The area of the false-alarm section will
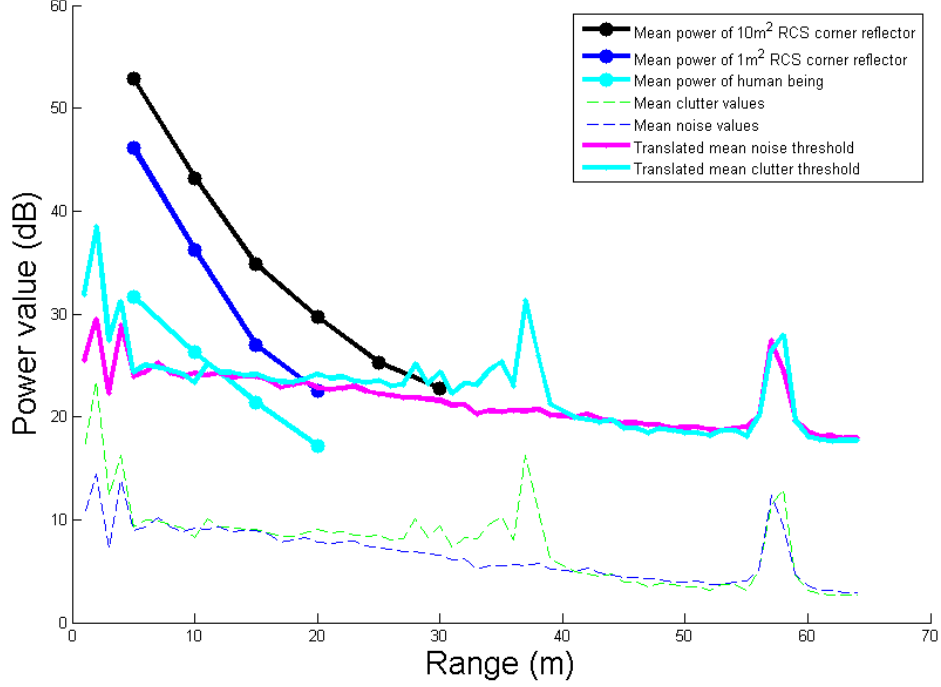
Figure 3.30: The mean power measurements of different-sized targets and the translated mean noise and clutter thresholds are plotted here. The mean values of noise or clutter are calculated for each bin, and are translated by a constant value of 15dB.

be equal to P_{fa}. This is easiest done by converting measurements from dB to a linear power scale and integrating with respect to power via the trapezoidal rule.

For bin *i*, the threshold can be calculated as:

$$DistributionThreshold[i] = \int_a^\infty noise[i] = 10^{-6} \qquad (3.15)$$

As with the mean noise power threshold, the clutter distribution can be used to calculate the threshold instead of from noise measurements.

This approach is grounded in empirical data of the radar's performance, but the results are strongly affected by the quality of the measurements that are used to calculate it. The larger the set of measurements the more reliable they will be, but this can be difficult to gather, particularly for clutter measurements; real clutter environments are rarely homogeneous and local clutter variations will be reflected in the calculated threshold unless a very large set of clutter samples are taken from multiple positions

and perspectives. Noise within the radar may not be consistent across all bins; for the CPD radar, radar frontend used has a defect causing a constant-frequency noise signal at bin 40.
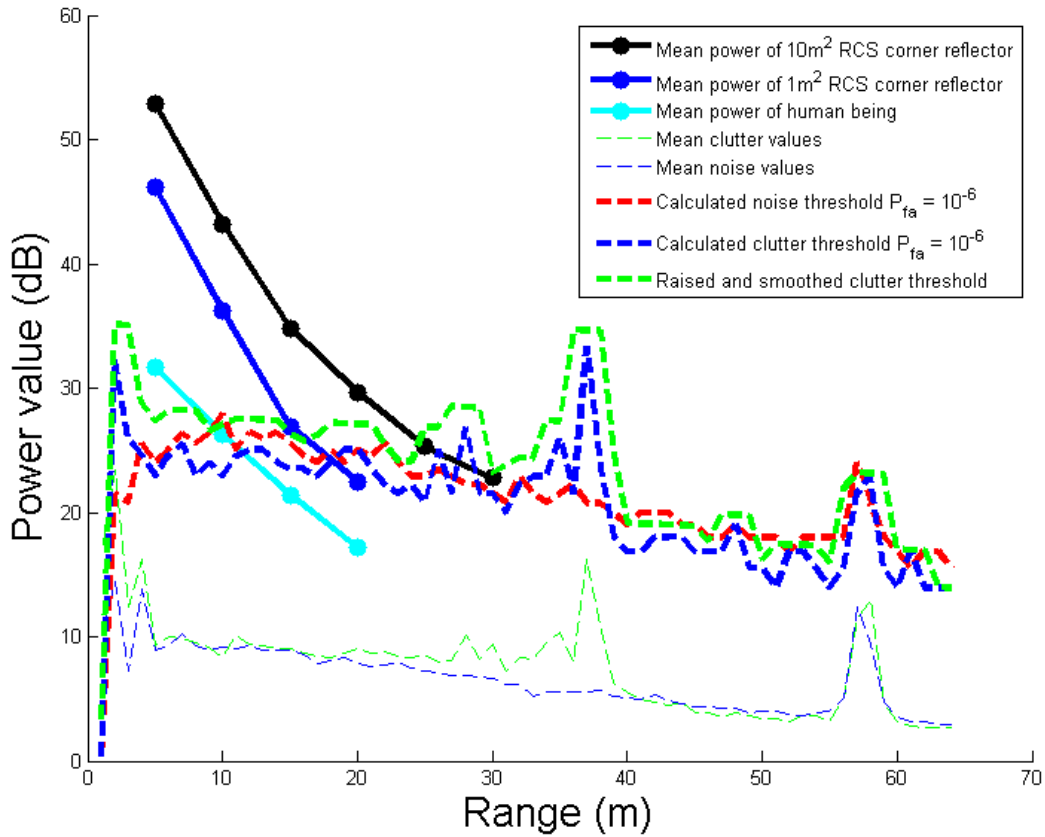
Figure 3.31: The mean power measurements of different-sized targets and the translated mean noise and clutter thresholds are plotted here. The mean values of noise or clutter are calculated for each bin, and are translated up a constant value of 15dB to produce a detection threshold.

3.4.1.6   Modified Noise/Clutter Power Distribution-based threshold

In practice, due to unevenness in the environment and additional noise within the sensor, the calculated distribution-based thresholds do not produce a smooth curve, but have substantial fluctuation from bin to bin. This will produce inconsistent results in

environments even slightly difficult from the one used to gather the calibration measurements. Modifications can be made to the calculated distribution-based threshold to improve the performance.

To reduce some of the signal variability in the early bins without lowering the threshold, a local maximum filter with a window of 3 bins is passed over the clutter thresholds. This had the effect of raising some bins and smoothing the overall curve.

$$RaisedDistributionThreshold[i] = max(DistributionThreshold[i-1]...DistributionThreshold[i+1])$$
$$(3.16)$$

The threshold was still found to be too low in some cases; the clutter environment used to generate the data was not as rough as some of the other environments where the radar would be used. This is a problem with generating a static threshold - it can be difficult to anticipate the clutter levels of all environments where it will be employed. The problem was that clutter was stronger at close range than previously encountered. To overcome this, a decreasing offset was added to all bins.

$$OffsetRaisedNoiseDistributionThresh[i] = RaisedNoiseDistributionThresh[i] + k_1 - k_2 * i$$
$$(3.17)$$

In bin 0, the offset was 3dB, declining to 0dB in bin 63; that is, $k_1 = 3$dB and $k_2 = \frac{-3}{63}$dB.

This threshold achieves good results but introduces ad-hoc approaches and multiple parameters in need of manual tuning (the smoothing filter used, the decreasing offset).

3.5   Comparison of Detection Thresholds

A number of different thresholds have been considered in Sec. 3.4: dynamic CFAR, and static uniform, decreasing logarithmic, translated mean noise, translated mean clutter, calculated noise, calculated clutter, and raised-and-smoothed calculated clutter thresholds. To select the threshold to use to develop a close proximity system, the performance of these different thresholds need to be compared under the same test conditions.

3.5.1   Testing Procedure

The CPD radar is used to scan different-size targets over a range of distances, starting from 0m to a maximum of approximately 30m, using a configuration much like in

Fig. 3.17. The tripod is either carried back continuously over this distance, or in 5m increments, pausing at each distance. The targets used are a 10m^2 and 1m^2 RCS corner reflector, and a human being. The mean measurement powers for these targets are plotted in Fig. 3.21. A lidar is also used in these tests to provide ground truth.

To compare the performance of different detection thresholds, the optimal subpattern assignment (OSPA) metric is used [60]. It calculates a value proportional to the "incorrectness" of a set of estimates compared to a "truth" set, which in this thesis will be the mean distance to a target measured by the lidar.

For two sets $X, Y \in W$, the set of all permutations on $\{1, 2, ..., k\}$ for any natural number $k > 1$ is calculated and denoted as Π_k. These permutations are used to calculate all possible assignments between the elements (x_i, y_i) to minimise the total distance between X and Y - the optimal subpattern is the subset of X that is closest to Y. That is, for a distance function $d(x, y)$ that calculates the distance between elements x and y, we calculate the association of elements between X and Y such that the assignment chosen has the minimum total distance between them. Assuming X has m elements and Y has n elements, and $m \leq n$, we wish to determine

$$\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d(x_i, y_i) \tag{3.18}$$

To calculate the OSPA metric, a cutoff c is implemented on the distance, to produce a saturated distance function:

$$d^{(c)} := min(c, d(x, y)) \tag{3.19}$$

Using a power term p as well, the OSPA metric of order p with cutoff c is defined as[60]:

$$d_p^{(c)}(X, Y) := \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}(x_i, y_i)^p + c^p(n - m) \right) \right) \tag{3.20}$$

Equation 3.20 assumes $m \leq n$; if $n > m$ then $d_p^{(c)}(X, Y) := d_p^{(c)}(Y, X)$.

The metric calculation requires two parameters to be set, the order parameter p and the cutoff parameter c. The order parameter determines the exponent to which errors are raised; larger p values punish outliers more heavily. The cutoff specifies the maximum distance error value to be assigned to any poorly matched (or unmatched) points in the

sets being compared [60]. The standard value of p is 2 [60], while c is implementation-dependent; at its most it should be the maximum localisation error possible, which will be very penalising for cardinality errors (that is, differences in the lengths of the sets being compared). Lower values of c place a greater emphasis on localisation than cardinality of the sets.

The order parameter p and the cutoff parameter c are both set to 2 in this thesis. A cutoff of 2m indicates that this is the maximum that a range measurement to a single target would be expected to reasonably deviate between sets of measurements. This is a function of the scale that the CPD radar operates at and its range resolution. The Matlab implementation of OSPA metric by Ba-Ngu Vo [71] is used throughout this thesis to quantify the performance of filters under different conditions.

The calculated thresholds are tested using six datasets of three different targets performing two different motions:

- A human being walking from a distance of 0m to 25m slowly and continuously directly in front of the radar.

- A human being walking from 0m to 25m, in rapid increments of 5m with long pauses after each increment (that is, most measurements will be at distances of 5m, 10m, 15m, 20m and 25m).

- The same two tests performed carrying the 1m^2 RCS corner reflector.

- The same two tests performed carrying the 10m^2 RCS corner reflector.

The radar scans 160° across the environment, but only the middle 20° is considered for this test, as this is where the targets are. All tests are performed at a scan speed of 500°/s.

Each threshold is subtracted from the measurements (or, for CFAR, a threshold is calculated for each measurement and then subtracted). Any bins with positive residual power values (that is, power values that were greater than the threshold for that bin) are counted as detections, and compared to the mean laser measurement of the target at the closest time to calculate the OSPA metric for that timestep.

Ideally a single power value will exceed the threshold, which would be a single detection corresponding to a point target at the position of the true target. Realistically, the slow scan speed and width of the beam will cause multiple consecutive detections

as the radar is scanned across the target. The finite range resolution, chirp nonlinearity and imperfect detection thresholds will result in some detections spanning multiple bins in a single measurement.

Fig. 3.32 shows the results of these tests for a human target moving continuously, and radar measurements being processed using a uniform threshold. The difficulty in designing a uniform detection threshold is clear; at close range, there are too many measurements above the threshold (as there are multiple radar measurements for each laser scan). At longer range, there are missed detections, evidenced by gaps in the red marks.

Fig. 3.33 shows the results for a $1m^2$ target moving in 5m increments, and radar measurements being processed using a smoothed and translated calculated clutter threshold. At close range, the detection threshold works well, as there is typically only one radar measurement above the detection threshold. At longer range, it is less effective, and there are missed detections.

Figure 3.32: This figure shows the mean laser measurements of a human being moving continuously away from the radar, in green. The radar measurements that exceed a uniform detection threshold are shown in red.

Fig. 3.34 shows the means of the OSPA values for each different detection threshold over a different range. A lower OSPA value is better, as it shows lower error in location

Laser and radar detections of a 1m^2 RCS target walking backwards in 5m increments.
Raised and smoothed calculated clutter threshold used for radar.
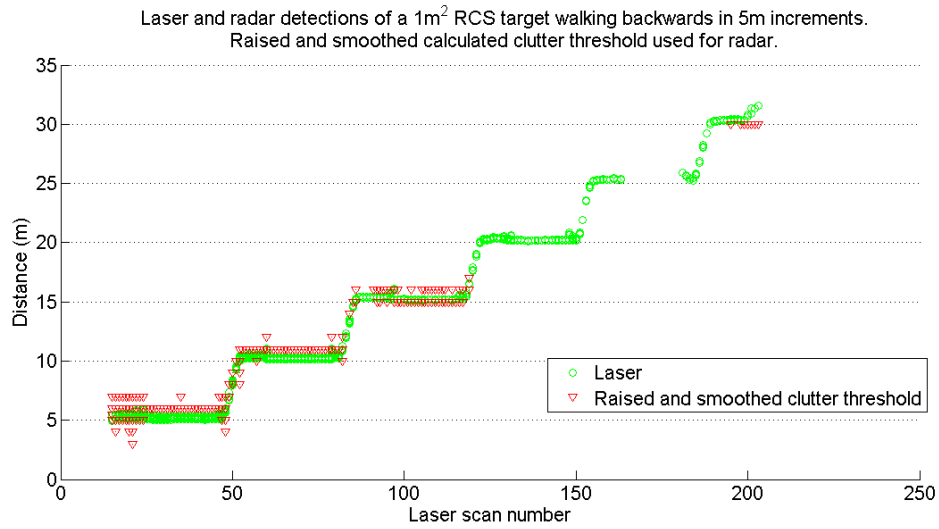
Figure 3.33: This figure shows the mean laser measurements of a 1m^2 RCS target moving away from the radar in increments of 5m, in green. The radar measurements that exceed a smoothed and translated calculated clutter threshold are shown in red.

and cardinality between the "truth" set (the laser) and the radar measurements. The smoothed and translated calculated clutter threshold produces the best result up to a distance of 15m, then its error increases - a result of the threshold being slightly too high and rejecting valid detections of weaker targets at longer range. On average, over 25m, it is still the best option, so it will be the one used as a detection filter for further development of the close proximity system using the CPD radar. If it were rejecting weak targets at close range, it would not be suitable for use due to the risk of collision in that scenario.

3.6 Discussion

This chapter's contributions are an analysis of the close proximity problem, the characterisation of a FMCW radar sensor to solve the close proximity problem, and the development of a method for calculating a static detection threshold for use with the radar in known environments. Understanding of the radar's performance and inner working are emphasised; much of the work described in this chapter would not be possible with a commercial radar, or the work would be much more difficult.
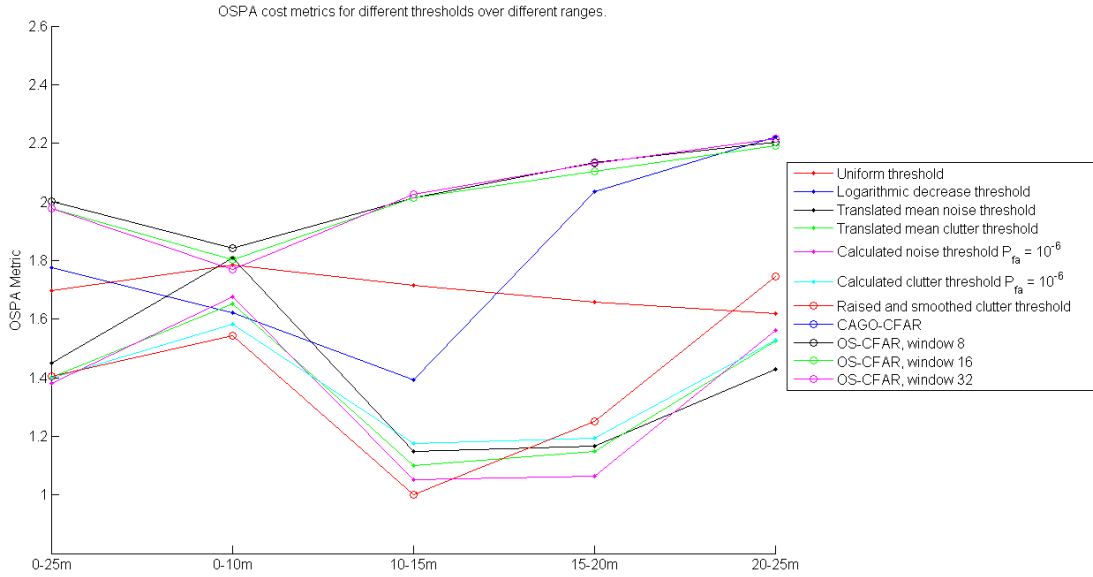
Figure 3.34: This figure shows the mean results of OSPA calculations for all targets using different thresholds, over different distance ranges. Each column is a specific distance range. Each colour corresponds to a different threshold. The lines between points are only present to help illustrate the variation in performance at different ranges for each detection threshold. The blue CAGO-CFAR line cannot be seen - it is beneath the magenta OS-CFAR, window 32 line.

3.6.1   Sensor requirements and characteristics

The analysis performed in Sec. 3.1 shows that the CPD radar is not quite ideal for solving the close proximity problem, but represents an advance over existing commercial radars in terms of bearing resolution and cost. It allows the investigation of linearisation, detection and mapping approaches that are hardware-agnostic and could be used on other FMCW radars. Experimentation must be performed to linearise the chirp in Sec. 3.3.1, but this allows the degree of linearisation to be known and controlled.

Sec. 3.3.2 described how to conduct tests to investigate a radar's sensitivity, beam width, sidelobe strength and range resolution, and offered further evidence of FMCW radar's superior all-weather performance. The measurement plots in this section showed that this sensor was able to detect light vehicles and human beings at close range, but also demonstrated the radar's sensitivity to clutter and the presence of sidelobes in the beam.

A thorough understanding of a sensor's characteristics is important for developing a close proximity system. The presence (or absence) of sidelobes, the width of the

beam at longer range, and the maximum detection range for smaller targets indicates the suitability (or unsuitability) of the radar for its intended role. One limitation of the CPD radar is that its first couple of bins are polluted by DC noise and cannot be used to detect any objects reliably. This imposes a minimum range of about 3m on the radar. This would be a problem in a close proximity application; a very small object (such as a human being) that was immediately adjacent might be difficult to detect. Larger objects, such as light vehicles, would extend into the detectable range of the sensor.

3.6.2 Detection Threshold Design

A method for designing a static detection threshold is one of the contributions of this thesis. The widely-used CFAR dynamic threshold approach underperforms in the scenario where targets are common and span multiple bins while the radar's measurement vector is fairly short. As illustrated in Sec. 3.4, there are many ways to design a static detection threshold. The method selected here, using empirical data of noise and clutter and a desired probability of false alarm and some heuristic smoothing, is not the final say in static threshold design but represents an improvement over a static uniform threshold, or a dynamic CFAR threshold in this application. Simply using the mean values of noise or clutter plus an offset (as in Sec. 3.4.1.4) provides a fast, simple method to develop a static detection threshold for a FMCW radar in a known environment. The thresholds calculated using the distributions of noise and clutter for a given false-alarm probability give marginally better performance, but the threshold curves are slightly uneven in appearance and require some heuristic modification (smoothing and raising). This could be improved by gathering more data to calculate the thresholds with.

Developing a clutter-based threshold in this way is highly contingent on the quality and quantity of the data gathered, which is itself a function of the environment that it is gathered in. Realistically, no environment has the perfectly uniform clutter distribution that is desired. Calibration data could however be gathered from many different positions in an environment totally free of targets, so that any minor differences in clutter distribution would cancel out. A rougher, more hilly, more heavily cluttered calibration environment would produce a threshold that is more robust to these conditions, but would be less able to detect smaller targets at longer ranges.

Some of the processing to modify the threshold - the raising and smoothing - is ad-hoc and based on empirical trials and error. Gathering more data from more positions

would mean that less ad-hoc smoothing and modification are needed as there would be less "overfitting" of the threshold to a limited calibration environment.

Using a static threshold trades versatility for reliability in a known environment. Designing a static detection threshold involves compromising between too many false alarms and too many missed detections. If just the "raw" data after detection is plotted (as in Fig. 3.9 and Fig. 3.26) there will probably still be some bad measurements due to clutter and noise, unless the threshold is so high that important targets can be rejected also.

3.6.3   Limitations of the Detection Threshold

The detection threshold allows the radar measurement (a vector of bins containing power values, produced from a FFT) to be converted into a list of binary values classifying each bin as either containing a target or empty space. A detection filter and the ability to plot raw measurement data are part of the solution to the close proximity problem, but still lacks some desirable features. The range and bearing to detected objects will be known, and the power value of the reflection. But a power value that is just barely over the detection threshold will be classified the same as a power value that is much, much higher than the threshold; it would make sense to make use of the amount by which the threshold is exceeded.

A single radar measurement is very limited in coverage, and measurements of the same location may provide conflicting data regarding the presence or absence of targets. The detection threshold provides no way to resolve such conflicts, so the most recent measurement is the one that will be plotted. Even with a judiciously-designed detection filter, some noise and clutter will still slip through, and some detections will be missed.

If a very fast response is required, and the radar measurements were very reliable and did not need to be scanned across a wide range (such as for a collision avoidance radar that only pointed straight ahead) then a detection threshold might be sufficient to improve situation awareness and even some level of automation (such as emergency braking). But the approach used to this thesis is focused on giving the driver more information about a wider area around the vehicle, with a high level of reliability, and a more extensive mapping approach is required to achieve this. A detection threshold is a prerequisite for the mapping approaches described in Chapters 4 and 5; the raised and smoothed clutter threshold is the one that will be used as it produced the best results so

far.

3.7 Summary

In this chapter,

- the requirements for a sensor to solve the close proximity problem have been analysed

- the close proximity detector radar has been designed to try to meet these requirements

- the CPD radar's chirp has been linearised

- the CPD radar's performance has been characterised

- a range of static detection thresholds have been designed and compared using the optimal subpattern assignment (OSPA) metric.

- one of the static detection thresholds has been selected for use in further development in solving the close proximity problem

A detection filter is useful for identifying the presence of targets within individual measurements, but sensors used for the close proximity problem will be required to produce many measurements quickly. Processing them will require some sort of memory or retention of information between measurements, a representation of the physical volume of the observed area, and a way to combine measurements of the same or different areas in a mathematically coherent way. Combining multiple measurements to represent the environment is a problem of mapping, which is discussed in Chapter 4.

Chapter 4

Static mapping with FMCW radar

The detection filter developed in Chapter 3 allows raw sensor data to be processed and displayed to the driver, but there is no modelling of the environment state beyond a list of positions that were most recently measured as containing targets or not. The filter's output is a binary value of detection or no detection, and newer observations of the same area simply overwrite older ones.

In this chapter, knowledge of the sensor's characteristics and raw sensor data are used to create a mathematical representation of the physical environment around the sensor [67]. Multiple measurements are made using the CPD radar, and they are converted into one coordinate frame and pieced together. Measurements covering the same area are fused together to filter out noise and increase confidence in the presence of detected objects using a Bayesian occupancy grid implementation. This provides richer information than a binary detection/no-detection filter, and greater reliability than a plot of raw data.

This chapter addresses the static mapping problem, where a map of the environment is created using the CPD radar when both the radar and the objects in the environment are static or moving at very low speed. This is the case when the vehicle is parked and the driver is about to start moving, or when the vehicle is being reversed slowly. The majority of close proximity accidents take place at low speeds or when the vehicle is starting from rest [40]. A close proximity system needs to provide the driver with a snapshot of the state of the environment around the vehicle in the moments before the vehicle begins to move. While information about the velocity of moving objects is useful, it is not strictly necessary; the driver is concerned with whether there is something present or not. The driver needs to be notified of the presence of any object in close

proximity, whether it is staying still, moving closer or moving away. For the purpose of this chapter, all changes in the map are expected to be gradual. The problem of dynamic mapping, when the sensor or environmental objects are moving, will be considered in Chapter 5.

The primary contribution of this chapter is a sensor model for a FMCW radar for use with a Bayesian occupancy grid in mostly static environments. The environment around the vehicle is discretised into a grid whose resolution is a function of the sensor's performance, and updated using a sensor model based on the detection threshold developed in Chapter 3. The radar beam's observed volume is projected into the 2D grid space, and measurements above the threshold are taken to indicate an object occupying the corresponding grid cell, while measurements below the threshold indicate empty space. This is implemented with the CPD radar and results are demonstrated of outdoor environments mapped in this way.

4.1 Occupancy grid mapping with the CPD radar

The first mapping approach considered is an occupancy grid, because it is fairly simple to implement and has been successfully used with other FMCW radars in the past [14].

In an occupancy grid, consecutive measurements can be fused together probabilistically, providing more accurate estimates of the environment than a single raw measurement. The state tracked in this implementation is the **occupancy** of the cell; that is, whether there is an object of interest within the region of space bounded by that cell, inclusive to the lower and left edge (relative to the radar). An "object of interest" is anything that the close proximity system should detect - other vehicles, buildings, people, obstructions such as large rocks, and anything else that the driver of a large vehicle would want to avoid driving into. The value stored in a cell is a function of the grid's estimate of the probability that the cell is occupied; for example, an occupancy value of 0.5 is maximum uncertainty (a 50/50 chance that the cell is occupied or empty), a value of 0.51 is very slight confidence of occupancy, and 0.10 is strong confidence of emptiness.

In this implementation, occupancy grid cells are updated after each CPD radar measurement using a Bayes filter, following the technique in [34]. A posterior over the cell

occupancy is calculated from the previous occupancy estimate and the newest measurement [67], as given in Eq. 2.28:

$$P(occ|R) = \frac{p(R|occ)P(occ)}{p(R|occ)P(occ) + (1 - p(R|occ))(1 - P(occ))} \tag{4.1}$$

The probability of cell occupancy $P(occ)$ will be bounded to prevent instabilities when it diverges to 0 or 1. Alternatively, the probability can be stored in log-odds form (logarithm of odds) to prevent these instabilities. The sensor model $p(R|occ)$ maps from sensor measurements (in decibels) to a probability of receiving that measurement given a particular state (as a probability between 0 and 1).

4.1.1   Assumptions

The following assumptions are made about the state of the environment for the occupancy grid and sensor model, in addition to the ones listed in Sec. 2.7.

- Since there is no prior information about the environment, all cells start with the uniform uncertainty occupancy value $P(occ) = 0.5$.

- One or no target is assumed to be detected per bin per measurement [50]. At close range this is reasonable, but at longer range bins can cover several square metres, making this assumption less accurate.

- The environment is assumed to be mostly static.

- It is assumed that

$$p(R|emp) = 1 - p(R|occ) \tag{4.2}$$

  This simplifies the Bayesian occupancy formula [34].

These are standard assumptions for grid mapping approaches in static environments. The assumption of a single target being detected per bin becomes less reasonable at longer range, where the beam diverges over a larger area. As an example, the CPD radar has a beam 6° wide in azimuth. At a range of 5m, the beam is 0.52m wide. At 25m, it is 2.6m wide, which is wide enough that the detection of multiple objects in one bin is becoming more probable.

4.1.2 Designing an Occupancy Grid for FMCW Radar

There are a number of considerations in designing an occupancy grid for use with the CPD radar:

- The grid resolution

- The bounds on the grid occupancy values

- The sensor model

4.1.3 Grid Resolution

Several factors influence and are influenced by grid resolution.

- Computational efficiency

- Aesthetics and useability (since this system is being designed for assisting a human driver)

- Dynamic response to changes in the environment

- Conflicts between adjacent bins

These will be discussed in greater detail.

4.1.3.1 Computational Efficiency

A coarser grid will allow faster calculation of measurement updates (due to the lower number of cells to update) and will require fewer calculations to project from the radar beam coordinates into the grid due to the lower number of beam-grid intersections.

Independent of all other considerations, computation requirements place a lower bound on grid cell size; if the map is to be computed in real time there is a limit to how small the cells can be. The number of cells increases quadratically with decreasing cell size, and it quickly becomes too computationally expensive to calculate the intersection and overlap (or lack of) between the bin footprint and all of the covered grid cells.

4.1.3.2   Aesthetics and useability

This system is designed to improve the situation awareness of a human driver. It needs to be as clear and easy to understand as possible. A finer grid will produce a sharper and useful map. If the grid cell size is too large, it negates the benefit of having a narrower-beam radar as the location of an object will not be known precisely enough to be useful to the driver of a large vehicle.

4.1.3.3   Dynamic response to changes in the environment

The environment is considered mostly but not totally static. New objects may appear or existing objects may move away. The grid needs to respond to this within a reasonable period. As only one object is assumed to be present within each measurement bin, that bin's update must be divided between all the cells that lie within its footprint. If the bin only intersects with a single cell, the full update will be performed on that cell. As the measurement is split between more cells, the amount that each cell is updated by decreases. This leads to a lower rate of change of the occupancy estimate. Particularly at longer range, where the wide beam ensures that the bin footprint will span multiple grid cells, a finer grid will be a grid that is slower to reflect change in the environment. With a coarser grid, the update is diluted between fewer cells and the map is more responsive to changed measurements.

Some inertia in the grid is a desirable thing, to filter out noise and clutter measurements. Too much will produce an unacceptable delay in the detection of potential threats.

Other than adjusting the cell resolution, another way to mitigate these inertia effects is to bound the cell occupancy probability, especially at longer range, as described in Sec. 4.1.4. This is only really an option when the radar is kept static, or the bounds for each cell would change as the sensor moves relative to them.

4.1.3.4   Conflicts between adjacent bins

If the grid cell size is larger than the bin length, or a cell straddles two consecutive bins, multiple bins from a single radar measurement will be projected into a single cell.

If the grid cell is large enough that multiple bins fit into it evenly, there are complications when the different bins have different occupancy states. If there is an object within one bin but another is empty, the final occupancy estimate of the cell becomes

difficult to predict and will depend heavily on the design of the perceptual model. If one bin dominates the cell, the outcome is clearer to anticipate. A cell size that is smaller than the bin length will have at most two adjacent bins within the one measurenent overlapping into it.

To some extent this phenomenon is unavoidable; at close range, a single bin will update a cell for several consecutive measurements as the angular velocity will be fairly low. But at close range the beam is more sensitive and sidelobes are stronger (see Fig. 3.18) so the object will likely be detected in multiple measurements. The number of updates from the same bin number in consecutive measurements that overlap into it are a function of the scan speed and the beam width as well as the cell size.

4.1.3.5 Selecting the grid resolution

Ideally each bin would map to one and only one grid cell. This would require every bin's footprint to be equal in length and width; unfortunately the only constant is the bin length. For the CPD radar, this is 1m.

A good range of resolution is between half a bin length and one bin length. Any larger and too much spatial information is thrown away. Any smaller and the map begins to become undesirably slow to update.

For short-range testing up to 16m, a resolution of half a bin length (0.5m) is adopted in this chapter. For longer range tests, a resolution of a full bin length (1m) is used.

4.1.4 Bounding Cell Occupancy Probability In Static Environments

The occupancy probability for each cell is bounded, for two reasons. The first is that the Bayes filter becomes numerically unstable when probabilities become very close to zero [8] [67] (unless the data is stored in log-odds form). The second is to improve the dynamic response of cells at longer range.

The update equation used in this implementation is commutative; updates of equal values of occupancy and emptiness cancel one another out (assuming the occupancy state has not saturated). This has the potential to develop a very large amount of inertia in the map; neglecting the effect of saturation, a measurement sequence of 50 measurements of emptiness followed by 50 measurements of occupancy would result in a state estimate of total uncertainty, even though the space has clearly become occupied.

Selecting the probability bounds is selecting how many measurements it takes for the occupancy or emptiness estimate to saturate, and how many it would take to reverse. If the cell occupancy/emptiness estimate cannot deviate as far from maximum uncertainty (0.5) before hitting the saturation value, it will not require as many measurements to change the other way if the environment changes.

For the purposes of static tests, the cell occupancy is bounded between 0.1 and 0.9 for all cells.

4.1.5 Sensor model design

In the occupancy grid, the sensor model will convert a radar power measurement to a probability of the measured cell being occupied for use in the Bayesian cell update.

The Bayes update for the occupancy was given in Eq. 4.1. The sensor model component is $p(R|occ)$; this is the probability of the radar producing a particular power value (R) from a bin covering a given cell, given that the cell is indeed occupied. This function treats occupancy and emptiness as a single binary state and converts a power value in dB to a value between 0 and 1, with 0 representing maximum confidence of emptiness, and 1 representing maximum confidence of occupancy. This function is necessary to map from the measurement space that the sensor operates in, to the probability space of the map.

This sensor model uses the detection threshold filter developed in Sec. 3.4 as a basis for updating the cell as empty or occupied, and extends this based on the expected measurements of the targets and the geometry of the radar beam:

- Rather than just separating a bin's power value into a classification of "object detected" or "no object detected", the sensor model needs to convert this into a probability of occupancy reflective of the level of confidence in this classification. The higher the power value above the detection threshold, the higher the probabilities of occupancy. Similarly, the lower the power value below the detection threshold, the lower the probability of occupancy. Values equal to the threshold should be 0.5 to reflect the maximum uncertainty.

- The sensor model must reflect the physical geometry of the beam and its gain profile. The detection of an object in a bin corresponds to the presence of an object within some region of space, with a larger area at increasing range. The

antenna gain follows a sync function [4], decreasing with bearing from the axis. Any detected object is most likely to lie on the mainlobe axis.

- The sensor model must consider the conversion from the polar radar coordinates to the Cartesian grid. The beam's footprint will intersect multiple grid cells in at least some bins. All cells within the footprint must be updated. When a single measurement covers several cells, the probability of occupancy must be divided among them proportionate to the areas of intersection between bin footprint and grid cell [50].

These three components are split into a power scaler, angle scaler and area scaler, as in [14]. The sensor model for updating an occupancy grid cell that lies in the footprint of the CPD radar is:

$$P_{occ}(R, \alpha, A, b|occ) = 0.5 + 0.5 * G_p(R, b)G_\theta(\alpha)G_a(A, b) \tag{4.3}$$

where $G_p(R, b)$ is the power scaler, $G_\theta(\alpha)$ is the angle scaler, and $G_a(A, b)$ is the area scaler.

4.1.5.1 Power Scaler

The power scaler is designed based on the assumption that the further a power value deviates from the detection threshold (above or below), the higher the probability of the corresponding state (occupancy or emptiness respectively). The detection threshold is used as a power level above which the cells in the bin footprint are considered to be probably occupied, and below which they are considered probably empty.

It is assumed that the detection threshold is not entirely reliable in discriminating clutter from targets, due to imperfect linearisation of the radar chirp, detections due to radar sidelobes, phase noise, the fluctuating radar cross-section of complex reflectors, imperfect data used to calculate the detection threshold and unpredictable variations in the test environment. Therefore every detection above the threshold is not assigned maximum probability of occupancy, nor is every detection below the threshold be assigned a minimum; the values scale based on how far above or below the threshold the value is.

Errors are assumed to be rare. Valid targets are expected to typically be substantially more powerful than the threshold, but due to noise and terrain that is rougher than was

calibrated for, clutter may also pass the threshold. Additionally, due to obstructions, specular measurements or interference between component reflectors, valid targets may fall below the threshold.

Because a missed detection could result in a close proximity accident, the decision is made that the occupancy filter will need to be slightly "conservative", and add objects to the map more readily than they are removed. In the region close to the threshold, the occupancy function will have a steeper gradient than the emptiness function; that is, targets will be added faster than they would be cleared if they were just below the threshold by the same amount. For measurements further from the the threshold, the emptiness function quickly ramps down to exceed the rate of addition, so targets well below the threshold will be removed swiftly. In decibel space, this can be efficiently implemented as two linear functions, one for occupancy and one for emptiness.

A maximum power threshold R_{max} and a minimum power threshold R_{min} are defined either side of the detection threshold for a bin, deviating above and below it by an occupancy delta δR_{occ} and emptiness delta δR_{emp} respectively.

$$R_t = DetectionThresholds(b) \tag{4.4}$$

$$R_{max} = R_t + \delta R_{occ} \tag{4.5}$$

$$R_{min} = R_t - \delta R_{emp} \tag{4.6}$$

These values are in decibels; in this thesis, $\delta R_{occ} = 6dB$ and $\delta R_{emp} = 12dB$ were found by trial and error to be effective values. In decibels:

$$G_p(R,b) = \begin{cases} -1 & : R \leq R_{min} \\ \frac{R-R_t}{2*(R_t-R_{min})} & : R \in (R_{min}, R_t) \\ 0 & : R = R_t \\ \frac{R-R_t}{2*(R_{max}-R_t)} & : R \in (R_t, R_{max}) \\ 1 & : R \geq R_{max} \end{cases} \tag{4.7}$$

The variation of probability of occupancy with varying power and range is shown in Fig. 4.1.

Adjusting the thresholds δR_{occ} and δR_{emp} adjusts the sensitivity of the occupancy
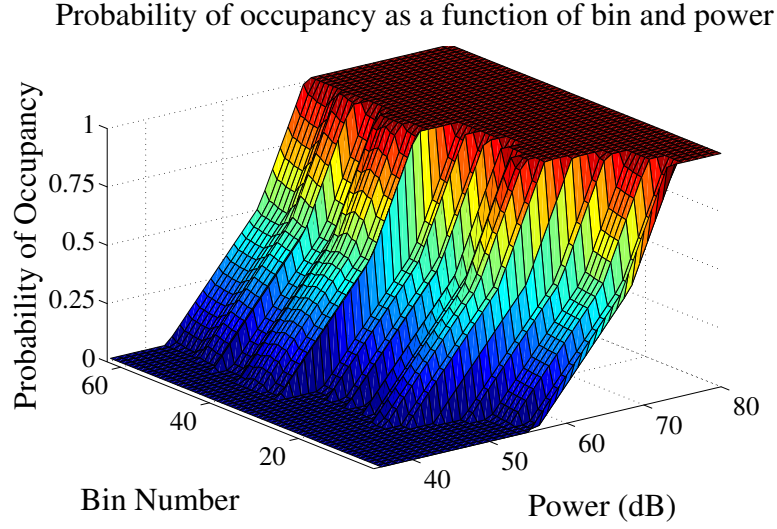
Figure 4.1: A surface plot of cell occupancy probability as a function of bin number and measurement power. This is solely a function of the power scaler G_p. Measurements are assumed to lie on the beam axis and the bin is contained completely inside the cell ($G_\theta = G_a = 1$)

or emptiness filters. Smaller δ values lead to steeper curves (visible as linear slopes in decibel power in Fig. 4.1) which give more extreme probabilities over smaller changes in power, adding or removing targets from the map faster. Larger values for these thresholds lead to gentler curves and lower rates of change.

4.1.5.2 Angle Scaler

The antenna gain pattern for a rectangular aperture is of the form of a sync function $sin(x)/x$ [4]. This can be approximated fairly accurately as an inverted parabola [14], which is less computationally expensive. The equation used is:

$$G_b(\alpha) = 1 - \frac{2\alpha^2}{\theta^2} \tag{4.8}$$

where α is the angle between the beam axis and the centroid of the intersection between the occupancy grid cell and the radar bin. θ is the mainlobe beamwidth. Fig. 4.2 also shows how it scales down P_{occ} at wider angles. The angle scaler has a range between 0.5 and 1.0, since the beam width is defined as the angle range over which the gain halves (dropping by 3dB). The gain pattern continues beyond that range, and may be able to detect particularly strong or nearby reflectors, resulting in false measurements,
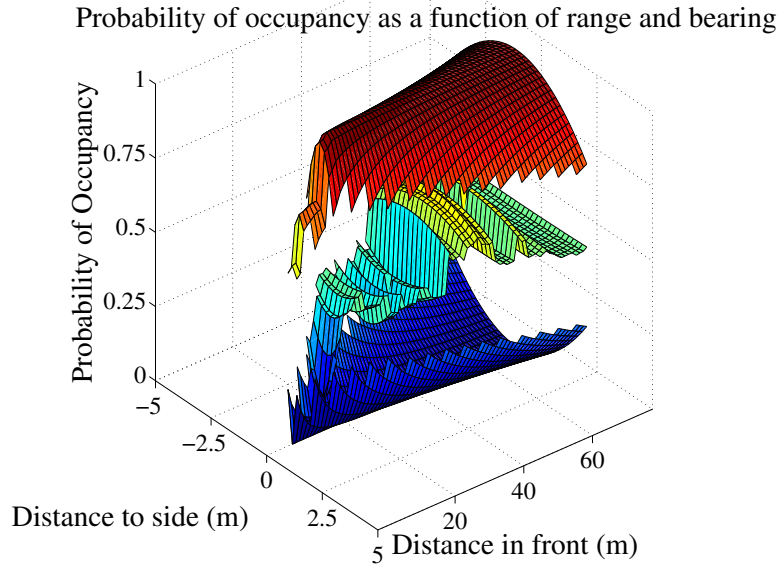
Figure 4.2: Surface plots of cell occupancy probability as a function of range and bearing to the target from the beam axis. The bottom plot is for constant power $P = 40$dB, the middle plot is for constant power $P = 55$dB and the top plot is for constant power $P = 70$dB. For all these plots, $G_a = 1$. Note that the distance axes use different scales. This is to exaggerate the effect G_θ in creating the cone-like shape of the sensor model; the beam is much narrower when the scales are equal.
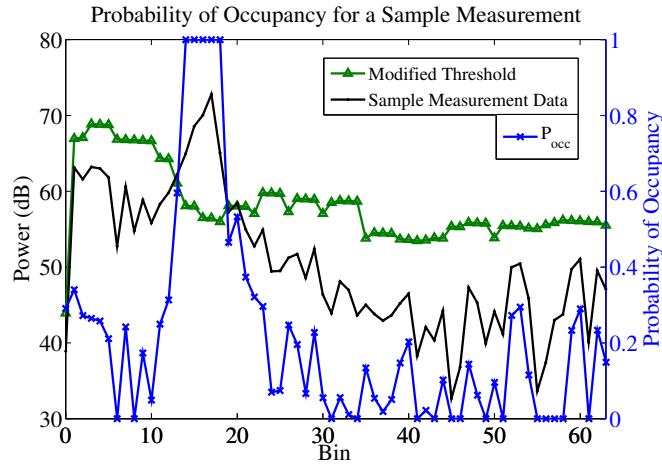
Figure 4.3: For every measurement, the occupancy probability for each bin is calculated based on the difference between the value of each bin and the modified detection threshold. In this measurement, there is a strong target centred on bin 18 but all other bins are below the detection threshold.

but these detections are too difficult to model. The calculation of the bearing to the centroid of the area of intersection is illustrated in Fig. 4.4.

4.1.5.3 Area Scaler

It is assumed that there is only one target in any bin, therefore the occupancy probability must be divided between all the covered grid cells and weighted proportionally to each area of bin-grid intersection and the total area of the bin [50]:

$$G_a(A, b) = \frac{A}{BinArea(b)} \tag{4.9}$$

where A is the area of intersection between the grid cell and bin footprint. *BinArea* is a function that calculates the area of a bin (which depends on beamwidth θ, the bin length and the bin number). The maximum value of G_a is 1.0, when a bin lies completely within one occupancy grid cell. The minimum value is slightly greater than zero, for very small intersections between the grid and the beam. The calculation of the area of intersection for the area scaler is illustrated in Fig. 4.4.
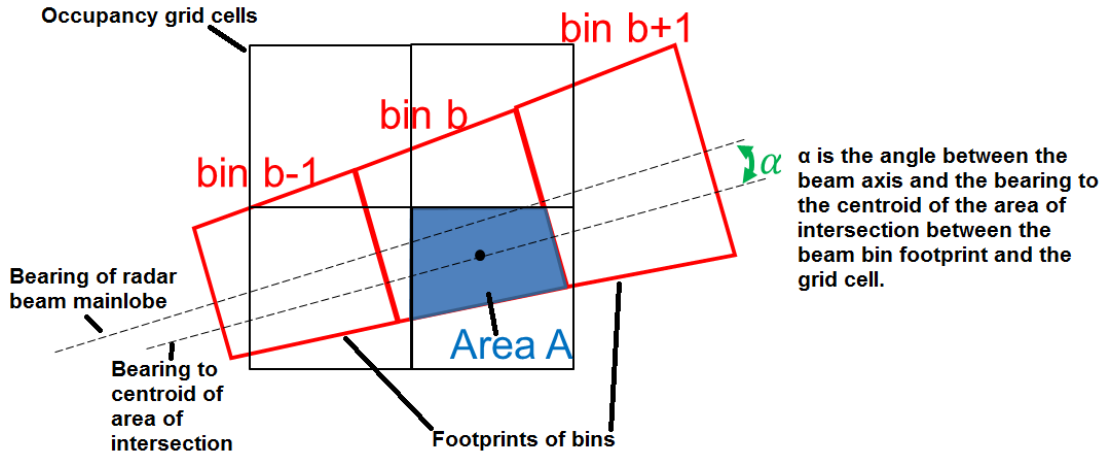
Figure 4.4: In the sensor model used, the area scaler is a function of the area of intersection between the bin footprint and a grid cell. The bearing scaler is a function of the angle α between the angle of the beam and the centroid of the area of intersection.

4.2 Experimentation and Results

Tests are performed over two ranges:

- Short-range tests in a carpark using corner reflectors and light vehicles as targets

- Long-range tests in a large open field using a corner reflector as a target

In the short-range tests, the radar is scanned across 160° but only the first 16 bins of the measurement vector are used, as the test area is hemmed in by buildings and areas of pedestrian and vehicle traffic. In the full-range tests, 64 bins are used but the radar is only swept across 90°, again to avoid buildings and trafficked areas.

The short-range tests are plotted with a grid resolution of 0.5m. The full-range test is plotted with a grid resolution of 1.0m.

4.2.1   Short-range tests

A photograph of the carpark test site is shown in Fig. 3.8, repeated here as Fig. 4.5. The radar measurements are plotted in Fig. 4.6, and an occupancy grid plot is made. There are notable differences between the raw radar plot and the occupancy grid produced from the same data. The occupancy grid is absent of the sidelobes that are present around the target in the centre. The four pillars (two either side of the sensor) appear much less elongated, and the distant one on the sensor's left is not merging into the background any more.
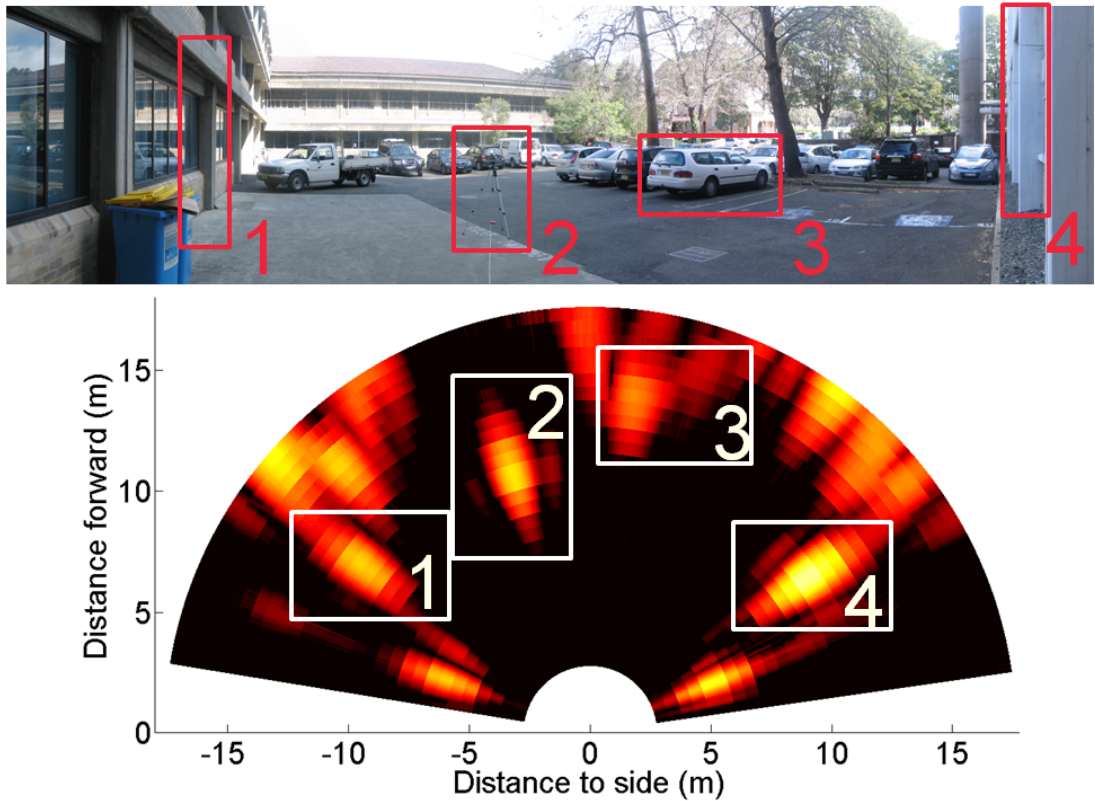
Figure 4.5: A photograph and radar scan of the environment are shown here. Some of the same key features are labelled in both images for ease of comprehension. Feature 1 is a column of the building, feature 2 is a corner reflector on a tripod, feature 3 is a parked car, and feature 4 is a column on another building. Additional features such as columns and vehicles are also present in both images. This is plotted using a uniform threshold.

4.2.2 Long-range tests

A panoramic photograph of the open field test area is shown in Fig. 4.7. It is is a large, open park space approximately 70m across, sloping down slightly in a shallow depression, with a large tree in the distance on the right of the radar, concrete walls and footpaths outside a building on the left, a bridge over a pond approximately 15m in front of the radar on the right, and a few people scattered under the tree in the distance. A 10m^2 RCS corner reflector is placed on a tripod in front of the CPD radar at a range of 5m, and a human being gradually carries it away from the radar in a straight line to a distance of 50m, pausing for a few seconds each metre. The person moving the target remains in the field of view during the measurements, but a human being's RCS is less than 1m^2 and will not significantly impact the results. This test is done with the CPD
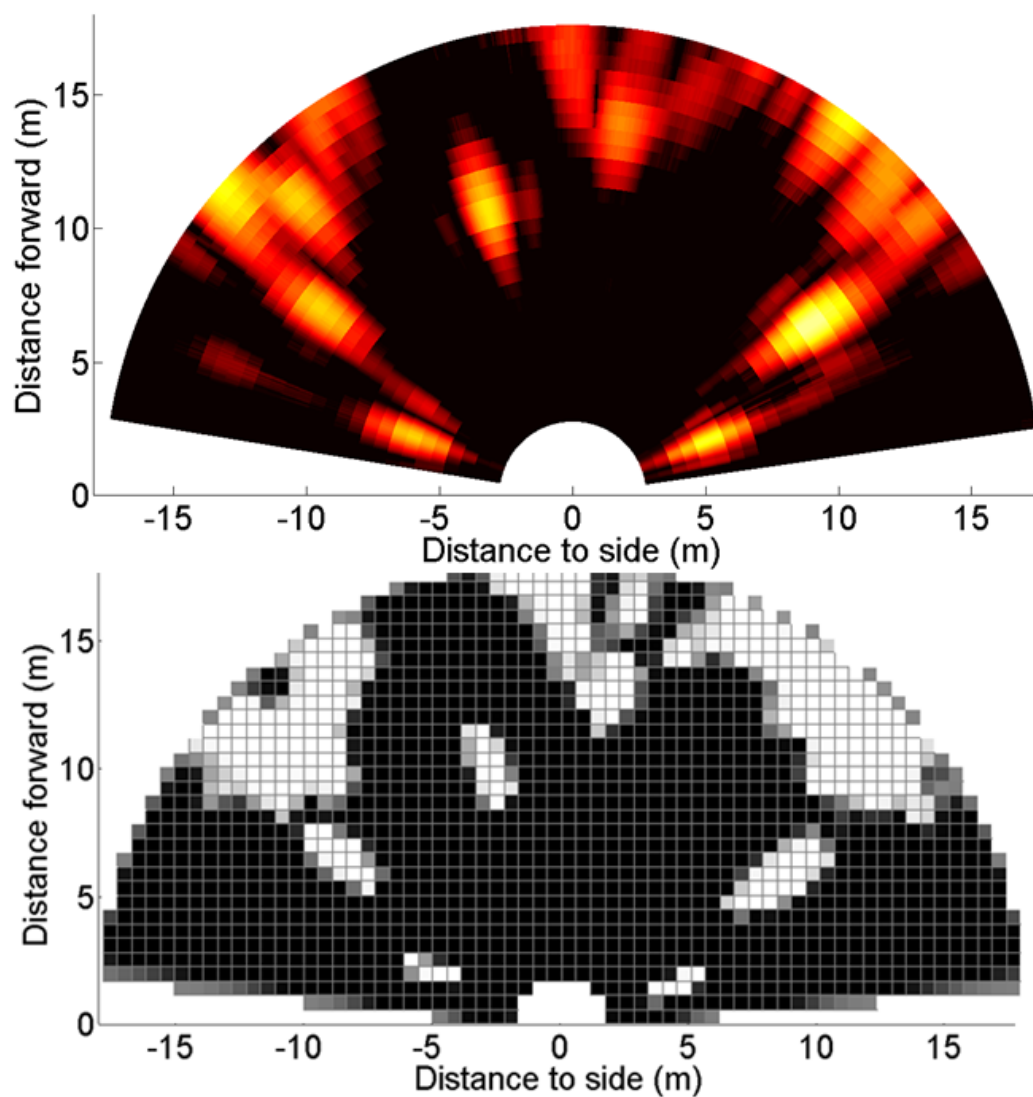
Figure 4.6: Measurement plot (top) and occupancy grid (bottom) made from scanning a carpark with a 100m^2 RCS target set up in it. This is the same data that is plotted in Fig. 4.5.

Figure 4.7: A panoramic photo of the field that long-range tests were performed in. The radar sits on the trolley in the middle of the picture. The tree, bridge and a few people can be seen in the occupancy grid map.

radar scanning across 90° to map the environment as well as without scanning, so that the radar constantly measures the range to the target. When the radar is not scanned, the target is measured continuously but its orientation to the radar is changed while it moves, leading to a noisy signal that stabilises at each 1m increment.

The occupancy grid map in Fig. 4.8 uses different colours to represent different cell states. This colour scheme is chosen due to the difficulty of distinguishing weak occupancy from weak emptiness in greyscale images at this resolution.

- A cell with a high confidence of occupancy is drawn as red.

- A cell with a lower confidence of occupancy is drawn as dark orange.

- A cell with a weak confidence of emptiness is dark purple.

- A cell with a stronger confidence of emptiness is drawn as dark blue brightening to lighter blue as confidence of emptiness increases.

The plots in Fig. 4.8 show maps made by scanning the entire area. The corner reflector is detected by the CPD radar at a range of approximately 36m, approaching its maximum detection range of 45m to 50m. The large tree is detected, as is the bridge on the right. The objects at the far left extremity of the measurements are most likely part of the low wall or concrete steps outside the building to the left. There was a group of people under the tree when the measurements were taken; they appear as isolated
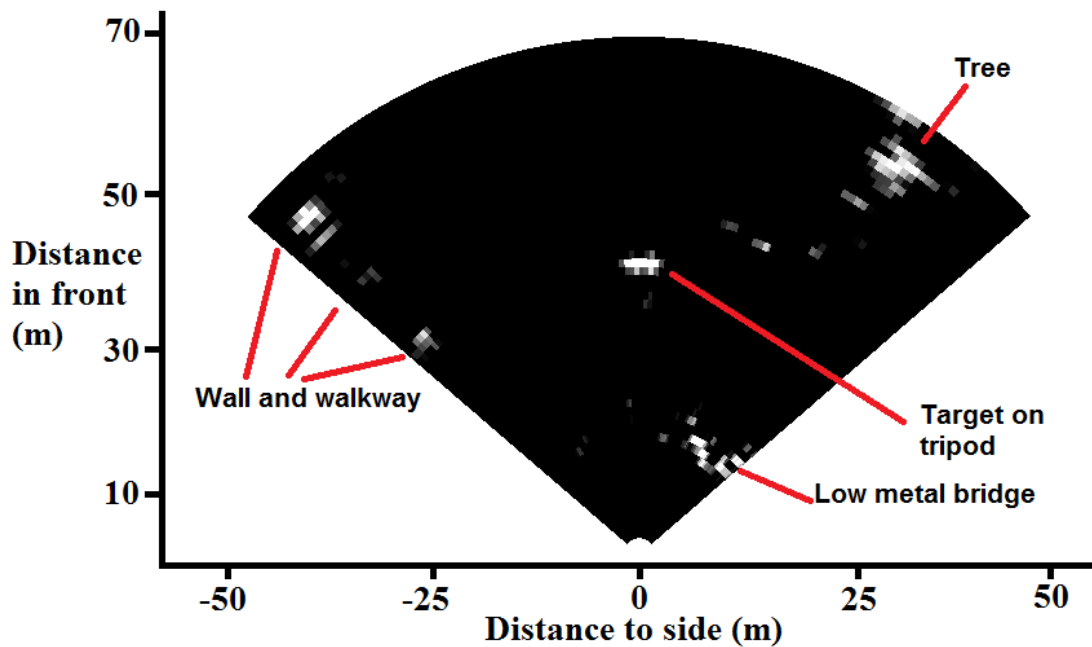
(a) Raw data plot.
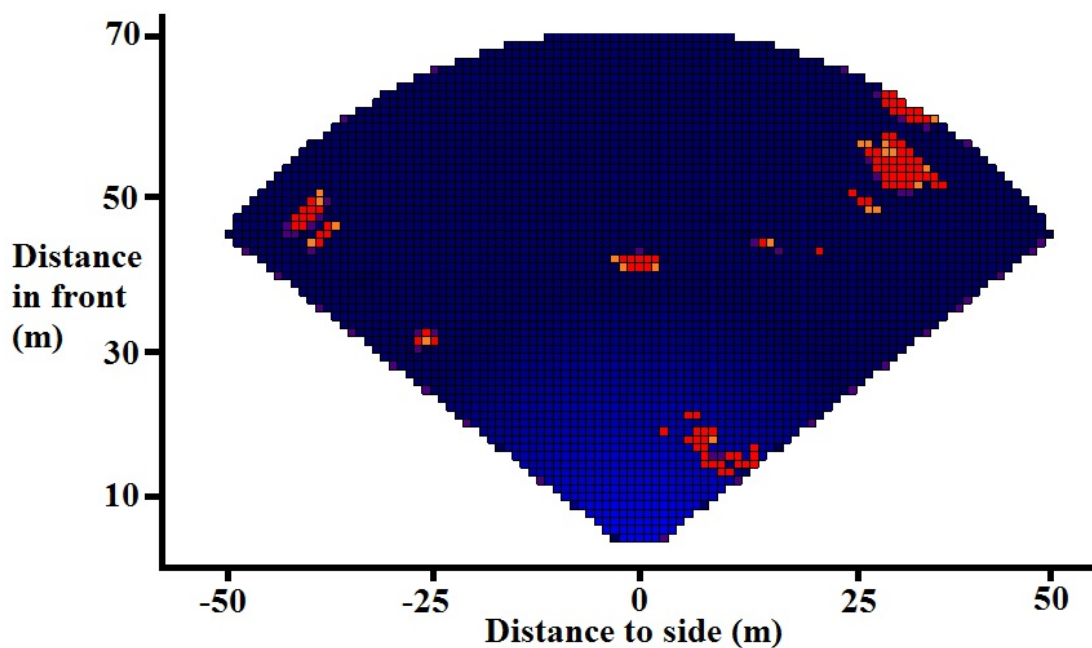
(b) Occupancy grid map, resolution 1.1m.

Figure 4.8: Raw measurements and occupancy grid map from scanning an open field with a 10m^2 RCS corner reflector approximately 36m from the radar.

detections between the target and the large tree on the right. There are measurements of clutter at a range of around 15m in plot Fig. 4.8(a) but these are filtered out in the occupancy grid in Fig. 4.8(b).

The plots in Fig. 4.9 show continuous measurements of a $10m^2$ corner reflector moved away from the CPD radar. The range to the target from the occupancy grid is usually slightly less than the raw measurement, probably due to poor range resolution; the raw measurement range is the peak power for one measurement, while the occupancy grid range is the nearest peak occupancy, which can saturate from several non-peak measurements. Once more, the occupancy grid consists of one cell per bin with resolution equal to the bin length, and angle and area factors are neglected. At times the target is lost due to poor orientation of the corner reflector to the radar, defaulting to a minimum range of approximately 3m. Occupancy grids are plotted using this data and four different detection thresholds in Fig. 4.10. Poor range resolution has the effect of stretching the target vertically. Clutter appears as additional targets in the column; clearly there is clutter around the 5m and 10m range from the sensor, which only the modified clutter threshold is able to filter out. The approximate shape of the target's movements (away from the sensor slowly, then rapidly) is visible in all the plots, but with clutter decreasing as the threshold increases.

4.3   Discussion

This chapter's contribution is a method of extending a static detection threshold for FMCW radar into a sensor model for use with a Bayesian occupancy grid. It also discusses the design of the occupancy grid itself, including the grid resolution and the bounding of the occupancy values. The occupancy grid is designed to be a method of representing the environment around a large vehicle to the driver in a way that is more reliable than just plotting the raw measurements.

Occupancy grid design involves a number of trade-offs: between robustness to noise and speed of map response to change (or just going from unknown to known when starting from scratch), and between geometric precision and processing speed. An implementation that is more resistant to noise will also be more resistant to change in the environment. An implementation that is very high-resolution and precise (via smaller
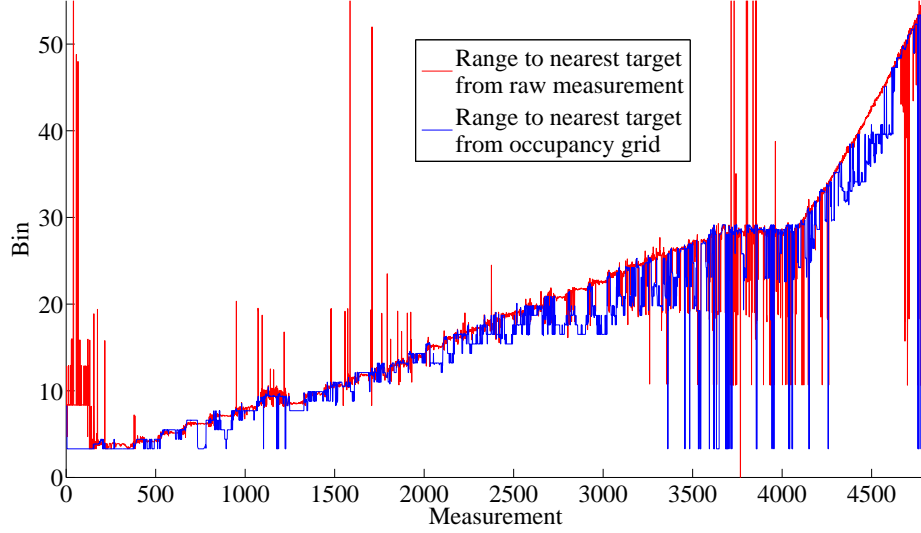
Figure 4.9: A $10m^2$ RCS corner reflector is placed on a tripod in front of the CPD radar at a range of 5m and gradually moved away from the radar in a straight line to a distance of 50m, pausing for a few seconds each metre. The range estimates to the target from the raw measurements and occupancy grid are extracted the same way as for the plots in Fig. 4.10. The detection filter uses the modified clutter threshold.
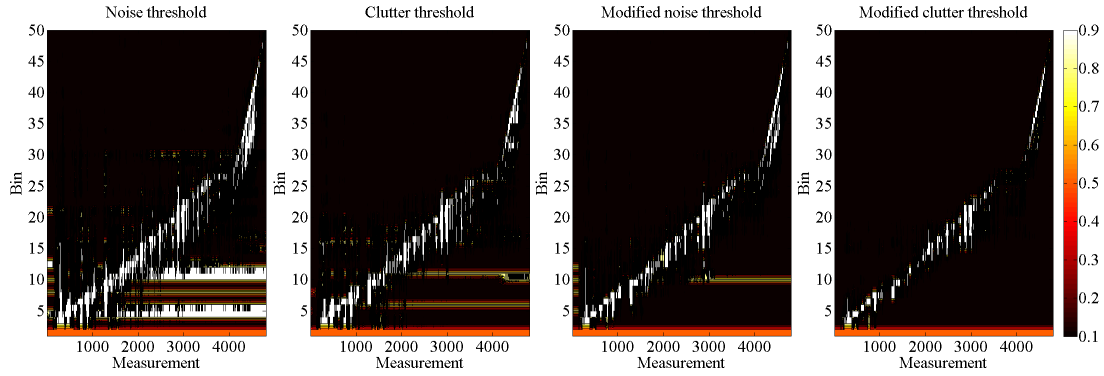
Figure 4.10: These plots show consecutive occupancy grid values for the same measurement data plotted in Fig. 4.9 for four different thresholds. This data is used to update a small occupancy grid, one column wide with one cell per bin, and beam angle and area effects are neglected. Each column of the plot shows the state of the occupancy grid based on all the measurements up to that one.

grid cells and more complex beam geometry modelling) will be much slower to compute. Human factors must influence the design. The driver is relying on the map to be produced quickly and accurately so that they can move the vehicle. If there are too many false alarms, or it takes too long to make a map, the driver will be annoyed and not use the system.

This chapter considers only static or near-static environments. The static assumption allows the sensor model to focus more on noise rejection rather than map response speed, since a static map will not have many changes for the map to respond to. The map can be reset between each use (that is, before movement, the map is cleared and several scans made to rebuild it from scratch) to prevent the "map inertia" that can result from many observations. If a faster-scanning sensor were used, faster-moving targets could be tracked using the occupancy grid, allowing this method to be used in dynamic environments and reducing or even removing the need for a map reset.

The parameters with the most influence on the reliability, responsiveness and update speed of the occupancy grid in the static case are the grid resolution and sensor model design. The bounding of the cell occupancy will also influence performance, but it has a larger impact in dynamic environments. Assuming the environment is static, bounding of the cell occupancy is not strictly necessary, even for numerical stability, as the probabilities can be converted to logarithm of odds form.

4.3.1 Grid Resolution

The grid cell size should be tailored to the bin length; not much less than half a bin length, and not much more than one. This reduces the dilution of updates between multiple cells, and cases where a single cell is updated with multiple conflicting measurements. There is also the possibility of varying the cell resolution with range, to reflect the varying footprint size of the beam. A quadtree implementation can be used to create a grid of different resolutions, and to change the size of cells during mapping.

4.3.2 Sensor Model

The performance of the sensor model developed in this chapter is dependent on the quality of the detection threshold calculated in Chapter 3. As described in Sec. 3.6, the quality of this threshold is dependent on the measurements that are made to calculate

it; a large number of measurements, made from many positions in an environment of fairly homogeneous clutter, should produce the best results. The approach used here - defining a detection threshold, δR_{occ} and δR_{emp} - could be used with any sort of detection threshold, such as a CFAR threshold if the sensor used had better performance with dynamic thresholds.

Other than the detection threshold, the sensor model used here has a number of variables influencing its performance:

- the probability saturation bounds of the sensor model output (currently 0.10 and 0.90)

- the maximum power threshold δR_{occ} and a minimum power threshold δR_{emp} (currently $\delta R_{occ} = 6dB$ and $\delta R_{emp} = 12dB$)

- The calculation of the angle scaler and area scaler.

4.3.2.1   Sensor Model Saturation Bounds

Widening the probability saturation bounds of the sensor model output will produce a more confident sensor model that can shift map estimates more rapidly. Narrow output bounds produces a less confident estimate that updates the map more slowly.

4.3.2.2   Maximum and Minimum Power Threshold

Choosing what values of δR_{occ} and δR_{emp} to use would depend on the performance of the radar (the radar's dynamic range would be a key value) and the desired performance of the occupancy grid. Small values of both δR_{occ} and δR_{emp} will produce a very confident sensor model that frequently generates maximum occupancy or emptiness estimates. Large values of both will produce a larger range of potential outputs. A large δR_{occ} and δR_{emp} will produce a less confident sensor model. Different values (as implemented here) produce an asymmetric sensor model that will either add or remove detected targets more rapidly than the reverse. For a collision avoidance application like this one, it is a good idea to be cautious and remove targets more slowly than they are added (as is done in this implementation). But in a particularly uneven terrain this may may cause too many false alarms. Determining the optimal values of δR_{occ} and δR_{emp} for a given environment is a problem that may be best solved by an optimisation or machine learning method; this is beyond the scope of this thesis.

4.3.3 Area Scaler and Angle Scaler

The occupancy grid's effectiveness as a map is limited by the need to strike a balance between geometric accuracy and speed of processing and update. A finer grid resolution will create a more accurate representation of the environment, but will require many more calculations to project the radar footprint onto the grid. A simple geometric model of the beam could be used; it would be possible to only update the grid cell with the largest area of intersection with the beam footprint, or the one closest to the beam axis, to reduce processing. This would remove the need for an area and angle scaler in the cell update. The detected target would be assumed to lie entirely within this cell. This would increase map response speed at longer range, but would represent the geometric uncertainty of the beam with less fidelity. Alternatively, a more complicated geometric model could be used, such as one that included the presence of sidelobes. In the next chapter, a different mapping approach will be used, sidestepping many of these issues (but introducing others).

4.4 Summary

In this chapter, the detection filter developed in Chapter 3 was used to develop a sensor model for a radar for use in occupancy grid mapping of a static environment. An occupancy grid was implemented using this sensor model and achieved good results. Tests performed in different environments have been used to produce maps with lower levels of noise or clutter than plots of the raw measurements. These can be used by the drivers of large vehicles to understand the state of the environment around them, in a more reliable way than just viewing a plot of the radar measurements.

The sensor modelling approach used here could be modified for use with other radar sensors, different threshold designs or environments. Although all the tests conducted here were for a static or low-speed environment, with a faster-scanning sensor this could be used in more dynamic cases. Most close proximity accidents occur at low speed, so solving for the static case first is a worthwhile approach. In Chapter 5, the close proximity problem in more dynamic environments is considered and a new mapping approach is introduced.

Chapter 5

Dynamic mapping with FMCW radar

Chapter 4 discussed occupancy grid mapping approaches under static conditions - that is, when both the sensor and the environment are either stationary or moving at very low speed, and there is little change in the environment over time. In a static or low-speed environment, the map will eventually converge to the most accurate representation of the environment that it can produce using the given sensor model and grid resolution; any noisy measurements will eventually be averaged out over time.

Many environments will contain moving objects, and the vehicles that sensors are fitted to can themselves move through the environment. In a dynamic scenario, the map needs to respond quickly to changes in the environment without being overly susceptible to noise and clutter. Inertia in the map can be a problem; too little inertia and the map cannot filter out noise, too much and it will be too slow to respond to changes such as objects moving, appearing and disappearing from view.

This chapter looks at extending the CPD radar occupancy grid mapping system introduced in Chapter 4 into scenarios where the sensor and objects in the environment are moving relative to one another.

After limitations in the occupancy grid implementation are identified, the probability hypothesis density filter is considered as an alternative for tracking and mapping in these environments using the CPD radar.

5.1 Mapping with Radar in Dynamic Environments

A **dynamic environment** is defined for this thesis as an environment where there is fast relative motion between the sensor and one or more objects in the environment. This may be due to the sensor being moved through the environment, or moving objects being present in the environment, or both.

The previous chapter assumed a "semi-static" environment; if the environment was not completely static, any relative speeds were slow, any changes in environment state between sensor scans were minor, and changes between scans were only gradual. If the sensor were being moved, it would be moved at low speed and only over short distances to remove any problems potentially caused by poor localisation of the vehicle between scans. This was modelled as a progression of static scenes with minor changes between them that the occupancy grid map was able to reflect. Moving targets were either human beings moving at very slow walking speeds, or trihedral corner reflectors being moved by human beings at very slow walking speeds.

In a mine or other industrial environment, objects close to a large vehicle may move at higher speeds. A close proximity system needs to be able to detect their presence swiftly and communicate it to the driver. Vehicles that are not in the immediate path of the large vehicle may be moving in a way that will soon place them in its path, or may be moving safely away from the large vehicle. One challenge is that changing measurements caused by moving objects will need to be distinguished from changing measurements due to noise or clutter corrupting measurements of empty space; in an environment assumed to be static, any random variations in measurements due to noise will be averaged out in the longer term, but in a dynamic environment there needs to be greater responsiveness to change, which can increase susceptibility to erroneous detections.

A challenge in using the CPD radar in a dynamic environment is the comparatively slow scan and measurement rate. The radar makes measurements at 100Hz, that is, with a measurement period of 10ms. This presents challenges in terms of angular range, angular resolution and scan period for a dynamic environment.

To reduce scan period, the scan range could be reduced from 160° (which can be done with existing hardware) or the measurement rate could be increased (which would require hardware changes). Regardless of the specific changes made, it becomes apparent that in its current form, the CPD radar will struggle to produce timely, accurate

maps in a dynamic environment. It can still be used to develop algorithms for use in such an application however. A fast-moving scene can be simulated by conducting tests involving slow-moving targets and compressing the time measurements.

5.2   Challenges for grid mapping in dynamic environments

Fusing multiple measurements of static obstacles together can produce a more accurate representation of the environment than raw, noisy measurements. But when the environment changes, the map needs to be able to respond quickly to these changes to provide a timely warning to the driver if necessary. Some mapping approaches in dynamic environments attempt to distinguish between static and dynamic objects and map or track them separately [8]. Occupancy grids can be effective with static objects but are not completely satisfactory for dynamic environments, on conceptual and practical levels.

The obvious danger of a dynamic environment is that an object that is not currently a hazard might become one in the near future. To better anticipate this, it helps to be able to understand the motion of the objects in the environment so that their future positions (and danger levels) might be better predicted. The occupancy grid has no concept of velocity, or indeed of individual objects. There are extensions to the classical implementation that model cell velocity, such as the Bayesian occupancy filter [9] or modelling cell transition properties with hidden Markov models [38]. Clustering and segmentation methods can be used to extract objects from the wider grid but these are external to the occupancy grid rather than inherent.

Additionally, there are a number of undesirable features of the occupancy grid's performance in changing environment that limit its usefulness. To borrow a phrase from software debugging, these are not bugs, they are features; they result from the original design and assumptions made for the occupancy grid for mapping static environments. But they combine to pose challenges for the application of an occupancy grid to problems in dynamic environments.

5.2.1 Repeated observations create inertia

Chapter. 4 used a Bayesian occupancy grid to illustrate how multiple repeated observations of a cell state of either occupancy or emptiness increase the confidence of the state estimate. The update is commutative; if the observations were to change (from occupancy to emptiness or vice-versa), a proportionate number of changed observations would be required to "undo" the previous ones and return the estimate to uncertainty. Only then would the estimates tend towards the new state. This "inertia" slows the response of the grid to changed environmental state. This was the problem in the creation of the occupancy grid in Fig. **??**. With a fast-measuring low-uncertainty sensor, such as lidar or camera in a low clutter-environment, this is less of an issue, as measurements will be produced fast enough for any latency in the map to be low. But radar typically has a lower measurement rate and a higher level of uncertainty in its measurements.

The practical effect of this is that the driver of a vehicle equipped with a radar using such a system would need to wait a longer time for the system to be sure that the map is accurate, trying their patience and rendering the system less useful. In a worst-case scenario, the environment would change too quickly for the map to reflect at all.

Some inertia is a good thing, as it makes the system robust to clutter and noise. But this robustness to clutter is in opposition to response to changing environments [8]. Too much noise-rejection will produce a cell that responds slowly to a genuine change; too little and the grid will be volatile and potentially inaccurate in high noise or clutter. A slow response is not desirable in a situation awareness context; if the occupancy grid is overconfident based on measurements made long ago, the map will respond slowly to changes and the driver of a heavy vehicle will not be informed within a useful timeframe. A high false-alarm rate is likewise unacceptable.

The sensor model described in Chapter 4 is designed partly to alleviate the problem of inertia. It is biased towards occupancy to ensure safety (see Sec. 4.1.5.1) but this only causes a faster shift towards saturation for weaker detections. Assuming updates are equally saturated (that is, maximum occupancy or emptiness value has been reached, via strong measurements of either occupancy or emptiness, and these values are equally strong towards occupancy or emptiness) then an equal number will be required to cancel one another out.

An example is given in [8] of how a series of an equal number of measurements of occupancy and emptiness $Z = \{O_1, O_2, ...O_n, E_1, E_2, ...E_n\}$ would produce the exact same

estimate of cell occupancy if it were reversed i.e. $Z_{rev} = \{E_1, E_2, ...E_n, O_1, O_2, ...O_n\}$ due to the commutative property (and assuming that the occupancy did not saturate).

As implemented, the cell occupancy is bounded to mitigate this; once it saturates at the maximum occupancy or emptiness estimate, further identical measurements do not effect it. This limits how confident the cell can become, allowing faster responsiveness when the cell state changes. But this also reduces the grid's ability to reject noise and clutter measurements and there is a greater risk of be occasional cell "flickers" due to clutter and noise. False alarms would annoy the driver.

5.2.2 Coupling between state estimate and confidence

A cell's occupancy estimate and confidence in the estimate are coupled; they are encapsulated in the occupancy value [8]. In terms of the occupancy state estimate, there is very little difference between occupancy probabilities of $p_{occ,1} = 0.99$ and $p_{occ,2} = 0.999999999$ (again, assuming cell occupancy were not bounded). However, $p_{occ,2}$ may be the result of many observations and $p_{occ,1}$ the result of comparatively few, depending on the sensor model. Assuming they were produced using the same sensor model, $p_{occ,2}$ would take more observations of emptiness to reverse under the standard Bayesian update, and so we would say it is a more "confident" estimate of occupancy. A less confident estimate would ideally exhibit greater sensitivity to differing measurements and so have faster dynamic response, while a more confident estimate would be less responsive to influence.

Another example illustrates the undesirable link between occupancy estimate and confidence and how it is caused by inertia. Suppose a cell in an occupancy grid in a dynamic environment is observed briefly and the measurement set Z_a consists of i consecutive measurements of occupancy, and the sensor model produces an estimate of occupancy of 0.9:

$$Z_a = \{O_1, O_2, ..., O_i\} \text{ where O = occupied and E = empty}$$

$$p_{occ,a} = 0.9$$

Suppose another cell in the dynamic environment is observed for a very long time, and the measurement set Z_b consist of a long sequence of j consecutive occupancy measurements (where $j > i$) followed by a short sequence of k consecutive emptiness

measurements (where $i \geq k$), and the sensor model produces an estimate of occupancy of 0.9:

$$Z_b = \left\{ O_1, O_2, ..., O_j, E_1, E_2, ..., E_k \right\} \text{ where O = occupied and E = empty and } j > i \geq k$$

$$p_{occ,b} = 0.9$$

Both cells have the same probability of occupancy, but $p_{occ,b}$ should be much less confident than $p_{occ,a}$; its most recent measurements are all of emptiness. Instead, they have the same state estimate and exhibit identical dynamic response.

5.2.3 Coupling between position uncertainty (resolution) and existence uncertainty

At longer range, the uncertainty in the estimated position of a detected object is translated to existence uncertainty as the update is split between more grid cells. This is based on the assumption that there is a single object detected per bin, and the total update must be split proportionally between the grid cells it covers. The power scaler value may be a maximum, but the area scaler will dramatically reduce the update for each cell. This has the effect of converting a high-confidence detection of an object to several low-confidence updates in adjacent cells. There is no easy way to address this. Increasing the grid size would reduce the number of cells covered by one bin increasing map responsiveness, but the grid size is limited by the bin size; it should not be much larger than one bin or multiple conflicting bins may update the one cell.

5.3 Bounding Cell Occupancy In Dynamic Environments

The factors described in Sec. 5.2 combine to make it difficult for an occupancy grid to track the velocity of moving objects unless they are moving slowly, or the sensor has a high measurement rate across the entire area. With a different radar that had a higher measurement and scan speed, this may be feasible.

For the CPD radar, different approaches are taken to overcome this problem. The simplest is to only use the occupancy grid to map static objects - but this is not so simple, as it would require the detection and tracking of moving objects to remove them from the grid. Another approach would be to recalculate the cell occupancy over

the measurements from a narrow time window [8], but this would require storing the update histories for every cell and be computationally taxing by requiring every cell to be re-updated at every timestep, as well as requiring thoughtful selection of the window length.

A simple approach is to bound the maximum occupancy/emptiness values. Once the cell's estimate reaches the saturation threshold, further updates of the same observation will not change the cell's occupancy, and so will not increase the response time should the observation change. This allows the maximum confidence to be limited and allows a faster response to change. With a fast enough measurement rate and a well-designed sensor model, the total response time can be small enough to be useable in dynamic environments.

Any bounding will need to consider the effect of target range on grid response rate. At close range, the beam is fairly narrow and the update will be divided between comparatively few grid cells. This allows a wider occupancy bounds to be used without sacrificing response speed. At longer range, the update will be divided between more grid cells due to the larger beam footprint. This slows the change in state estimate, especially when the cell has reached a high confidence about its state only to have the true state change [8].

As discussed in 4.1.4 the cell occupancy is bounded to prevent numerical instability and enhance dynamic response. At close range, the bounds are 0.10 and 0.90.

In a scenario where the radar is held static but objects in the environment can move, cell occupancy can be more narrowly bounded with increasing range. A simple approach is to scale the maximum resulting occupancy of a cell by a linear function of the last bin to update it. To bound the occupancy of cells i between 0.4 and 0.6 in bins b greater than 32, with a linear decrease for bins nearer than this, the equations shown in Eq. 5.1 and Eq. 5.2 can be used. Bounds of 0.4 and 0.6 are quite low confidence levels; the occupancy grid information will not be providing very strong evidence of either occupancy or emptiness, and noise and clutter filtering ability will be reduced.

$$P_{max}(occ_i, b) = \begin{cases} 0.9 - 0.3 * \frac{b}{32} & : b \leq 32 \\ 0.6 & : b > 32 \end{cases} \tag{5.1}$$

$$P_{min}(occ_i, b) = \begin{cases} 0.1 + 0.3 * \frac{b}{32} & : b \leq 32 \\ 0.4 & : b > 32 \end{cases} \tag{5.2}$$

This is a heuristic approach and requires manual adjustment of the maximum and minimum bounds and the rate at which they are set for cells at increasing range. Longer-range cells saturate to less "confident" estimates and are much more responsive to conflicting observations. One downside of this is that these cells are more vulnerable to the effects of false measurements due to clutter, noise or missed detections. This is partly reduced by the update being split between multiple cells at range, but multiple "bad" measurements will be reflected in the map faster than they would at closer range.

This approach is dependent on the sensor being held still, as the bounds of each cell is a function of range to the sensor; if this changes, the bounds will need to be changed.

5.4 Mapping with the probability hypothesis density filter

The limitations on the use of occupancy grids with radar in dynamic environments are becoming clear. It is worthwhile to investigate an alternative mapping paradigm to occupancy grids, the feature-based mapping approach, for use with radar in a close proximity system. The probability hypothesis density (PHD) filter is growing in popularity for mapping and tracking with radar due to its robustness to clutter. Rather than updating the occupancy values of map cell, it models the multi-object state of the environment as a single distribution, and individual environmental features can be extracted as local maxima of the distribution.

It estimates the state of objects as individual features in the environment, s. The dynamic environment can be represented more responsively than with the occupancy grid, while still exhibiting desirable features such as noise and clutter rejection. This chapter's contribution is an implementation of the PHD filter with FMCW radar for solving the close proximity problem.

The PHD filter is often approached from a tracking perspective as opposed to a mapping or filtering one. A distinction should be made between filtering and tracking. Filtering can be defined as "the estimation of the state of a signal at each point in time based on a discrete set of noisy measurements, where a new measurement is received at each time-step" [10]. Tracking is interested in estimating both the position of targets in the state space *and* their identification over time. "Tracking algorithms output tracks, and a track represents a labeled temporal sequence of state estimates, associated with the

same target" [56]; it is filtering plus consistent target labelling over multiple timesteps. For the close proximity application, tracking is desirable, but not necessary; it is left as future work.

The PHD filter is capable of outperforming the popular Multiple Hypothesis Tracker (MHT), however the PHD filter does not produce labelled tracks the same way as the MHT by default; an additional track labelling heuristic is required for it to function as a tracker [51] [52].

The algorithm for the extended Kalman filter Gaussian mixture PHD algorithm is provided in [74]. To allow its performance with the CPD radar to be evaluated, an implementation of it is made in Matlab and made available online by the author of this thesis [11]. Fig. 5.1 shows a plot of the PHD filter output for the simulated problem described in [74].
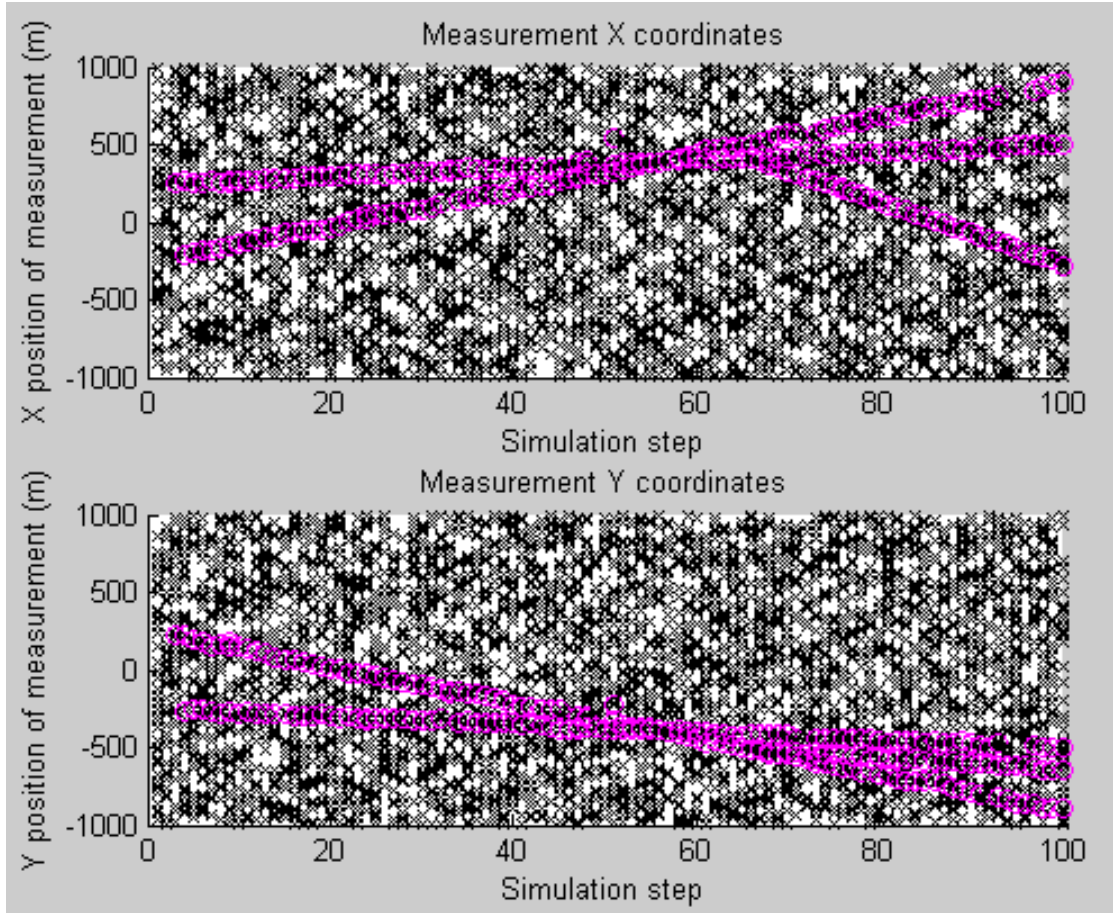
Figure 5.1: The GM-PHD filter used for the CPD radar is based on the GM-PHD algorithm developed by Vo & Ma [74]. A direct implementation of the algorithm and simulator (shown here) has been released online at [11]. The black 'x' values are measureements (mostly noise) while the magenta 'o' values are the targets being tracked correctly by the filter.

There are a number of terms in the GM-PHD filter that are application-specific and need to be defined for use with the CPD radar for collision avoidance [74]. These are:

- The state model $\mathbf{x}$

- The initial state, weight and covariance of new targets as they are added to the filter via birthing ($m_{\beta,k}$, $w_{\beta,k}$ and $P_{\beta,k}$) or spawning ($m_{\gamma,k}$, $w_{\gamma,k}$ and $P_{\gamma,k}$)

- The probability of target survival $p_{S,k}$ and detection p_D

- The motion model $\psi_k(x)$ and motion noise covariance Q_k for the prediction step

- The measurement model $h_k(z)$, measurement noise covariance R_k

- The clutter density model, including surveilled volume V and expected number of clutter detections λ

Some of these parameters are not considered at all in other mapping or filtering methods, are bundled implicitly into other terms, or have additional heuristics or ad-hoc methods to handle them. For example, in a Kalman filter, the motion model and noise covariance and observation model and covariance would be the same as in a Gaussian mixture PHD filter, but the probabilities of detection and survival, clutter modelling and weight vector are not considered.

Additionally, the PHD filter requires several heuristics for map management, including:

- The way that features are extracted from measurements and used to generate new targets in the filter

- The pruning threshold T, merging threshold U, maximum number of targets J_{max} for the pruning step

- The target extraction threshold W_E for the feature extraction step

A heuristic for track labelling is desirable to produce a GM-PHD tracker rather than filter, however this is not essential for this application; it is left as future work. An extended target representation would be similarly useful but is nonessential and is also left for future work.

Changing these parameters requires prior knowledge of the sensor, the targets and the environment [1, 74, 75] but will have major influence on:

- How accurately the object position is estimated,

- How accurately the object velocity is estimated,

- How quickly track weight increases from first detection (or, how many detections are necessary for a tracked object's weight to increase above 0.5),

- How resistant a track is to the effect of clutter or noise (whether is is shifted, substantially by a bad measurement, or ignores it),

- How quickly and accurately a track responds to changes in an object's trajectory,

- How substantially track weight decreases after a missed detection,

- The false alarm rates.

For the close proximity problem, it is important that hazards are detected quickly, with a low false-alarm rate, and with good accuracy in their current position. Therefore the most important criteria are how quickly target weight increases, the target position accuracy, and the false alarm rate.

An important value is the number of observations until a target's weight exceeds the feature extraction threshold W_E. Any tracked target with a weight less than W_E is treated as a potential object in the environment; any target with a weight greater than or equal to W_E is extracted as a genuine feature of the environment, displayed to the user, and depending on its position relative to the vehicle may be treated as a hazard or not (triggering a warning or alarm).

Assuming an object is detected at every timestep (that is, assuming there are no missed detections) and neglecting any effects from merging multiple nearby detections together, the rate at which an object goes from first detected to being assumed to exist (that is, its target track has a weight greater than 0.5) is a function of:

- birth weight (higher is faster)

- sensor measurement covariance (lower is faster)

- motion model covariance (lower is faster)

- observation model covariance (lower is faster)

- probability of survival (higher is faster)

- probability of detection (higher is faster)

- expected number of clutter measurements (lower is faster)

- merging threshold U (higher is faster)

- extent of the target (larger is faster as there will be multiple detections which merge together)

Adjusting these parameters will adjust the rate of weight increase. The filter's output - the estimated positions of objects detected by the sensor - is largely dependent on the sensor characteristics: sensitivity, resolution, accuracy and susceptibility to noise. Improving the position estimate itself is done by adjusting the target and observation covariance. Depending on how noisy the sensor's observations are and how cluttered the environment is, the observations may need to be filtered more heavily or less. The influence that a new observation has on an existing target is a function of:

- observation noise covariance (lower means more influence for the observation)

- prediction model covariance (higher means more influence for the observation)

- birth covariance (higher is more influence for the observations)

- merging threshold (higher will merge more distant targets, potentially creating inaccurate estimates if the merged targets are unrelated)

A balance needs to be struck between resilience to noisy observations and responsiveness to change. Additionally, the target covariance is used in the calculation of weights, and if the covariance is too high, bad matches between measurement and target will be given weights higher than they should be, and good matches will have much lower weight.

5.4.1   General Parameters

5.4.1.1   Point target assumption

All objects in the environment are modelled as point targets; that is, they are assumed to be of zero length and width, and to return only one measurement each.

When operating at a larger scale, such as an aircraft-tracking radar with a bin length hundreds of metres long tracking aircraft hundreds of kilometres away, it might be reasonable to approximate a large target as a point. At the scale the CPD radar operates, almost any target large enough to be detected (such as human beings, light vehicles and buildings) has the potential to span multiple bins and measurements as an extended target.

The point target assumption is an approximation, as most targets that the CPD radar can detect will act as extended targets. Even a very small target, shorter than one bin length, has the potential to cause multiple detections within a single measurement if:

- it sits on the border of two range bins, or the side edges of two consecutive measurements;

- it is a strong reflector with wide phase noise skirts;

- the scan speed is slow enough that the surveillance volume overlaps between consecutive measurements

For the sake of simplicity, the fact that multiple measurements often originate from a single target is ignored; all detected features will be detected and tracked independently as point targets by the PHD filter.

Because there are multiple target estimates output by the filter, analysis of the filter's performance will require an additional step of clustering these estimates together into different groups. Sec. 5.5.2 describes how this is done and why it is necessary for analysing filter performance.

5.4.1.2   State models

The pose state of the sensor $\mathbf{x}_S$ is

$$\mathbf{x}_S = \begin{bmatrix} x_S \\ y_S \\ \theta_S \end{bmatrix} \tag{5.3}$$

with coordinates (x_S, y_S) relative to some global coordinate frame in X-Y, and with heading θ_S bounded between π and $-\pi$ in the standard form with a bearing of 0 rad being along the X-axis and positive θ_S being counter-clockwise.

The sensor is assumed to be static and all objects in the environment move relative to it. The movement of the sensor can be modelled if needed.

The coordinate frame used is such that the positive-X axis extends to the sensor's right, while the positive-Y axis extends straight in front of the sensor. Positive bearings are counter-clockwise from the X-axis, negative bearings are clockwise.

The state of a tracked object or "feature" in the environment is:

$$\mathbf{x}_F = \begin{bmatrix} x_F \\ y_F \\ \dot{x}_F \\ \dot{y}_F \end{bmatrix} \tag{5.4}$$

Each object has a position (x_F, y_F) in the global coordinate frame, and velocities $(\dot{x}_F, \dot{y}_F)$, to represent its velocity vector in m/s.

Since the CPD radar is not configured for Doppler, it only measures the position of an object. Spawned targets will be initialised with a velocity calculated from the change of position.

5.4.1.3   Sensor range and field of view

The CPD radar's effective range is approximately 10-20m for small targets (objects with RCS $< 1\text{m}^2$) and 20-30m for larger targets. In this thesis, the most commonly used test environment, on the roof of the Seymour Centre parking garage at the University of Sydney, provides a maximum range of approximately 35m. One scan consists of a bearing range of up to $160°$, with inter-measurement angle difference varying based on the desired scan speed. Due to the way the CPD radar is mounted inside the bumper of a light vehicle, the effective range is limited to $130°$ ($80°$ to the right, $50°$ to the left). To simplify the PHD update process, all CPD radar detections from a scan are assembled into one measurement vector and the update is performed at the end of each scan. This reduces the need to recalculate the probability of detection for every object outside the beam volume of each measurement; the entire scan can be considered at once. However, the time that each measurement is made is recorded and is used to perform speed prediction, rather than using the same time for all measurements.

There is the option of running the PHD filter more often, after every measurement or after some fraction of the scan. This could give faster updates in the position of tracked objects, but at the cost of increased procesing. Care would need to be taken to ensure that detection updates are only performed on objects within the volume of the beam; this would involve varying the probability of detection and survival for each depending on proximity to the beam axis. At the scale of speed and time that these tests are performed at, the performance improvement would be small, so for simplicity's sake, this is not

done.

5.4.1.4   Target existence weight threshold W_E

At any given time, the list of targets whose states are being estimated by the filter may contain newly added targets, very faint targets that will soon be pruned out, targets caused by temporary noise or clutter, and targets that have been tracked reliably for many timesteps. This "target strength" i is represented by its weight w_k^i, which is between 0 and 1. Some form of heuristic target extraction is needed to extract which estimated targets are considered "truly existant" rather than "potentially existant". A simple and effective method is a weight threshold for existence W_E. Any tracked target whose weight exceeds W_E is extracted as an existant feature in the environment. The standard value is $W_E = 0.5$; this implementation uses this value.

This approach is not without flaws. The PHD propagates the intensity function of the target estimates, which is composed of a number of weighted Gaussians throughout the state space. The relative peaks of the intensity function are the locations of the most likely features; these peaks' heights and positions are a function of the weights, the covariances, and the relative positions of the means. Feature extraction is attempting to identify the locations of those peaks, which are not necessarily the same as the means of the targets with the highest weights; a target may have a large weight but also a large covariance that lowers its peak compared to a target of lower weight, or multiple nearby targets may have superposition effects that shift the peak. Assuming the merging function is properly configured, the chance of this should be greatly reduced and taking the targets with highest weights is a reasonable strategy.

5.4.2   Parameters for Birthing and Spawning New Objects

5.4.2.1   Generation of new targets to be birthed and spawned

Birthed targets are those that appear spontaneously in the environment. Spawned targets are those that split off from existing targets.

In a close proximity application, both scenarios are possible; vehicles and objects will appear spontaneously from behind occluding objects or from outside the field of view (birthing) and tight clusters of objects (such as crowds of people) may split (spawning).

In the first example given in [74], spawned targets are assigned a weight based on their proximity to an existing target. Spawning of targets is necessary as in that implementation, birthing only occurs at one local point in a wide volume of space. In this implementation, the CPD radar is initially scanned across an environment that is totally unknown. New objects could be detected at any point in this surveilled area. When the radar is static and has been mapping for some time, new objects might be expected to appear from outside the bounds of the surveilled region - but this will not always be the case. Objects within the scan region may be occluded and move out of the occlusion. Small objects that are undetectable at the limit of the radar's range may move closer and be detected. It is difficult to determine which locations have a higher probability of birthing objects than others, if any of them do.

At the scale that the CPD radar is working at, birthing is probably a better way to treat the appearance of new objects. The environment can be quite densely occupied, and determining which objects are spawned from those nearby and which are just moving in close proximity to others is difficult. More often than not, newly appearing objects will have no relationship between their motion and the motion of other objects. The most likely scenario for new target creation is surveillance of a previously unknown area, or a new object moving into the surveilled area. Even when one object spawns another - such as a driver exiting a parked vehicle, or a pedestrian splitting off from a crowd of other pedestrians - there is little relationship between their velocities, only their positions.

Therefore, in this implementation, all new targets are treated as being birthed and unrelated to any other tracked object.

5.4.2.2   Feature birth/spawn state

Targets to be birthed or spawned are created from the newest measurements Z_k at the end of the PHD filter update step. As the CPD radar cannot detect the speed of objects via Doppler, the velocity is unknown from a single measurement. Therefore the velocity components $\dot{x}$ and $\dot{y}$ will be set to zero and further measurement updates will be necessary to set a velocity.

To birth objects with an initial velocity, two consecutive sets of radar scans Z_k and Z_{k-1} must be used. Object velocities can be estimated by treating every combination of one measurement each in Z_k and Z_{k-1} as a detection of a distinct moving object. Since

the timestamp of each measurement is known, the velocity between the two points can be calculated.

For $j = 1..|Z_k|$, static tracks are formed as:

$$\mathbf{x}_{static} = \begin{bmatrix} x_k^{(j)} \\ y_k^{(j)} \\ 0 \\ 0 \end{bmatrix} \tag{5.5}$$

where $(x_k^{(j)}, y_k^{(j)})$ are the coordinates of the centre of the corresponding bin of the measurement after conversion to the global Cartesian frame.

Dynamic tracks obtain their latest positions from the measurements in the last scan in the same way static objects do. Their velocities are calculated by taking every combination of one measurement from the new scan and one measurement from the previous scan. That is, for $i = 1..|Z_{k-1}|$ and $j = 1..|Z_k|$ and $\Delta t = t_k - t_{k-1}$

$$\mathbf{x}_{dynamic} = \begin{bmatrix} x_k^{(j)} \\ y_k^{(j)} \\ \frac{x_k^{(j)} - x_{k-1}^{(i)}}{\Delta t} \\ \frac{y_k^{(j)} - y_{k-1}^{(i)}}{\Delta t} \end{bmatrix} \tag{5.6}$$

Therefore, there will be $|Z_{k-1}| * |Z_k|$ dynamic targets birthed, with $|Z_{k-1}|$ at the location of each measurement from Z_k.

The advantage of birthing targets with velocities already instantiated is that for moving objects; a track would be instantiated with a velocity after being observed twice, and so its prediction step would keep it closer to the true position of the moving object. A static track's position estimate will lag behind the measurements of a moving object, and if the object moves too quickly and its initial position covariance is too low, its updated weight may be very low.

Using two sets of measurements to create tracks with some initial velocity is attractive in theory, but the combinatorial explosion in the number of tracks created substantially slows down processing; instead of $|Z_{k-1}|$ new objects birthed/spawned at time k, there are $|Z_{k-1}| + |Z_{k-1}| * |Z_{k-2}|$ added (the first $|Z_{k-1}|$ term is for static objects). This makes the number of new tracks created in the update step substantially larger, and

requires more processing in the pruning and merging step. The vast majority of these will have their weight reduced substantially in the update step, and be removed in the pruning step.

To avoid this extra processing, tracks are birthed with zero velocity but very large velocity covariance. After one update, the estimated velocity will be close to the observed one, and should track quite closely in the next prediction update. Results of experiments with varying birth velocity covariances with simulated data are given in Sec. 5.5.1.

When birthing dynamic objects, a static object is also birthed at the location of each measurement from Z_k, in case the newly detected object is unrelated to any previous one.

5.4.2.3   Feature birth/spawn intensity

Once a scan has been made of the environment, and assuming it has detected all the objects present, new objects are most likely to be birthed at the perimeter of the surveilled area, as objects move into it. But it is also possible for objects to move out from behind occlusion, or for fast-moving objects to move a fair way into the centre of the field of view between scans.

The birth intensity is therefore implemented as a uniform birth weight of $w_\beta = 10^{-5}$ across the entire environment. It could be modelled as an extremely large number of Gaussians distributed across the environment, but this is needlessly complex.

Since all tracked objects are assumed to be birthed, the spawn intensity w_γ will be zero and no objects will be initialised as having been spawned.

The value of the birth intensity is chosen out of consideration of two factors - the pruning threshold, and how many observations it should take before the weight of a track is $W_E = 0.5$ or greater. A range of values can be used, between $w_\beta = 10^{-3}$ to $w_\beta = 10^{-1}$; larger values will cause more objects to be treated as real by the filter, more quickly. Lower values will cause a slower increase in weight of fewer objects. Through trial and error a birth weight of $w_\beta = 5 * 10^{-3}$ was selected, with slowly moving targets reaching a weight of 0.5 after 3 or 4 consecutive observations (see Sec. 5.5.1).

5.4.2.4 Feature birth/spawn covariance

The spawn covariance is of the form

$$
Q_\beta = \begin{bmatrix}
\sigma_{\beta_x}^2 & 0 & 0 & 0 \\
0 & \sigma_{\beta_y}^2 & 0 & 0 \\
0 & 0 & \sigma_{\beta_{v_x}}^2 & 0 \\
0 & 0 & 0 & \sigma_{\beta_{v_y}}^2
\end{bmatrix}
\tag{5.7}
$$

The values of the terms in Q_β are a function of the sensor accuracy and resolution, and the expected behaviour of tracked objects. The CPD radar has a range resolution of one bin, approximately 1m. The cross-range resolution will vary with range, and the resolution in the global X-Y frame will vary with the heading of the vehicle and the relative bearing to a detected object. For the sake of simplicity, these effects are neglected and a constant value of Q_β is used; at longer range, less reliable target detection is likely to have a much larger impact on state estimation than the birth covariance.

The values of $\sigma_{\beta_x}^2$ and $\sigma_{\beta_y}^2$ are based on the CPD radar's characteristics - namely, the bin length and susceptibility to phase noise. It also functions as an effective ceiling on the maximum velocity of targets that the filter can track; if $\sigma_{\beta_x}^2$ and $\sigma_{\beta_y}^2$ are too low, a moving target will not be correctly matched to their previous location and will be treated only as a newly spawned target. A larger covariance will result in a higher weight at positions further from the previous position - but also lower weight if the position is unchanged (and so slower weight increase for a static object) and the risk of matching more distant measurements incorrectly. $\sigma_{\beta_x}^2 = \sigma_{\beta_y}^2 = 1m^2$ provides sufficient uncertainty.

$\sigma_{\beta_{v_x}}^2$ and $\sigma_{\beta_{v_y}}^2$ are more difficult to assign due to the range in potential target velocities. Values of $\sigma_{\beta_{v_x}}^2 = \sigma_{\beta_{v_y}}^2 = 25m^2/s^2$ and $\sigma_{\beta_{v_x}}^2 = \sigma_{\beta_{v_y}}^2 = 100m^2/s^2$ are trialled in Sec. 5.5.1.

The velocity covariance is very high since the initial velocity estimates are typically very inaccurate. If the target track is instantiated from a single measurement, the initial velocity is completely unknown; a high covariance means that the velocity observed in the next update step will be given high credulity and the estimated velocity will quickly converge to it.

5.4.3   Parameters for the Prediction Step

5.4.3.1   Probability of object survival, p_S

The probability of survival is used to scale down the weights of tracked targets in the prediction step. For this implementation, it is assumed that tracked objects have a very high probability of survival, as in [74]. In the controlled tests that will be performed, the targets will not vanish without trace. At worst they will move outside of the surveilled region or be occluded by other objects. Failed detections because of this will see a reduction in weight, regulated by the probability of detection p_D.

Therefore the value of p_S chosen will be high.

$$p_S = 0.99 \tag{5.8}$$

Noise or clutter measurements may cause temporary tracks due to being detected once and then vanishing, but this phenomenon is represented by the clutter model λ in the update step, which will reduce the weight according to how many clutter measurements are expected.

5.4.3.2   Motion model ϕ_k

A constant velocity model ϕ_k is used for all tracked objects. This is assumed to be reasonable if target speed is fairly low and measurement rate is fairly high. Due to the large range of different targets that might be encountered - human beings, light vehicles, large vehicles, buildings and other static objects - it is difficult to develop a target model that would be applicable for all of them. The resolution of the CPD radar is not high enough to allow a lot of information about target shape and orientation to be obtained, so a more complicated motion model using target heading and turn rate cannot be used.

The predicted state $x_{k|k-1}^{(j)}$ given posterior state $x_{k-1}^{(j)}$ and noiser term v_k is given by:

$$\mathbf{x}_{k|k-1}^{(j)} = \phi_k(\mathbf{x}_{k-1}^{(j)}, v_k) \tag{5.9}$$

In this thesis this is implemented as:

$$\mathbf{x}_{k|k-1}^{(j)} = \begin{bmatrix} x_{k-1}^{(j)} + \dot{x}_{k-1}^{(j)} * \Delta t \\ y_{k-1}^{(j)} + \dot{y}_{k-1}^{(j)} * \Delta t \\ \dot{x}_{k-1}^{(j)} \\ \dot{y}_{k-1}^{(j)} \end{bmatrix} \tag{5.10}$$

where Δt (s) is the time difference between this measurement and the last update of the PHD filter.

This linear transform function ϕ_k can be written as a matrix F [13]:

$$\mathbf{x}_{k|k-1}^{(j)} = F x_{k-1} + Q_k \tag{5.11}$$

$$F = \begin{bmatrix} 1 & \Delta T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.12}$$

5.4.3.3 Motion noise covariance Q_k

The motion noise covariance Q_k is of the form

$$Q_k = \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 \\ 0 & 0 & \sigma_{v_x}^2 & 0 \\ 0 & 0 & 0 & \sigma_{v_y}^2 \end{bmatrix} \tag{5.13}$$

This is used to increase the target covariance in the prediction step, representing the increase in target state uncertainty due to the estimated movement between observations. If Q_k is too low, the covariance's lower bound is controlled solely by the observation noise covariance R_k.

The accuracy of the velocity estimate is a function of the sensor noise, as the position estimate is used to calculate the velocity, which is used to predict the position of the object. It is also linked to the sensor measurement rate; too low a measurement rate will give a poor estimate of the instantaneous velocity of an object, too high and the

magnitude of the measurement noise will be large compared to any change in position.

For these tests we model a slow rate of movement, or a faster rate of movement with a hypothetical sensor that measures faster. The selected prediction standard deviation values are $\sigma_x = \sigma_y = \sigma_{v_x} = \sigma_{v_y} = 0.1$, determined through trial and error.

5.4.4   Parameters for the Update Step

5.4.4.1   Extraction of features from measurements

The detection threshold developed in Chapter 3 is used to extract features from measurements. Any bin containing a power value above the threshold is treated as a detection of an object. It is converted to a range-bearing measurement, with the object assumed to lie at the same range as the centre of the range bin, and at the same bearing as the beam axis. If there is more than one bin above the threshold in a given measurement, there will be multiple range-bearing measurements with the same bearing.

5.4.4.2   Measurement model h

Objects are detected by the sensor, which returns a set of observations Z_k of the form:

$$\mathbf{Z}_k = \{\mathbf{z}_k^{(1)}, \mathbf{z}_k^{(2)}, ... \mathbf{z}_k^{(n)}\} \tag{5.14}$$

$\mathbf{z}_k$ are in sensor-relative polar coordinates:

$$\mathbf{z}_k^{(i)} = \begin{bmatrix} r_k^{(i)} \\ \theta_k^{(i)} \end{bmatrix} \tag{5.15}$$

where $r_k^{(i)}$ is the range measured to the ith feature relative to the sensor at time k, and $\theta_k^{(i)}$ is the bearing. The sensor model h is used in the update step to simulate the generation of polar position measurements from the Cartesian state vector $\mathbf{x}$.

The sensor is unable to detect velocity, as the CPD radar is not configured for

Doppler. Therefore the sensor model $h(\mathbf{x}_S, \mathbf{x}_F)$ will only detect the first two dimensions of object state.

$$h(\mathbf{x}_S, \mathbf{x}_F) = \begin{bmatrix} \sqrt{(x_F - x_S)^2 + (y_F - y_S)^2} \\ tan^{-1} \frac{y_F - y_S}{x_F - x_S} - \theta_S \\ 0 \\ 0 \end{bmatrix} \tag{5.16}$$

To include velocity in the measurement estimate, the velocity is calculated using the measured position and the position estimate prior to the prediction step.

$$h_{vel}(\mathbf{x}_S, \mathbf{x}_F, m_{k-1}) = \begin{bmatrix} \sqrt{(x_F - x_S)^2 + (y_F - y_S)^2} \\ tan^{-1} \frac{y_F - y_S}{x_F - x_S} - \theta_S \\ \frac{x_F - m_{k-1,x}}{\Delta t} \\ \frac{y_F - m_{k-1,y}}{\Delta t} \end{bmatrix} \tag{5.17}$$

Although the measurement is in polar coordinates, the simulated velocity measurement is Cartesian.

For ease of plotting the measurements, the inverse sensor model $h^{-1}(\mathbf{z}_k^{(i)}, \mathbf{x}_S)$ is

$$h^{-1}(\mathbf{z}_k^{(i)}, \mathbf{x}_S) = \begin{bmatrix} r_k^{(i)} * cos(\theta_k^{(i)} + \theta_S) + x_S \\ r_k^{(i)} * sin(\theta_k^{(i)} + \theta_S) + y_S \end{bmatrix} \tag{5.18}$$

It is used to convert from sensor-relative polar coordinates to global-frame Cartesian coordinates. This can be used to plot the raw measurements on the same axes as the state vector $\mathbf{x}$.

5.4.4.3 Measurement noise covariance R_k

The EKF makes an assumption of a Gaussian distribution. If it is Gaussian, we expect 68% of all measurements to lie within one standard deviation of the mean, and 95% to lie within two. This influences the values selected for R_k.

The measurement noise covariance R_k is of the form

$$R = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix} \tag{5.19}$$

The values chosen for σ_r and σ_θ should be dependent on the characteristics of the sensor. Low measurement covariances will mean the object position estimate will more closely track the measurement, and the position covariances will be lower. High covariances indicate a less trustworthy measurement; the measurements will not influence the state estimates as strongly, and the covariances will be higher.

5.4.4.4   Measurement range noise covariance σ_r

The CPD radar returns measurements with a range resolution equal to its bin length. If a bin contains a power value above the threshold, it is converted to a range measurement equivalent to the distance to the midpoint of the bin.

Assuming it is only detected in one bin, the maximum range error possible for a small or flat target (such as corner reflectors, or the surface of walls) should be half a bin length (either positive or negative). Neglecting the difference in received power with range over the length of one bin, a target will be detected almost identically regardless of its position in the bin. Since objects are equally likely to exist at any range within the bin, a uniform distribution across this range would be the best model.

Larger objects, phase noise, poor linearity and clutter effects may contribute to detections that appear in multiple bins. Additional processing could be used to try to improve the range accuracy for such objects (for example, polynomial interpolation between the power values of consecutive bins of one measurement to determine the precise location of a peak in power). But multiple or extended objects may truly be present in multiple bins, and so there should be detections in multiple bins, not just the peak. The error in range estimate will therefore be dependent on the structure of the environment and the type of objects present. Testing would be required to determine the shape of the error distribution.

The range error distribution for small single-bin objects is approximately uniform between $\frac{BinLength}{2}$ and $\frac{-BinLength}{2}$, and for multiple-bin objects it is environment-dependent and would require testing. The smaller errors will mostly be due to the limited range resolution of the sensor; larger errors will be due to poor chirp linearity, scintillation, phase noise, clutter or other negative influences. Since the Kalman filter requires a zero-mean Gausssian distribution, we are forced to approximate the overall noise distribution of the measurement as a Gaussian defined by its covariance.

Trial and error shows that a value of σ_r between $\frac{BinLength}{4}$ and $\frac{BinLength}{2}$ work well;

these values imply between 5% and 32% of range measurements are in the wrong bin.

$$\sigma_r = \frac{\text{BinLength}}{4} \tag{5.20}$$

produces reasonable results.

5.4.4.5   Measurement bearing noise covariance σ_θ

The radar's bearing resolution is a function of the beam width and the antenna sweep pattern. The CPD radar beam is $6°$ wide in azimuth and all measurements are assumed to lie on the beam axis. Although the gain drops off, particularly strong targets may be detected at angles beyond the 3dB width, but this is difficult to model as it is a function of range, target geometry and target aspect, and will vary wildly between environment and target type. The gain profile of the beam follows a sync function. This is not the same as a Gaussian, but it makes the Gaussian assumption appear to be more reasonable than it is for the uniform distribution of range error; both the Gaussian and sync distributions have a peak in the middle decreasing either side of the mean.

Similar to σ_r, trial and error show that a σ_θ value of between $\frac{\text{BeamWidth}}{4}$ and $\frac{\text{BeamWidth}}{2}$ produces acceptable results.

5.4.4.6   Probability of object detection, p_D

The assumption in [74] that probability of detection is independent of state can be relaxed to allow a state-dependent detection probability that more accurately reflects the characteristics of the sensor, the environments and the objects to be detected.

Objects may pass behind other objects and be occluded from the sensor. If the probability of detection is not updated to reflect this, these objects' weights will be reduced substantially- more substantially than they should; the PHD filter has less memory than it should in the case of a missed detection [13].

It has been shown that excessively low values of p_D leads to very large increases in the weights of targets [49] [20] [13] which produces inaccurate estimates of the number of targets present in the scene; this is a recognised flaw in the PHD filter. The cardinalised PHD filter is less susceptible to this effect, but is beyond the scope of this research.

A state-dependent probability of detection is designed and implemented in [20] for a

laser rangefinder. In that implementation, the probability of detection is scaled down for objects at greater range but similar (or equal) bearing. This is reasonable for a sensor such as a LIDAR or camera, where only one object is detectable along a particular bearing in 2D space. This is *slightly* less of a concern with radar; the wider beam makes it more difficult for nearer objects to totally occlude further ones, so multiple objects can be detected on a single measurement's bearing. It must still be recognised that it is likely that physically larger objects will totally occlude smaller objects behind them, but the CPD radar does not provide information about the physical size of an object, only its effective radar cross section. Larger objects tend to have larger radar cross-sections, but this is not guaranteed. The probability of detection decreases with range for most objects, due to reduced reflective power, but this is not the case for all. Modelling and testing a varying p_D for the CPD radar is beyond the scope of this thesis, therefore a constant p_D will be utilised.

The chosen value is $p_D = 0.66$ for the CPD radar. This is based on visual inspection of plotted data; a more thorough approach would involve counting the number of scans where a known object was detected or not, but this would involve selecting a specific size and type of target as a baseline. Generally, probability of detection decreases at range, dropping off for an "average" target (approximately 0.1m^2 RCS) approximately two-thirds of the way to maximum range, with detection only sporadic beyond that range.

5.4.4.7   Clutter density model

The clutter model is a Poisson random finite set K_k, with an intensity of κ_k [74]. The mean number of clutter detections are uniformly distributed across the surveilled area in each measurement.

$$\kappa_k(z) = \lambda_c V u(z) \tag{5.21}$$

V is the surveilled area in m^2 (or volume, in a 3D implementation), calculated using the range of the radar and scanned bearing range. λ_c is the expected number of clutter measurements per unit area, in measurements/m^2. $\lambda_c * V$ returns the average number of measurements expected in a single scan, so λ_c can be set as simply $\frac{N_{clutter}}{V}$ where $N_{clutter}$ is the expected number of clutter measurements per scan. $u(z)$ is the uniform distribution across the observed environment, chosen because clutter is assumed to be evenly

distributed across the environment. For ease of implementation, the uniform distribution is modelled across the rectangular region rather than fan-shaped sector scanned by the radar.

λ_c is highly dependent on the characteristics of the sensor and the environment. In the clear environment of a carpark where the ground is flat, fairly smooth concrete, there are only a small number of clutter or noise measurements, so a value of

$$\lambda_c = \frac{2}{V} \tag{5.22}$$

works well. Increasing this will slow the rate of weight increase but allow the filter to exclude clutter measurements more reliably, but possibly exclude some valid objects also. Decreasing λ_c will increase the rate of weight increase, allowing objects to be identified faster, but allowing more clutter to slip through.

5.4.5 Parameters for Merging and Pruning

5.4.5.1 Pruning threshold T

The pruning threshold T is the weight minimum below which tracked objects are deleted from the PHD filter after the update step. Every update step produces a combinatorial explosion of new tracks added, and this will grow exponentially if the number of tracks is not trimmed down. The majority of new tracks added are the result of incorrect association between measurements and tracked objects, and so will have very low weights.

Some tracked objects will fail to be detected due to occlusion or moving out of the field of view, and some tracks will originate from clutter or noise that exceeded the detection threshold. These objects will have their weights reduced in the prediction step and the missed detection stage of the update step. After enough failed detections, their weight will fall below T and they will be pruned from the filter.

T needs to be low enough not to delete a valid track after just one bad or missed detection, and high enough that the filter does not "clog up" with too many spurious tracks. [74] used a value of $T = 10^{-5}$ which works well in this implementation; a newly spawned track must fail to be detected twice before it is pruned out, since the missed detection stage of the update scales the weight as:

$$w_k = (1 - p_{D,k})w_{k|k-1} \tag{5.23}$$

and the prediction step scales the weight as

$$w_{k|k-1} = p_S \, w_{k-1} \tag{5.24}$$

With $p_D = 0.99$, $w_{k|k-1} = 5 * 10^-3$ (newly spawned) and $p_S = 0.99$, the track's weight will be $5 * 10^{-5}$ after one missed detection, $4.95 * 10^{-5}$ after the prediction step, and $4.95 * 10^{-7}$ after two missed detections. At this stage it is below the threshold and will be deleted. Objects that have been detected several times and have stronger weights will require more missed detections before they are deleted.

5.4.5.2   Merging threshold U

The merging threshold U is the Mahalanobis distance threshold within which two nearby tracked objects will be merged into one, after the weakest tracks have been pruned out. This threshold is important in that it prevents there being repeated tracks for the same object. An object that is detected in multiple measurements multiple times will spawn new objects, and create new objects when the measurement is matched against existing tracks in the update step. When these separate tracks are almost identical, merging them reduces the number of objects that the filter must process. By combining the weights of the merged tracks, the weights of existant objects increase faster.

If the value of U is too low, there will be multiple objects tracked in nearly the exact same position, slowing the filter's processing and inflating the estimated number of tracked objects if they are a concern. If the value of U is too high, tracks corresponding to objects that are quite distant from each other will be merged incorrectly, reducing filter accuracy and potentially deleting valid objects entirely.

At this application's scale, the environment consists almost entirely of extended objects which return more than one detection per scan. Strong reflectors can be detected in multiple bins per measurement due to imperfect linearisation and phase noise, and in multiple consecutive measurements due to the width of the beam. Large objects will naturally be detected in multiple measurements as the beam sweeps across them. This will cause several tracks to be instantiated in close proximity to one another. This has negative effects on filter performance via increased processing load and some unwanted effects in the update step (described in Sec. 5.5.1).

In some situations, such as a when tracking a small trihedral corner reflector, it would be ideal for the all these tracks to be merged into one at the centroid of the

object. But for other objects, such as walls or light vehicles, the object extends over such a large area that it is both impossible and undesirable to represent it with a single point; the object's true extent should be represented, and if the value of U were high enough to combine all the points into one, it is likely that separate objects would also be merged together. More sophisticated partitioning and merging algorithms exist but none of these will address the real problem: a point target representation is not ideal for tracking extended objects.

The merging threshold is difficult to calculate. The Mahalanobis distance is a function of both covariance and the difference between means, making it difficult to visualise:

$$MahalanobisDistance(i, j) = (m_k^{(i)} - m_k^{(j)})^T * (P_k^{(i)})^{-1} * (m_k^{(i)} - m_k^{(j)}) \qquad (5.25)$$

In this calculation, dimensions with lower covariances are more strongly weighted in the difference calculation. Because tracks are spawned with a high velocity covariance compared to the position covariance, newly spawned targets have the position coordinates dominate the Mahalanobis distance calculation.

The value of U used in this implementation $U = 10$, determined using experimentation. Lower values were found to lead to too many targets almost superimposed on one another, while large values lead to separate objects being merged together.

5.4.5.3   Maximum number of tracked objects J_{max}

In this implementation no limit is placed on the number of tracked objects, therefore J_{max} is not used. The environment is unknown and could contain any number of targets.

5.5   Results of Radar Mapping with the PHD Filter

The GM-EK-PHD filter was implemented for this thesis in Matlab with the parameters described in Sec. 5.4. Its performance was analysed with both simulated and real logged data from the radar.

5.5.1 Simulated Data

A number of synthetic radar datasets are generated to test the GM-EK-PHD filter implementation performance before using noisier, real datasets. The synthetic data consists of simplified, noise-free measurements of the same structure as the real CPD data, simulating simple test environments:

- a point target moving away from the sensor

- two point targets moving away from the sensor

- a static extended target

- an extended target moving away from the sensor

These simulations are used to verify the performance of the implemented filter. For simulated point targets they function as expected. However there are issues for the extended target which will now be discussed.

The primary problem is that of static extended targets being believed to have velocities perpendicular to the beam axis. A simple explanation for this is illustrated in Fig. 5.2 which will be referred to in this discussion. In the simulations, the static extended target is represented as four consecutive detected objects at bearings $-4°$, $-6°$, $-8°$ and $-10°$ relative to the radar axis, all in bin 10. The moving extended object starts in a nearer bin and increases its range by one bin each scan. At longer range this would correspond to a larger object than at close range (due to the longer arc length) but this inconsistency is neglected.

The filter treats the extended target as four independent point targets, shown in Fig. 5.2 (a). Every scan there are four new measurements (Fig. 5.2 (b)), and all four are used to update each of the four known targets. This spawns 16 new targets, 4 of which correspond to correct matches between measurement and known target, and 12 which are incorrect (Fig. 5.2 (c)). The updated targets are shown in the same column as the target that generated the measurement used to perform the update. Their position in the X and Y dimensions is a function of their covariance and the noise covariance of the measurement; for the sake of this discussion they are assumed to move to the exact same position as the measurement.

There are also four targets based on the assumption the known targets may not be observed in the newest scan; their weights are typically very low if there is a high

probability of detection p_D. These are omitted from Fig. 5.2 for simplicity.

In this implementation, the sensor does not detect the target's velocity, only its position. The incorrect associations between known target and observation will be given a velocity estimate in the direction from $m_{k-1}^{(i)}$ to $m_k^{(j)}$. The further the separation between these two states, the larger the velocity will be; in Fig. 5.2, matches between the purple target and blue measurement result in a large velocity as the purple measurement was furthest away from the location of the measurement.

Assuming the true target is stationary, and there is no noise on the measurements, each measurement will spawn a number of new targets with some velocity component moving in the direction of the new measurement, based on the assumption that a measurement associated with target $m(i)_k$ is associated to a target that was in position $m(j)_{k-1}$ (i.e. incorrect association) . The further the separation between $m(i)_k$ and $m(j)_{k-1}$, the larger the velocity.

Reweighting of targets is then performed (Fig. 5.2 (d)). Targets with a lower velocity will better match the previous target state of zero velocity, and so will have larger weights.

Pruning takes place next and will remove the worst matches with the lowest weights. The new targets generated from the "less wrong" but still incorrect associations between measurements and targets will be merged with the "correctly" updated targets. The final result will be targets in the correct position, but with a net velocity (Fig. 5.2). The direction and magnitude of a given point target's velocity is a function of the object's shape and size, and the position of the point target within it. The points on the extremities of the extended target will have a larger velocity, due to their greater maximum displacement from the other targets and the fact that the velocities will all be in the same direction. This creates a sort of "splitting" effect where the component points of the target are moving apart. This is illustrated in Fig. 5.3, which is a plot of filter output for the static extended target case. A larger extended object will have larger velocities at its extremes.

(a) Four static targets currently in the filter, located side by side in a horizontal line

(b) Four new measurements of these targets

(c) 16 new targets produced by updating the known targets with the new measurements. Each column contains the results of updates from one measurement. The colour of the centre of each circle shows which target was updated. The coloured ring around each circle shows which measurement was used to perform the update. The length of the white arrow shows the direction and relative magnitudes of the newly estimated target velocities (not to scale).

(d) Targets reweighted. Larger circles mean a higher weight (not to scale).

(e) Targets pruned and merged (down the column). The new estimates retain some velocity.

Figure 5.2: If there are multiple objects detected in close proximity to one another, but not close enough to be merged, the PHD filter's update process can cause them to be estimated to have velocities even when this is not the case.

Figure 5.3: A static extended target's component targets will appear to have velocities attempting to "split" the object. The arrows show the magnitude and direction of the target velocity. The red plus signs are the most recent measurements, and the cyan triangles are the target mean estimates, and the blue ellipses are the one standard deviation covariance ellipses.

This effect is present in the moving target case also, adding an additional component to the velocity vector. In the simulated case where a wide 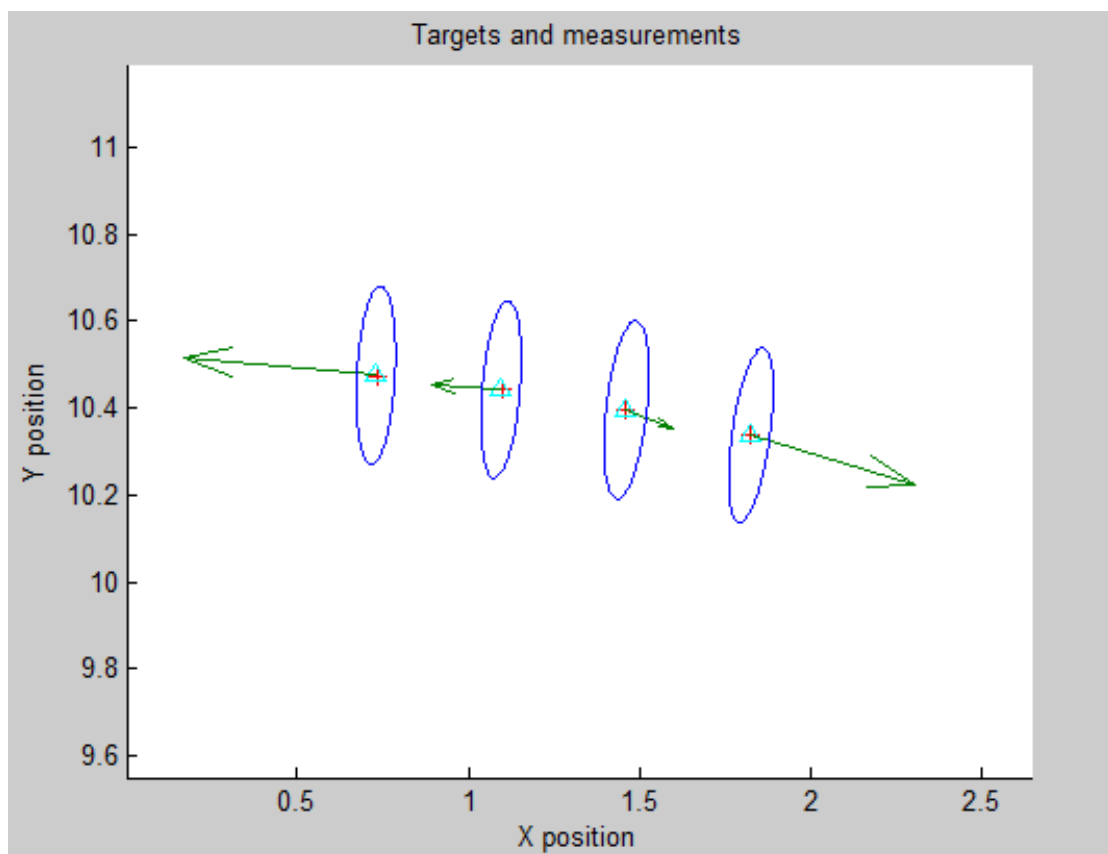target moves away from the sensor at constant speed, this has the effect of widening the range of directions of motion - creating a wider "fan" of velocity vectors (see Fig. 5.4) than would truly be the case.
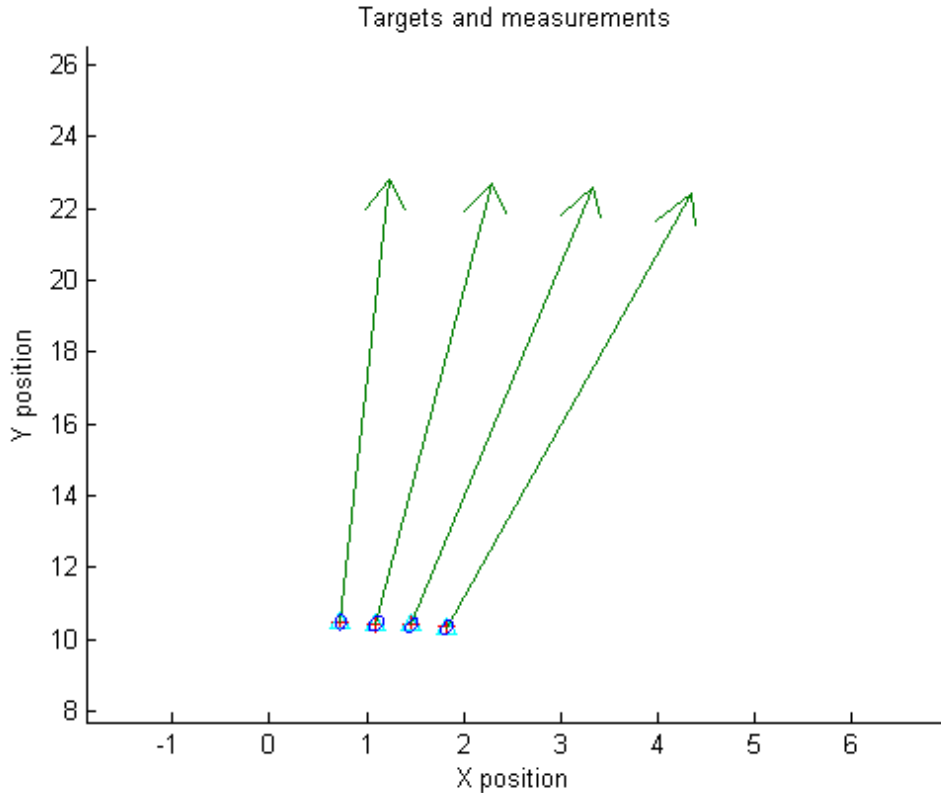
Figure 5.4: A moving extended target's component targets will appear to have a larger component that is perpendicular to the beam axis. The true velocity headings are along the axes of -4° , -6° , -8° and -10° . The estimate approximates these but with an enlarged cross-range component (see Fig. 5.4); that is, the estimated "fan" shape of the four velocity vectors is wider than the true velocity vectors.

The velocity estimate will eventually converge towards a more accurate speed (see Fig. 5.7) and bearing (see Fig. 5.8). The rate at which this occurs can be manipulated by changing the birth speed covariance. Both these figures show the results from the static and dynamic test cases with different birth speed covariances ($\sigma_{V_x}^2 = \sigma_{V_y}^2 = 25m^2/s^2$ and $\sigma_{V_x}^2 = \sigma_{V_y}^2 = 10m^2/s^2$).

Figure 5.5: The PHD filter estimates of target position in the X-dimension, for static and dynamic objects, with different values for the speed covariance at target birth. This is generated in simulation using noise-free measurements.
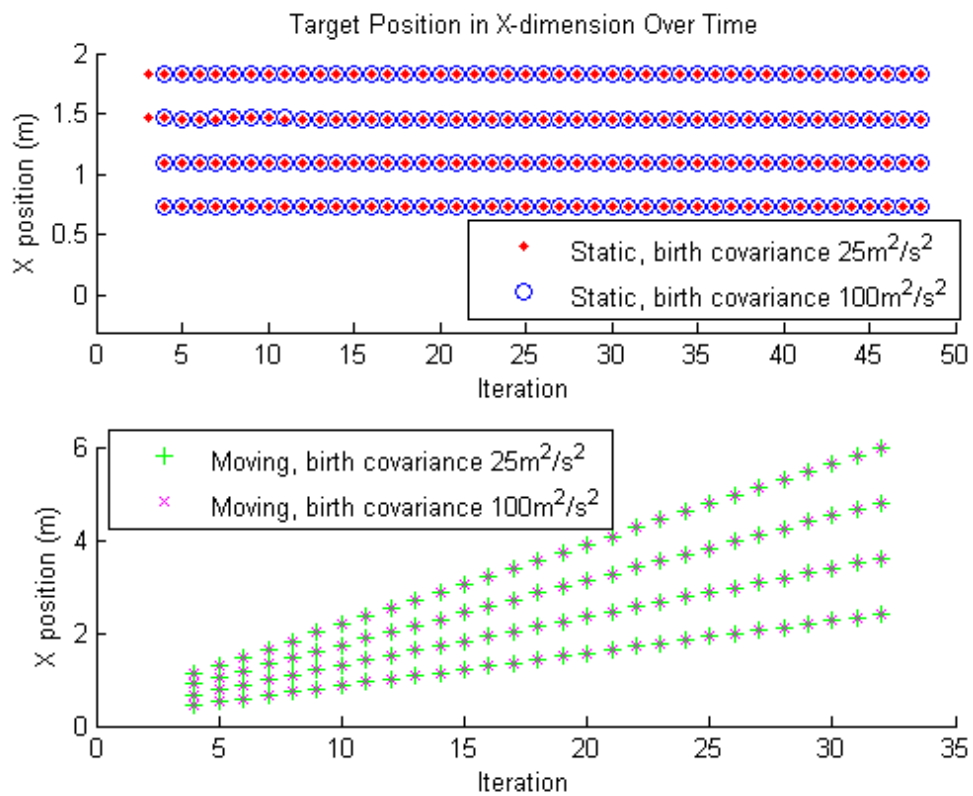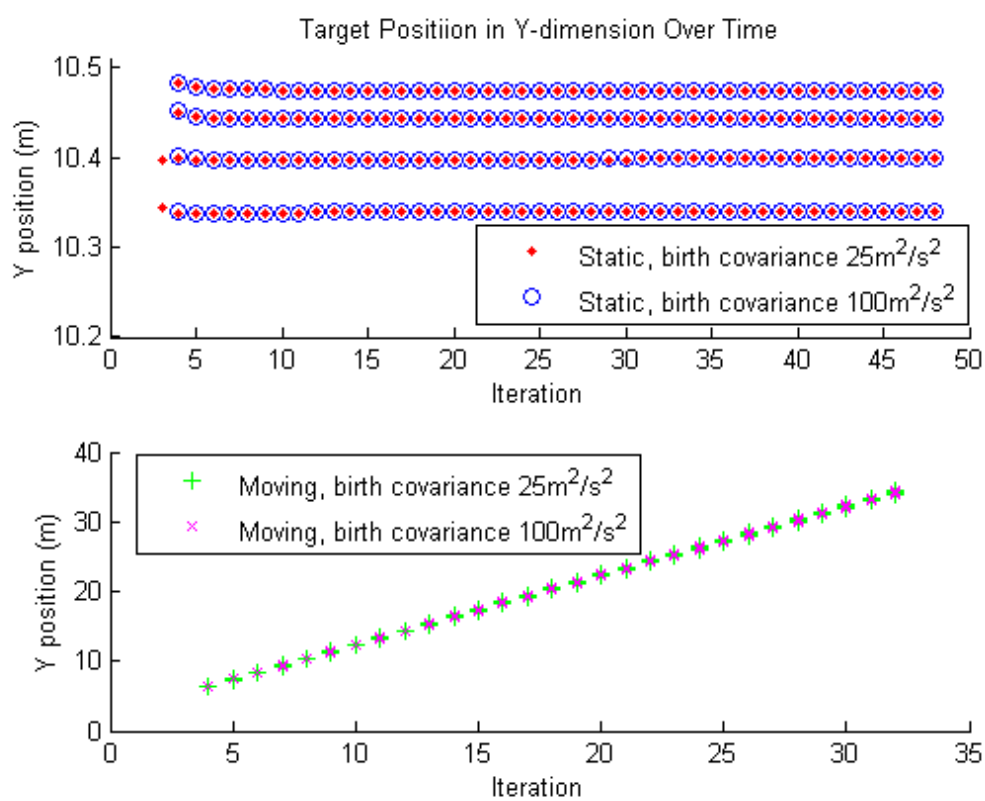
Figure 5.6: The PHD filter estimates of target position in the Y-dimension. This is generated in simulation using noise-free measurements.

Fig. 5.5 and Fig. 5.6 show that in this noise-free simulation, varying the speed covariance at birth does not influence the accuracy of target state estimation in X or Y. However, in the static target case, a lower speed covariance causes the target's weight to increase more rapidly; it is classified as a real target (with weight greater than 0.5) after only 3 iterations rather than 4.
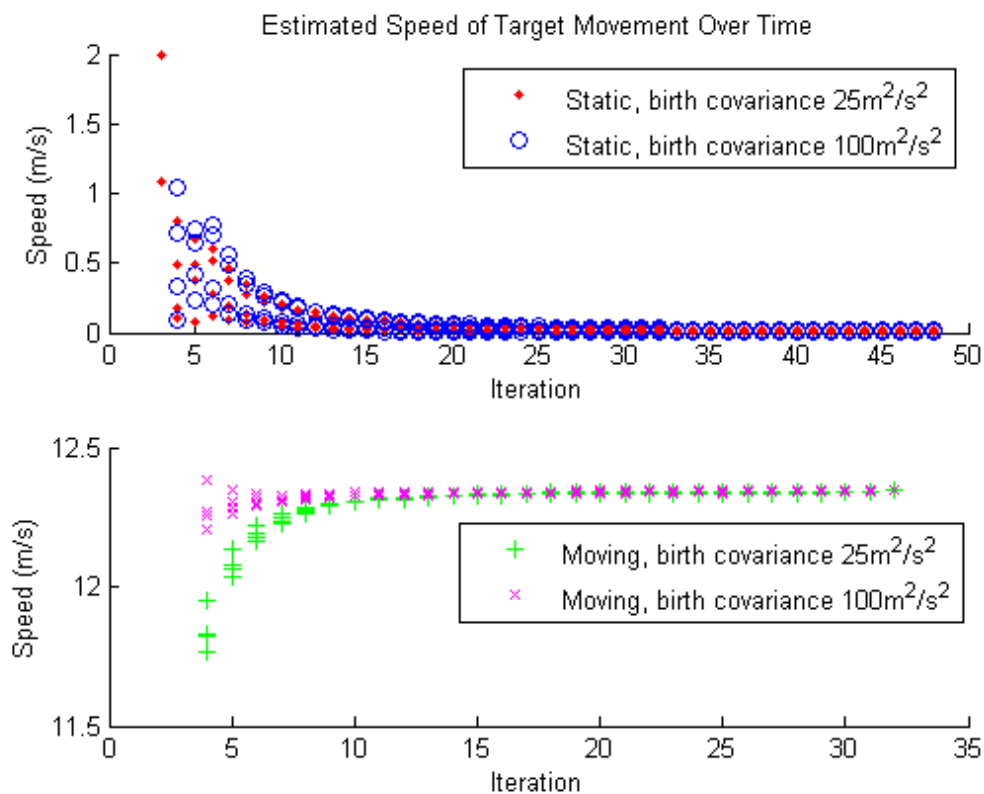
Figure 5.7: The estimated speeds of a static or dynamic extended object's component targets in one simulation. In both static and moving cases, both the speed error and the range of the errors reduce over time. A lower birth covariance for the speed increases the rate of convergence to the true speed for a static target, but slows it for a moving one.

Fig. 5.7 shows that in the static test case, a lower speed covariance will cause the filter to converge to a more accurate speed estimate faster than for a higher birth speed covariance. For the moving target case, the reverse is true - a larger birth speed covariance allows the speed estimate to converge more rapidly to the correct value.
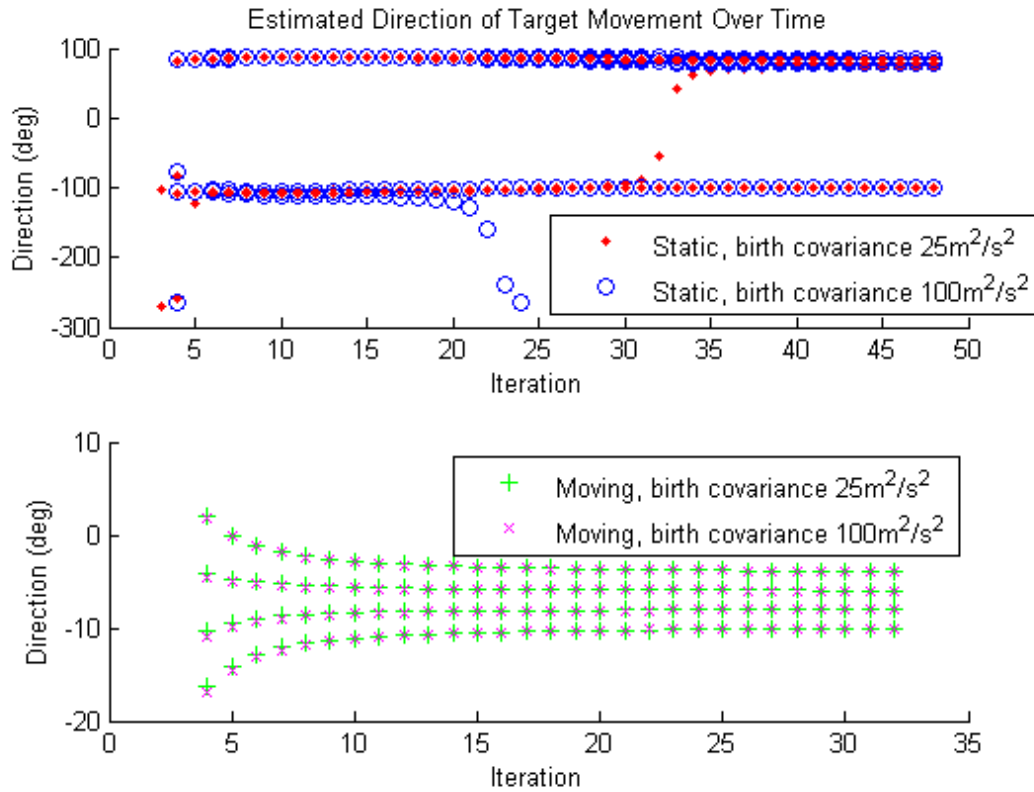
Figure 5.8: The estimated direction of movement of an extended object's component targets in simulation. In the static case, the component targets are typically in two clusters moving in opposite directions (with a difference of 180° between them). One point changes direction in the course of the simulation, but by this timestep, the estimated velocity is so low that this will have very little negative effect (see Fig. 5.7). Changing the birth covariance influences the update to the extent of changing which direction one of the tracked targets moves in, but the overall effect of this is small. For the moving target there are less drastic changes; the headings of all target stabilise quickly, but with a clear separation between each.

The change in estimated heading is an artifact of slight oscillations in the estimated position of the target within the filter. The estimated range to the target changes very little (see Fig. 5.9) and its speed is very low (see Fig 5.7).
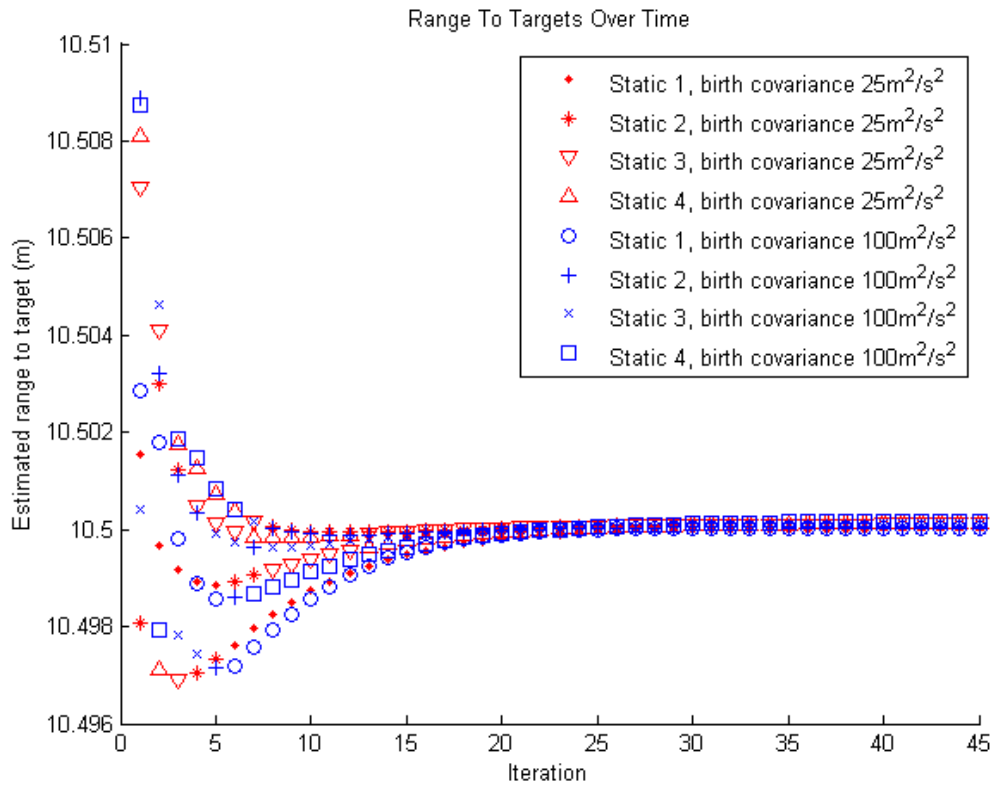
Figure 5.9: The estimated ranges of a static extended object's component targets in one simulation. The labels ("Static 1", "Static 2", etc) refer to the target's position in the PHD filter target listing and are not specific to individual targets; the labels can and do swap between targets (eg: the red static 1 and 2 swap at iteration 6). Note that the absolute change in range estimate is very small, less than 0.01m, but this is enough to produce an apparent change in heading in Fig. 5.8.

5.5.2 Real Data

The next set of tests uses the CPD radar to gather data about a real environment. In all real tests, ground truth measurements are made using a SICK LMS200 2D scanning laser rangefinder. The radar is kept static, and a human being moves in the environment in front of it. In some tests, corner reflector targets are carried to simulate larger targets.

The tests are performed in an open environment on the Seymour Centre car park roof, with no other dynamic targets.

The lidar produces a very different representation of the scene to the radar. Fig. 5.10 shows a plot of a lidar scan of the test environment. The extended targets - the long wall (green) on the left of the sensors, and the shorter wall (blue) on the right - are not detected as clearly by the radar. The presence of the laser and the vehicle's bumper bar prevents the radar from sensing more than 50° to the left, so it is only able to detect the long wall at a further range, where it is a weaker reflector. The radar detects it as a "blob", rather than a long line. It is a similar case with the shorter wall on the right, though a much larger proportion of the wall is detected.
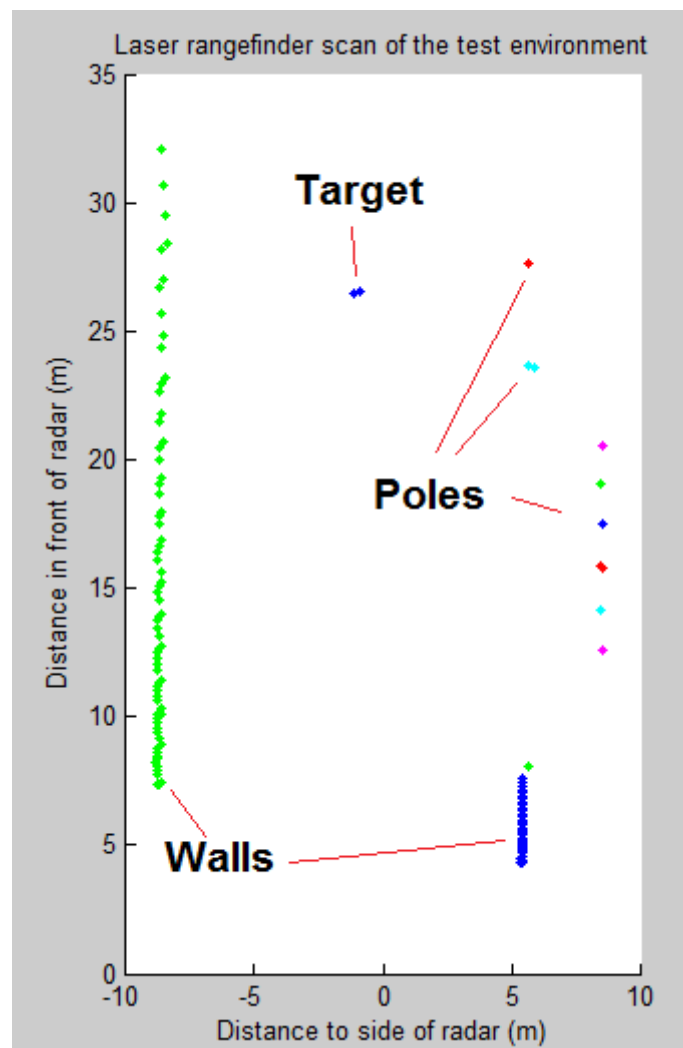
Figure 5.10: A laser scan of the test environment. Laser measurements are clustered together, shown as adjacent points of the same colour. The long wall on the left and a shorter wall on the right are clearly visible. A series of six poles are detected on the right. The small blue cluster approximately 26m in front of the radar is a human being acting as the target.

Due to the extended nature of many objects in the environment, both the radar and laser will frequently return multiple measurements from a single target. When these radar measurements are passed into the PHD filter, they are used to form target state estimates. Some will be merged or deleted, but not all. The PHD target state estimate at any given time will usually include multiple point estimates corresponding to a single target.

For the purpose of this analysis, these measurements are clustered together, with the goal of measurements of one object being in a distinct cluster from measurements of a separate object. Clustering is not necessary to use the PHD filter output, but it makes it much easier to quantify target performance in this case; rather than examining a large number of points, we reduce it down to much fewer and can analyse the using the OSPA performance metric (see Sec. 3.5.1). There are many different clustering approaches available and a thorough review is outside the scope of this work. A simple distance heuristic is used: any two radar measurements within a fixed distance threshold $d_{cluster}$ of one another are treated as being spawned by the same object. A more advanced heuristic might produce better results, but as this test environment is controlled and the ground truth is known, a visual inspection can be used to verify performance.

The value of $d_{cluster}$ is adjusted manually until satisfactory performance is achieved; it needs to be slightly longer than one bin length (so that consecutive detections in the one measurement are clustered together) but not so large that distinct objects are conjoined. A value of $d_{cluster} = 1.5m$ produces good results.

The raw laser measurements are clustered together in a similar manner, but with a clustering threshold that is a function of the range and bearing between points; this is more feasible than with radar due to the high angular resolution of the lidar.

Simply sorting measurements into clusters is not sufficient for performance analysis; this has not reduced the total number of measurements to be analysed, it has only split them into groups. Though the full lidar and radar measurement sets are produced by the same objects in the environment, they do not correspond well due to the difference in resolution and modality of the sensors. Processing the radar measurements within the PHD filter does not improve the correspondence.

To use the laser data as ground truth to analyse the CPD radar's performance via the OSPA metric, each cluster of laser measurements and PHD filter estimates needs to be converted to a representation by a single point. Otherwise, the high angular resolution of the lidar would lead to a mismatch in the cardinality of the set of laser measurements

and the set of target state estimates; this would inflate the resulting OSPA calculation to the point of uselessness.

The question, then, is exactly how to convert many points in a cluster to a single representative point. This is not difficult for small, tightly concentrated clusters, but for the more extended ones (such as the walls) it presents a major challenge; a single point cannot effectively represent a wall 20m long. Rather than have to implement an extended target representation, the decision was made to exclude the large wall on the left from the processed area.

Two regions are considered for performance metric analysis:

- The region of space within 5m of the left of the radar and 10m of the right of the radar. This removes the long wall to the left but retains the shorter wall on the right, as well as several metal poles forming a barrier.

- The region of space within 3m to the left or right of the radar. This excluded everything except the human being and any corner reflector they may be carrying directly in front of the sensors.

The narrower region is useful for investigating the maximum accuracy of the PHD filter with the CPD radar. The wider region is useful for investigating a more general scene, containing a mix of small and large, static and dynamic targets.

Within these regions, the clusters of radar and laser points are all converted to a single point. The laser clusters are all small enough that the centroid of each cluster is used. The radar clusters can be spread over a larger area. Two different single-point representations are considered for them:

- The centroid of the cluster. This is the most intuitive from a target tracking perspective. The target centroid should move in a fairly consistent manner relative to the sensor, centred on the object, even as the total radar cross-section fluctuates.

- The point of the cluster nearest to the front of the vehicle. This is the most intuitive from a safety perspective, as it is difficult to distinguish between a strong reflector far away but with wide phase noise skirts, and a strong reflector far away with a weak reflector in front of it.

The performances using these two types of reference points will be compared.

The tests being analysed are the same that are used to develop the detection thresholds in Chapter 3, namely

- A human being as a target

- A $1m^2$ RCS corner reflector as a target

- A $10m^2$ RCS corner reflector as a target

In all tests the target moves back in a fairly straight line in front of the static laser and radar, out to a distance of approximately 35m. The ranges beyond which the target is undetected by the radar are not considered. Fig. 5.11 shows measurements from the laser rangefinder and radar.

For each update of the PHD filter, the last laser scan made before the update was performed is used as ground truth.

For the narrower test where only the one target was observed, the range from the laser rangefinder, the mean to the target centroid, and the range to the nearest point on the target were all plotted. Fig. 5.12 to Fig. 5.20 show the outputs from the PHD filter tracking different sized targets in the narrow region directly in front of the radar. The laser rangefinder measurements, drawn as black crosses, are treated as the "ground truth" and are assumed to be accurate to within 10cm. The lidar range estimate is drawn as a black '+'. The radar generates multiple detections for each measurement, each drawn as a green 'x'. The centroid of these is drawn as a blue dot, and the one closest to the radar is a red 'o'.

The CPD radar has a minimum range of 3 bins (3m), which is substantially further than the minimum range of the lidar (a few centimetres). When the true range is less than about 4m, both the minimum and centroid target representations overestimate the range to the target cluster. Beyond this range, the centroid representation begins to track the true range quite well. The minimum range estimate is approximately as accurate at longer range, when the target cluster is smaller and phase noise effects do not stretch the target towards the radar as much.

In these tests, the coordinate frame used is of **positive X being to the right** and **positive Y being up**. Since the tests conducted are of the target moving in a fairly straight line away from the sensors, the X coordinates should be close to zero, and the Y coordinates should have a similar shape to the total range estimate.

(a) Human being at 5m

(b) Human being at 10m

(c) 1m^2 corner reflector at 5m.

(d) 1m^2 corner reflector at 10m.

(e) 10m^2 corner reflector at 5m.

(f) 10m^2 corner reflector at 10m.
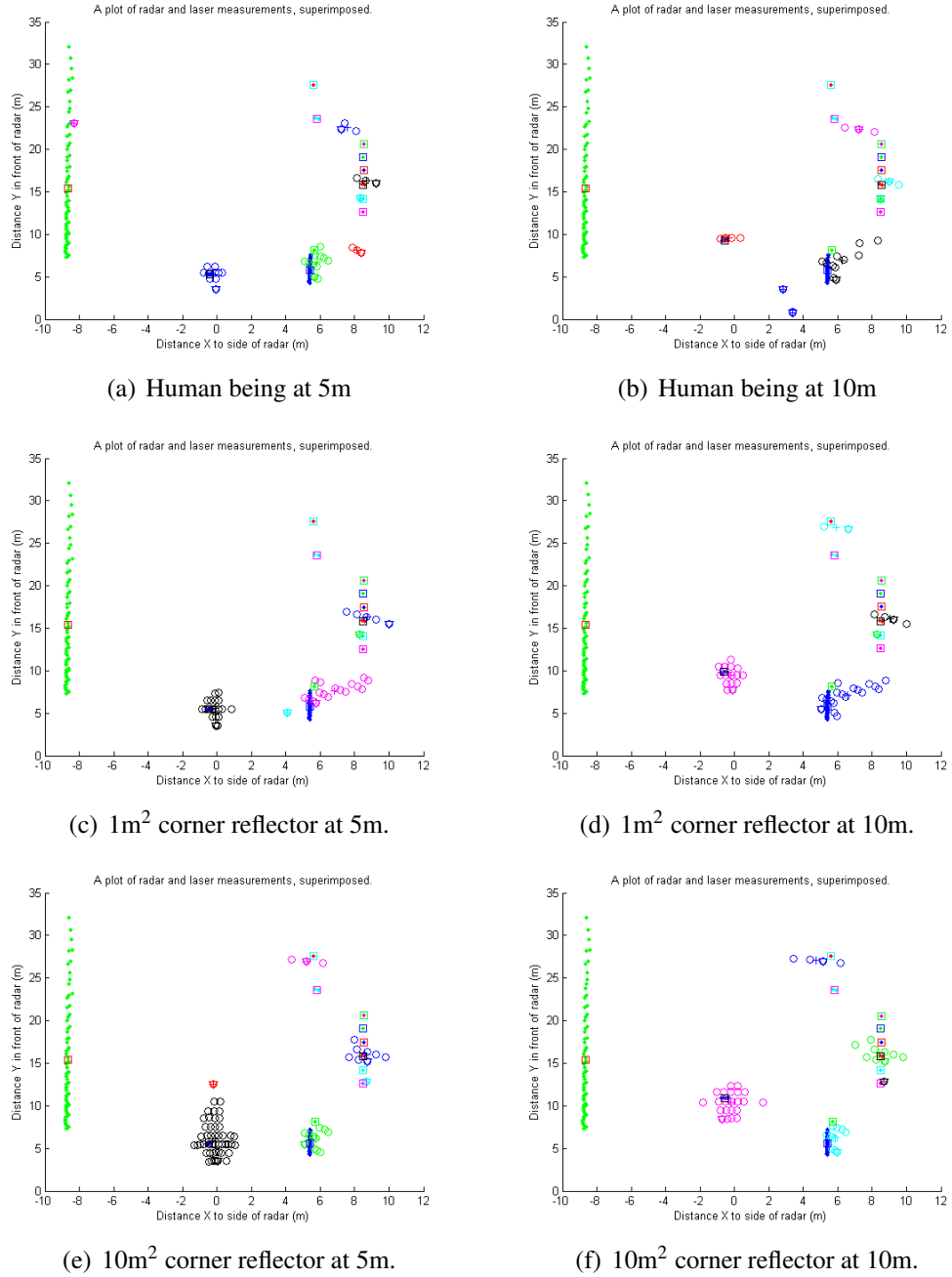
Figure 5.11: Different-sized targets are detected by the CPD radar and lidar at different ranges. Lidar measurements are drawn as dots. PHD filter target estimates from radar are drawn as circles. The centroid of a lidar cluster is drawn as a square. The centroid of a radar cluster is drawn as a '+'. The point in a radar cluster nearest the lidar is drawn as a downwards triangle.

5.5.2.1   Target Range Estimation With The PHD Filter

The CPD radar can only detect the human being reliably to a range of approximately 12m. Beyond this, there are increasing gaps between radar detections. Beyond 20m, the human being is not detected at all by the radar.

Fig. 5.12 shows that both the nearest and the centroid of the target cluster will overestimate the range to the target up to approximately 5m, when the centroid begins tracking more accurately, with the minimum not far behind. There is some phase noise at close range, particularly around iteration 40, where there are additional targets created around the 4m range while the true target is at a range closer to 6m.
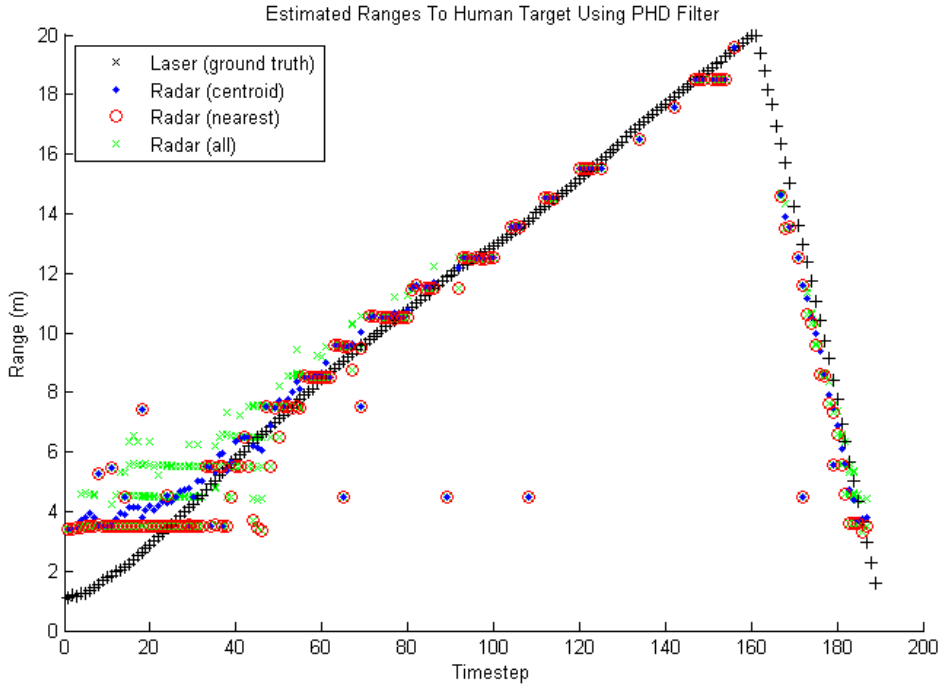
Figure 5.12: The PHD filter is used to estimate the range to a human being scanned by the CPD radar.

The 1m^2 RCS corner reflector target is detected reliably out to a range of approximately 20m, with detection failing beyond 25m. At close range, there is significant phase noise, leading to the target being detected in multiple bins; spans of 5-6m are noticed around a true range of 5m. Fig. 5.13 shows that nearest-point clustering approach (red circles) produces significant target range underestimates out to ranges of approximately 7.5m due to phase noise skirts extending towards the nearer bins of the

measurement. These detections are clustered around the true range, however, so the centroid method produces much more accurate results.

Figure 5.13: The PHD filter is used to estimate the range to a 1m^2 RCS corner reflector being scanned by the CPD radar.

The 10m^2 RCS corner reflector is a very large target. It is detectable reliably to ranges up to 30m, with more intermittent detections possible to 35m and possibly further; there was not enough space in the test region to try any further

Fig. 5.14 shows that at close range there are major phase noise effects; when the target is at a true distance of 9m from the target (iteration 50), the detection range of the phase noise skirts can span 5m either side of the true target. There are also secondary peaks formed behind the target, manifesting as additional targets with ranges of 10-15m while the true target is at ranges of 5-10m. The phase noise is significant enough to cause the target range to be slightly underestimated by the centroid between iterations 50 and 75 (between approximately 7.5m and 10m).
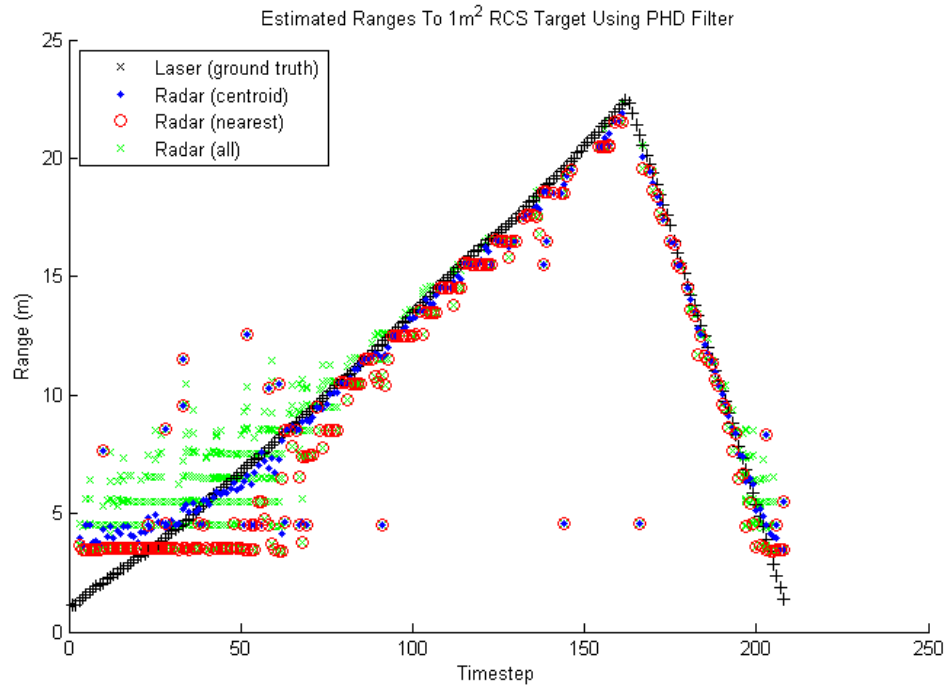
Figure 5.14: The PHD filter is used to estimate the range to a $10m^2$ RCS corner reflector being scanned by the CPD radar.

5.5.2.2 Target X-position Estimation With The PHD Filter

The X-coordinate of the human target is tracked fairly accurately, within 0.5m for the majority of detections. Fig. 5.15 shows that although there is a fairly wide spread of detections at close range (iterations 0-80, corresponding to ranges of 0m to the maximum reliable range of 12m), the centroid estimate is consistently less than 0.5m above the ground truth. At longer range, the beam is wider and detections more sporadic; the spread of the X-coordinate estimates increases, even though there are fewer detections. As there are fewer detections to average out variations in detected position, one or two further to the left or right than usual will shift the estimate much more.

Figure 5.15: The PHD filter is used to estimate the X-coordinate of a human being scanned by the CPD radar.

The X-dimension estimate of the $1m^2$ RCS corner reflector, shown in Fig. 5.16 demonstrates a dense and broad spread of detections at close range. The target is a strong enough reflector that it is detected across the entire width of the beam, and by the sidelobes at close range. The targets are detected in a broader spread at longer range, with errors up to 2m, due to the target remaining detectable at a range where the beam becomes quite wide.

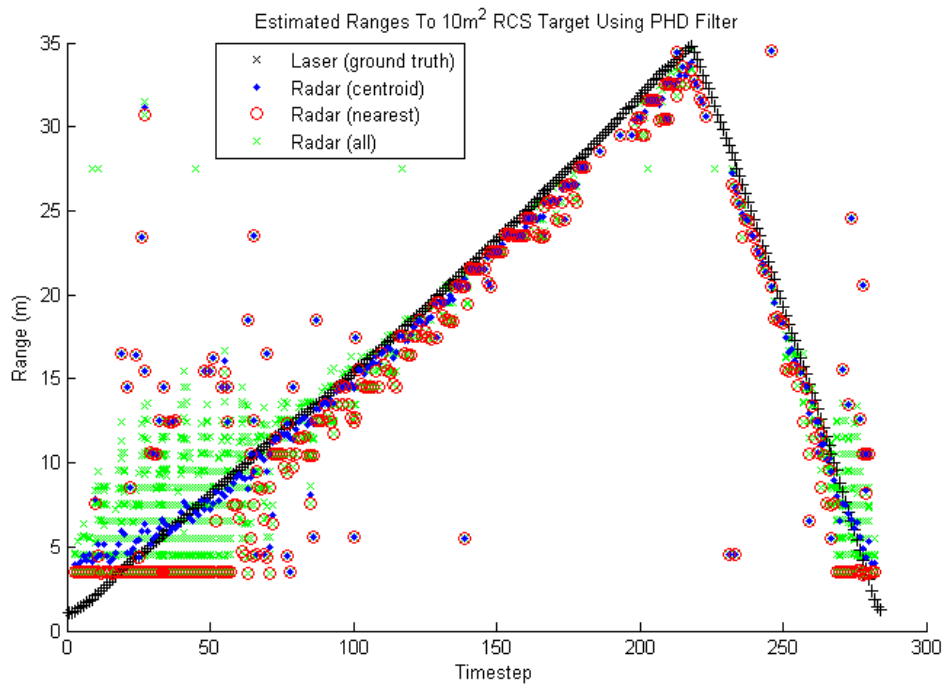Figure 5.16: The PHD filter is used to estimate the X-coordinate of a 1m^2 RCS corner reflector being scanned by the CPD radar.

As we would expect, Fig. 5.17 shows that the 10m^2 RCS target is detected with even stronger sidelobe effects than Fig. 5.16 does for the 1m^2. It even caused the PHD filter to birth additional targets between iterations 20 and 40, appearing below the ground truth. At longer ranges there is significant spread of the targets, but between iterations 50 and 150 (distances of 5-20m) the majority of centroid estimates are within approximately 1m of the ground truth.

Figure 5.17: The PHD filter is used to estimate the X-coordinate of a $10m^2$ RCS corner reflector being scanned by the CPD radar.

5.5.2.3 Target Y-position Estimation With The PHD Filter

Since the X-coordinate is close to zero, the Y-dimension contributes most of the range. The human target Y-distance estimate in Fig. 5.18 therefore approximates the range estimate in Fig. 5.12.
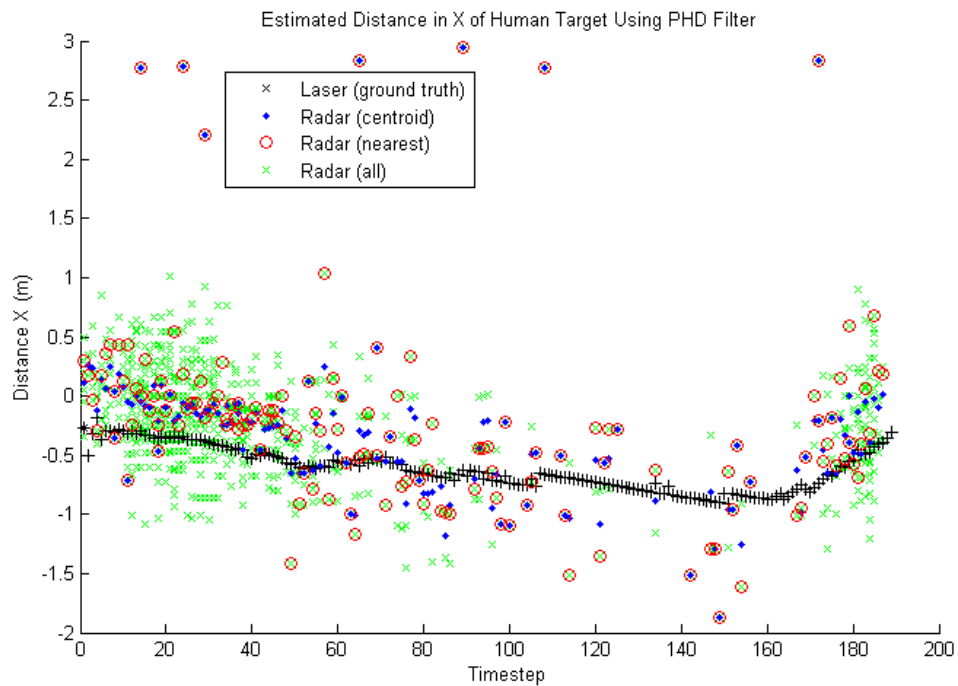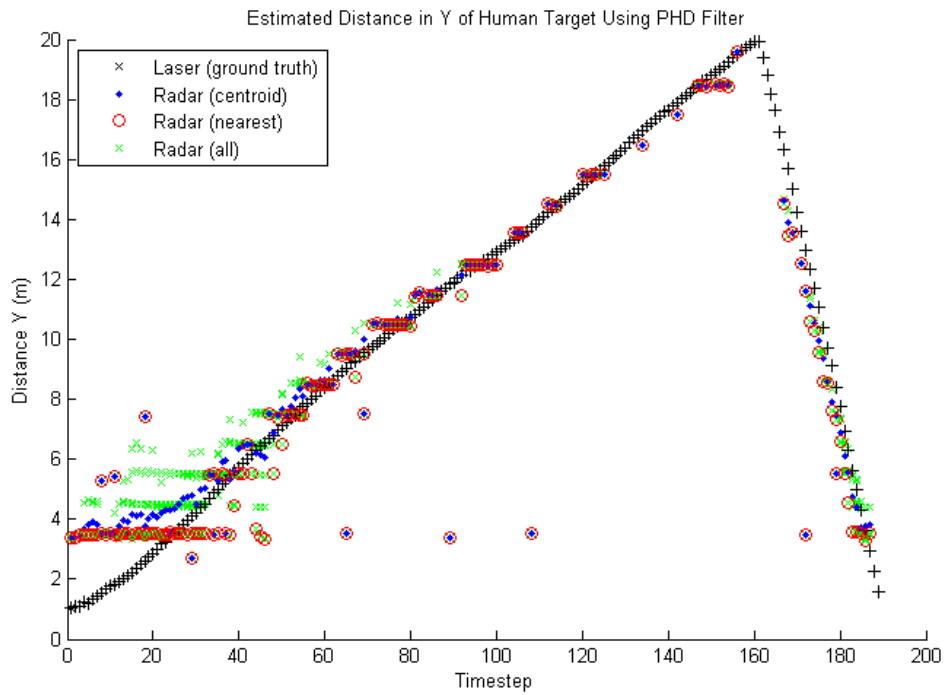
Figure 5.18: The PHD filter is used to estimate the Y-coordinate of a human being scanned by the CPD radar.

And the 1m^2 target Y-estimate in Fig. 5.19 is of the same shape and characteristics as the full range estimate in Fig. 5.13.
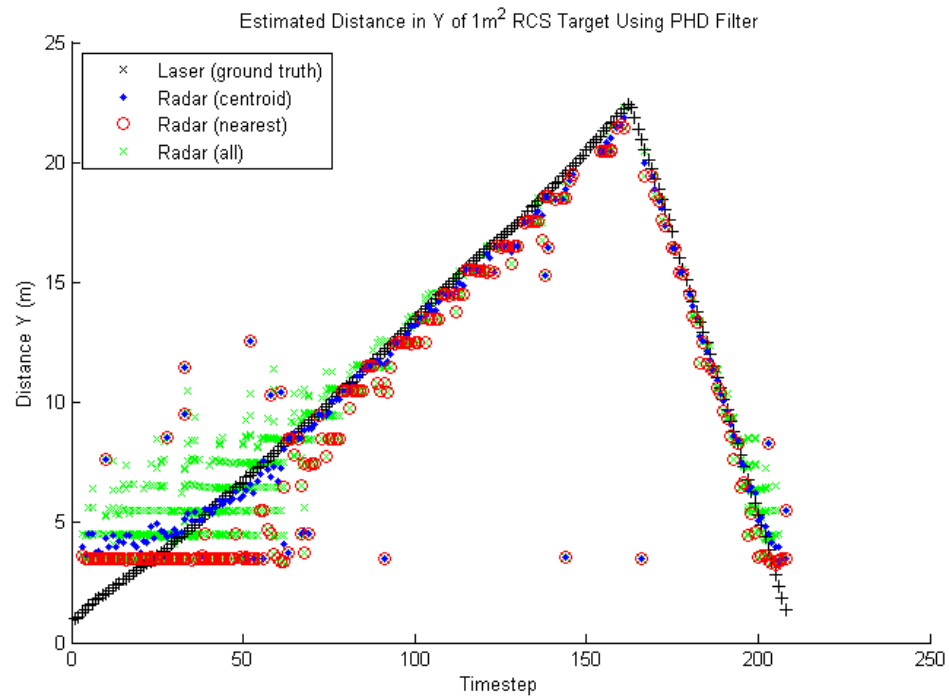
Figure 5.19: The PHD filter is used to estimate the Y-coordinate of a $1m^2$ RCS corner reflector being scanned by the CPD radar.

And once more, Fig. 5.20 shows that the $10m^2$ RCS target's Y-estimate is very similar to the range estimate in Fig. 5.14.
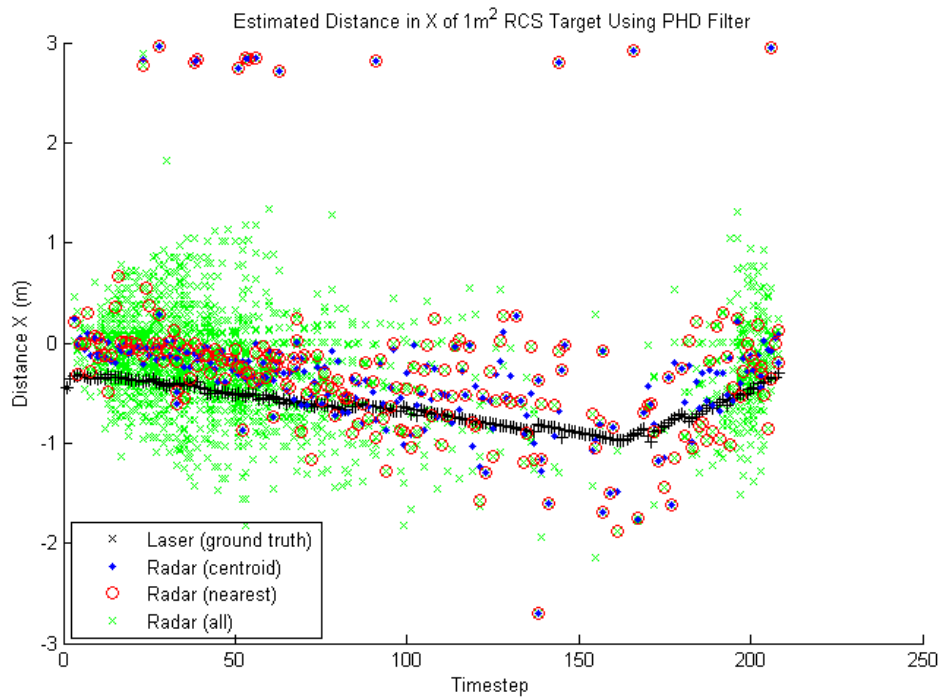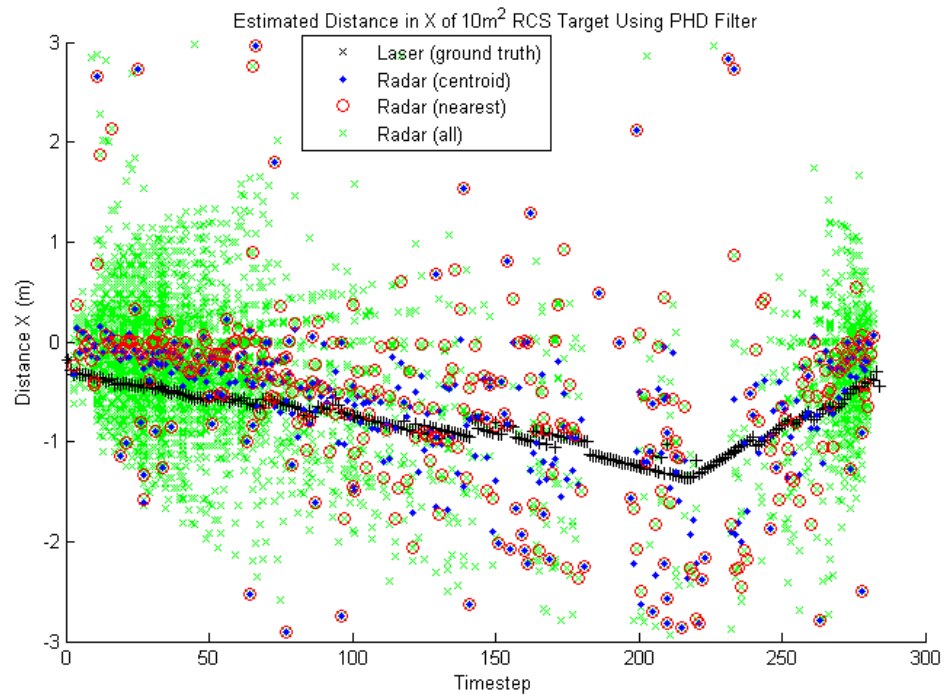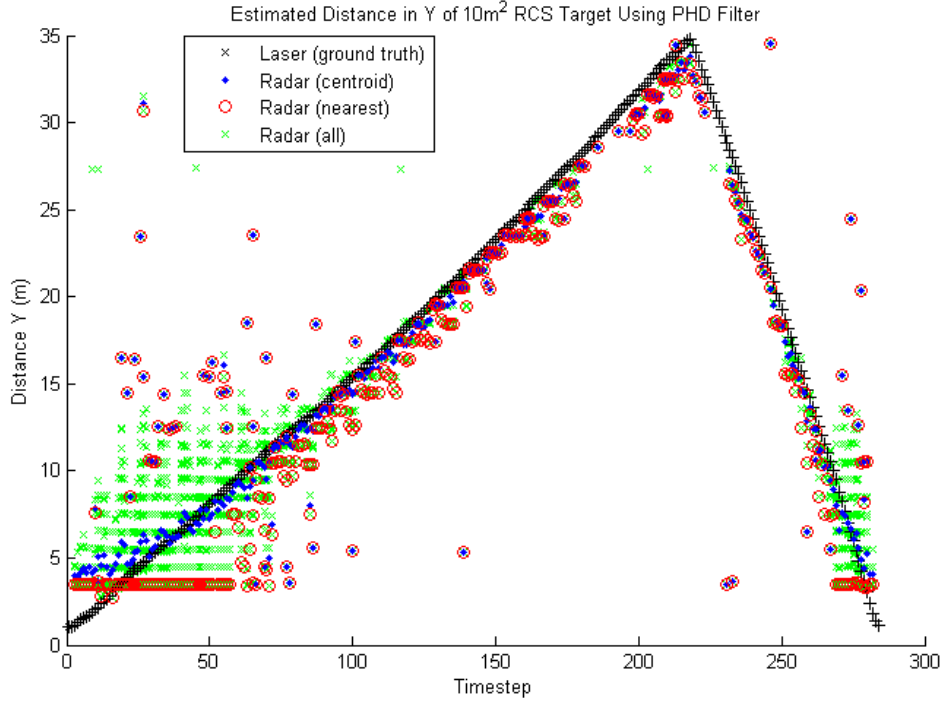
Figure 5.20: The PHD filter is used to estimate the Y-coordinate of a 10m^2 RCS corner reflector being scanned by the CPD radar.

5.5.2.4   OSPA Metric Results Using Different Size Targets

The OSPA metric calculation is a quantified method to compare the performance of a tracking estimate versus a ground truth. It produces a score based on the mismatch between the positions of targets, and the number of targets, in two datasets. A lower OSPA score indicates a lower error, and in this case a better match between the PHD filter's estimates of target position using radar measurements and the lidar ground-truth. The control parameters of the OSPA are power value $p = 2$ and a cutoff $p = 2$, as was used in comparing the detection thresholds. In all these tests, two OSPA values are calculated: one using the centroid of each radar cluster, and one using the shortest-ranged measurement in each radar cluster. The clusters are divided using a simple distance heuristic, as described in Sec. 5.5.2. It should be noted that a different clustering threshold could produce a different number of clusters, which would influence the OSPA calculation significantly. Given the difference in measurement structure and resolution between the lidar and radar, it is difficult to determine an ideal segmentation approach.

The OSPA values are calculated over two sets of data - one narrow set, containing

only the desired target and any ground clutter in a narrow band in front of the radar, and one broader set including static objects in the environment (concrete walls and steel poles and fences).

The OSPA metrics for the human test are plotted in Fig. 5.21. This figure shows that using the nearest measurement in a radar target cluster provides a better estimate to the target position in around iterations 15 to 25, but not by a large margin. Examining Fig. 5.12, this corresponds to true ranges of 2-4m, or the very lowest threshold of the radar's detection range. For iterations 25 to 90, corresponding to ranges of 4m to 12m, using the cluster's centroid is as good as or better than using the nearest point. Beyond this range, the performance of the two methods is near-identical.

Figure 5.21: The OSPA performance metric is calculated using a human being as the target, and tracking the centroid or the nearest point of the PHD estimate cluster. There are no other targets in the test environment. Using the centroid of the radar target is usually as good as or better than using the nearest measurement, except at very close range - but the difference is small.

The OSPA metric calculation for the $1m^2$ target in Fig. 5.22 shows that for a similarly broad swathe of measurements (iterations 30-130, covering ranges 5m-15m),

using the cluster's centroid is much more accurate. As seen in Fig. 5.13, phase noise causes the nearest detection to precede the true position of the target.

Figure 5.22: The OSPA performance metric is calculated using a $1m^2$ corner reflector as the target, and tracking the centroid or the nearest point of the PHD estimate cluster. There are no other targets in the test environment.

The effect of phase noise on estimated target range is even more obvious in Fig. 5.23 for the $10m^2$ RCS target. At almost any range, the centroid of the cluster of measurements is a more accurate estimate than the nearest point in the cluster. However, the average value is higher using the centroid as well; Fig. 5.14 shows how phase noise shifts the cluster's centroid closer to the radar, causing it to underestimate the true range to the target.

Figure 5.23: The OSPA performance metric is calculated using a 10m^2 corner reflector as the target, and tracking the centroid or the nearest point of the PHD estimate cluster. There are no other targets in the test environment.

5.5.2.5 OSPA Metric Calculation Over A Wider Test Area

The region of consideration is now expanded to cover a wider region, including the static objects on the right of the radar seen in Fig. 5.10. This is to simulate a more crowded environment, and present more opportunity for clutter and noise measurements to foul the target estimates. In all scenarios, plotted in Fig. 5.24 to Fig. 5.26, using the centroid to estimate target position produces lower error at almost all ranges.

The overall OSPA metric for each test is higher, on average, than for the corresponding graphs with a narrower environment with only one target. Fig. 5.21 and Fig 5.22 have regions (approximately iterations 30-100) where the target is in the prime detectable range and the OSPA error value is around 0.5 (using the centroid of the cluster). The OSPA in Fig. 5.24 and Fig. 5.25, which track the same human and 1m^2 targets respectively, average around 1.5 in these regions.

Performance is similarly degraded for the 10m^2 reflector in F.g 5.26, though it already had a higher mean OSPA and the effect is not quite as stark.
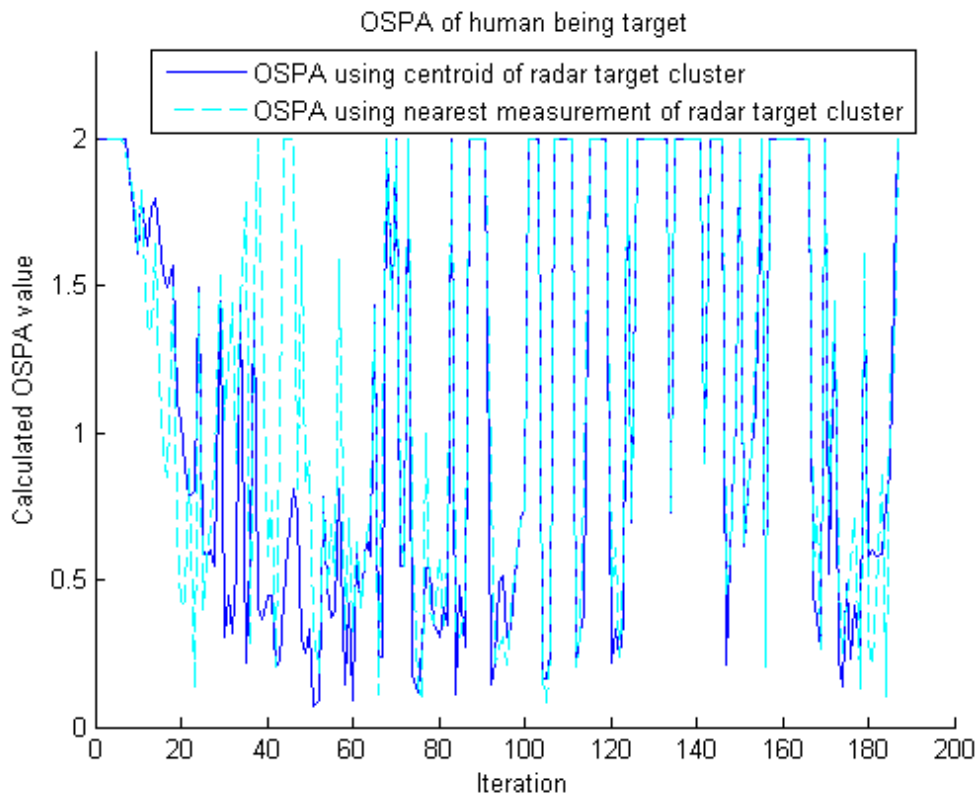

Figure 5.24: The OSPA performance metric is calculated using a human being as the target, and tracking the centroid or the nearest point of the PHD estimate cluster, as it moves through an environment containing static targets. Once more, the difference in performance between the different reference points is small.
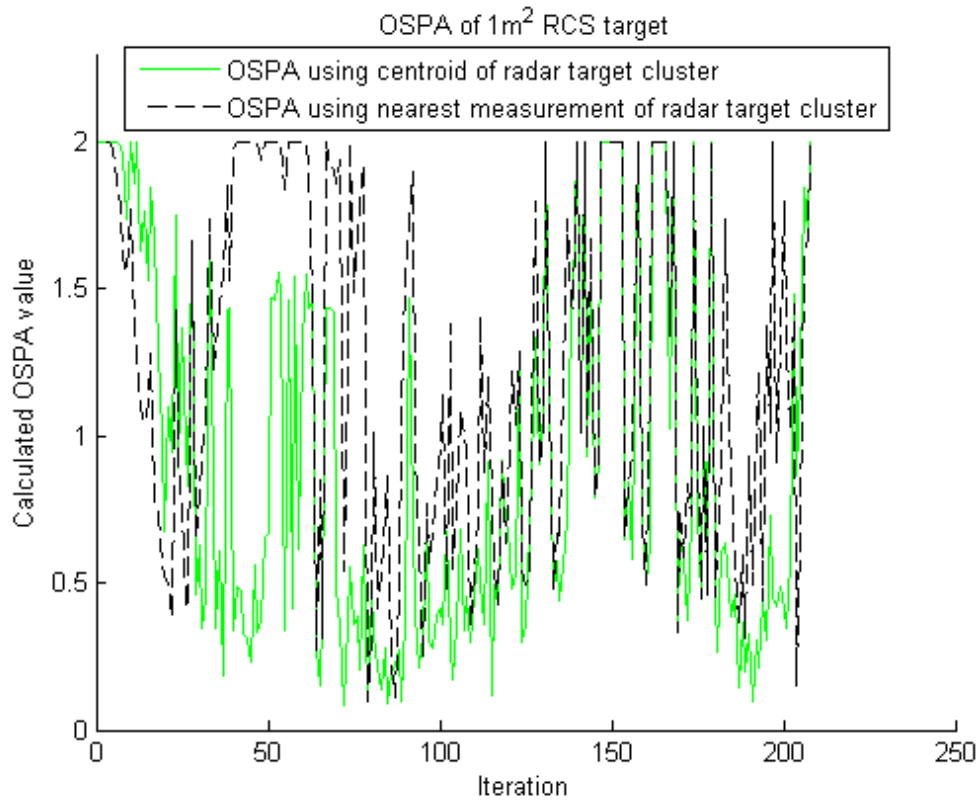
Figure 5.25: The OSPA performance metric is calculated using a $1m^2$ corner reflector as the target, and tracking the centroid or the nearest point of the PHD estimate cluster, as it moves through an environment containing static targets.
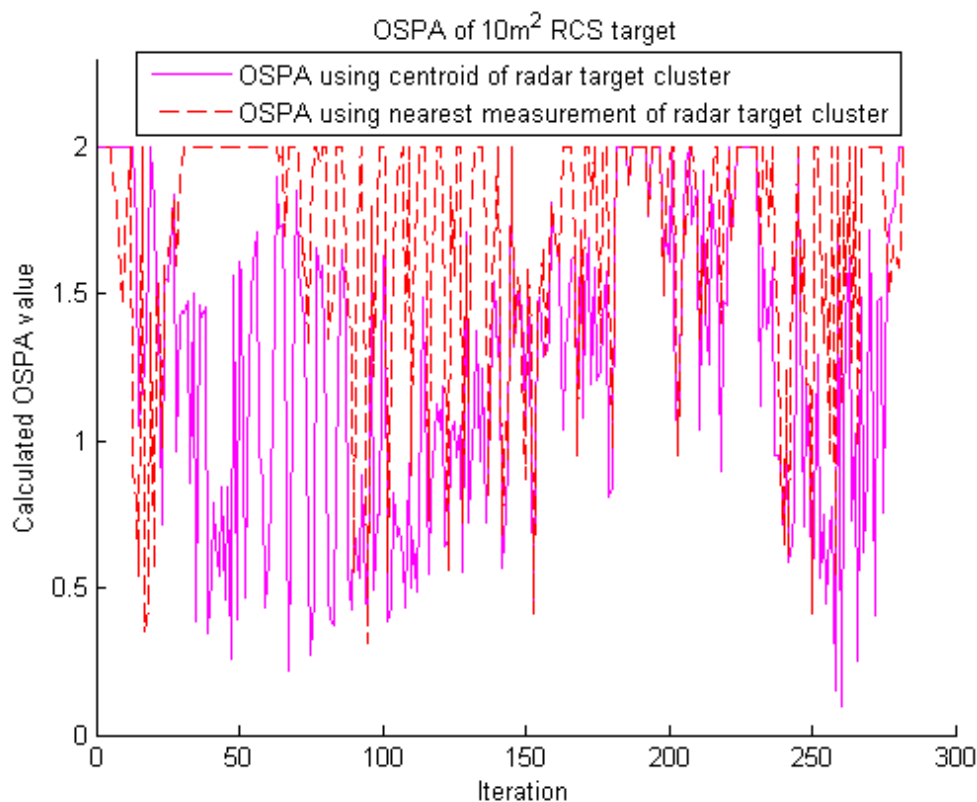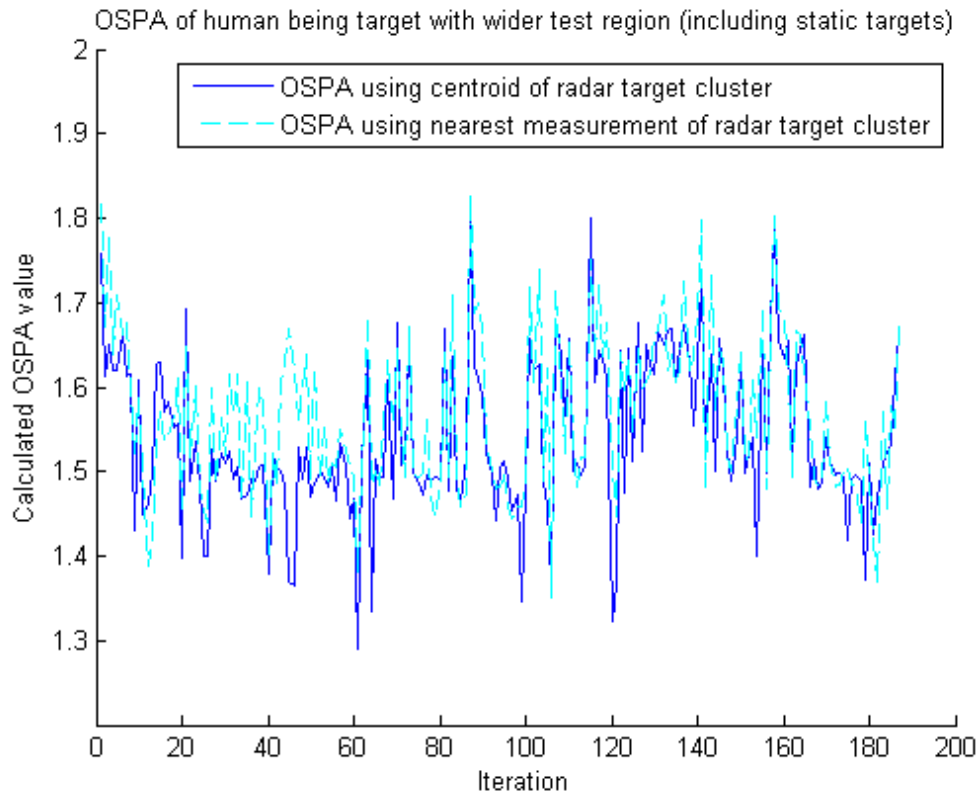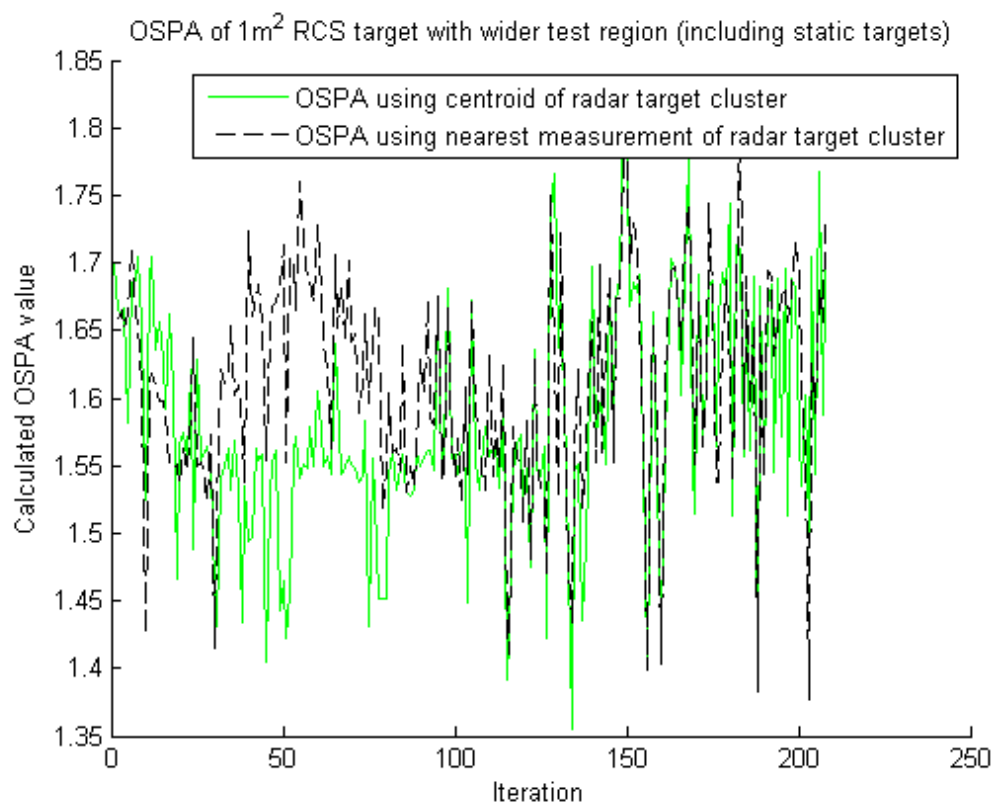
Figure 5.26: The OSPA performance metric is calculated using a 10m^2 corner reflector as the target, and tracking the centroid or the nearest point of the PHD estimate cluster, as it moves through an environment containing static targets.
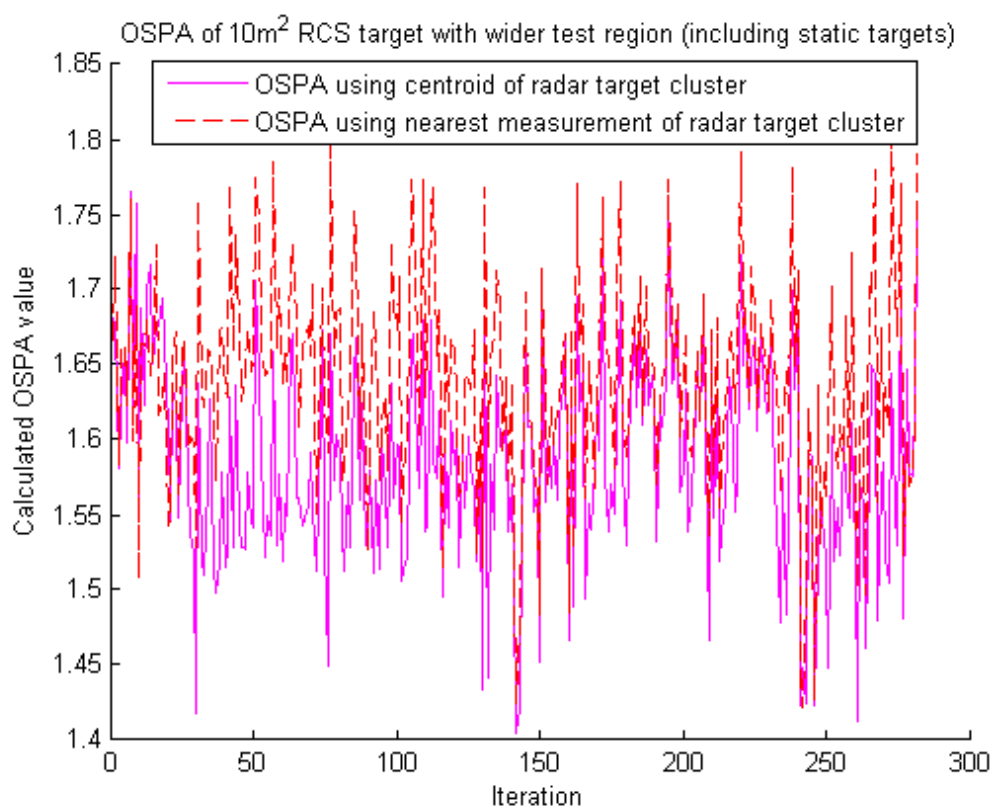
5.6 Discussion

This chapter focused on the problem of informing the driver of the state of the environment when the state of the environment relative to the vehicle was changing. Although consideration was given to using the same occupancy grid-based approach developed in Chapter 4, most emphasis was given to the feature-based PHD filter approach and this chapter's contribution was the development of a EK-GM-PHD filter implementation for mapping in dynamic environments using FMCW radar. Grid-based methods have been used in mapping dynamic environments, with different approaches available; some remove dynamic objects from measurements while others explicitly model the movement of objects in the environment [9]. These methods have not been given heavy consideration here due to the comparatively slow measurement and scan speed of the CPD radar; with faster radars they may be more effective. They still suffer some of the limitations of occupancy grids outlined in Sec. 5.2.

The PHD filter was investigated as an alternative method of producing a map to display to the driver. The EK-GM-PHD filter was discussed in this chapter with emphasis on what values are used for certain parameters and what the effects of varying them, with the goal that the system could be adapted to a different FMCW radar or other modality.

The results of Sec. 5.5 must be analysed in the context that the CPD radar is intended to be used - that is, situation awareness for collision avoidance in large vehicles. Sec. 1.4 outlines the main requirements for a close proximity system as being high reliability, good coverage and an intuitive user interface.

The EK-GM-PHD filter in this implementation produces some very inaccurate velocity estimates, but that is less important than quickly detecting targets within the danger zone. Fig. 5.2 shows that birth velocity covariances $\sigma^2_{\beta_{v_x}} = \sigma^2_{\beta_{v_y}} = 25m^2/s^2$ and a birth weight of $5 * 10^{-3}$ lead to targets being estimated to have weights above 0.5 after 3 iterations of the filter. This is in line with the desired balance of noise rejection and change responsiveness. If targets were to appear existent after a single detection, there would be zero resistance to noise. Two detections still has a high risk of consecutive close noise/clutter measurements being treated the same as a target. Three detections is still quite responsive but with better noise resistance. The amount of time (in seconds) of radar scanning that it relates to will vary with the hardware.

Every radar and environment will be different and will require careful selection of

the parameters described in Sec. 5.4. Any one facet of filter performance is often influenced by several of these terms, and it probably will not be feasible to perform exhaustive parametric analysis to test every potential combination of values at high resolution in all potential environments. There will need to be a balance between tractability and desired performance.

Fig. 5.11 shows that the EK-GM-PHD filter is able to generate maps of the environment containing a useful depiction of the closest hazards. While not as clear as the lidar map, it is able to detect a potential hazard that might be obscured by blind spots or bad weather. Additionally, Fig. 5.11 shows some of the downsides of using radar - targets appearing elongated along the range-dimension due to phase noise, and broadened in the cross-range dimension due to the wide beam. In (f), two small sidelobe-detected targets are visible either side of the main one.

Looking more closely at the estimated ranges to the targets in Fig. 5.12, Fig. 5.13 and Fig. 5.14 and the calculated OSPA metrics for each (Fig. 5.21, Fig. 5.22 and Fig. 5.23), it is apparent that there are numerous false detections at close range when a strong reflector is used. Compared to the corner reflectors, the human being is a preferable target for the CPD radar in some ways; although it is more difficult to detect and has a lower maximum range, it is not so strong a reflector as to induce heavy phase noise. It should be noted that the $1m^2$ and especially $10m^2$ RCS reflectors are *very* strong targets, on the order of a light vehicle parked close to the sensor (see Fig. 3.24). If a target of that size is close enough to induce strong phase noise, there will probably be no safe way to manoeuvre the large vehicle around it anyway. Further testing should be done using a light vehicle as a target, but time constraints prevented this in this thesis.

Even the human being generates multiple detections at close range, leading to the target being modelled as a "cloud" of point targets by the filter. A better approach might be to sidestep this clustering of filter output entirely and use an extended target representation; this would be rich grounds for future work, adapting the research in [19] [20] and others to the close proximity problem. This would have the advantage of tracking a single target with a single covariance matrix and velocity, rather than many.

The clustering approach used in Sec. 5.5 is simple, and would perform less effectively in a more crowded or cluttered environment, where it would probably merge together targets that should be treated separately. Clustering can only be used here because the environment is known and controlled, and there is only one isolated target that is known to have a small physical size but the potential for a large RCS.

In these tests, the most accurate point of a cluster of PHD filter estimates for estimating the true target position was the centroid. This provides a "best case" estimate of CPD radar performance with this EK-GM-PHD filter. But in an unknown environment, one cluster could be representative of a reflector with large physical size, or several smaller reflectors, or one reflector of large RCS but small physical size. Despite its superior performance, the centroid of a cluster of points could not be used to estimate target position. There is too much uncertainty about the true extent of a target, and the risk of overlooking a small, valid target on the assumption that there is only one target present further back. If clusters of PHD filter targets must be used using the CPD radar, the nearest-point target estimate is the one that would need to be used as an estimate of the true range to target, not the centroid.

It is clear that the nearest-point estimate frequently detects objects as being far closer than they truly are. But even though this produces a less accurate position estimate for nearby targets, it must be remembered that the objective is not to obtain an accurate position estimate for all targets in the environment - it is to provide situation awareness to the driver, and help avoid collisions. Clustering is not strictly necessary outside of the benchmarking done here; the driver could be warned of the presence of one target or several in close proximity. Whether the hazard is 6m away or 10m away, the driver needs to be aware of it. If the large vehicle cannot be manoeuvred around the hazard safely, something must be done to remove it.

The Gaussian mixture extended Kalman PHD filter implemented here is a starting point for more advanced solutions. Features that could be added in future research include:

- Conversion to the unscented Kalman filter (UKF) for better performance in non-linear applications.

- Improved models for the target prediction step; the constant velocity model used here may be the best that the CPD radar is capable of, but with higher resolution data a better model tracking target heading, angular velocity or other states, might be possible. The covariances for each state dimension could be better calculated also.

- Implementation of the Cardinalised PHD filter (CPHD) for better performance in environments where the number of targets fluctuates.

- An implementation of an extended target representation, to reflect the environment being composed of targets of varying size. This would allow the filter to output an estimate of the size and orientation of targets, not just a cloud of point targets. For the CPD radar, which detects nearly everything as an extended target, this could lead to major performance improvements. Care would need to be taken to prevent too many targets merging together into one great blob.

- Track labelling to allow the state of specific objects to be monitored over time, and better predict moving objects that might soon move into a danger zone.

- More development, modelling and simulation to investigate the optimal values of all the control parameters discussed in Sec. 5.4.

- Further use of radar measurement characteristics in target birthing, such as using detected power value to determine the birth weight of a target, or the width of the beam at target range to influence the birth covariance.

- Better calculation and measurement of target velocity, such as the use of Doppler to allow new targets to be instantiated with accurate velocities.

5.7 Summary

This chapter has considered the close proximity problem for large vehicles in dynamic environments. The occupancy grid mapping approach developed in Chapter 4 has modifications made to improve performance in dynamic environments, by bounding the cell occupancy as a function of range from the static sensor. For applications when the sensor is moved, and for other reasons such as representation of target position uncertainty, the EK-GM-PHD is considered as an alternative approach. This chapter's contribution is an implementation of the EK-GM-PHD filter for use with FMCW radar. The detection threshold developed in Chapter 3 is used for feature extraction, and the parameters that control filter performance have been analysed and calculated based on the sensor's performance.

Chapter 6

Conclusion

6.1 Contributions

This thesis has presented contributions towards robust, all-weather solution for the close proximity problem for large vehicles through the use of a low-cost FMCW radar. The close proximity problem has been identified as being a result of:

- blindspots, caused by the size and design of the vehicle, and the placement and construction of the driver's cabin;

- bad weather, which impairs the driver's visibility in the limited region where they are able to see;

- the necessity of complex, close-range interactions between the large vehicle and other vehicles or static structures in the applications where large vehicles are used.

As a result of these factors, the drivers of large vehicles possess inadequate situation awareness to operate safely. Complete automation of a mine operation is not technologically or economically feasible at this point in time. Therefore, a solution to this problem must come in the form of a driver assistance tool that warns the driver of any objects within their blind spots, regardless of bad weather conditions, and prevents collisions.

The only sensor with acceptable all-weather performance is shown to be radar, but at the time this research commenced there was no commercially available radar capable of solving this problem adequately; existing systems have wide beams that possess poor cross-range resolution and therefore provide very vague information about the location

of any detected object, and are susceptible to false detections. The challenges of mapping with FMCW radar are identified: complex measurement structures, unpredictable detection characteristics, wide beams and sidelobes, and susceptibility to environmental clutter and noise.

The contributions of this thesis are:

- An analysis of the close proximity problem and the requirements of any FMCW radar sensor that will be used to solve it, in terms of the sensor's reliability, accuracy, coverage and ease of use. Sensor properties such as sensor size, power, beam width, scan rate and range resolution are considered.

- The close proximity detector (CPD) radar is presented as a radar designed to meet these criteria. It provides a platform for the development of an in-depth understanding of FMCW measurement and detection characteristics for use in the development of a system for processing radar measurements. It also allows demonstration of the importance of having access to the sensor design process to maximise the capabilities of the sensor for a particular application. In this case, precise localisation of detected objects was important to allow vehicle operators to determine if it was possible to move safely; therefore linearisation of the FMCW chirp needed to be conducted to improve target range resolution and decrease noise.

- The development of a method for designing and testing static detection thresholds that are used to identify targets as being present within radar measurements corrupted by clutter and noise. The traditional CFAR dynamic thresholds perform poorly in applications with short measurement vectors and frequent targets spanning multiple range bins. This thesis introduces a method of calculating a static detection threshold using measurements of the background clutter and heuristic methods. These provide fast and relatively simple methods for developing static detection thresholds for a FMCW radar for a known environment. This allows clutter and noise to be excluded from raw radar measurements, and the targets extracted from the measurements can be plotted and displayed to the driver. This provides them with more information of the state of the environment around their vehicle.

- An implementation of an occupancy grid mapping approach for mapping static

environments using FMCW radar, using a heuristic sensor model developed from the detection threshold designed earlier. This allows a raw radar measurement to be converted into a probability of target presence in a given location, rather than just a binary classification of target presence or absence in each bin. Successive measurements are fused together for improved accuracy. Results of this contribution are shown using data from the CPD radar, and show cleaner, less cluttered maps of the environment than raw data plots.

- An implementation of a Gaussian-mixture extended Kalman probability hypothesis density filter (GM-EK-PHD) was presented for solving the close proximity problem in dynamic environments. It allowed for better mapping in a moving environment using a slow-scanning, noisy sensor like FMCW radar than an occupancy grid is capable of. It possessed good clutter filtering while still allowing fast response to changes. The static detection threshold developed earlier was used for feature extraction, creating target detections that were used to add new targets and update known ones. A thorough investigation was conducted of the necessary modelling and calculation of control parameters for using data from the CPD radar and their influence on the filter's performance. The performance of this filter was demonstrated with both simulated and real data.

The FMCW radar detection threshold design, sensor model and EK-GM-PHD filter implementation are all developed with the use of measurements and properties of the CPD radar, but the methods introduced by this thesis are hardware-agnostic. They could be applied to mapping using other FMCW radars, and parts could be used with different sensing modalities. The focus is on possessing a thorough knowledge of the sensor's physics and performance, and exploiting this in implementing signal processing technology.

6.2 Future Work

There is broad scope for further research in the field of radar mapping and tracking. When this research was commenced, the CPD radar was to the best of our knowledge unique in its combination of capabilities and size for its cost. Since then, new radars have continuously been developed, and the price and size of commercially available radars continued to decrease. The techniques shown in this thesis could be implemented

using newer commercially-available FMCW radars; while there would be some loss of flexibility due to not having as much knowledge or control over the inner workings of the radar, newer sensors may have improved characteristics, lower cost and smaller size. This could allow the potential use of multiple off-the-shelf radars on a single vehicle without substantial increase in cost or configuration. Faster-scanning radars would provide improved performance in dynamic environments.

The detection thresholds generated using gathered clutter data can be improved by a larger dataset from which the thresholds are generated. By moving the sensor around to different positions in the environment, local variations in the environment would be evened out and a more portable detection threshold would be produced. Tests would need to be done to see how well this approach works for radars other than the CPD radar; the method would need to be rethought if its portability between radar systems were low. Newer radars may have a larger measurement vector, which would would produce better performance using CFAR approaches and require less tuning.

There are terms within the occupancy grid that could be further optimised, such as the cell resolution, the upper and lower power thresholds in the sensor model, and bounds on probabilities for the sensor model or the grid cells. The use of logarithm of odds form can provide performance improvements, and a quadtree implementation could provide varying grid resolution at longer range.

The EK-GM-PHD filter analysed here is only one of many implementations of the PHD filter, and has many potential improvements. The cardinalised PHD filter is an extension on the standard PHD filter in that the number of targets is propagated in the update step, as well as their weights, means and covariances. It provides superior performance when the number of targets fluctuates. Additionally, the unscented Kalman filter (UKF) has superior performance to the EKF in highly nonlinear applications. Recent implementation of extended target PHD filters demonstrate accurate tracking of targets that generate multiple returns per measurement, as was the case here. Rather than an ad-hoc clustering of many point targets output by the filter, a systematic approach to tracking them as a group within the filter would provide clearer and more accurate information. The implementation of track labelling would allow specific targets to be consistently identified, providing more useful information to the drivers. It would also allow the speeds of individual targets to be more easily tracked, and warnings generated when a nearby target that was not currently a hazard was identified as moving towards a dangerous position. The use of a sensor capable of measuring target speed as well as

position, such as Doppler radar, would allow new targets to be birthed with an accurate velocity estimate, and make initial prediction of the target's position much more accurate. The cost of such sensors should decrease as they become standard on consumer vehicles. A wider set of tests in more cluttered environments, typical of large vehicle operations, using light vehicles as targets, would provide better benchmarking data for analysing filter performance. The combination of the PHD filter and occupancy grid could be investigated, such as using the occupancy grid to estimate the size and shape of extended targets.

The situation awareness system developed here has been focused on assisting human drivers, but there is no reason that a similar system could not be a component of a fully automated system. FMCW radars will be used more widely in unmanned ground vehicles as sensor cost and size decrease. The approaches used here to provide improved situation awareness to a human driver in a large vehicle could also be applied as part of a fully autonomous system in a large or light vehicle. FMCW radar's all-weather performance makes it attractive in any automotive application where reliability is demanded.

Bibliography

[1] Adams, Martin, John Mullane, Ebi Jose & Ba-Ngu Vo. *Robotic Navigation and Mapping with Radar*, Artech House, 2012. 1.5.1, 1.5.2, 2.2.1.3, 2.3.3, 2.4.1.2, 2.4.1.3, 2.4.1.5, 2.6.1, 2.6.2.1, 2.6.2.2, 2.6.2.3, 2.6.3.1, 2.6.3.2, 2.6.3.3, 2.8, 2.8.1, 2.8.1, 2.8.1, 2.8.1, 2.8.2, 2.8.2, 5.4

[2] Agamennoni, Gabriel, Juan Nieto, and Eduardo Nebot. "Mining GPS data for extracting significant places." In *IEEE International Conference on Robotics and Automation, 2009*, pp. 855-862. IEEE, 2009. 1.2.5

[3] Australian Communications Authority, Spectrum Planning Team, Radiofrequency Planning Group, "A review of Automotive Radar Systems - Devices and Regulatory Frameworks", April 2001 1.5.1, 2.3

[4] Brooker, Graham, *Introduction to Sensors for Ranging and Imaging*, SciTech Publishing, Raleigh, NC; 2009. 1.5.1, 2.2.1.1, 2.2.1.2, 2.2.1.3, 2.2.3.1, 2.1, 2.2.3.3, 2.3, 2.3.2, 2.3.2.1, 2.7, 2.3.2.2, 2.3.2.3, 2.8, 2.3.4, 2.3.4, 2.3.4, 2.3.5, 2.3.5, 2.3.6.1, 2.3.6.1, 2.3.6.2, 2.3.8, 2.10, 2.3.9.1, 2.4.1, 2.4.1.1, 2.4.1.1, 2.4.1.2, 2.4.1.4, 2.4.1.4, 2.4.1.6, 3.1.1, 3.3.2.3, 3.3.2.5, 4.1.5, 4.1.5.2

[5] Brooker, Graham, "Long-Range Imaging Radar for Autonomous Navigation". PhD diss., University of Sydney, 2005. 1.5.1, 2.3.2.2, 2.3.6.1, 2.3.6.3, 2.3.9, 2.4.1, 2.4.1.3, 2.4.1.5, 2.4.1.6, 2.7.1

[6] Brooker, Graham. "Millimetre wave radar for robotics". In Penelope Probert Smith (Eds.), *Active Sensors for Local Planning in Mobile Robotics*, (pp. 137-162). United States: World Scientific Publishing. (2001) 2.2, 2.2.3.3, 2.2.4, 2.3

[7] Brooker, Graham M. "Understanding millimetre wave fmcw radars." In *1st international Conference on Sensing Technology*, pp. 152-157. 2005. 2.2.1.3

[8] Canas, Jose M., and Vicente Matellan. "Dynamic gridmaps: comparing building techniques." *Mathware and Soft Computing* 13, no. 1 (2006): 5. 2.6.2.1, 2.6.2.3, 4.1.4, 5.2, 5.2.1, 5.2.2, 5.3

[9] Chen, Cheng, Christopher Tay, Christian Laugier, and Kamel Mekhnacha. "Dynamic environment modeling with gridmap: a multiple-object tracking application." In *9th International Conference on Control, Automation, Robotics and Vision, 2006* , pp. 1-6. IEEE, 2006. 2.6.2.1, 5.2, 5.6

[10] Clark, Daniel Edward. "Multiple target tracking with the probability hypothesis density filter." PhD diss, Heriot-Watt University, 2006. 5.4

[11] Clarke, Bryan. Matlab implementation of Gaussian Mixture Probability Hypothesis Density Filter (GM-PHD) (version 1.10) (2014). [Software]. Available from `http://au.mathworks.com/matlabcentral/fileexchange/42769-gaussian-mixture-probability-hypothesis-density-filter--gm-phd-` 5.4, 5.1

[12] Endsley, Mica R. "Design and evaluation for situation awareness enhancement." In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 32, no. 2, pp. 97-101. SAGE Publications, 1988. 1.1, 1.2.5

[13] Erdinc, Ozgur, P. Willett, and Yaakov Bar-Shalom. "Probability hypothesis density filter for multitarget multisensor tracking." In *8th International Conference on Information Fusion, 2005*, vol. 1, pp. 8-pp. IEEE, 2005. 5.4.3.2, 5.4.4.6

[14] Foessel, Alex, "Scene modeling from motion-free radar sensing." PhD diss., Robotics Institute, Carnegie Mellon University, January 2002. 2.2.1.3, 2.3.2.1, 2.3.6.1, 2.3.6.1, 2.3.7, 2.3.8, 2.4.1, 2.4.1.3, 2.6.2.1, 2.6.2.2, 2.6.2.3, 2.7, 2.7, 2.7.1, 2.9, 3.4, 4.1, 4.1.5, 4.1.5.2

[15] Glynn, Patrick. "Review of Vehicle Collision Matrix for the Australian Constructors Association", CSIRO Earth Science and Resource Engineering, November 2011. Accessed online at `http://www.constructors.com.au/safety/ACA%20Report%20on%20Collision%20Avoidance%201%20Matrix.doc`, last accessed 16 September 2014 2.2.1

[16] Gonzalez, Rafael C., Richard E. Woods, "Digital Image Processing Third Edition", Pearson Prentice Hall, Upper Saddle River, New Jersey, 2008. 2.2.2

[17] Google Maps. (2014). [Australian Centre for Field Robotics, the University of Sydney] [Satellite photograph]. Retrieved from `https://www.google.com.au/maps/place/Australian+Centre+for+Field+Robotics/@-33.8901531,151.1936568,195m/data=!3m1!1e3!4m2!3m1!1s0x6b12b1d6b0df9ea1:0x173de9d1b097ca4` 2.5

[18] Gindele, T., Brechtel, S., Schroder, J., & Dillmann, R. (2009, June). "Bayesian occupancy grid filter for dynamic environments using prior map knowledge". In *Intelligent Vehicles Symposium*, 2009 IEEE (pp. 669-676). IEEE.

[19] Granstrom, K., Christian Lundquist, and Umut Orguner. "A Gaussian mixture PHD filter for extended target tracking." In *13th Conference on Information Fusion, 2010* , pp. 1-8. IEEE, 2010. 5.6

[20] Granstrom, Karl, Christian Lundquist, and Umut Orguner. "Extended target tracking using a Gaussian-mixture PHD filter." *Aerospace and Electronic Systems, IEEE Transactions on* 48, no. 4 (2012): 3268-3286. 2.8.2, 5.4.4.6, 5.6

[21] Hillier, N., J. Ryde, A. Denman, G. Brooker, and J. Martinez. "Performance of Laser and Radar Ranging Devices in Adverse Environmental Conditions." *ACARP Report, July* (2008): 712727. 2.2.1.2, 2.2.1.3, 2.2.3.3, 3.3.2.7

[22] Hinze, Jimmie W., and Jochen Teizer. "Visibility-related fatalities related to construction equipment." *Safety science* 49, no. 5 (2011): 709-718. 1.1, 1.2.1, 2.1.2, 2.2.3.2

[23] International Organization for Standardization, "ISO 16001:2008 Earth-moving machinery – Hazard detection systems and visual aids – Performance requirements and tests", last reviewed 2011. 2.1.3, 2.2.1, 2.2.2

[24] International Organization for Standardization, "ISO 5006:2006 Earth-moving machinery – Operator's field of view – Test method and performance criteria", last reviewed in 2013. 2.2.1, 2.2.2

[25] IMST. (2015) [sR-1200] [Website] Retrieved from `www.radar-sensor.com/products/radar-modules/sr-1200/` 3.2

[26] Johnson, David. "Complex Target Reconstruction Using Near-field Synthetic Aperture Radar", PhD diss., University of Sydney, 2009. 2.3.9, 2.4.1, 2.4.1.6

[27] Kloos, Gerold. "Radio-Frequency Signal Strength Based Localisation in Unstructured Outdoor Environments." PhD diss., University of Sydney, August 2007. Retrieved from `http://ses.library.usyd.edu.au/handle/2123/2242` 1.1, 1.1, 1.1(a), 1.2.1, 1.2, 1.6, 1.2.3, 1.3, 2.2.1, 2.2.1.1

[28] Knee, M. J. "Significant incident report No. 132 Mine haul truck runs over a light vehicle following a driver change", Safety and Health Division, Western Australia Department of Industry and Resources, 3 Feb 2005. Accessed online at `http://www.dmp.wa.gov.au/documents/Significant_Incident_Reports/MSH_SIR_132.pdf` 1.1, 1.1(c), 1.2.1

[29] Komatsu, "Dump Trucks-Rigid - 730E-7" [Website], accessed August 2014 at `http://www.komatsu.com.au/Equipment/Pages/DumpTrucks-Rigid/730E-7.aspx` 1.2.1

[30] Konolige, Kurt. "Improved occupancy grids for map building." *Autonomous Robots* 4, no. 4 (1997): 351-367. 2.6.2.1, 2.7, 2.7.1

[31] Lamard, L. Boyer, J. P., PAscal, I., and Sas, R. *Multi target tracking with CPHD filter based on asynchronous sensors*. In International Conference on Information Fusion, pages 892-898, 2013. 3.2

[32] MacNeill, Patrick. "International Mining Fatality Database Project Report", April 2008. Accessed at `http://www.resourcesandenergy.nsw.gov.au/__data/assets/pdf_file/0009/182484/International-Mining-Fatality-Database-project-report.pdf` 2.1.1, 2.1.2

[33] Mahler, Ronald PS. "" Statistics 101" for multisensor, multitarget data fusion." *Aerospace and Electronic Systems Magazine, IEEE* 19, no. 1 (2004): 53-64. 2.6.2.3, 2.8.1, 2.8.1

[34] Matthies, Larry, and Alberto Elfes. "Integration of sonar and stereo range data using a grid-based representation." In *Robotics and Automation, 1988. Proceedings.,*

1988 IEEE International Conference on, pp. 727-733. IEEE, 1988. 2.6.2.1, 2.7, 4.1, 4.1.1

[35] Magaz, Boualem, and Adel Belouchrani. "Automatic threshold selection in OS-CFAR radar detection using information theoretic criteria." *Progress In Electromagnetics Research B* 30 (2011): 157-175. 3.4.1.1

[36] Meinel, Holger H. & Juergen Dickmann. "Automotive Radar: From Its Origins to Future Directions", *Microwave Journal*, Vol. 56, No. 9, September 2013. 2.3

[37] Meta, Adriano, Peter Hoogeboom, and Leo P. Ligthart. "Range non-linearities correction in FMCW SAR." *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*. IEEE, 2006. 2.4.1.6

[38] Meyer-Delius, Daniel, Maximilian Beinhofer, and Wolfram Burgard. "Occupancy Grid Models for Robot Mapping in Changing Environments." In *Twenty-Sixth Conference on Artificial Intelligence*. 2012. 5.2

[39] Miller, Wayne K., "Analysis of Haulage Truck Visibility Hazards at Metal and Nonmetal Surface Mines-1975," MESA Information Report 1038 1.3, 1.4

[40] Miller, Steve, and Antonio Nieto. "In-rest Vehicle Proximity Systems With Real-Time Transmission Lock." 32nd International Symposium on Application of Computers and Operations Research in the Mineral Industry (APCOM*), Tucson, AZ, pp. 405-411. (2005) 1.1, 3.1.2, 4

[41] Moravec, Hans P., and Alberto Elfes. "High resolution maps from wide angle sonar." In *Robotics and Automation. Proceedings. 1985 IEEE International Conference on*, vol. 2, pp. 116-121. IEEE, 1985. 2.6.2.1

[42] Mullane, John, Martin D. Adams, and Wijerupage Sardha Wijesoma. "Robotic mapping using measurement likelihood filtering." *The International Journal of Robotics Research* 28, no. 2 (2009): 172-190. 2.7.1

[43] Mullane, John, Ebi Jose, Martin D. Adams, and Wijerupage Sardha Wijesoma. "Including probabilistic target detection attributes into map representations." *Robotics and Autonomous Systems* 55, no. 1 (2007): 72-85. 2.2.1.3, 2.4, 2.5, 2.6.2.1, 2.6.2.2, 2.7.1

[44] Mullane, John, Ba-Ngu Vo, Martin Adams, and Ba-Tuong Vo. "Random finite sets for robot mapping and slam." *Springer Tracts in Advanced Robotics* 72 (2011). 1.5.2, 2.3.1, 2.8.1, 2.8.1, 2.8.2

[45] National Institute for Occupational Safety and Health, Division of Safety Research, "Highway Work Zone Construction Equipment Visibility NIOSH Manual Method", last reviewed June 30th 2014, accessed online at http://www.cdc.gov/niosh/topics/highwayworkzones/BAD/manualmethod.html, page last visited 16 September 2014. 2.1.3

[46] National Institute for Occupational Safety and Health, Division of Safety Research, "HIGHWAY WORK ZONE SAFETY Construction Equipment Visibility Euclid EH4500" [Internet], Accessed online at http://www.cdc.gov/niosh/topics/highwayworkzones/bad/artictrucks/euclid4500.html on 15th Oct 2014, last modified June 30, 2014. 2.1(a), 2.1(b)

[47] Nebot, Eduardo, Jose Guivant, and Stewart Worrall. "Haul truck alignment monitoring and operator warning system." *Journal of Field Robotics* 23, no. 2 (2006): 141-161. 2.2.1.3

[48] Orchansky, David, Stewart Worrall, Andrew Maclean, and Eduardo Nebot. "Designing a user interface for improving the awareness of mining vehicle operators." In *Proceedings of the 13th International IEEE Annual Conference on Intelligent Transportation Systems*, pp. 19-22. 2010. 1.2.5

[49] Orguner, Umut, Christian Lundquist, and K. Granstrom. "Extended target tracking with a cardinalized probability hypothesis density filter." In *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, pp. 1-8. IEEE, 2011. 5.4.4.6

[50] Pagac, Daniel, Eduardo M. Nebot, and Hugh Durrant-Whyte. "An evidential approach to probabilistic map-building." In *Reasoning with Uncertainty in Robotics*, pp. 164-170. Springer Berlin Heidelberg, 1996. 1.5.2, 2.6.2.1, 2.7, 4.1.1, 4.1.5, 4.1.5.3

[51] Panta, Kusha, Ba-Ngu Vo, Sumeetpal Singh, and Arnaud Doucet. "Probability hypothesis density filter versus multiple hypothesis tracking." In *Defense and Security*, pp. 284-295. International Society for Optics and Photonics, 2004. 5.4

[52] Panta, Kusha, and D. E. Clark. "An efficient track management scheme for the Gaussian-mixture probability hypothesis density tracker." In *Intelligent Sensing and Information Processing*, 2006. ICISIP 2006. Fourth International Conference on, pp. 230-235. IEEE, 2006. 5.4

[53] Patterson, Jessica M., and Scott A. Shappell. "Operator error and system deficiencies: analysis of 508 mining incidents and accidents from Queensland, Australia using HFACS." *Accident Analysis & Prevention* 42, no. 4 (2010): 1379-1385. 2.2.3.2

[54] Resources Safety, Department of Mines and Petroleum, Western Australia, "Safety performance in the Western Australian mineral industry - accident and injury statistics 2009-10", 2011. 2.1.2

[55] Rohling, Hermann. "Radar CFAR thresholding in clutter and multiple target situations." *Aerospace and Electronic Systems, IEEE Transactions on* 4 (1983): 608-621. 2.4, 2.5

[56] Ristic, Branko, Ba-Ngu Vo, Daniel Clark, and Ba-Tuong Vo. "A metric for performance evaluation of multi-target tracking algorithms." *Signal Processing, IEEE Transactions on* 59, no. 7 (2011): 3452-3457. 5.4

[57] Ruff, Todd, "Report of Investigations 9672 Recommendations for Evaluating and Implementing Proximity Warning Systems on Surface Mining Equipment", Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Spokane Research Laboratory, June 2007. 1.3, 1.5.1, 2.1.1, 2.2.1, 2.3, 2.3

[58] Rohling, Hermann. "Radar CFAR thresholding in clutter and multiple target situations", *IEEE Transactions on Aerospace and Electronic Systems*, Volume 19, Issue 4, pp 608-621, 1983. 1.5.2, 2.4.1, 2.5

[59] Ruff, Todd M. *Evaluation of devices to prevent construction equipment backing incidents*. No. 2004-01-2725. SAE Technical Paper, 2004. 2.2.1, 2.2.3.3, 2.3, 3.1.2

[60] Schuhmacher, Dominic, Ba-Tuong Vo, and Ba-Ngu Vo. "A consistent metric for performance evaluation of multi-object filters." *Signal Processing, IEEE Transactions on* 56, no. 8 (2008): 3447-3457. 3.5.1, 3.5.1, 3.5.1, 71

[61] Sivak, Michael. "The information that drivers use: is it indeed 90% visual?." *Perception* 25 (1996): 1081-9. 2.1.3

[62] Skolnik, Merrill I. "Introduction to Radar Systems", McGraw-Hill, 2003 (Third edition) 2.3.2.1

[63] Smith, Randall, Matthew Self, Peter Cheeseman. "A stochastic map for uncertain spatial relationships." In *4th International Symposium on Robotic Research*, pp. 467-474. 1987. 1.5.2, 2.6.3.1

[64] Stevenson, Richard. "Long-Distance Car Radar." IEEE Spectrum (2011). Accessed online at `http://spectrum.ieee.org/green-tech/advanced-cars/longdistance-car-radar/` 1.5.1

[65] Swerling, Peter. "Probability of detection for fluctuating targets." *Information Theory, IRE Transactions on* 6, no. 2 (1960): 269-308. 2.3.3

[66] Thrun, Sebastian B. "Exploration and model building in mobile robot domains." In *Neural Networks, 1993., IEEE International Conference on*, pp. 175-180. IEEE, 1993. 2.6.2.1

[67] Thrun, Sebastian. "Robotic mapping: A survey." *Exploring artificial intelligence in the new millennium* (2002): 1-35. 1.5.2, 2.6.1, 2.6.1, 2.6.1, 2.6.2.2, 2.6.3.1, 2.6.3.2, 2.7, 2.7, 4, 4.1, 4.1.4

[68] Thrun, Sebastian, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics*. MIT press, 2005. 1.5.2, 2.6.1, 2.6.2.1, 2.6.3.1, 2.6.3.2

[69] United States Department of Labor, Mine Safety and Health Administration (MSHA), "Mine Safety and Health Administration Fatalgram #12 - June 20, 2010". Accessed online at `http://www.msha.gov/FATALS/2010/FAB10m12.asp` 1.1(b)

[70] United States Department of Labor, Mine Safety and Health Administration, "2013 Metal and Nonmetal Fatal Accident Review" Accessed online at `http://www.msha.gov/stats/statinfo.htm`, last accessed 11 Sept 2013 2.1.2

[71] Vo, Ba-Ngu. Matlab implementation of optimal subpattern assignment (OSPA) metric, as described in [60]. (2007) [Software] Available from `http://people.eng.unimelb.edu.au/bnvo/vo/ospa_dist.zip` 3.5.1

[72] Vo, Ba-Ngu, Sumeetpal Singh, and Arnaud Doucet. "Sequential Monte Carlo implementation of the PHD filter for multi-target tracking." In *Proc. Int'l Conf. on Information Fusion*, pp. 792-799. 2003. 2.8.2

[73] Vo, Ba-Tuong, Ba-Ngu Vo, and Antonio Cantoni. "The cardinalized probability hypothesis density filter for linear Gaussian multi-target models." In *Information Sciences and Systems, 2006 40th Annual Conference on*, pp. 681-686. IEEE, 2006. 1.5.2

[74] Vo, Ba-Ngu, and Wing-Kin Ma. "The Gaussian mixture probability hypothesis density filter." *Signal Processing, IEEE Transactions on* 54, no. 11 (2006): 4091-4104. (document), 1.5.2, 2.8.1, 2.8.2, 2.8.2, 2.8.2.3, 2.8.2.3, 5.4, 5.1, 5.4, 5.4.2.1, 5.4.3.1, 5.4.4.6, 5.4.4.7, 5.4.5.1

[75] Wood, Trevor. "Tracking in dense clutter with the PHD filter." *Proc. IMA Mathematics in Defence* (2009). 2.6.3.3, 2.8.1, 2.8.1, 2.8.2, 2.8.2.1, 5.4

[76] Worrall, Stewart. "Providing situation awareness in complex multi-vehicle operations." PhD diss., University of Sydney, 2009. 1.1, 1.5(a), 1.5(b), 1.5(c), 1.2.2, 1.3, 1.4, 2.1.1, 2.1.2, 2.2.1.1, 2.2.1.3, 2.2.3.1, 2.2.3.2, 2.2.3.3

[77] Worrall, Stewart, and Eduardo Nebot. "A probabilistic method for detecting impending vehicle interactions." In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pp. 1787-1791. IEEE, 2008.