A UNIFIED RANDOM FINITE SET THEORETIC APPROACH TO AUTONOMOUS UNDERWATER VEHICLE NAVIGATION

BHARATH KALYAN

School of Electrical and Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

2011

I would like to dedicate this thesis to my parents who have always stood by me and instilled the importance of hard work.

Acknowledgements

First and foremost I'd like to thank my ever supporting parents for all of the love and support they have given me through the years, without whom, nothing would have been possible. I would like to dedicate this work to them.

I would next like to sincerely express my gratitude to my two supervisors - Dr. Sardha Wijesoma (WJ) and Dr. Arjuna Balasuriya for their excellent supervision through the course of my PhD, their consistent support and guidance has been invaluable. Hopefully WJ has managed to imbibe in me his most used idiom " Devil lies in the details". I'm also indebted to Professor Tamaki Ura from University of Tokyo for funding my visit to work with him in Japan, and the invaluable experience gained working with his team and Kondo Sensai during the field trials. I would also like to thank Dr. Ba-Ngu Vo from University of Melbourne for his valuable inputs on the work of this thesis.

I am grateful to NTU for providing me the financial support and a conducive environment to carry out this research work. I'd also like to thank my fellow team member Kwang Wee for his immense help in form of various technical discussions. Also many thanks to DSO, Singapore, and, in particular, Alvin and Soon Lee for their interest in this work and assisting us in conducting field trials. I would also like to thank our friendly technicians Chia and Neo from Intelligent Robotics Lab at NTU for all the technical assistance.

Apart from the research aspects, there are a number of people who have helped me along this long never ending course. Panda, my dive, travel, letching buddy with whom I've spent innumerable hours discussing about everything under the sun and beyond. Ebi, my flatmate for god knows how many years now, and my friends Nona, Amit murgawalah, Bihari, Chodu, Jamla, Talku, Pinky, Sylvia for all their love and support. Thanks to Mandar for his endless thought provoking discussions on topics varying from science, philosophy, photography, diving, wine, chillies, etc., John Potter for introducing me to world of rebreather diving, Mehul

and Koay, my ARL buddies and Mohan for his sumptuous Onam meals. Finally, I'd also like to thank my brother and all members of my family for all of the love and support they have given me through the years.

Abstract

This thesis addresses the problem of navigation, localization and map-building in an unknown and unstructured environment and its application to an autonomous underwater vehicle (AUV). Given the dynamic nature of the underwater environment and the limitations of traditional baseline based systems, the inability to use GPS underwater, and the range limited optical vision systems, our approach for overcoming these navigation limitations has been to embrace a feature based simultaneous localization and mapping (SLAM) framework while explicitly exploiting the available navigational sensor suite and rich blazed array sonar imagery that is commonly used in underwater explorations. The most common formulation of the feature based SLAM problem is founded on a vector based stochastic framework, where the sensor models and the vehicle models are represented in state space form and the joint posterior or its statistics are obtained based on recursive Bayesian estimation. All the SLAM solutions leading from the stochastic vector state-space approach require that we solve certain parallel problems in each recursion. These include effective solutions to the problems of data association, feature detection and extraction, clutter filtering and landmark or map management. In this thesis, we devise an alternative unifying framework based on finite set statistics (FISST), where the SLAM problem is reformulated so that the landmark map and the measurements are represented using random finite sets and the landmark map is jointly estimated with the vehicle state vector, whilst explicitly accounting for measurement detection uncertainty, data association uncertainty, false alarms and map management in the SLAM filter framework. The central thesis of this dissertation is that the random finite set framework is not only a natural, elegant and a unifying framework, but also leads to practical, efficient and reliable algorithms for autonomous vehicle navigation. The first key contribution is a mathematically rigorous multi-landmark map evolution model that accounts for new incoming landmarks and a sensor measurement model that accounts for sensor's detection uncertainty, false alarms and clutter for autonomous vehicle navigation.

The second key contribution is a novel algorithm for autonomous vehicle navigation where a clutter rejection algorithm based on moment approximations to the posterior density of the random finite set state is naturally integrated within the conventional SLAM framework. The performance of the proposed algorithm is demonstrated using synthetic data and real offshore data, and shown to considerably outperform conventional SLAM approaches in underwater environments. The third key contribution is a mathematically principled unifying random finite set theoretic SLAM framework, where the SLAM problem is reformulated so that the landmark map and the measurements are represented using random finite sets and the landmark map is jointly estimated with the vehicle state vector. Results of this proposed algorithm are presented both in simulation and real data collected during deployment of an AUV equipped with a blazed array sonar in an offshore environment.

Contents

List of Figures

List of Tables

List of Acronyms

ASC	Autonomous Surface Craft
AUV	Autonomous Underwater Vehicle
BAS	Blazed Array Sonar
CCDA	Combined Constraint Data Association
CFAR	Constant False Alarm Rate
CLSF	Constrained Local Submap Filter
CPHD	Cardinalized Probability Hypothesis Density
DR	Dead Reckoning
DVL	Doppler Velocity Log
EAP	Expected a Posteriori
EKF	Extended Kalman Filter
EM	Expectation Maximization
FISST	Finite Set Statistics
FOG	Fiber Optic Gyroscope
FOV	Field Of View
GM	Gaussian Mixture
GPS	Global Positioning System
GRBI	Generic Range Bearing Intensity
ICP	Iterative Closest Point
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
JCBB	Joint Compatibility Branch and Bound

JPDA	Joint Probabilistic Data Association
KF	Kalman Filter
LBL	Long Baseline
MAP	Maximum a Posteriori
MCM	Mine Counter Measure
MFDA	Multiple Frame Data Association
MFT	Multiple Feature Tracking
MHT	Multiple Hypothesis Tracking
OS	Order Statistics
PF	Particle Filter
PGFl.	Probability-generating functionals
PHD	Probability Hypothesis Density
RBPF	Rao-Blackwellization Particle Filter
RCD	Region of Constant Depth
RFS	Random Finite Set
ROV	Remotely Operated Vehicle
SBL	Short Baseline
SIS	Sequential Importance Sampling
SLAM	Simultaneous Localization and Mapping
SMC	Sequential Monte Carlo
SNN	Standard Nearest Neighbor
SONAR	Sound Navigation and Ranging
UKF	Unscented Kalman Filter
USBL	Ultra Short Baseline

1

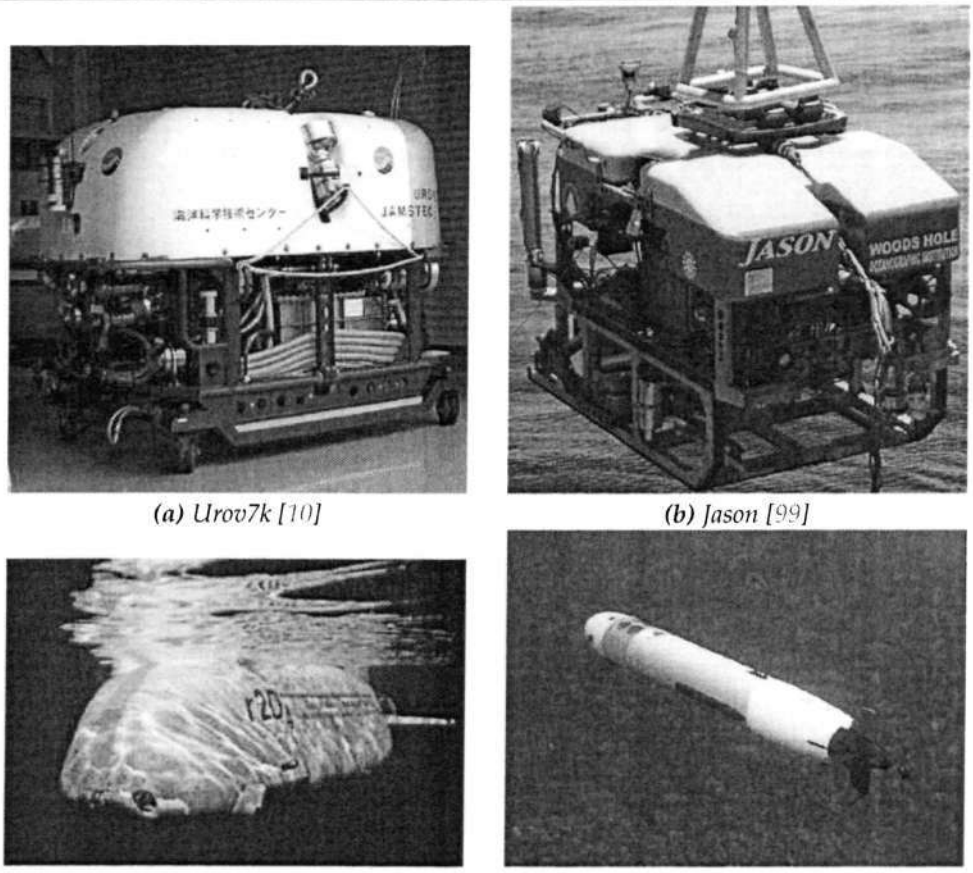
Introduction

1.1 Motivation

Underwater is rich with mineral resources, oil and gas reserves. Depletion of the renewable resources on land makes underwater exploration an awful necessity. With increase in population, underwater man made structures such as oil and water transportation pipelines, telecommunication cables, underwater structures etc., are rapidly increasing. These man made systems have to be routinely checked for safe and reliable functioning. Homeland security and mine counter measures (MCMs) are some of the important concerns of the present times. Thus to explore the deep and shallow underwater environments, maintenance of underwater structures and to provide safe waters, we would need underwater vehicles. Figure 1.1 illustrates different classes of underwater vehicles. The tethered remotely operated vehicles (ROVs) are currently the most widely used underwater vehicles, serving a range of military, commercial, and scientific needs. The tether supplies power and communication to the ROV and is controlled directly by a remote human operator. The

umbilical is one of the vehicles main assets, and at the same time, one of its main drawbacks. Since the ROV is physically connected to a surface vessel, ample power can be sent to the vehicle and large amounts of data can be received. Working against this, however, is drag on the umbilical and the increasing power required as depth or speed of operation of ROV is increased. Also twisting, breaking of the tether or entangling with the objects is of high concern, when operating the ROVs. For ROVs, which must operate in great depths or in high currents, a substantial cable winch and power generator are required, which necessitate a sizable surface support ship. Autonomous underwater vehicles (AUVs) that complement the ROVs, have been growing in popularity due to their unmanned and untethered design which makes them suited for large scale exploratory surveys with minimal human intervention and surface support. Also with the advent of new hardware technologies and powerful processors, these underwater vehicles have shown the potential to revolutionize our access to the oceans to address critical problems facing the marine community such as underwater search and mapping [100], climate change assessment, under ice exploration [20], geological mapping [84], marine habitat monitoring, man-made structure maintenance [17] and shallow water mine countermeasures [28], [47]. Accurate navigation is a crucial aspect of each of these missions.

Navigation, which is the combination of precise positioning and guidance, is a fundamental challenge in mobile robotics community. Good navigation information is essential for safe operation and recovery of any mobile robot and this especially true for AUVs. For the data collected by an AUV to be of value, the location from which it was acquired must be precisely known. While the underwater terrain itself presents its own challenges to autonomous vehicle navigation such as the effects of



(a) Urov7k [10] *(b) Jason [99]*

(c) r2d4 [107] *(d) Remus [28]*

Figure 1.1 *State of the art Unmanned Underwater Vehicles*

acoustic propagation, backscatter and marine snow for vision, the lack of obtaining precise navigational information is not limited to underwater vehicles. Global positioning systems (GPS) which have been used extensively used to provide precise positioning for land and air-based systems cannot be used underwater due to rapid attenuation of electromagnetic signals in water. As acoustic energy propagates well in the underwater environments, acoustic transponders can be used as "beacons" to guide the motion of an AUV. The most common techniques employed for tracking underwater vehicles by acoustic base line systems in form of either long base line

(LBL), short base line (SBL) and ultra short base line (USBL) systems [57]. However, the major drawback with these systems being the line of sight operation and expensive and elaborate deployment and recovery procedures [113]. Within the last decade, the development of commercially available, precise, high update rate navigation sensors such as Doppler sonars, optical gyro-compasses, and inertial measurement units (IMUs), have served to complement the old school of underwater sensors such as acoustic positioning systems, magnetic compasses, and pressure depth sensors [57]. Data from these sensors, along with data from side scan sonars and optical cameras, have served in the development of novel navigation methodologies. Many of these methods aid sensor data with information from dynamic or kinematics models. IMUs offer high update rate, excellent strap down capabilities due to which it has been applied in numerous underwater applications. The development of high-frequency (600 kHz to 1.2 MHz), multi-beam Doppler sonars that provide bottom velocity measurements with a precision of 0.3 m or less and update rates up to 5 Hz provide researchers with velocity measurements for near-bottom (2 - 100 m) navigation [61]. This has enabled the development of a wide variety of Doppler-based navigation techniques. However these systems suffer from an unbounded-error position estimate based on integrating velocity or acceleration. While the level of precision obtained by such systems can often be satisfactory for shorter duration missions, but is typically intolerable for larger area missions. The past two decade of underwater navigation literature shows that, the community has been focusing significantly on developing terrain aided navigational algorithms using on-board sensing systems, that eliminate the need for additional infrastructure and bound the position error growth. The solution to the problem of mapping an environment and simultaneously using this map to localize is considered as a key

prerequisite in the synthesis of truly autonomous vehicles. The question of how to use such a methodology for mobile robot navigation was theoretically addressed in a seminal paper series by [102] and [80], now popularly referred to as simultaneous localization and mapping (SLAM) problem. Thus, a SLAM framework seems like a natural choice for overcoming the current navigation limitations of the underwater vehicles. The following section gives a detailed overview of underwater sensing and current practices for underwater navigation.

1.2 A Overview of Underwater Vehicle Sensing & Navigation

In the context of Autonomous Vehicles, navigation involves answering three distinct questions: *"Where am I?"*,*"Where am I going?"*, and *"How do I get there?"* [34]. The first question is simply to know where the vehicle is, at all times. The second question is mainly concerned with vehicle path planning. The third question looks at the control of the vehicle in the desired path. The focus of this thesis is on the first question, which is the vehicle positioning or the localization problem for underwater vehicles. A recent survey paper also highlights the challenges dominating the underwater vehicle navigation [57].

1.2.1 Dead-Reckoning Navigation

Dead reckoning (DR) is the most widely used navigation method for mobile robot positioning. The primary sensors used in underwater environments for dead reckoning are the inertial navigation systems (INS) and Doppler velocity logs (DVL).

Inertial Navigation

INS uses gyroscopes and accelerometers to measure rate of rotation and acceleration, respectively. Measurements are integrated to yield position and orientation(pose) of vehicle with respect to some inertial frame of reference. Inertial navigation systems have the advantage that they are self-contained, that is, they don't need external references. However inertial sensors are prone to errors from many sources - bias, random noise, temperature, aging. These lead to unbounded errors in position and velocity upon integration.These systems also require accurate calibration regularly. Thus inertial sensors are mostly unsuitable for accurate positioning over an extended period of time.

Doppler Velocity Log

DVLs calculate vehicle velocities, relative to the ocean floor or the water column based on the principle of Doppler shift [108]. Most DVLs have four transducers, each aligned at an angle of 30^o from the vertical and arranged in a Janus configuration (facing opposite directions) as shown in the fig. 1.2 [6]. Each transducer transmits sound waves (referred to as ping) in a narrow beam at a known frequency (usually high frequency in the order of $300 - 1200kHz$). The pings forming an acoustic beam strike the particulate matter suspended in the water column or the seabed. When the pings strike scattering centers, some of the sound energy is reflected along the acoustic beam to the transducer. The returned sound has a frequency (Doppler) shift proportional to the velocities of the scattering centers and the water they are traveling in, along the acoustic beam.

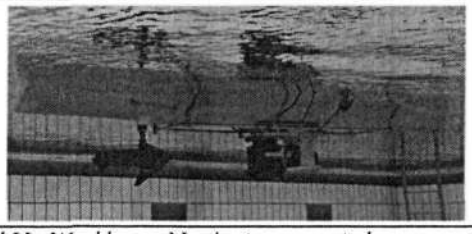


Figure 1.2 *RDI 600kHz Workhorse Navigator mounted on an autonomous surface craft*

1.2.2 Compasses, Inclinometers and Depth Sensors

These sensors form the core of any underwater navigation system due to their drift free nature. A compass measures the orientation with respect to earth's magnetic field. However they are fairly susceptible to errors due to local magnetic disturbance, geographical magnetic anomalies or due to improper mounting. The inclinometers are used to measure the roll and pitch based on measuring the direction of acceleration due to gravity. The depth sensor computes the depth based on the direct measurements of the ambient water pressure using either strain gauges or quartz crystals.

1.2.3 Acoustic Time of Flight Navigation

As acoustic energy propagates effectively in the underwater environment, acoustic transponders are commonly used as "beacons" to guide the motion of an AUV. There are three standard techniques available for tracking/positioning underwater vehicles by acoustic means; long base line (LBL), short base line (SBL) and ultra short base line (USBL) acoustic navigation [12], [20].

In LBL navigation an array of acoustic transponders are deployed on the bottom at known locations that reply to transmissions from the vehicle yielding direct range

measurements from which the position of the AUV can recovered from the process of triangulation. The array is usually geodetically calibrated, i.e., its geographical x, y, and z coordinates are known resulting in "absolute" fixes. Most LBL systems work at a frequency of around 10 kHz and provide position accuracy to within a few meters over an operating area of the order of few kilometers. Higher frequency LBL systems operating around 300 kHz are expected to provide position accuracy of the order of centimeters, with an operating area of 100 m [113].

In SBL navigation an array of hydrophones mounted on a vessel track the AUV with direct range measurements in relative coordinates while converting to an absolute coordinate system with the aid of typical surface navigation systems and compass [12].

In USBL navigation a multi-element array of hydrophones are mounted on the vehicle that measures direct range and bearing to an acoustic beacon. By measuring the phase difference of the sonar returns between the hydrophones, the bearing from the vehicle to the beacon can be determined. Range is determined by measuring the travel time between the vehicle and the acoustic beacon, as in LBL systems [20].

Figure 1.3 illustrates a typical Long Base Line array setup. The main disadvantage with these systems are that they are very expensive and deployment of the equipments are cumbersome. Also a significant difficulty in acoustic navigation is due to reflection and multi-path errors, which will result in incorrect time-of-flight values and hence erroneous position fixes [75]. Another major drawback of such systems is that the transponders and the vehicle's receivers must be on line of sight at all times.

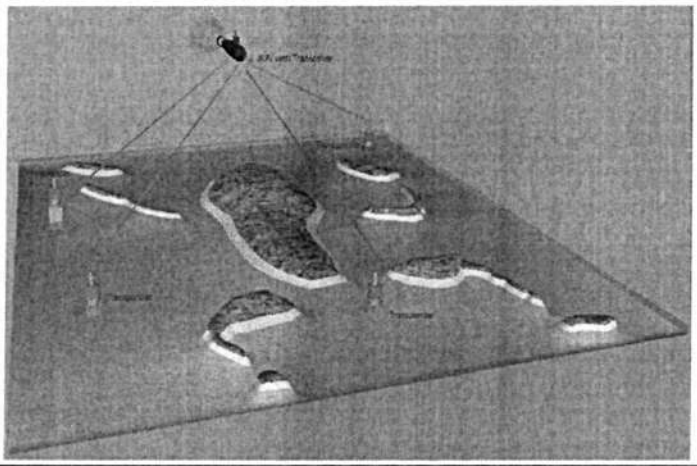


Figure 1.3 *An artistic view of Long Base Line array setup with transponders positioned on the sea bed and AUV navigation based on triangulation*

1.2.4 GPS aided Navigation

The global positioning systems (GPS) system of satellites is extensively used to provide precise positioning data for land-based and air-based systems. Unfortunately, electromagnetic signals attenuate rapidly in water, and hence the use of GPS in the underwater regime is severely limited. One approach taken generally is to equip an AUV with a GPS receiver system and bring it to surface at periodic intervals. The GPS measurements are then be used to reset the drift errors of the INS measurements [9]. Similarly, several organizations have successfully integrated GPS with LBL or USBL networks to perform tasks such as station keeping [103], [9]. Since these techniques rely on the previous setup of an acoustic network, it is not possible to explore completely new environments.

1.2.5 Terrain Map aided Navigation

Map aided navigation aims at correcting vehicle navigation errors with the help of the available terrain map. Most of the work has been concentrated on obtaining the position estimate of the vehicle given an *apriori* map of the environment. If an accurate *apriori* environmental map such as bathymetry, magnetic field, or gravitational anomaly are available, one of the approaches could be to match the sensor data with these maps to estimate the position of the vehicle [43], [56]. With the advent of sophisticated sensor technologies and efficient sensor fusion algorithms, terrain map aided systems have been used for underwater navigation [30], [88]. However in practice, an up-to-date, high quality environmental map may be unavailable in the operating area of interest as there is no access to an external reference to provide position information for navigation, especially for unchartered territories underwater or otherwise.

1.2.6 Simultaneous Localization and Mapping

The SLAM problem examines whether a vehicle can build a map of its unknown environment incrementally by using the uncertain information extracted from its on-board sensors whilst simultaneously using that map to position itself in real time. However, after two decades of research and having established a comprehensive framework to the SLAM problem researchers still face several challenges in realizing fully deployable SLAM solutions especially in complex unstructured underwater environments. This is mainly attributed to inaccuracies in sensor perception and modeling, data association related issues in SLAM [35]. Had there been ideal sensors capable of providing accurate measurements and ideal models

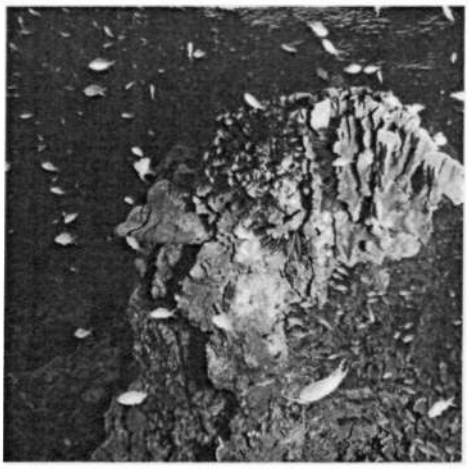
for vehicles and sensors, which can be used to exactly predict the vehicle motion and relative landmark positions there would not have been a localization and a map building problem. The general solutions to the SLAM problem have been trying to exploit the perceptual sensing capabilities of these vehicles that are equipped with on-board sensors such as sonars, optical cameras, to correct for the large inaccuracies that accumulate over time in DR sensors (INS and DVL) due to integration of noisy information, to achieve accurate positioning in vehicle navigation applications. However, the use of these perceptual sensors introduces problems in the form of feature extraction, data association, clutter filtering, and landmark or map management that have to be addressed in order to determine the vehicle's state and knowledge it currently possesses about its environment in form of a map.

1.2.7 Underwater SLAM

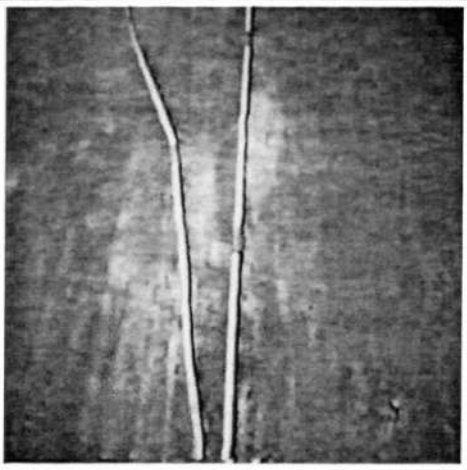
Apart from the problems discussed in trying to incorporate a SLAM methodology for navigation, a number of further constraints come to the fore in the underwater environments due to sensing limitations and the challenging constrains imposed by the natural unstructured terrain [57].

A large portion of the work in the SLAM literature has relied upon high-precision laser scanners or optical vision as the perceptual sensor of choice for constructing accurate maps. However, strong attenuation of the electromagnetic waves in the underwater environment limits the use of optical cameras and laser scanners to very short ranges $\mathcal{O}(1-5m)$ and its performance is further affected in adverse conditions such as murky water [51]. However, this has not deterred a number of researchers who have examined the use of optical cameras as a primary navigation sensor [40],

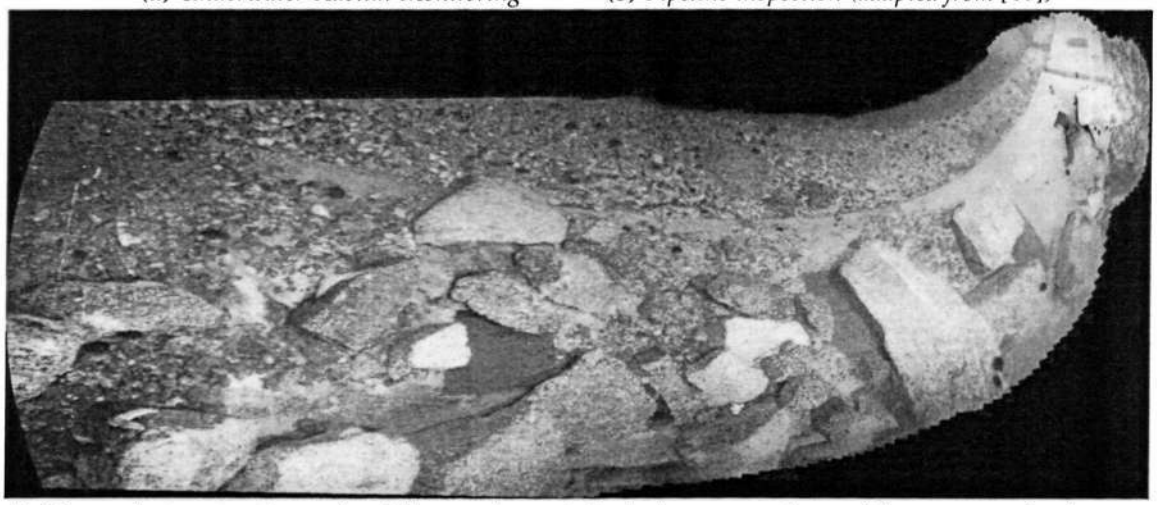
[85], [116], [39] apart from its use for near bottom applications; photomosaicing [73], [85], pipeline & cable following [16]. See fig. 1.4 for a few of these applications. The



(a) Underwater Habitat Monitoring



(b) Pipeline inspection (adapted from [16])



(c) Man made structure inspection. This mosaic was built during a research expedition to survey breakwaters along the coast of Kamaishi, Japan using Tridog-1 AUV

Figure 1.4 *Applications of Underwater Optical Vision*

optical cameras have also been used to perform feature based SLAM by extracting features from images of the seabed [42], [39] when the vehicle is operating very close to the sea-bed under good visibility conditions.

Sound navigation and ranging (SONAR) is a system that uses acoustic signals to explore the environment and is most commonly used in water to detect, classify, and localize underwater objects. They can be broadly categorized into passive and active sonars [108]. Passive sonar senses the sound radiated from the underwater environment. Active sonar involves transmitting an acoustical signal from a source and receiving reflected echoes from the object of interest. Passive sonars are mainly used in military applications such as sonars on submarines where covertness is essential and certain applications that are directed to studying the ambient properties of the underwater environment. While the active sonars have been most commonly used in civilian applications such as in fisheries, offshore oil and gas industries and in AUVs and ROVs for the purpose of obstacle avoidance and navigation. Hence given the task at hand of positioning AUVs and mapping the underwater environment, active sonar is the most appropriate choice as a perceptual navigation sensor. Active Sonars can be broadly categorized as follows:

- **Sidescan sonar**

  Sidescan sonar is generally used for mapping the sea-floor. It emits fan-shaped beam down toward the sea-floor across a wide angle perpendicular to the path of the sensor through the water, which may be towed from a surface vessel or attached to an underwater vehicle. The intensity of the acoustic reflections from the sea-floor of this fan-shaped beam is recorded in a series of cross-track slices. When stitched together along the direction of motion, these slices form an image of the sea bottom within the swath of the beam. The sound frequencies used in side-scan sonar is in the order of 100 to 500 kHz, with higher frequencies yielding better resolution but lower range of operation. A sidescan sonar composite image built during a ship wreck survey in coastal

waters of Singapore is illustrated in fig. 1.5.

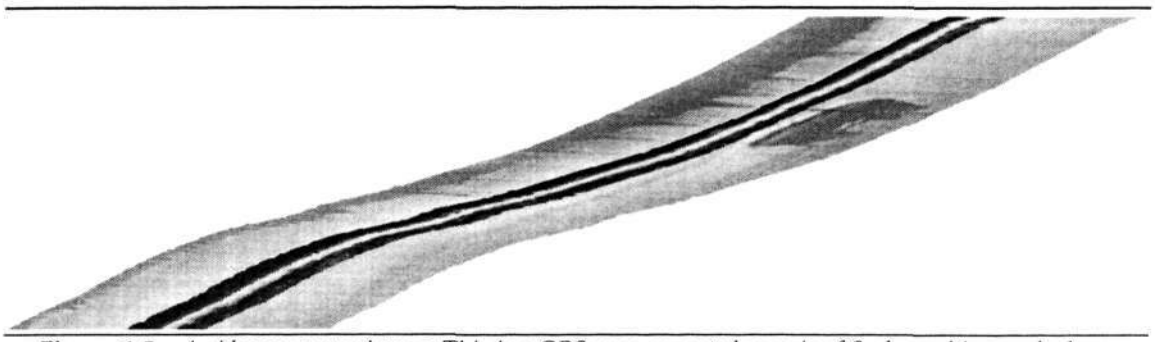


Figure 1.5 *A sidescan sonar image. This is a GPS pose corrected mosaic of Sudong ship wreck along coastal water of Singapore built using starfish sidescan sonar [8] mounted on a sea-kayak.*

- **Forward looking sonar**

  The forward-looking sonar, sometimes also called as a sector-scan sonar, is extensively used in obstacle and terrain detection for underwater vehicles. The same sonar, pointed down instead of ahead and with beams formed athwartship, are referred to as bathymetric sonar and are used to map the sea beds.

  - **Mechanically steered sonar**

    A mechanically steered sonar generally scans a two dimensional plane by steering a fan-shaped sonar beam in discrete angular steps using a stepper motor. For each transmitted beam, an intensity profile is returned from the environment that is accumulated to form an acoustic image. The frequency of operation is in the order of 500 to 900 kHz with an operating range of up to 200 m. However, mechanically steered sonars have very slow data update rates ($\sim$ 0.2 Hz) that limit its applications to very slow moving vehicles. Figure 1.6 shows an acoustic image obtained with a

mechanically steered sonar, Tritech Minking, during experiments in the swimming pool environment.

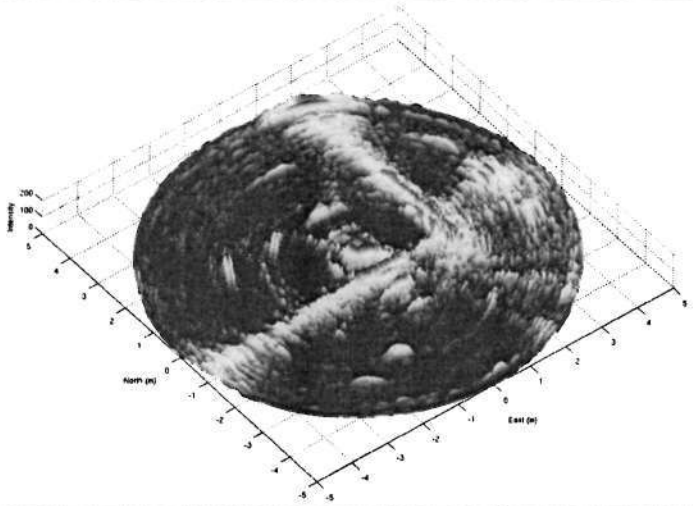


Figure 1.6 *An acoustic image formed from a Tritech miniking sonar [8]. This illustrates an acoustic image of the swimming pool walls with artificial beacons (cylinders) in the environment*

- **Blazed array sonar**

  A blazed array sonar (BAS) generally transmits a broadband pulse through a transmitting array that creates a frequency dispersed sound field. It interacts with the underwater environment and the sound field is received by a receiver array which is then combined to form a single signal. An acoustic image is generally generated through time scale processing. The frequency of operation is in the order of 900 to 1200 kHz with an operating range of up to 55 m. The main advantage of these sonars over to their mechanically steered counterparts is the faster data update rates ($\sim$ 10 Hz) and a higher resolution due to overlapping beams. Thus given these advantages the BAS is used as a primary navigation and mapping sensor in this thesis for demonstrating the efficacy of the algorithms, de-

veloped. Figure 1.7 illustrates an acoustic image formed with a Blueview BAS P900E during experiments along the coastal water of Singapore.

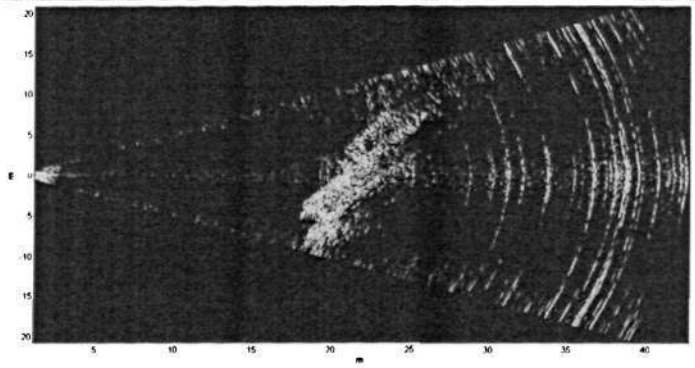


Figure 1.7 *An acoustic image formed from a Blueview blazed array sonar [1]. The acoustic image shows the Sudong wreck along coastal waters of Singapore as seen by BAS*

For range sensors such as sonar, geometric primitives such as lines, corners that correspond to walls are generally used as features to perform SLAM in structured indoor environments. In contrast to the indoor environments, which are dense and resemble proper geometric primitives like corners and straight lines, unstructured natural underwater environments are sparse in detail and do not conform to any simple geometric model. The common approach used in such scenarios is to use segmentation followed by a simple clustering algorithm to group raw sensor measurements into data clusters and extract descriptors that sufficiently describe the cluster. Cluster centroids have been extensively used as feature descriptors to perform EKF based underwater SLAM, given the difficulty in modeling the terrain and processing cost involved [94], [115]. Carpenter has used a series of morphological operations on a sonar scan image to form objects (clusters), and retrieved perimeter, area, area-to-perimeter ratio and radial signature (distance from objects centroid to its boundary as a function of angle) as geometric descriptors to characterize a land-

mark and used these landmarks within an EKF based SLAM framework [23]. An alternative representation was introduced by Majumder [72], where Sum of Gaussian (SOG) distribution is used to build a feature map of landmarks to be used within the stochastic SLAM framework. The landmarks were initially characterized based on the intensity returns from sonar signal returns and a distance criterion was used to group measurements as blobs (clusters), and its peak amplitude, variance and the associated area were retrieved as descriptors. Ribas et.al have used a simple line extraction technique from sonar image when operating in man made structured underwater environments and incorporated them within the EKF based SLAM framework [92].

1.2.8 Limitations of current approaches

Acoustic transponder based navigation systems using LBL, SBL offer bounded position measurements, but at high deployment costs. Moreover, such systems limit the vehicle navigation to the deployed network, which is not conducive in a long term when multiple site surveys are essential. The usage of GPS, underwater is severely limited due to the attenuation of electromagnetic signals underwater. Alternatively, the usage DVL and other INS systems reduces the need for additional infrastructure by improving the dead-reckon navigation capabilities. However, the open-loop nature of these systems result in an error that is unbounded as a function of distance traveled.

The optical vision based systems are limited by environmental visibility and limited range of operation. Moreover, most of the prominent vision based navigational approaches rely on building 2D mosaics of a 3D environment severely violate the

planar assumption of mosaicking [42],[85].

Given the dynamic nature of the underwater environment and the limitations of traditional LBL and DVL methods, the inability to use GPS underwater, and the range limited optical vision systems, our approach for overcoming these navigation limitations has been to embrace a SLAM framework while explicitly exploiting the available navigational sensor suite and rich blazed array sonar imagery that is commonly used in underwater explorations.

1.3 Thesis Objective and Approach

This thesis addresses the problem of simultaneous localization and map-building in an unknown and unstructured environment using a novel unified framework based on Random Finite Sets. The most common formulation of the SLAM problem is founded on a vector based stochastic framework, where the sensor models and the vehicle models are represented in state space form and the joint posterior or its statistics are obtained based on recursive Bayesian estimation. All the SLAM solutions leading from the stochastic vector state-space approach require that we solve certain parallel problems in each recursion. These include effective solutions to the problems of data association, feature extraction, clutter filtering, and landmark or map management. In this thesis, an alternative unifying framework based on finite set statistics (FISST) is devised, where the SLAM problem is reformulated so that the landmark map and the measurements are represented using random finite sets and the landmark map is jointly estimated with the vehicle state vector, whilst explicitly accounting for measurement detection uncertainty, data association uncertainty, false alarms and map management in the SLAM filter framework.

1.4 Contributions

The key contributions of the thesis can be summarized as follows:

- Development of a mathematically rigorous multi-landmark map evolution model that accounts for new incoming landmarks and a sensor measurement model that accounts for sensor's detection uncertainty, false alarms and clutter for autonomous vehicle navigation based on random finite set theory and FISST calculus.

- Analysis of the landmark (feature) detection problem for generic range-bearing imaging (GRBI) sensors (with blazed array sonar in particular) in an autonomous navigation framework. The introduction of adaptive statistical detectors and their applicability to autonomous underwater navigation, whose efficacy is shown through statistical analysis on blazed array sonar data.

  Development of a novel clutter rejection algorithm for autonomous vehicle navigation. The algorithm is based on the fusion of the moment approximation of the posterior density within the random finite set framework with the conventional SLAM framework.

- Development of a novel algorithm for autonomous vehicle navigation where a clutter rejection algorithm based on moment approximations to the posterior density of the random finite set state is naturally integrated within the conventional SLAM framework.

- Development of a mathematically principled unifying random finite set theoretic SLAM framework, where the SLAM problem is reformulated so that the

landmark map and the measurements are represented using random finite sets and the landmark map is jointly estimated with the vehicle state vector.

- Demonstrate the feasibility of the proposed solutions through the use of extensive simulations and experiments using AUVs in real offshore underwater environments.

1.5 Thesis Structure

This section gives an overview of the approach taken in this thesis and provides the structure of the thesis. The work carried out and the structure of the thesis are outlined here.

- **Chapter 2** examines and reviews the state of the classical vector based SLAM algorithms. A Bayesian filter navigation framework is derived and the approximations to the optimal Bayes navigational filter are discussed. The chapter concludes by emphasizing on the standard properties and oversights of classical vector based SLAM solutions.

- **Chapter 3** briefly introduces the concepts of random finite set theory. A map evolution model that accounts for new incoming landmarks and a sensor measurement model that accounts for sensor's detection uncertainty, false alarms and clutter are developed to be applied in an autonomous vehicle navigation framework. These models are adapted for autonomous vehicle navigation based on FISST calculus, originally developed by Mahler.

- **Chapter 4** outlines the key role played by the feature extraction in the feature based environmental mapping and vehicle localization context. Here we

also analyze the landmark (feature) detection problem for GRBI sensors and introduce adaptive statistical detectors in the form of order statistics CFAR detectors and their applicability to autonomous underwater navigation. The efficacy of these detectors are shown through statistical analysis on the blazed array sonar data. A clutter rejection algorithm based on probability hypothesis density filter is presented that simplifies the data association process and obviates the need for *ad hoc* multi-landmark map management strategies. The proposed algorithm is initially validated for performance using simulations and further using blazed array sonar. The effectiveness of the proposed clutter rejection strategies is further demonstrated in conjunction with the entire SLAM framework using a standard EKF-SLAM filter without any external map management strategies.

- **Chapter 5** proposes a unified framework based on random finite sets referred to as FISST-SLAM, where the SLAM problem is reformulated so that the landmark map and the measurements are represented using random finite sets and the landmark map is jointly estimated with the vehicle state vector, whilst explicitly accounting for measurement detection uncertainty, data association uncertainty, false alarms and map management in the SLAM filter framework. Similar to FastSLAM, the proposed formulation involves factorization of the full SLAM posterior into a product of the vehicle trajectory posterior and the landmark map posterior conditioned on the vehicle trajectory. The vehicle trajectory posterior is then estimated using a particle filter and the map posterior conditioned on the vehicle trajectory via a probability hypothesis density (PHD) filter. Two different implementations of the FISST-SLAM algorithm in-

form of sequential Monte-Carlo and Gaussian mixture implementations are discussed. Simulations and offshore underwater field trial results are presented and benchmarked against FastSLAM to demonstrate the effectiveness and improved performance of the FISST-SLAM in the presence of significant clutter.

- **Chapter 6** concludes by summarizing the contributions of this thesis, critiquing the results and providing recommendations for future research in this domain.

2

Classical Vector based SLAM

In this chapter we will present an overview of the classical vector state space based Simultaneous Localization and Mapping (SLAM) problem, along with the most common approaches employed to solve this problem. It has often been recognized as one of the key problems in building truly autonomous systems, in the mobile robotics literature. The SLAM problem examines whether a vehicle in an unknown environment can build an incremental map of its surroundings by using uncertain information extracted from its on-board sensors whilst concurrently using the same map to position itself and navigate in real time. The SLAM problem was first proposed by Smith et. al. [102], [101], where they introduced the Extended Kalman Filter (EKF) based stochastic framework for joint estimation of the vehicle state and feature based point landmark (map) locations. Since then, variants of EKF-SLAM [63], [45], [117] and several other alternative approaches including probabilistic [72] and particle filter based methods [77], have been presented. SLAM is a highly active field of research within the mobile robotics community as presented in the recent survey papers [35], [15].

This chapter is organized as follows. In Section 2.1, a probabilistic framework to the SLAM problem along with a Bayes recursive solution is discussed. Section 2.2 deals with the estimation theoretic approximations in general to overcome the computational intractability involved in calculation of Bayes SLAM posterior. It also highlights the potential drawbacks of the standard EKF-SLAM filter and variants that have been developed by researchers to overcome these problems. In Section 2.3, sequential Monte-Carlo approximation techniques are discussed, followed by Rao-Blackwellization methods. We then discuss the properties of all these classical vector based SLAM solutions highlighting the drawbacks in Section 2.4 and conclude by proposing an alternative finite set theoretic framework to the SLAM problem in Section 2.5.

2.1 Probabilistic Formulation of the SLAM Problem

Let $\mathrm{U}^{k-1} = [\mathrm{u}_0, \mathrm{u}_1, ..., \mathrm{u}_{k-1}]^T$, denote the time sequence of controls applied to the vehicle up to and including time $k-1$, where u_k is the control input at time k. Let the map containing N_k number of static landmarks be represented by $\mathrm{M}_k = [m_{k,1}, \ldots, m_{k,n_k}]^T$ at time k. Let $\mathrm{Z}_k = [\mathrm{z}_{k,1}, \ldots, \mathrm{z}_{k,l_k}]^T$ be measurements taken by the perceptual sensor on-board at time k. Let $\mathrm{Z}^k = [\mathrm{Z}_1, \mathrm{Z}_2, \ldots, \mathrm{Z}_k]^T$ denote the history of all the measurements taken up to time k. Let $\mathrm{X}^k = [\mathrm{x}_1, \ldots, \mathrm{x}_k]$ denote the vehicle's complete pose history, where x_k represents vehicle's pose at time k w.r.t a global navigation frame.

The following probabilities are of interest for the formulation of the problem,

- $f_{k|k-1}(\mathrm{x}_k|\mathrm{x}_{k-1}, \mathrm{u}_k)$: Probability of evolving to a vehicle state x_k at time k given the vehicle was at a state x_{k-1} and a control input of u_k was applied at time

k. It is commonly referred to as state transition probability and constitutes the vehicle motion model [105]. The motion model is intrinsic to the vehicle parameters and its construction must also reflect the modeling of other uncertainties related to the vehicle dynamics.

- $f_k(z_k|x_k, M_k)$: Probability of making an observation z_k at time k, given the vehicle state is x_k and map state M_k. It is commonly referred to as observation probability (or sensor likelihood function) and constitutes the measurement model. The measurement models must reflect the measurement uncertainties as well as modeling errors, if any.

From a probabilistic perspective, there are two main forms of the SLAM problem. Determining the map state and the pose of the vehicle during the operation of the vehicle requires, finding the current vehicle state x_k and the map state M_k. This is generally referred to as an *online SLAM problem* and requires that the probability distribution

$$f_{k|k}(x_k, M_k|Z^k, U^k, x_0) \tag{2.1}$$

be computed for all times k with x_0 being the initial state [1] of the vehicle. This probability distribution describes the joint posterior density of the map state and vehicle pose at time k given all the past observations and control inputs along with the initial vehicle state. This formulation does not use the new information obtained over time to update the old vehicle pose estimates, which in fact can be used to minimize past vehicle pose estimate errors. The *full SLAM problem* estimates the joint posterior over the entire vehicle path X^k along with the entire map M_k instead

[1] Initial state of the vehicle x_0 is assumed to be known and hence will not be indicated.

of just the current vehicle pose x_k

$$f_{k|k}(X^k, M_k|Z^k, U^k, x_0) \tag{2.2}$$

Thus in fact, online SLAM is the result of removing the past poses from the full problem using integration.

$$f_{k|k}(x_k, M_k|Z^k, U^k, x_0) = \int_{x_1} \int_{x_2} \cdots \int_{x_{k-1}} f_{k|k}(X^k, M_k|Z^k, U^k, x_0) dx_1 dx_2 \ldots dx_{k-1} \tag{2.3}$$

Here, we derive the Probabilistic Bayes recursion for the online SLAM problem, which can be easily extended to the full SLAM problem. The problem is to recursively calculate the probability of the vehicle pose x_k and the map state M_k at time k, given the vehicle state, x_{k-1} and the map state M_{k-1} at time $k-1$, and the set of all past control inputs U^k and the set of all past observations taken up to time k, Z^k. The solution to this problem requires appropriate probabilistic models of the vehicle motion, map evolution and the observation process. To derive a recursive two step prediction (time-update) and correction (measurement-update) form, we proceed as follows. The joint distribution of vehicle state, map and current observation, conditioned on the past measurements and history of controls when expanded in terms of the vehicle pose and map state is,

$$\begin{aligned} f_{k|k}(x_k, M_k, Z_k|Z^{k-1}, U^k) &= f_{k|k}(x_k, M_k|Z_k, Z^{k-1}, U^k) f_k(Z_k|Z^{k-1}, U^k) \\ &= f_{k|k}(x_k, M_k|Z^k, U^{k-1}) f_k(Z_k|Z^{k-1}, U^k) \end{aligned} \tag{2.4}$$

and then in terms of the observation,

$$\begin{aligned} f_{k|k}(\mathrm{x}_k, \mathrm{M}_k, \mathrm{Z}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k) &= f_k(\mathrm{Z}_k|\mathrm{x}_k, \mathrm{M}_k, \mathrm{Z}^{k-1}, \mathrm{U}^k) f_{k|k-1}(\mathrm{x}_k, \mathrm{M}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k) \\ &= f_k(\mathrm{Z}_k|\mathrm{x}_k, \mathrm{M}_k) f_{k|k-1}(\mathrm{x}_k, \mathrm{M}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k) \end{aligned} \tag{2.5}$$

Equating eqn. 2.4 and eqn. 2.5 and rearranging gives

$$f_{k|k}(\mathrm{x}_k, \mathrm{M}_k|\mathrm{Z}^k, \mathrm{U}^k) = \frac{f_k(\mathrm{Z}_k|\mathrm{x}_k, \mathrm{M}_k) f_{k|k-1}(\mathrm{x}_k, \mathrm{M}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k)}{f_k(\mathrm{Z}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k)} \tag{2.6}$$

From the total probability theorem the second term in the numerator can be rewritten in terms of the vehicle model and the joint posterior from time-step $k-1$

$$f_{k|k-1}(\mathrm{x}_k, \mathrm{M}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k) = \int f_{k|k-1}(\mathrm{x}_k|\mathrm{x}_{k-1}, \mathrm{u}_k) f_{k-1|k-1}(\mathrm{x}_{k-1}, \mathrm{M}_{k-1}|\mathrm{Z}^{k-1}, \mathrm{U}^{k-1}) d\mathrm{x}_{k-1} \tag{2.7}$$

Substituting eqn. 2.7 into eqn. 2.6 gives

$$f_{k|k}(\mathrm{x}_k, \mathrm{M}_k|\mathrm{Z}^k, \mathrm{U}^k) = \frac{f_k(\mathrm{Z}_k|\mathrm{x}_k, \mathrm{M}_k) \int f_{k|k-1}(\mathrm{x}_k|\mathrm{x}_{k-1}, \mathrm{u}_k) f_{k-1|k-1}(\mathrm{x}_{k-1}, \mathrm{M}_{k-1}|\mathrm{Z}^{k-1}, \mathrm{U}^{k-1}) d\mathrm{x}_{k-1}}{f_k(\mathrm{Z}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k)} \tag{2.8}$$

where,

$$f_k(\mathrm{Z}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k) = \int f_k(\mathrm{Z}_k|\mathrm{x}_k, \mathrm{U}^k) f_{k|k-1}(\mathrm{x}_k|\mathrm{Z}^{k-1}, \mathrm{U}^k) d\mathrm{x}_k \tag{2.9}$$

is the normalization factor which is independent of the map state and the vehicle state.

Equations 2.7 and 2.8 provide a recursive procedure for calculating the joint

posterior $f_{k|k}(x_k, M_k|Z^k, U^k)$, given the initial vehicle position x_0, for the vehicle state x_k and the map M_k at time k based on all the observations Z^k and all control inputs U^k. The recursion is a function of the vehicle motion model $f_{k|k-1}(x_k|x_{k-1}, u_k)$ and the measurement (sensor) model $f_k(Z_k|x_k, M_k)$.

2.2 Estimation Theoretic Approximations

The calculation of the full posterior as presented in the Section 2.1 is infeasible owing to the huge complexity of representing the probability distributions and their update operations. However, estimation theoretic approximations have resulted in closed form SLAM solutions that have been developed by restricting the form of the SLAM posterior, the motion model, and the measurement model. Most present day SLAM solutions originate from a seminal paper by Smith et. al. [101], which proposed the use of the Extended Kalman Filter (EKF) to estimate the SLAM joint posterior. Most of these algorithms are feature based in the sense that they represent maps using sets of features, which are also commonly referred to as landmarks [35].

2.2.1 The EKF-SLAM Filter

The EKF represents the SLAM joint posterior as a multivariate Gaussian distribution parameterized by second order statistics i.e., the mean and the covariance. The mean describes the most likely pose of the vehicle and landmarks, and the covariance encompasses their uncertainties. The SLAM joint posterior of the vehicle pose and the map (landmark) state is represented by an augmented state vector

$$\pi_k = [x_k \quad M_k]^T \tag{2.10}$$

The vehicle motion model $f_{k|k-1}(\mathrm{x}_k|\mathrm{x}_{k-1}, \mathrm{u}_k)$ and the measurement model $f_k(\mathrm{Z}_k|\mathrm{x}_k, \mathrm{M}_k)$ are of the form,

$$\begin{aligned} f_{k|k-1}(\mathrm{x}_k|\mathrm{x}_{k-1}, \mathrm{u}_k) &= \mathcal{N}(\mathrm{x}_k; \mathrm{f}_k(\mathrm{x}_{k-1}, \mathrm{u}_k), \mathrm{Q}_k) \\ f_k(\mathrm{Z}_k|\mathrm{x}_k, \mathrm{M}_k) &= \mathcal{N}(\mathrm{Z}_k; \mathrm{h}_k(\mathrm{x}_k, \mathrm{M}_k), \mathrm{R}_k) \end{aligned} \tag{2.11}$$

where, $\mathrm{f}_k(\cdot)$ models vehicles kinematics, $\mathrm{h}_k(\cdot)$ models measurement process and $N(0, \mathrm{Q}_k)$, $N(0, \mathrm{R}_k)$ refer to zero mean uncorrelated Gaussian process and measurement noise with a variance of Q_k and R_k respectively.

If at time $k-1$, the joint posterior density is a Gaussian of the form,

$$f_{k-1|k-1}(\pi_{k-1}|\mathrm{Z}^{k-1}, \mathrm{U}^{k-1}) = \mathcal{N}(\pi_{k-1}; \hat{\pi}_{k-1}, \mathrm{P}_{k-1}) \tag{2.12}$$

then the time-updated density to time k is also a Gaussian distribution

$$f_{k|k-1}(\pi_k|\mathrm{Z}^{k-1}, \mathrm{U}^k) = \mathcal{N}(\mathrm{x}_k; \hat{\mathrm{x}}_{k|k-1}, \hat{\mathrm{P}}_{vv,k|k-1}) \tag{2.13}$$

where,

$$\hat{\mathrm{x}}_{k|k-1} = \mathrm{f}_k(\hat{\mathrm{x}}_{k-1|k-1}, \mathrm{u}_k) \tag{2.14}$$

$$\hat{\mathrm{P}}_{vv,k|k-1} = \nabla \mathrm{f}_k . \mathrm{P}_{vv,k-1|k-1} . \nabla \mathrm{f}_k^T + \mathrm{Q}_k \tag{2.15}$$

and $\nabla \mathrm{f}_k$ is the Jacobian of f_k as a result of the linearization, evaluated at the vehicle pose estimate $\hat{\mathrm{x}}_{k-1|k-1}$. No time update is performed on the landmarks on the assumption that the landmarks are static in nature.

The data-updated joint density at time k is also approximated as a Gaussian

distribution

$$f_{k|k}(\pi_k|Z^k, U^k) = \mathcal{N}(\pi_k; \hat{\pi}_k, P_k) \tag{2.16}$$

where,

$$\hat{\pi}_k = \hat{\pi}_{k|k-1} + K_k \left[Z_k - h_k(\hat{x}_{k|k-1}, \hat{M}_{kk-1}) \right], \tag{2.17}$$

$$P_{k|k} = [I - K_k \nabla h_k] P_{k|k-1} - K_k, \tag{2.18}$$

$$K_k = P_{k|k-1} \nabla h_k^T S_k^{-1}, \tag{2.19}$$

$$S_k = \nabla h_k P_{k|k-1} \nabla h_k^T + R_k \tag{2.20}$$

and ∇h_k is the Jacobian of h_k evaluated at $\hat{x}_{k|k-1}$ and $\hat{M}_{kk-1}$.

Naive EKF based SLAM algorithms are adversely affected by their vulnerability to inconsistent treatment of nonlinearities, quadratic computational complexity and data association failures. The linearization process in EKF can lead to inconsistencies in the filter when the time step intervals are not sufficiently small. In order to address these limitations, Julier et.al [52] developed the unscented Kalman filter (UKF). Instead of linearizing using Jacobian matrices, the UKF uses a set of discretely sampled points to parameterize the mean and covariance of the estimates. The idea behind the unscented transform is that it is easier to approximate a Gaussian distribution than an arbitrary non-linear function and it was shown that the performance is analytically superior to the EKF [52].

The computational complexity of these KF based SLAM solutions lies in the quadratic nature of the joint covariance matrix $P_{k|k}$ that represents a correlation between individual state variables (with state variables being the vehicle pose and the map (landmark) state). This is attributed to the data-update step where all the

landmarks and the joint covariance matrix are updated every time an observation is made. Naively, this means computation grows quadratically with the number of landmarks $N_{k|k}$. Hence this limits the number of landmarks that can be handled by the naive EKF to only a few hundreds. Over the years, a number of alternatives based on EKF-SLAM algorithms have been proposed to increase the scalability of the SLAM algorithms. Notable works include decoupled stochastic mapping [63], constrained local submap filter (CLSF) [117] where a global map is decomposed into a network of smaller, more manageable submaps. An optimization for the SLAM algorithm was provided in [45], that provides an approximate factorization of the covariance matrix. Thus, these techniques scale to larger environments encompassing thousands of landmarks, however, still incur an $\mathcal{O}(N^2)$ burden and at a cost of slower overall rate of convergence [105].

However, in spite of all these vulnerabilities they are widely used because of their uncomplicated treatment of vehicle and map uncertainties in the estimation process, simple recursion in the solution phase yielding good results in general. For this reason, the EKF is often used to benchmark newly proposed SLAM solutions.

2.3 Sequential Monte-Carlo Approximations

Both EKF and UKF estimation theoretic frameworks mainly rely on the use of approximations to ensure mathematical tractability and computational efficiency. These come at the cost of restricted representational power, where only unimodal distributions are represented and that too only when the uncertainty and nonlinearity in the system is not very high. However non-parametric stochastic filters based on Sequential Monte Carlo (SMC) techniques have proved to be successful

in approximating non-unimodal posterior distributions. They are also commonly referred to as particle filters or bootstrap filters [33]. As the number of samples N_s becomes very large (i.e., $N_s \to \infty$), this Monte Carlo approximation approaches the true posterior density [33].

The principle behind this approach is to approximate the joint posterior distribution function (eqn. 2.8) by a set of random samples and to compute point mass estimates based on these samples and their associated weights. The weights are chosen based on the idea of importance sampling. The idea of using importance sampling to recursively construct point mass approximations to the posterior density, known as sequential importance sampling (SIS) filter or particle filter is as follows:

At time $k-1$, the joint posterior density $f_{k-1|k-1}(\cdot)$ is approximated by a set of weighted particles $\{w_{k-1}^{[i]}, \pi_{k-1}^{[i]}\}_{i=1}^{N_s}$ i.e.,

$$f_{k-1|k-1}(\pi_{k-1}|\mathbf{Z}^{k-1}, \mathbf{U}^{k-1}) \approx \sum_{i=1}^{N_s} w_{k-1}^{[i]} \delta(\pi_{k-1} - \pi_{k-1}^{[i]}) \tag{2.21}$$

Given a proposal distribution (importance) function $q_k(\cdot|\pi_{k-1}^{[i]}, Z_k)$, the particle filter approximates the posterior at time k by a new set of weighted particles $\{w_k^{[i]}, \pi_k^{[i]}\}_{i=1}^{N_s}$ i.e.,

$$f_{k|k}(\pi_k|\mathbf{Z}^k, \mathbf{U}^k) \approx \sum_{i=1}^{N_s} w_k^{[i]} \delta(\pi_k - \pi_k^{[i]}) \tag{2.22}$$

where,

$$\pi_k^{[i]} \sim q_k(\pi_k|\pi_{k-1}^{[i]}, Z_k) \tag{2.23}$$

$$w_k^{[i]} = w_{k-1}^{[i]} \frac{f_k(Z_k|\pi_k^{[i]}) f_{k|k-1}(\pi_k^{[i]}|\pi_{k-1}^{[i]}, u_k)}{q_k(\pi_k|\pi_{k-1}^{[i]}, Z_k)} \tag{2.24}$$

$$\tilde{w}_k^{[i]} = w_k^{[i]} / \sum_{i=1}^{N_s} w_k^{[i]} \tag{2.25}$$

The basic SIS algorithm described above is well known to be subjected to degeneracy phenomenon, where after a few iterations, all but one particle will have negligible weight. In [33], it has also been shown that the variance of the importance weight increases with time, thereby making it difficult to avoid the degeneracy phenomenon. However, this phenomenon is usually overcome with a proper choice of proposal distribution function and the use of re-sampling techniques. Several re-sampling techniques have been developed as discussed in [33]. In this dissertation, systematic re-sampling is adopted for its computational efficiency and effectiveness in minimizing the sample variation and is used wherever required for numerical studies [11].

The key advantage of SMC approximation discussed is their ability to represent arbitrary probability distributions. This makes them a viable alternative when faced with problems to which the Kalman filters and their variants are not well suited. Compared to other multi-modal approaches, particle filters are very efficient since they automatically focus their resources (particles) on the regions in the state space that have high probability. However, a straightforward implementation of the particle SLAM filter described above would be computationally impractical in large environments due to large number of variables involved in describing a map, as the

number of dimensions of the estimation problem scale exponentially.

Despite this, the use of Rao-Blackwellization particle filters within the probabilistic SLAM framework exploiting the structure of the problem marked a conceptual shift in the design strategies involving computational burden reduction.

2.3.1 Rao Blackwellized Particle Filter

The key idea of Rao-Blackwell theorem is to make use of the "structural information" inherent in the problem itself to analytically infer part of state parameters conditional upon other state components which will be estimated by the Sequential Monte Carlo approaches. The process of transforming the particle filter using the Rao-Blackwell theorem is called Rao-Blackwellization process and the filter as Rao-Blackwellization particle filter (RBPF). Thus, RBPF use particles to represent the posterior over some variables along with a parametric probability distribution function to represent all the other variables.

Influenced by an earlier occupancy grid mapping experiments using RBPFs by Murphy [82], Montemerlo et al. [77] exploited the Rao-Blackwellization strategy to represent the vehicle path by a set of particles and features (landmarks) by Gaussian distribution that led to formulating the SLAM problem in an efficient manner. They exploited the fact that given the knowledge of the path of the vehicle, the individual landmark locations become conditionally independent i.e.,

$$f_{k|k}(\mathrm{X}^k, \mathrm{M}_k|\mathrm{Z}^k, \mathrm{U}^k) = f_{k|k}(\mathrm{X}^k|\mathrm{Z}^k, \mathrm{U}^k) f_k(\mathrm{M}_k|\mathrm{X}^k, \mathrm{Z}^k, \mathrm{U}^k) \tag{2.26}$$

The resulting approach known as the FastSLAM uses particle filters to estimate the vehicle's path and an EKF for each landmark. Compared to the quadratic complex-

ity of the naive EKF-SLAM filter, FastSLAM has an attractive linear complexity of $\mathcal{O}(NP)$ where N is the number of features in the map and P is the number of particles used in the process. Further, efficient implementations based on structured trees reduced the computational time of FastSLAM to $\mathcal{O}(P \log N)$ [77].

2.4 Properties of Classical Vector based SLAM Filters

All the classical vector state-space based SLAM solutions (EKF-SLAM, FastSLAM and its variants) discussed above require that we solve certain parallel problems in each recursion. These include effective solutions to the problems of data association, feature extraction, clutter filtering and landmark or map management. Here, the measurement model is represented as a non-linear functional mapping $g_k(\cdot)$ of the vehicle pose vector x_k and the map state vector $[\mathrm{m}_{k,1}, \ldots, \mathrm{m}_{k,n_k}]$ into the measurement vector $[\mathrm{z}_{k,1}, \ldots, \mathrm{z}_{k,n_k}]$

$$[\mathrm{z}_{k,1}, \ldots, \mathrm{z}_{k,n_k}]^T = g_k(\mathrm{x}_k, [\mathrm{m}_{k,1}, \ldots, \mathrm{m}_{k,n_k}]^T) + w(k) \tag{2.27}$$

where $w(k)$ is the additive measurement noise vector. If all the measurements are generated from a single sensor with measurement likelihood function $f_k(\mathrm{z}|m, \mathrm{x}_k)$, and the measurements are independent, then the measurement likelihood function of the measurement model (eqn. 2.27), can be expressed as,

$$f_k(\mathrm{z}_{k,1}, \ldots, \mathrm{z}_{k,n_k}|\mathrm{x}_k, \mathrm{m}_{k,1}, \ldots, \mathrm{m}_{k,n_k}) = f_k(\mathrm{z}_{k,1}|\mathrm{m}_{k,1}, \mathrm{x}_k), \ldots, f_k(\mathrm{z}_{k,n_k}|\mathrm{m}_{k,n_k}, \mathrm{x}_k) \tag{2.28}$$

It is evident from the classical vector based multi-landmark map measurement model (eqn. 2.27) used in conventional SLAM that the effects due to detection

uncertainty of sensor, filtering of the raw sensor measurements to remove clutter, and association of the filtered measurements with landmarks in the map need to be handled before performing a SLAM measurement update. As a consequence, in the conventional approach it is necessary to address the issues arising from detection uncertainty (missed detections), clutter or false alarms (measurements not due to landmarks) and data association (landmark to measurement association) by other explicit means before a measurement update is executed.

2.4.1 Data Association

Before incorporating data into the map, new measurements are associated with existing map landmarks. Thus data association has always been a critical issue for any practical SLAM implementations and most of them perform it using a statistical validation gating (a nearest neighbor gating), a technique inherited from the target-tracking literature [18].

The standard nearest neighbor technique uses a statistical discriminator based on the squared normalized measurement innovation. This approach does not consider temporal and spatial correlations among landmarks and measurements in arriving at the decision. Data association inaccuracies give rise to large errors in the estimated states. The impact of data association errors is severe in the case of naive EKF-SLAM filter where only a single association hypothesis is maintained and the miss associations invariably induce divergence into the map estimates and further leading to the failure of vehicular localization. Given the vulnerability of individual gating techniques in densely populated natural environments and noisy sensor measurements, batch gating techniques were introduced, where multiple as-

sociations are considered simultaneously. The concept of batch gating is based on the idea that if multiple measurements are gathered per control, the data association ambiguity is partially caused by vehicle pose uncertainty, and hence the data associations of multiple measurements per control are correlated. The two popular forms of batch gating are the combined constraint data association (CCDA) [14], based on graph search and the joint compatibility branch and bound (JCBB) [86] method based on tree-search. These methods consider the spatial correlations existing among the landmarks of the SLAM state vector into consideration in decision making and therefore are more robust to increased vehicle uncertainty and measurement noise. Nevertheless, all these data association algorithms cannot effectively filter out spurious measurements (clutter) and returns from moving objects.

The data association algorithms discussed so far, have only used a single association hypothesis as an input to the SLAM filter. However in cluttered and dense environments, to overcome the data association ambiguity, multiple association hypothesis are needed for robust tracking of landmarks. One such approach that is commonly used in the target tracking literature is multiple hypothesis tracking (MHT) [91]. It resolves the data association ambiguities by deferred decision making in conflicting situations where in, it forms a tree of all possible association hypotheses, which are propagated in time with a belief that new information is likely to resolve the conflicts. Thus MHT is capable of automatically dealing with clutter or spurious measurements and missed detections. While MHT has been applied to a wide array of target tracking problems, maintaining separate map estimates for each association hypotheses in the SLAM context is computationally cumbersome. Thus, the major drawback of MHT is that the hypothesis tree grows exponentially in time requiring exponential memory and computational resources

making it impracticable for real-time implementation.

2.4.2 Clutter Rejection & Map Management

Making classical vector state space based SLAM solutions robust in practice often requires certain additional steps to be performed in the form of clutter filtering and map management. Many map management techniques have been devised to deal with spurious measurements and outliers in the measurements. The most common practice to reject outliers is to maintain a tentative landmark list. Instead of augmenting the map by a new landmark once a measurement indicates the existence of a new landmark, it is first added to a tentative list of landmarks. Once a landmark has been consistently observed, it is transferred from the tentative list to the permanent map [32]. Another common map management technique, where the landmark existence probability is implemented as a log-odds ratio [36]. Whenever, the $i-$th landmark m_i is observed, the log odds for its existence o_i is incremented by a fixed value or else its value is decremented. If the log odds exceeds a predefined threshold τ_{lo}, the landmark is incorporated into the map.

Thus in practice, these techniques tend to reduce the number of landmarks in the map by a significant factor, reducing spurious measurements. However, it is important to emphasize that the tentative list to be maintained, (i.e., the number of times the landmarks have to be observed before being incorporated into the map) is an *ad hoc*, problem dependent parameter, that needs to be fine tuned based on environmental conditions.

2.5 Chapter Summary

Based on the discussion and careful examination of related work in SLAM reveals that the need to solve the parallel problems such as clutter rejection, data association and map management outside the Bayes recursion primarily stems from the limitations of the representation of the map model, and map measurement model used. This thesis is motivated to directly address these limitations. The approach presented in this thesis is inspired by recent developments in the target-tracking community where random finite sets have been perceived as a more natural representation of the multisource-multitarget problems [70]. It is formulated based on an exhaustive mathematical formulation presented by Mahler based on finite set statistics (FISST) [69]. By adopting a random finite set approach, we can construct a natural multi-landmark (map) measurement model that explicitly accounts for detection uncertainty, clutter and unknown data association and a multi-landmark map evolution model that accounts for the new landmarks appearing in the field of view of sensor for autonomous vehicle navigation.

3

Random Finite Set Modeling: An Autonomous Vehicle Navigation Perspective

In this chapter, models are constructed for estimating unknown varying number of landmarks/features from a moving vehicle using a random finite set theoretic approach. This approach is chosen, as random finite sets model the problem in an extremely natural and very flexible fashion. The estimated multiple landmark state M_k can be naturally represented as a random finite set that consists of unknown varying number of landmarks based on the sensors field of view. Similarly, for the sensors, the number of measurements Z_k that are obtained at any time instance is varying and can be naturally considered as random. The models derived in this chapter cater to a generic range-bearing (GRB) sensor (e.g., Sonar, Radar, Laser). However, these models can be easily generalized to accommodate any type of sensor. A brief background on finite set statistical (FISST) calculus is provided in

Appendix A, that would assist us in developing these models.

This chapter is organized as follows. The following section provides a taxonomy on the feature tracking literature outlining the general problem of tracking multiple unknown number of features. A dynamic map evolution model that caters for limited FOV of the GRB sensor mounted on a moving autonomous vehicle is developed in Section 3.2, based on FISST calculus [68]. In Section 3.3, a random finite set measurement model is formulated. Based on these models, a recursive optimal multi-landmark Bayes filter is then presented in Section 3.4. Due to combinatorial computational complexity associated with the optimal Bayes filter, its approximations in form of its first order statistical moment filter (PHD filter) have been by developed in [69], which has been incorporated within the autonomous vehicle navigation framework in Section 3.5. Section 3.6 concludes, highlighting the important properties of the models developed.

3.1 Taxonomy of Multiple Feature Tracking

The multiple feature tracking (MFT) involves automatic tracking of unknown, independent point features from a single sensor. A point feature is modeled as having neither dimensions nor having any resolvable entities. It exists purely in a dynamic state space usually consisting of either position if it is static, or position, velocity and acceleration if dynamic. The basic MFT problem concerns the estimation of the feature states given noisy measurements from a sensor which is mounted on a mobile platform, whose field of view includes the features of interest. In real world applications, spurious measurements called false alarms or clutter are also present, which generally create ambiguity in the origin of the measurements. This problem

is compounded by the presence of multiple features and the detection uncertainty in real world sensors. All the factors combined to produce data association ambiguity, which is largely to blame for the complexity of the MFT problem [22], [18], [90]. Figure 3.1 illustrates the sequence of steps that are generally involved in tracking multiple features in an environment. The basic approaches to the MFT problem can

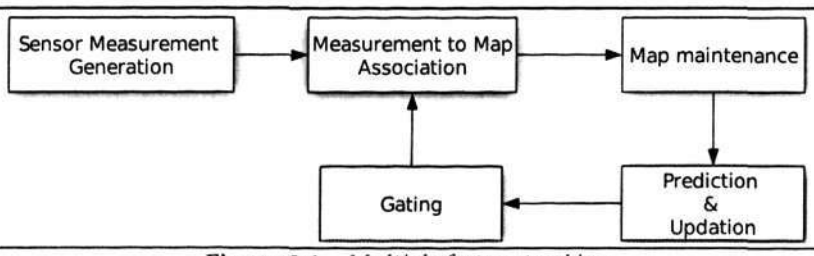


Figure 3.1 *Multiple feature tracking*

be broadly classified based on the manner in which they arrive at data association decisions (see [18], [22]). Algorithms that use only one scan of measurements at a time to arrive at data association decisions are known as single scan methods or hard decision making methods. Association algorithms that use several scans of measurements received at several time steps in the process of decision making are known as deferred decision logic methods. The simplest and most commonly used hard decision technique is the standard nearest-neighbor (SNN). It allows only the measurement closest in statistical distance to the predicted track to be used to update the feature state. Another Bayesian approach that provides a good solution to the clutter or spurious measurements and ambiguities of the SNN approach is the joint probabilistic data association (JPDA), where all the measurements that are close to the predicted track are used in the Bayesian update. The overall Gaussian mixture probability density function is then condensed into a single Gaussian by retaining its mean and covariance. Both SNN and JPDA are recursive single-feature

tracking algorithms. Global nearest-neighbor or 2D assignment algorithm is a so called 'true MFT' that accounts for distances between all the tracks and the measurements and propagates the single most likely hypothesis in each scan. Among the deferred decision logic techniques, the most popular is the multiple hypothesis tracking (MHT) in which alternative data association hypotheses are formed whenever measurement-to-feature conflict occurs. Then, rather than combining the hypotheses as in the JPDA method, the hypotheses are propagated in anticipation that subsequent data will resolve the uncertainty.

In most of the real-life scenarios, the number of features to be estimated and tracked are generally unknown. There are two basic approaches to deal with such problems. The first one being the most common, is the "algorithmic" approach, where an algorithm, which is a combination of Bayes theorem and an *ad hoc* logic (heuristics) that somehow caters for the varying numbers of features is employed. Once the number of features are known, standard MFT algorithms discussed earlier can be used conditioned on the number of features determined by heuristics. While the second approach is based on development of rigorous models for varying number of features [24], [83], [68]. Thus, true multiple feature tracking is the joint estimation of the unknown varying number of features along with their states and measurement-to-feature association. The following sections describe the construction of the mathematical models based on the random finite set framework for estimating unknown number of landmarks and their states for the purpose of autonomous vehicle navigation.

3.2 Formulation of Random Finite Set Map Evolution Model

This section discusses the process of constructing a dynamic map motion (evolution) model using a random finite set framework. Due to the limited field of view (FOV) of the perceptual sensors and occlusions, the landmark map will not remain static over time. The motion of the vehicle causes new landmarks to appear and some landmarks to disappear in the limited FOV of the sensor, resulting in the number of landmarks in the map to change randomly over time. A natural and an effective representation of such a behavior is using a random finite set (RFS), M_k,

$$\mathrm{M}_k = \{\mathrm{m}_{k,1}, \ldots, \mathrm{m}_{k,i}, \ldots, \mathrm{m}_{k,n_k}\} \in \mathcal{F}(\mathcal{M}), \tag{3.1}$$

where $\mathcal{F}(\mathcal{M})$ is a collection of all finite subsets of the space of maps $\mathcal{M}$ and $n_k = |M_k|$ is the number of elements or the cardinality of M_k representing the number of landmarks detected and observed at time k and $\mathrm{m}_{k,i}$ representing the state of the i^{th} landmark at time k. The dynamic model for unknown and varying number of landmarks is constructed using Markov transition model given by $f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1}, \mathrm{x}_k)$, where x_k is the state of the vehicle at time k as explained in Chapter 2.

The finite set dynamic map model that accounts for evolving persistent landmarks and appearing/disappearing of landmarks from the FOV of the sensor (due to vehicle motion) is constructed as follows.

- Let $\mathrm{m}_{k-1} \in \mathrm{M}_{k-1}$[1] represent an individual landmark state at time $k-1$. Now this landmark may or may not appear in the sensors view in the next time

[1] $\mathrm{m}_{k-1} = \mathrm{m}_{k-1,i}$, where $i = 1, \ldots, n_{k-1}$. Subscript $_i$ is dropped for notational simplicity

instant k, due to the limited FOV of the sensor and vehicle motion. The landmark $\mathrm{m}_{k-1} \in \mathrm{M}_{k-1}$ will appear in the next time instant k with a probability $p_S = p_{S,k}(\mathrm{m}_{k-1}, \mathrm{x}_k)$ or disappear with a probability $1 - p_{S,k}(\mathrm{m}_{k-1}, \mathrm{x}_k)$. For a landmark that is currently (time $k-1$) in view this probability represents its probability of persistence or survival, and for a landmark not in the current field of view, it represents the probability of reappearance.

- Conditional on the persistence of the landmarks at time k, the probability density of a transition from state m_{k-1} to m_k is given by a Markov process, $f_{k|k-1}(\mathrm{m}_k|\mathrm{m}_{k-1}, \mathrm{x}_k)$ (Single landmark motion model). For a landmark $\mathrm{m}_{k-1} \in \mathrm{M}_{k-1}$ at time $k-1$, its state at the next time instant k is modeled as an RFS $S_{k|k-1}(\mathrm{m}_{k-1})$. Now $S_{k|k-1}(\mathrm{m}_{k-1})$ will take on either $\{\mathrm{m}_k\}$ with probability $p_S(\mathrm{m}_{k-1}, \mathrm{x}_k)$ or $\emptyset$ with probability $1 - p_S(\mathrm{m}_{k-1}, \mathrm{x}_k)$.

- The limited FOV sensor coupled with vehicle motion also results in new landmarks coming into view at time k. Let us denote these new appearances by an RFS Γ_k.

Thus, the evolution of the map of landmarks is given by the union of persistent landmarks and the new landmarks in the FOV:

$$\mathrm{M}_k = \left[\bigcup_{\mathrm{m}_{k-1} \in \mathrm{M}_{k-1}} S_{k|k-1}(\mathrm{m}_{k-1}) \right] \cup \Gamma_k \tag{3.2}$$

The dynamic models of sets Γ_k and S_k and are constructed using case studies 1 and 2 respectively. The density of the entire dynamic map model is constructed with the assistance of case study 3.

Case 1. *No landmarks from state* M_{k-1} *evolve to state* M_k. *New landmarks appear in the FOV.*

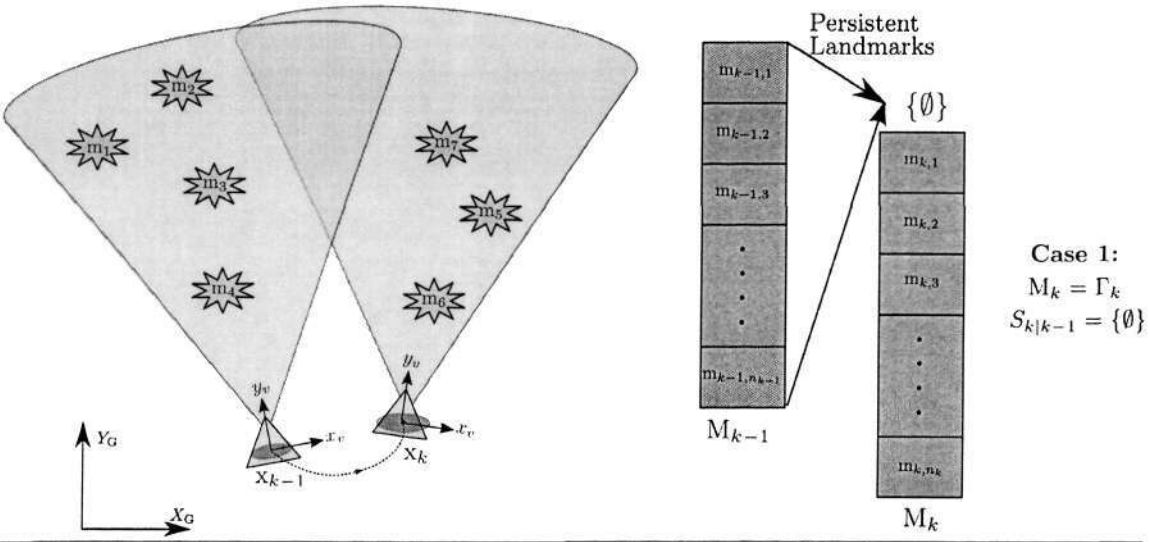


Figure 3.2 *Illustration of the new landmarks* (m_5, m_6, m_7, *in this case) that appear in the FOV of the sensor when vehicle (yellow triangle) moves from state* x_{k-1} *to* x_k. *The landmarks* ($m_1 \ldots m_4$, *in this case) that are in FOV of the sensor at time* $k-1$, *are not re-observed by the sensor at time k due to vehicle motion.* X_G, Y_G *represent some global coordinate frame.*

In this case, the dynamic map model in eqn. 3.2 evolves to

$$\mathrm{M}_k = \Gamma_k \tag{3.3}$$

The appearance of landmarks into the FOV of the sensor is assumed to be Poisson distributed in time with expected value μ and distributed in the FOV according to an arbitrary density $b(\mathrm{m}_k)$. Hence,

$$\Gamma_k = \{\Gamma_{k,1}, \ldots, \Gamma_{k,B}\} \tag{3.4}$$

where, $B = |\Gamma_k|$ is a random non-negative integer with probability distribution

$$p_B(\tau) = \frac{e^{-\mu}\mu^{\tau}}{\tau!} \tag{3.5}$$

and where, conditioned on $B = \tau$, $\Gamma_{k,1}, \ldots, \Gamma_{k,\tau}$ are independent and identically distributed random vectors with probability density $b(\mathrm{m}_k) = f_{\Gamma_k}(\mathrm{m}_k)$.

In order to determine the probability density function, a belief mass function of the model (eqn. 3.3) is constructed (For details on belief mass function, refer to Appendix A, Section A.3.). Let T be the sensor FOV. The probability that Γ_k will be in T (given that $\mathrm{S}_{k|k-1} = \emptyset$), is denoted by $\beta_{k|k-1}(T|\emptyset)$ and given by,

$$\begin{aligned}
\beta_{k|k-1}(T|\emptyset) &= P(\Gamma_k \subseteq T|\emptyset) \\
&= \sum_{b=0}^{\infty} P(\Gamma_k \subseteq T, |\Gamma_k| = \tau|\emptyset) \\
&= e^{-\mu} \sum_{b=0}^{\infty} \frac{\mu^{\tau}}{\tau!} P(\Gamma_k \in T)^{\tau} \\
&= e^{-\mu} \sum_{b=0}^{\infty} \frac{\mu^{\tau}}{\tau!} p_{\Gamma}(T)^{\tau} \\
&= e^{\mu(p_{\Gamma}(T)-1)}
\end{aligned} \tag{3.6}$$

Performing a set derivative on the belief mass function and setting $T = \emptyset$ (See Appendix A, Sections A.3 and A.5 for more details), the probability density function is obtained as follows,

$$\begin{aligned}
f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1}, \mathrm{x}_k) = f_{\Gamma_k}(\mathrm{M}_k) &= \frac{\delta}{\delta \mathrm{M}_k} \beta_{k|k-1}(\emptyset|\emptyset) \\
&= e^{-\mu} \prod_{\mathrm{m}_k \in \mathrm{M}_k} \mu.b(\mathrm{m}_k)
\end{aligned} \tag{3.7}$$

Case 2. *Some landmarks evolve from state* M_{k-1} *to* M_k*, while others fall out of the FOV of the sensor*

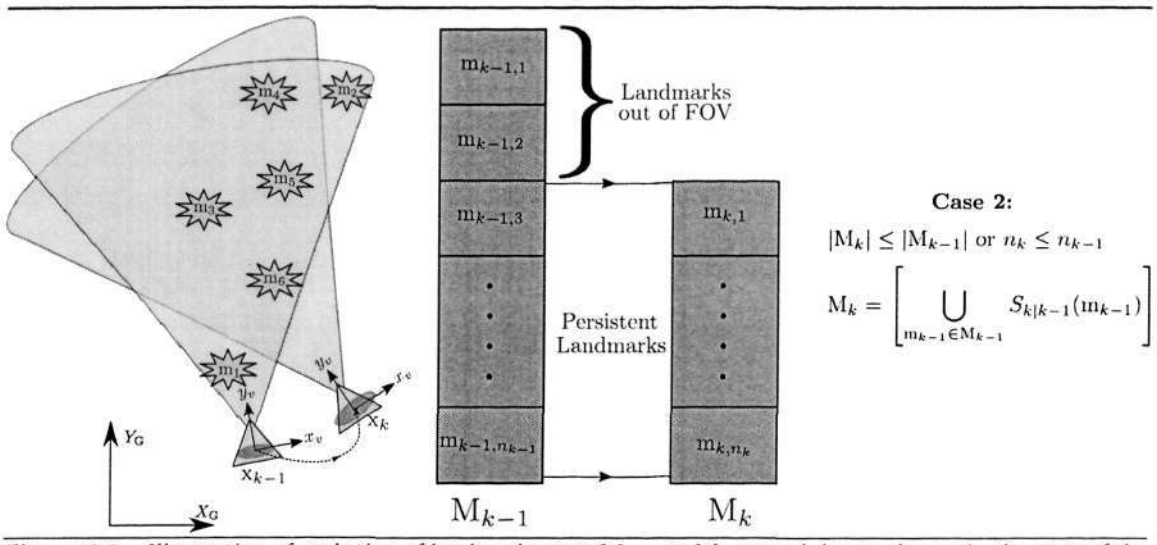


Figure 3.3 *Illustration of evolution of landmarks state* M_{k-1} *to* M_k *caused due to change in the state of the vehicle from* x_{k-1} *to* x_k*. Some of the landmarks evolve* (m_3, m_4, m_5, m_6*, in this case) while others fall out* (m_1, m_2*, in this case) of the FOV of the sensor.*

In this case, the dynamic map model in eqn. 3.2 has the form

$$\mathrm{M}_k = \left[\bigcup_{\mathrm{m}_{k-1} \in \mathrm{M}_{k-1}} S_{k|k-1}(\mathrm{m}_{k-1}) \right] \tag{3.8}$$

where, the landmark $\mathrm{m}_{k-1} \in \mathrm{M}_{k-1}$ will appear in the next time instant k with a probability $p_S = p_{S,k}(\mathrm{m}_{k-1}, \mathrm{x}_k)$ or disappear with a probability $1 - p_{S,k}(\mathrm{m}_{k-1}, \mathrm{x}_k)$. For a landmark that is currently (time $k-1$) in view this probability represents its probability of persistence or survival, and for a landmark not in the current field of view, it represents the probability of reappearance. Conditional on the persistence of the landmarks at time k, the probability density of a transition from state m_{k-1} to m_k is given by a Markov process, $f_{k|k-1}(\mathrm{m}_k|\mathrm{m}_{k-1}, \mathrm{x}_k)$ (Single landmark motion

model).

In order to derive for the probability density function for the above case, a belief mass function of the model (eqn. 3.8) is constructed as follows,

$$\begin{aligned}
\beta_{k|k-1}(T|\mathrm{M}_{k-1}) &= P(\mathrm{S}_{k|k-1} \subseteq T|\mathrm{M}_{k-1}, \mathrm{x}_k) \\
&= P(\mathrm{m}_{k,1} \in T, \ldots, \mathrm{m}_{k,n_k} \in T|\mathrm{M}_{k-1}, \mathrm{x}_k) \\
&= P(\mathrm{m}_{k,1} \in T|\mathrm{m}_{k-1,1}, \mathrm{x}_k), \ldots, P(\mathrm{m}_{k,n_k} \in T|\mathrm{m}_{k-1,n_{k-1}}, \mathrm{x}_k) \\
&= \int_T f_{k|k-1}(\mathrm{m}|\mathrm{m}_{k-1,1}, \mathrm{x}_k)d\mathrm{m}, \ldots, \int_T f_{k|k-1}(\mathrm{m}|\mathrm{m}_{k-1,n_{k-1}}, \mathrm{x}_k)d\mathrm{m} \\
&= p_1(T), \ldots, p_{n_{k-1}}(T)
\end{aligned} \tag{3.9}$$

Based on convolution formula (refer Appendix A, SectionA.6 for more details),

$$\frac{\delta\beta_{k|k-1}}{\delta\mathrm{M}_k}(T|\mathrm{M}_{k-1}, \mathrm{x}_k) = \sum_{W_1 \uplus \ldots \uplus W_{n_{k-1}} = \mathrm{M}_{k-1}} \frac{\delta p_1}{\delta W_1}(T) \ldots \frac{\delta p_{n_{k-1}}}{\delta W_{n_{k-1}}}(T) \tag{3.10}$$

Setting $T = \emptyset$,

$$\begin{aligned}
f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1}, \mathrm{x}_k) &= \frac{\delta\beta_{k|k-1}}{\delta\mathrm{M}_k}(\emptyset|\mathrm{M}_{k-1}, \mathrm{x}_k) \\
&= \sum_{W_1 \uplus \ldots \uplus W_{n_{k-1}} = \mathrm{M}_{k-1}} \frac{\delta p_1}{\delta W_1}(\emptyset) \ldots \frac{\delta p_{n_{k-1}}}{\delta W_{n_{k-1}}}(\emptyset)
\end{aligned} \tag{3.11}$$

However,

$$\frac{\delta p_i}{\delta \emptyset}(\emptyset) = p_i(\emptyset) = 1 - p_S(\mathrm{m}_{k-1,i}) \tag{3.12}$$

$$\frac{\delta p_i}{\delta \mathrm{m}_k}(\emptyset) = p_S(\mathrm{m}_{k-1,i}) f_{k|k-1}(\mathrm{m}_k|\mathrm{m}_{k-1,i}) \tag{3.13}$$

$$\frac{\delta p_i}{\delta \mathrm{M}_k}(\emptyset) = 0 \qquad \text{if } |\mathrm{M}_k| > 1 \tag{3.14}$$

$$\tag{3.15}$$

The only surviving terms in the eqn. 3.11 are those for which $|W_1| \leq 1, \ldots, |W_{n_{k-1}}| \leq 1$. If $\mathrm{M}_k = \emptyset$ then,

$$\begin{aligned} f_{k|k-1}(\emptyset|\mathrm{M}_{k-1}, \mathrm{x}_k) &= \frac{\delta p_1}{\delta \emptyset}(\emptyset) \ldots \frac{\delta p_{n_{k-1}}}{\delta \emptyset}(\emptyset) \\ &= \prod_{i=1}^{n_{k-1}} (1 - p_S(\mathrm{m}_{k-1,i})) \end{aligned} \tag{3.16}$$

If $|\mathrm{M}_k| = n_k \leq n_{k-1}$ then

$$\begin{aligned} f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1}, \mathrm{x}_k) &= f_{k|k-1}(\emptyset|\mathrm{M}_{k-1}, \mathrm{x}_k) \sum_{W_1 \uplus \ldots \uplus W_{n_k}{}_1 = \mathrm{M}_k} \frac{\frac{\delta p_1}{\delta W_1}(\emptyset) \ldots \frac{\delta p_{n_{k-1}}}{\delta W_{n_{k-1}}}(\emptyset)}{\prod_{i=1}^{n_{k-1}} (1 - p_S(\mathrm{m}_{k-1,i}))} \\ &= f_{k|k-1}(\emptyset|\mathrm{M}_{k-1}, \mathrm{x}_k) \\ &\cdot \sum_{1 \leq i_1 \neq \ldots \neq i_{n_k} \leq n_{k-1}} \frac{\frac{\delta p_{i_1}}{\delta \mathrm{m}_{k,1}}(\emptyset) \ldots \frac{\delta p_{i_{n_k}}}{\delta \mathrm{m}_{k,n_k}}(\emptyset)}{(1 - p_S(\mathrm{m}_{k-1,i_1})) \cdots (1 - p_S(\mathrm{m}_{k-1,i_{n_k}}))} \\ &= f_{k|k-1}(\emptyset|\mathrm{M}_{k-1}, \mathrm{x}_k) \\ &\cdot \sum_{1 \leq i_1 \neq \ldots \neq i_{n_k} \leq n_{k-1}} \prod_{j=1}^{n_k} \frac{p_S(\mathrm{m}_{k-1,i_j}) . f_{k|k-1}(\mathrm{m}_{k,j}|\mathrm{m}_{k-1,i_j}, \mathrm{x}_k)}{1 - p_S(\mathrm{m}_{k-1,i_j})} \end{aligned} \tag{3.17}$$

For each n_k-tuple define, $\pi : \{1, \ldots, n_{k-1}\} \longmapsto \{0, 1, \ldots, n_k\}$, such that n_k-tuples are in one to one correspondence with the associations π, then,

$$f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1}, \mathrm{x}_k) = f_{k|k-1}(\varnothing|\mathrm{M}_{k-1}, \mathrm{x}_k) \sum_{\pi} \prod_{i:\pi(i)>0} \frac{p_S(\mathrm{m}_{k-1,i}).f_{k|k-1}(\mathrm{m}_{k,\pi(i)}|\mathrm{m}_{k-1,i}, \mathrm{x}_k)}{1 - p_S(\mathrm{m}_{k-1,i})} \tag{3.18}$$

Case 3. *Some landmarks evolve from state M_{k-1} to M_k, while some fall out from the FOV and some new landmarks appear in the FOV of the sensor*

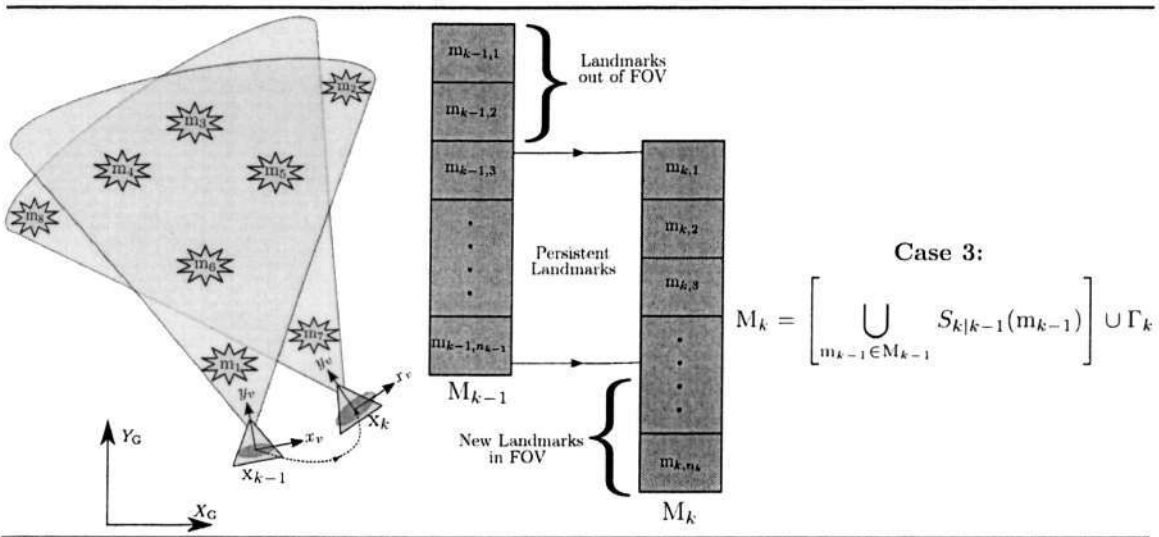


Figure 3.4 *Illustration of the evolution of landmarks from state M_{k-1} to M_k caused due to change in the state of the vehicle from x_{k-1} to x_k. Some landmarks evolve ($m_3, \ldots m_6$ in this case), while some fall out from the FOV (m_1, m_2, in this case) and some new landmarks (m_7, m_8, in this case) appear in the FOV of the sensor.*

In this case it is the complete dynamic map model as in eqn. 3.2

$$\mathrm{M}_k = \left[\bigcup_{\mathrm{m}_{k-1} \in \mathrm{M}_{k-1}} S_{k|k-1}(\mathrm{m}_{k-1})\right] \cup \Gamma_\mathrm{k} \tag{3.19}$$

In order to derive for the probability density function, as in the previous cases, a

belief mass function of the model (eqn. 3.8) is constructed. Due to the independence of disappearance the landmarks from the FOV, and appearance of new landmarks in the FOV, one can write the conditional belief as follows.

$$\begin{aligned}\beta_{k|k-1}(T|\mathrm{M}_{k-1},\mathrm{x}_k) &= P(\mathrm{S}_{k|k-1}\cup\Gamma_k\subseteq T|\mathrm{M}_{k-1},\mathrm{x}_k)\\ &= P(\Gamma_k\subseteq T)P(\mathrm{S}_{k|k-1}|\mathrm{M}_{k-1},\mathrm{x}_k))\\ &= \beta_{\Gamma_k}(T)\beta_{S_k}(T|\mathrm{M}_{k-1},\mathrm{x}_k)\end{aligned} \tag{3.20}$$

From the FISST product rule (Appendix A, Section A.6 for more details),

$$\frac{\delta\beta_{k|k-1}}{\delta\mathrm{M}_k}(T|\mathrm{M}_{k-1},\mathrm{x}_k)=\sum_{W\subseteq\mathrm{M}_k}\frac{\delta\beta_{S_k}}{\delta W}(T|\mathrm{M}_{k-1},\mathrm{x}_k).\frac{\delta\Gamma_k}{\delta(\mathrm{M}_k-W)}(T) \tag{3.21}$$

Setting $T=\emptyset$, eqn. 3.21 becomes

$$\begin{aligned}f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1},\mathrm{x}_k) &= \frac{\delta\beta_{k|k-1}}{\delta\mathrm{M}_k}(\emptyset|\mathrm{M}_{k-1},\mathrm{x}_k)\\ &= \sum_{W\subseteq\mathrm{M}_k} f_{S_k}(W).f_{\Gamma_k}(\mathrm{M}_k-W)\end{aligned} \tag{3.22}$$

where, $f_{\Gamma_k}(\mathrm{M}_k)$ and $f_{S_k}(\mathrm{M}_k)$ are the probability density functions of new landmark appearance Γ_k (as derived in Case 1) and landmark persistence $\mathrm{S}_{k|k-1}$ (as derived in Case 2) respectively. Thus applying the results from Case 1 and Case 2, eqn. 3.22 becomes,

$$\begin{aligned}f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1},\mathrm{x}_k) &= e^{\mu}f_{\Gamma_k}(\mathrm{M}_k).f_{k|k-1}(\emptyset|\mathrm{M}_{k-1},\mathrm{x}_k)\\ &\cdot\sum_{\pi}\prod_{i:\pi(i)>0}\frac{p_S(\mathrm{m}_{k-1,i}).f_{k|k-1}(\mathrm{m}_{k,\pi(i)}|\mathrm{m}_{k-1,i},\mathrm{x}_k)}{(1-p_S(\mathrm{m}_{k-1,i})).\mu b(\mathrm{m}_{k,\pi(i)})}\end{aligned} \tag{3.23}$$

where the summation in eqn. 3.23 is taken over all association hypotheses π :

$\{1,\ldots,n_{k-1}\} \longmapsto \{0,1,\ldots,n_k\}$. For every $i = 1,\ldots,n_{k-1}$, if $\pi(i) > 0$ then the landmark state at time k, $\mathrm{m}_{k,\pi(i)}$ is uniquely associated with landmark at time $k-1$, $\mathrm{m}_{k-1,i}$. However, if $\pi(i) = 0$ then no landmark at time k is associated with $\mathrm{m}_{k-1,i}$ indicating that the landmark $\mathrm{m}_{k-1,i}$ is out of the FOV of the sensor at time k. If $\pi(i) = 0, \forall i$ then every landmark $\mathrm{m}_k \in \mathrm{M}_k$ is a result of a new landmark in the FOV, as in Case 1. If $\pi(i) > 0, \forall i$ then every landmark from time $k-1$ evolved into a new state at time k.

3.3 Formulation of Random Finite Set Measurement Model

In the case of unknown number of landmarks at each time instant, the measurement model has quite an important role. In addition to giving information on the state (position, in this case) of the landmarks, the measurements should also give information of the number of landmarks in the FOV of the sensor. Thus conceptually, the construction of multi-landmark measurement model reduces to the definition of the conditional probability $f_k(\mathrm{Z}_k|\mathrm{M}_k,\mathrm{x}_k)$.

In a multi-landmark scenario, one can get multiple landmark detections. At any time instant k the result of the detection process is, say l_k measurements, $\mathrm{z}_{k,1},\ldots,\mathrm{z}_{k,l_k} \in \mathcal{Z}$. This vector has a fixed dimension, but the number of detections can vary such that $0 \leq l_k < \infty$ and do not have any inherent physical order. Consequently, the measurements can be represented as a *random finite observation set*, whose elements are individual detections, given by,

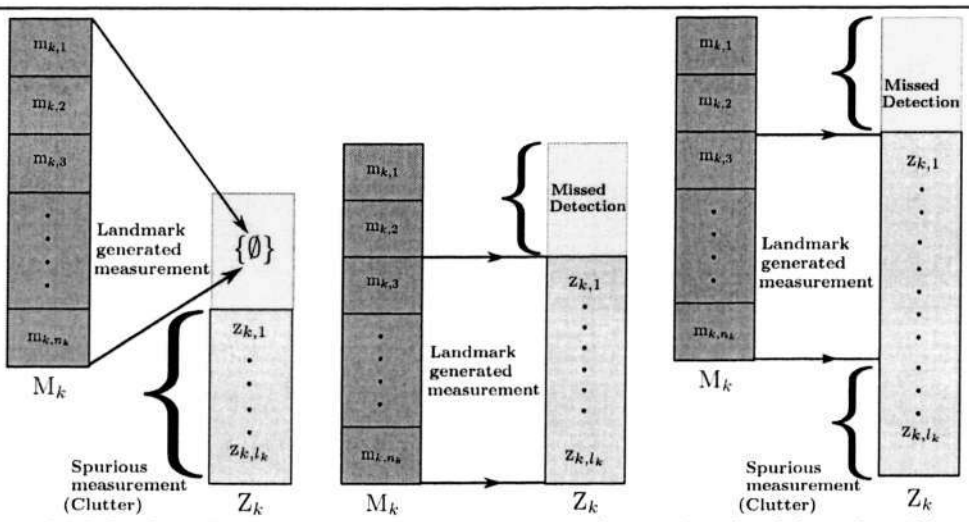
$$Z_k = \{z_{k,1}, \ldots, z_{k,i}, \ldots, z_{k,l_k}\} \in \mathcal{F}(\mathcal{Z}) \tag{3.24}$$

where $\mathcal{F}(\mathcal{Z})$ is a collection of all the random finite subsets of $\mathcal{Z}$ and $l_k = |Z_k|$ is the number of elements or the cardinality of Z_k representing the number of measurements at time k and $z_{k,i}$ representing the state of the i^{th} measurement at time k. The finite set measurement model accounts for the detection uncertainty and spurious detections (clutter) and is constructed as follows.

- Let $p_D = p_{D,k}(m_k)$ be the probability of detection of a particular sensor. i.e., a landmark in the FOV of the sensor $m_k \in M_k$ is detected with a probability $p_{D,k}(m_k)$ or missed with the probability of $1 - p_{D,k}(m_k)$.
- If a landmark is detected, the probability density of obtaining a measurement is given by the sensor likelihood function. Thus at time k, each state $m_k \in M_k$ generates a RFS $\Theta_k(m_k)$ that can take on either $\{z_k\}$ when the landmark is detected or $\emptyset$ when the landmark is not detected.
- In addition to landmark originated measurements, the sensor also receives a set of spurious (false) measurements called as clutter, at time k represented by a RFS Ω_k.

Given the multi-landmark state at time k, the multiple landmark measurement Z_k received at the sensor is formed by the union of landmark generated measurements and the false alarms given by

$$Z_k = \Omega_k \cup \left[\bigcup_{m_k \in M_k} \Theta_k(m_k) \right] \tag{3.25}$$



(a) False detections reported by the sensor

(b) Landmark generated measurements along with a few landmarks missed to be detected by the sensor

(c) Missed detections, false detections along with landmark generated measurements reported by the sensor

Figure 3.5 *Possible scenarios involved in sensor measurement modeling*

As in the case of dynamic map evolution model formulation, we construct the finite set model for the Ω_k and Θ_k using case studies.

Case 4. *No landmarks generated measurements.*

In this case, all the measurement received by the sensor are spurious and no landmarks are detected by the sensor as in fig. 3.5a. Hence the measurement model is

$$Z_k = \Omega_k \tag{3.26}$$

If we assume that all the spurious measurements are Poisson distributed with expected value λ_c and the spatial distribution governed by the probability density $c(z)$, then as in the dynamic model's case (refer Case 1), it can be shown that the

probability density function is,

$$f_k(Z_k|M_k, x_k) = f_{\Omega_k}(Z_k) = e^{-\lambda_c} \prod_{z_k \in Z_k} \lambda_c . c(z_k) \tag{3.27}$$

Case 5. *Missed detections and no clutter*

In this case, we consider that all the measurements received by the sensor are generated due to the presence of landmarks. However, all the landmarks are not detected by the sensor, as depicted in the fig. 3.5b. These are referred to as missed detections. Hence the measurement model is

$$Z_k = \left[\bigcup_{m_k \in M_k} \Theta_k(m_k) \right] \tag{3.28}$$

Again, drawing parallels from the case study for dynamic map model (refer to Case 2), to build the probability density function it can be shown that,

$$f_k(Z_k|M_k, x_k) = f_{\Theta_k}(Z_k) = f_k(\emptyset|M_k, x_k) \sum_{\vartheta} \prod_{i:\vartheta(i)>0} \frac{p_D(m_{k,i}) . f_k(m_{k,\pi(i)}|m_{k,i}, x_k)}{1 - p_D(m_{k,i})} \tag{3.29}$$

Case 6. *Missed detections along with clutter and landmark generated measurements*

Having obtained the probability density function for the clutter models and missed detections, the density function for the entire measurement can now be constructed.

$$\begin{aligned} f_k(Z_k|M_k, x_k) &= L_Z(x_k) \\ &= \sum_{W \subseteq Z_k} f_{\Theta_k}(W|M_k, x_k) f_{\Omega_k}(Z_k - W|M_k, x_k) \end{aligned} \tag{3.30}$$

$$L_Z(x_k) = e^{\lambda} f_{\Omega_k}(Z_k) f_k(\emptyset|M_k, x_k) \cdot \sum_{\vartheta} \prod_{i:\vartheta(i)>0} \frac{p_D(m_{k,\vartheta(i)}, x_k).f_k(z_{k,\vartheta(i)}|m_{k,i}, x_k)}{(1 - p_D(m_{k,\vartheta(i)}, x_k))\lambda c(z_{k,\vartheta(i)})} \tag{3.31}$$

where the summation in eqn. 3.31 is taken over all association hypotheses $\vartheta : \{1, \ldots, n_k\} \longmapsto \{0, 1, \ldots, l_k\}$. For every $i = 1, \ldots, n_k$, if $\vartheta(i) > 0$ then the observation $z_{k,\vartheta(i)}$ is uniquely associated with landmark $m_{k,i}$. However, if $\vartheta(i) = 0$ then no observation is associated with $m_{k,i}$ indicating that the landmark $m_{k,i}$ was not detected (miss detection). If $\vartheta(i) = 0, \forall i$ then every measurement in Z_k is a result of a spurious measurement (clutter), as in Case 4. If $\vartheta(i) > 0, \forall i$ then every landmark gave rise to a measurement.

3.4 The Bayes Random Finite Set Multi-landmark Filter

The mathematical foundations to the Bayes random finite set filter is due to [44], [68]. Conceptually the multi-landmark estimation problem reduces into a recursive Bayesian estimation problem. Thus given the map evolution model $f_{k|k-1}(M_k|M_{k-1}, x_k)$ and measurement model $f_k(Z_k|M_k, x_k)$, the multi-landmark Bayes recursion, in the form of random finite set densities is a standard two step process, viz. **time-update**

and **data-update**, as follows[1],

$$f_{k|k-1}(\mathrm{M}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k) = \int f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1},\mathrm{x}_k).f_{k-1|k-1}(\mathrm{M}_{k-1}|\mathrm{Z}^{(k-1)},\mathrm{x}_{k-1})\delta\mathrm{M}_{k-1} \tag{3.32}$$

$$f_{k|k}(\mathrm{M}_k|\mathrm{Z}^{(k)},\mathrm{x}_k) = \frac{f_k(\mathrm{Z}_k|\mathrm{M}_k,\mathrm{x}_k)f_{k|k-1}(\mathrm{M}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k)}{f_k(\mathrm{Z}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k)} \tag{3.33}$$

where, $f_k(\mathrm{Z}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k)$ is the Bayes normalization factor.

$$f_k(\mathrm{Z}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k) = \int f_k(\mathrm{Z}_k|\mathrm{M}_k,\mathrm{x}_k)f_{k|k-1}(\mathrm{M}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k\delta\mathrm{M}_k) \tag{3.34}$$

where the integral in eqns. 3.32 and 3.34 are the set integrals (refer to Appendix A, Section A.6 for more details on set integrals). Likewise, the multi-landmark map evolution function (Markov density) and the multi-landmark likelihood function are constructed from the belief-mass function as explained in Sections 3.2 and 3.3 respectively. The recursion in eqns. 3.32 and 3.33 involves multiple integrals, which are computationally intractable. Direct sequential Monte Carlo implementations of the optimal Bayes filter can be found in [98]. However, these optimal multi-landmark Bayes filters are computationally intensive due to the combinatorial nature of the densities, especially when the number of landmarks is large. Nonetheless, there have been some approaches that have been proposed to bring this recursion into a computationally feasible form. The probability hypothesis density (PHD)[2] filter developed by Mahler [69] is the among the most widely accepted approaches, which has been extensively used in this thesis.

[1] $\mathrm{Z}^{(k)}$ denotes the history of all the measurement sets taken up to time k

[2] The concept of PHD based on point process literature and FISST theory is provided in Appendix A

3.5 The Probability Hypothesis Density (PHD) Filter

The probability hypothesis density (PHD) filter is an approximation to alleviate the computational intractability of the optimal multi-landmark Bayes filter, proposed in [69]. Instead of propagating the entire multi-landmark posterior density, PHD propagates only the first order statistical moment of the state as follows,

$$D_{k|k}(\mathrm{m}_k|\mathrm{x}_k, \mathrm{Z}^{(k)}) = \int f_{k|k}(\mathrm{M}_k|\mathrm{x}_k, \mathrm{Z}^{(k)})\delta\mathrm{M}_k \tag{3.35}$$

Thus, PHD filter can be seen as an analogous statistic to the mean of the classical single landmark constant gain Kalman filter (α-β-γ filter). Also, the PHD filter can be characterized as compression of the multi-landmark posteriors into their corresponding first order multi-landmark moments as illustrated in the fig. 3.6 (illustration adapted from [69]).

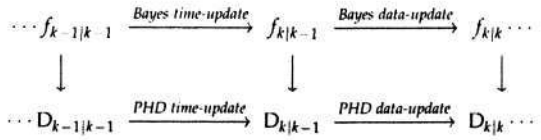


Figure 3.6 *Illustration of optimal Bayes filter recursion compressed as PHD filter recursion*

It is important to emphasize that the first order multi-landmark statistical moment (PHD) $\mathrm{D}_{k|k}$ of $f_{k|k}$ is neither a vector nor a probability density, but rather a density function and is uniquely characterized by the following property. Given any region A, the integral of the density function $\mathrm{D}_{k|k}$ over the region A gives the

expected number of elements in A [69].

$$N_{k|k} = \int_A D_{k|k}(\mathrm{m}_k|\mathrm{x}_k, \mathrm{Z}^{(k)})d\mathrm{m}_k \tag{3.36}$$

In [69], Mahler derived the update equations for the PHD for multitarget tracking. Here, we apply the same equations in the autonomous vehicle navigation framework, given that the following assumptions are satisfied.

A-1. The motion of the vehicle is Markovian in nature and hence the evolution of stationary landmarks induced by this vehicle motion is also Markovian.

A-2. The probability of each landmark's survival (p_S) is independent of the other landmarks.

A-3. Appearance of new landmarks in the FOV of the sensor is independent of the states of the existing landmarks in the FOV. The new landmarks appearing in the FOV of the sensor are modeled as a RFS Γ_k and its PHD is given by (refer Appendix A, Section A.8 for more details)

$$\gamma_k(\mathrm{m}_k) \triangleq \int \Gamma_k(\{\mathrm{m}_k\} \cup W)\delta W \tag{3.37}$$

A-4. The probability of detection (p_D) of each landmark is independent.

A-5. The landmark measurement produced by the sensor is given by the sensor likelihood function and is modeled as a RFS Θ_k.

A-6. False measurements from the sensor are Poisson distributed with expected value λ_c and the spatial distribution governed by the probability density $c(\mathrm{z}_k)$ and independent of landmark originated measurements.

A-7. The predicted multi-landmark distribution is approximated to be Poisson distributed with average number of landmarks μ, i.e.,

$$f_{k|k-1}(\mathrm{M}_k|\mathrm{Z}^{(k)},\mathrm{x}_k) \approxeq e^{-\mu} \prod_{\mathrm{m}_k \in \mathrm{M}_k} \mu.f_{k|k-1}(\mathrm{m}_k|\mathrm{Z}^{(k)},\mathrm{x}_k) \tag{3.38}$$

The assumptions A-1 to A-3 are satisfied by the dynamic multi-landmark map evolution model that was presented in Section 3.2 while the assumptions A-4 to A-6 are satisfied by the vehicle measurement model discussed in Section 3.3. It has been shown in [70] that assumption A-7 is completely satisfied when the RFS M_{k-1} and Γ_k are Poisson distributed and when there are no interactions between landmarks (no spawning). Under assumptions A-1 to A-7, it can be shown (using FISST [69]) that the PHD time-update (prediction) and data-update (correction) equations are as follows:

- **PHD time-update**

$$D_{k|k-1}(\mathrm{m}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k) = \overbrace{\gamma_k(\mathrm{m}_k)}^{\text{new landmarks}} + \int \overbrace{p_{S,k-1}(\mathrm{m}_{k-1}) f_{k|k-1}(\mathrm{m}_k|\mathrm{m}_{k-1})}^{\text{existing landmarks}} .D_{k-1|k-1}(\mathrm{m}_{k-1}|\mathrm{Z}^{(k-1)},\mathrm{x}_k) d\mathrm{m}_{k-1} \tag{3.39}$$

The predicted number of landmarks is therefore,

$$N_{k|k-1} = \int D_{k|k-1}(\mathrm{m}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k) d\mathrm{m}_k \tag{3.40}$$

- **PHD data-update**

$$D_{k|k}(\mathrm{m}_k|\mathrm{Z}^{(k)},\mathrm{x}_k) = (1-p_D)D_{k|k-1}(\mathrm{m}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k) + \sum_{\mathrm{z}_k\in\mathrm{Z}_k}\frac{p_D D_k(\mathrm{z}_k|\mathrm{m}_k,\mathrm{x}_k)}{\lambda_c c_k(\mathrm{z}_k)+p_D D_k(\mathrm{z}_k|\mathrm{m}_k,\mathrm{x}_k)}D_k(\mathrm{m}_k|\mathrm{z}_k,\mathrm{x}_k) \quad (3.41)$$

where,

$$D_k(\mathrm{z}_k|\mathrm{m}_k,\mathrm{x}_k) = \int L_{\mathrm{z}_k}(\mathrm{x}_k)D_{k|k-1}(\mathrm{m}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k)d\mathrm{m}_k \quad (3.42)$$

with $L_{\mathrm{z}_k}(\mathrm{x}_k) = f_k(\mathrm{z}_k|\mathrm{x}_k,\mathrm{m}_k)$ is the sensor likehood function and,

$$D_k(\mathrm{m}_k|\mathrm{z}_k,\mathrm{x}_k) = \frac{f_k(\mathrm{z}_k|\mathrm{m}_k,\mathrm{x}_k)D_{k|k-1}(\mathrm{m}_k|\mathrm{Z}^{(k-1)},\mathrm{x}_k)}{D_k(\mathrm{z}_k|\mathrm{m}_k,\mathrm{x}_k)} \quad (3.43)$$

Hence, the total expected number of landmarks is,

$$N_{k|k} = \int D_{k|k}(\mathrm{m}_k|\mathrm{Z}^{(k)},\mathrm{x}_k)d\mathrm{m}_k \quad (3.44)$$

The PHD filter operates on a single-landmark state space and there by completely avoiding the computational complexity involved in data association. However, the PHD recursion involving eqns. 3.39 and 3.41 have multiple integrals, that have no closed form solutions, in general. There are various approaches presented for the practical implementation of the PHD recursion with Sequential Monte Carlo approximation [97], [110], [65] and Gaussian mixture approximation [109] being the most prominent techniques. A study on the convergence properties of these techniques conducted in [26] and [25] give a theoretical justification for the use of these algorithms.

3.6 Chapter Summary

By adopting a finite random set approach, a multi-landmark map evolution model that accounts for the new landmarks appearing in the field of view of sensor and a natural multi-landmark (map) measurement model that explicitly accounts for detection uncertainty, clutter and unknown data association for autonomous vehicle navigation has been constructed. An optimal multi-landmark Bayes filter and its first order statistical moment filter (PHD filter) have been incorporated within the autonomous vehicle navigation framework. This effectively enables to estimate the unknown varying number of landmarks and their states from a GRB sensor mounted on any mobile platform.

4

Landmark Detection in Presence of Clutter

This chapter examines the detection of landmarks in the presence of false measurements from a range-bearing-intensity sensor in general and an underwater blazed array sonar in particular based on the random finite set models developed in the previous chapter. This chapter provides a mechanism and a detailed work-flow for processing the underwater sonar data to obtain a reliable feature (landmarks) set in underwater environments. It also presents a novel clutter rejection algorithm for autonomous vehicle navigation. The algorithm is based on the fusion of the moment approximation of the posterior density within the random finite set framework with the conventional SLAM framework.

Section 4.1 outlines the key role played by the feature extraction in the feature based environmental mapping and vehicle localization, highlighting the need for a robust clutter rejection strategy from the noisy measurements. A clutter rejection algorithm based on RFS models developed in Chapter 3 is presented in Section 4.4,

that simplifies the data association process and obviates the need for *ad hoc* feature management strategies. The algorithm is validated for its clutter filtering ability using simulated sonar scans in Section 4.5. A novel way of fusing the clutter filter with the standard EKF-SLAM filter is presented in Section 4.6. The effectiveness of the proposed clutter rejection strategies along with the standard EKF-SLAM filter without any external map management strategies is demonstrated with offshore underwater field trials and benchmarked against a standard EKF-SLAM filter in Section 4.7.

4.1 Background

In vehicle localization and mapping context, the representation of an environment can be done in two ways, namely

- Sensor centric approach
- Feature centric approach

Raw sensor information is acquired and processed in its original form in the sensor centric approach as in scan matching [66], [46] where locally consistent pose estimates are recovered. Scan matching techniques are based on the iterative closest point (ICP) algorithm that determines a relative translation and rotation to align two point sets with minimal disparity in an iterative manner by employing a reasonable initial guess of the translation and rotation terms. This can be equivalent to considering an entire sonar scan at any time instant as a *composite feature* and performing nearest neighbor data association pixel wise for each range-bearing measurement.

The grid based method is a variant of the sensor centric approach where environmental mapping is accomplished using occupancy grids [36], [58]. An occupancy grid represents a robot's environment by a two or three dimensional grid, where each cell stores the evidence (probability), based on the sensor readings, that a particular space is occupied. Bayesian estimation techniques are generally used to incrementally update the occupancy grids using sensor readings from multiple vantage points to create a dense map of the environment [79]. Since occupancy grid based approaches generate dense environmental models, correlation based map matching approaches are commonly employed for vehicle localization as in [36].

The main attribute of these approaches is that they eliminate the need for extraction of features. In a feature centric approach [13], the raw measurements are processed by a feature extraction algorithm in order to identify certain well defined entities referred to as *features* that can be used in the state estimation. Both the above approaches have their advantages and disadvantages. The sensor centric approaches have the luxury of ignoring all the aspects of environmental modeling and represent the environment with raw sensor scans which is an advantage in representing highly unstructured environments. However, the computational complexity of matching and estimating the state of each raw sensor point drastically limits the use of this approach and shifts the balance towards feature based environmental representation and modeling, which ensures a high degree of data compression. Henceforth, we would be focusing only on feature centric approaches. The literature offers a wide range of feature descriptors and their extraction techniques for a wide variety of sensors.

For a typical range sensor, geometric primitives such as lines, corners that correspond to walls, are used as features to perform navigation in structured indoor

environments. Most notable among the extraction of geometric primitives is the *region of constant depth* (RCD) technique that tries to fit an arc over a sequence of individual range returns, which has been successfully implemented in indoor environments [62]. In contrast to the indoor environments, which are dense and resemble proper geometric primitives like corners and straight lines, unstructured natural underwater environments are sparse in detail and do not conform to any simple geometric model. The common approach used in such scenarios is to use segmentation followed by a simple clustering algorithm to group raw sensor measurements into data clusters and extract descriptors that sufficiently describe the cluster. Cluster centroids have been used extensively as feature descriptors in underwater sonar systems, given the difficulty in modeling the terrain and processing cost involved [94], [115]. Moran has used RCDs as feature descriptors for underwater shape reconstruction [78]. Carpenter has used a series of morphological operations on a sonar scan image to form objects (clusters), and retrieved perimeter, area, area-to-perimeter ratio and radial signature (distance from objects centroid to its boundary as a function of angle) as geometric descriptors to characterize a landmark [23]. Majumder has used intensity returns from sonar signal returns and a distance criterion to group measurements as blobs (clusters), and retrieved peak amplitude, its variance and the associated area as descriptors to characterize a landmark [72]. Ribas et.al have used a simple line extraction technique from sonar image when operating in man made structured underwater environments [92].

Thus the goal of feature extraction is to identify well defined entities from the raw measurements, thereby reducing the dimension and complexity of the data association process. However, if feature extraction is not done optimally, it would result in an increase in false alarms and missed detections in the data association

process and also an increase in its computational complexity. In most of the feature extraction techniques discussed above, there are no concrete steps taken to reject the clutter from noisy sensor measurements, apart from basic pre-filtering. The main rationale behind the use of such feature extraction strategies in SLAM is that only stable features (landmarks) sustain in the long run by an *ad hoc* map management and track maintenance function [32], of the data association algorithm. Hence there has been very little importance given to clutter filtering during the feature extraction stage. However in this work, it is argued that if we are able to perform an efficient feature extraction process that can handle clutter and sensor's detection uncertainty, the data association phase would be much simpler and accurate. Thus by paying more attention to sensor modeling and devising a robust feature extraction algorithm that is capable of rejecting clutter from the noisy measurements, the data association process is simplified, obviating the need for ad-hoc map management strategies.

4.2 Sensor Measurement Analysis

Sensors such as imaging sonars, millimeter wave radars, imaging lidars return intensity valued measurements which can be grouped as generic range bearing intensity (GRBI) sensors. A GRBI sensor returns are in the form of $s(r, \theta)$, i.e., an intensity-valued function of range r and bearing θ. More specifically, at any time instant k, the set of raw measurements Z_k that forms an image as in fig. 4.1, can be

represented as,

$$Z_k = \begin{bmatrix} s(r_{k,1}, \theta_{k,1}) & \dots & s(r_{k,R}, \theta_{k,1}) \\ \vdots & \ddots & \vdots \\ s(r_{k,1}, \theta_{k,B}) & \dots & s(r_{k,R}, \theta_{k,B}) \end{bmatrix} = \begin{bmatrix} R_{k,1} \\ \vdots \\ R_{k,B} \end{bmatrix} \tag{4.1}$$

where, $R_{k,t}$ refers to the range intensity spectrum (fig. 4.2) at the $t-$ th bearing, at time instant k. R is the total number of range bins in a single intensity spectrum, and B is the total number of bearing angles sampled for a given measurement. Hence, for a particular range/bearing setting, the total number of raw measurements received for each scan (image) is a constant.

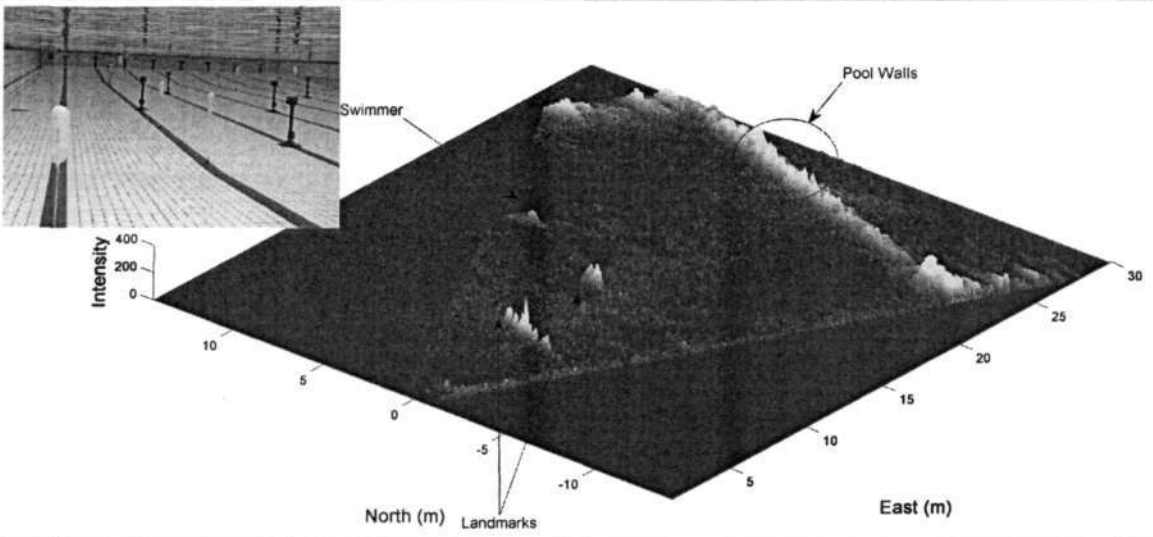


Figure 4.1 *Acoustic image of the swimming pool as obtained from a blazed array sonar. (Inset: An Optical image of the swimming pool environment)*

Specifically underwater, sound (sonar) generally has superior propagation characteristics to electromagnetic (visual) signals. Figure 4.1 shows a typical acoustic image formed by an active blazed array sonar along with an optical image of the swimming pool environment with artificial beacons. Sonar signals, either echoes of

active sonars or target sounds of passive sonars, must always be observed amid a background noise or reverberation. Before any system function is performed, the sonar system must first detect the signal in this background. However it can be very difficult to interpret sonar data due to the presence of multiple reflections, the presence of specular or diffuse features and objects, and the effects of high-power side lobes in the transmitted waveform. All these factors significantly influence the quality of the sonar image and make it difficult to determine object shape and size from the sonar data. A good look into the sonar process would give us a better understanding of the sonar return signals. It would also help us in determining a suitable feature extraction process from the sonar returns.

4.2.1 Sonar Process

Sound energy in water suffers two types of losses; spreading and absorption. The combination of these two losses is commonly referred to as transmission loss TL. Spreading refers to the broadening of the emitted acoustic wave with distance. There are mainly two types of spreading based on range r ; spherical spreading ($\mathrm{r} < 1000m$) and cylindrical spreading ($\mathrm{r} > 1000m$) [74]. Spherical spreading results in a reduction in signal level proportional to the inverse of the range squared, that is:

$$TL = 20 \log \mathrm{r} \tag{4.2}$$

Absorption occurs through viscous friction at high frequencies and molecular resonance at low frequencies. The absorption coefficient as stated by Urick is given by [108],

$$\alpha = \frac{0.036 f^2}{f^2 + 3600} + 3.2\mathrm{x}10^{-7} f^2 \tag{4.3}$$

where f is the frequency of operation of the sonar. From the above equation it can be noted that absorption coefficient increases with increase in frequency, thereby limiting the range of operation. Thus total transmission loss is:

$$TL = 20\log \mathrm{r} + \frac{0.036f^2}{f^2 + 3600} + 3.2\mathrm{x}10^{-7}f^2 \qquad \forall \mathrm{r} < 1000 \tag{4.4}$$

Like all other sensors, sonar systems also have a minimum required signal level to result in detection. Noise results from many sources; caused by the receiving platform (self noise) and unavoidable noise present in the environment (ambient noise). Self noise is mainly due to platform machinery, propellers of the vehicle on which the sensor is mounted, if any. Ambient noise is mainly due to the BOSH (Biological Ocean traffic Seismic Hydrodynamic) effects [108]. Thus total noise level NL is given by the sum of self noise NL_S and ambient noise NL_A.

$$NL = NL_S + NL_A \tag{4.5}$$

NL_A is generally determined by Wenz curve [108].

The sonar equation ties together the acoustic medium, the target (landmark) and the equipment in order to predict detection ranges and thereby relating it to the signal to noise ratio SNR. It is the ratio of the received echo from the target to background noise produced by everything else. Detection threshold D_t is the measure of the return signal required for any system to detect the presence of the landmark. Given that the source level of the transducer is SL, the target strength of the object being imaged is TS (which is a function of the material properties and the angle of incidence to the object) and the directivity index of the transducer is

DI (all quantities are in decibels referenced to standard units), the SNR required for detection is given by [108],

$$D_t = SL - 2TL + TS - NL + DI \tag{4.6}$$

4.3 The Landmark Detection Problem

Thus, detection concerns the recovery of features (landmarks[1]) from the noisy sensor measurements, such as that in fig. 4.2b. Feature extraction then concerns the reduction of the dimensionality of the prospective landmark detections, into landmark models. The landmark detection problem using a classical Bayes criterion, can be alternatively stated as a simple binary hypothesis testing problem involving two hypotheses:

- Null hypothesis $\mathcal{H}_0$: Signal is not present (Feature absence)
- Alternate hypothesis $\mathcal{H}_1$: Signal is present (Feature presence)

Thus at the receiver, $\mathcal{H}_0$ relates to the noise only case and $\mathcal{H}_1$ to the signal plus noise case and the likelihood ratio, L, is the ratio of probability that hypothesis $\mathcal{H}_1$ is true to probability that hypothesis $\mathcal{H}_0$ is true. A decision (detection) is made by comparing the likelihood[2] to a predetermined threshold D_t. That is, if $L = \frac{\mathcal{H}_1}{\mathcal{H}_0} > D_t$, a decision is made that the signal (feature) is present.

[1] Landmarks and features are loosely used through out the thesis, however, landmarks are specifically used to signify static features / targets.

[2] the log of this likelihood is referred to as log-likelihood ratio, which is equivalent to the sonar eqn. 4.6

4.3.1 Fixed Detection

Intensity thresholding has always proven to be an effective technique to eliminate low energy noise and to reduce the computational load on the feature tracking system. In an ideal case, if the landmark has a significant size and reflectivity, with the returns having a low SNR, the landmark-generated reflection-intensity will be significantly stronger than the surrounding noise floor of the sensor. As shown in a range-intensity spectra from a blazed array sonar in fig. 4.2a, the landmark can be easily detected from the intensity returns by setting the detection threshold D_t that is high enough to avoid the noise. The probability of such an event is referred to as probability of detection, denoted by p_D. Thus, detection indicates the presence of a landmark with range r and bearing θ. However if the landmark-generated signal is too small to exceed the set threshold, in which case the landmark is not detected (miss detection) and the probability of such an event is called the probability of miss detection, denoted by $1 - p_D$. If the SNR of the intensity returns are low, as depicted in fig. 4.2b, the landmark-generated intensities are not so obviously separated from the noise. In such a case, the event $s(r, \theta) > D_t$ will not only detect the landmarks, but also a large number of false alarms. If D_t is set too large, it might not be possible to detect the landmarks. If D_t is set small enough to ensure landmark detection, will result in large number of false alarms λ_c. Thus the biggest challenge is to set this threshold value D_t, which is a compromise between the probability of detection p_D and the probability of false alarm p_{FA}, which is commonly referred to as receiver operating characteristic (ROC) [108]. Thus an ideal detector with a fixed threshold is extremely sensitive to the total noise (measurement noise plus clutter) variance. In fact, it has been shown that a small increase in the total noise power results

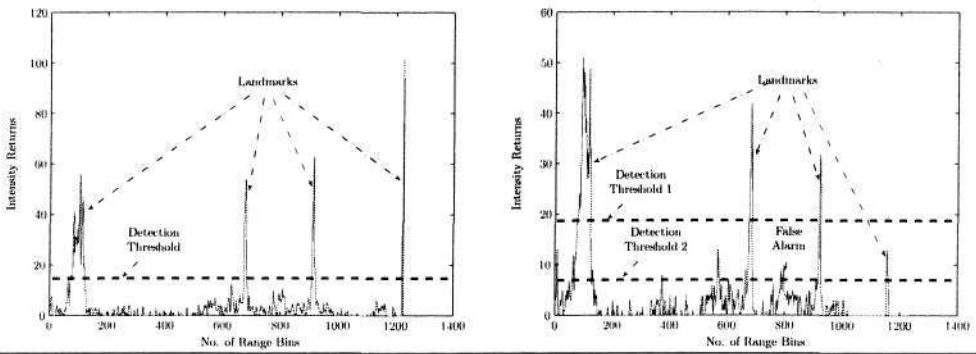


Figure 4.2 *Range-intensity spectra from a blazed array sonar is depicted for two different scenarios. The SNR is large, and landmark has a significant size and reflectivity in (a). The landmark can be deduced by applying a suitable threshold D_t to the returns. Case (b) illustrates the intensity returns with a low SNR. In such a scenario any value of D_t small enough to detect a landmark will also detect a large number of false detections, as in the case of detection threshold 2. Setting it high enough as in detection threshold 1 will result in miss detections.*

in a corresponding increase of several orders of magnitude in the probability of false alarm [41]. Therefore, adaptive threshold techniques are needed to maintain a constant false alarm rate (CFAR).

4.3.2 Adaptive CFAR Detection

In CFAR processors the detection threshold is adaptively set based on local information of the total noise power in order to regulate the false alarm probability [93]. The detection threshold D_t in a CFAR detector is set on a range bin by range bin basis using estimated noise power by processing a group of reference range bins surrounding the range bin under investigation. Of all the CFAR schemes tested on underwater sonar [53], ordered-statistics (OS) approach [93] has proven to be most robust in situations of high clutter, and is therefore adopted in this work. The working principles of OS-CFAR are explained further, leaving the mathematical specifics which are detailed in [41], [96]. As stated earlier, a range-intensity spectrum[1] for a

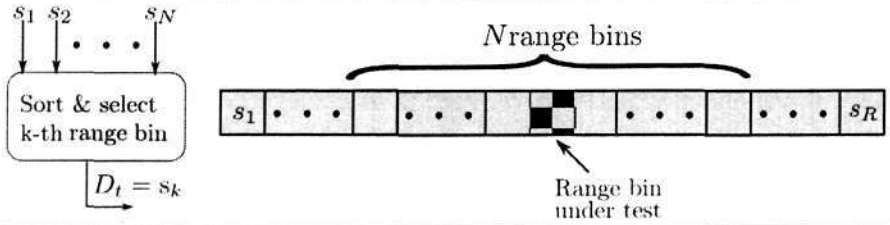


Figure 4.3 *Block diagram of order statistics constant false alarm rate detector.*

particular bearing is given by $R = [s_1, \ldots, s_R]$, as depicted in fig. 4.3. In OS-CFAR, N range bins are selected as reference cells around the range bin under test, as illustrated in fig. 4.3, and are ranked according to their intensity level. A k^{th} ordered sample s_k is chosen to represent the noise level estimate of the range bin under test. This selected statistic is then multiplied by the scale factor α to achieve the desired probability of false alarm. A detection is made by comparing the intensity of the range bin under test with the adaptive threshold $D_t = \alpha s_k$.

The output of an OS-CFAR detector, with parameters $N = 32$, $k = 3N/4$, $p_{FA} = 1\text{x}10^6$, on the same blazed array sonar range intensity spectra, as in fig. 4.2b, which has a low SNR can be seen in fig. 4.4. The adaptive threshold scheme using OS-CFAR manages to detect all the features from the data, reporting a few false alarms. However, these CFAR schemes are not without drawbacks. 1. The presence of noise power transition within a single window may result in severe performance degradation in an adaptive threshold scheme leading to excessive false alarms or landmark masking. 2. In the presence of closely spaced landmarks, a CFAR scheme may raise the threshold unnecessarily resulting in missed detection of the landmarks.

A unifying framework that incorporates the vehicle motion model into the land-

[1] $s(r, \theta)$ is represented as s for simplicity

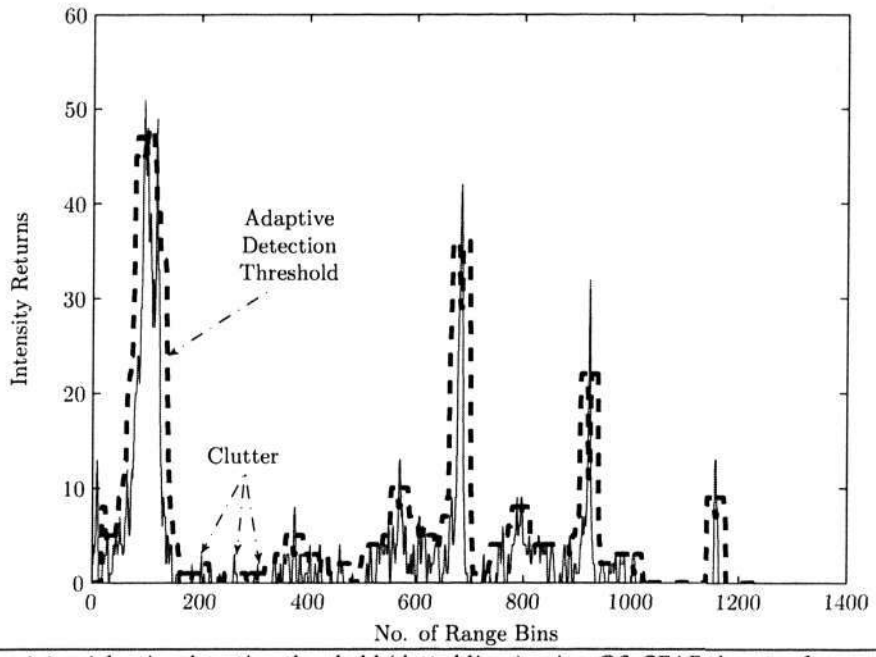


Figure 4.4 *Adaptive detection threshold (dotted lines) using OS-CFAR detector for a constant $p_{FA} = 1\text{x}10^{6}$ for a range-intensity spectra of a blazed array sonar.*

mark detection process, effectively removing the false alarms is discussed in the next section.

4.4 Clutter Filter Mechanics

This section describes the method for tracking multiple unknown number of features from the blazed array sonar returns in presence of false alarms (clutter). To achieve this, we use the random finite set models developed in the previous chapter. At any time instance k the landmarks present in the environment are modeled as a RFS given by Y_k and the result of the detection process from the sensor are modeled

as a RFS Z_k given by

$$\begin{aligned} Y_k &= \{y_{k,1}, \ldots, y_{k,i}, \ldots, y_{k,N_k}\} \in \mathcal{F}(\mathcal{Y}) \\ Z_k &= \{z_{k,1}, \ldots, z_{k,i}, \ldots, z_{k,l_k}\} \in \mathcal{F}(\mathcal{Z}) \end{aligned} \tag{4.7}$$

where $N_k = |Y_k|$ is the number of landmarks present at time k with $y_{k,i}$ representing the state of the i^{th} landmark and $l_k = |Z_k|$ is the number of elements or the cardinality of Z_k representing the number of measurements at time k with $z_{k,i}$ representing the state of the i^{th} measurement[1]. A probability hypothesis density (PHD) filter founded on the principles of FISST as discussed in Chapter 2, is effectively used as a multi-landmark clutter rejection filter within the Bayesian framework by constructing multi-landmark densities for the above described RFS.

4.4.1 The PHD Clutter Filter

The aim of this filter would be to eliminate most of the measurements that are likely to be false measurements so that the data association complexity is drastically reduced. The clutter filtering scheme is due to [89]. The incorporation of the PHD filter introduced in Chapter 2, as a clutter rejection filter for the blazed array sonar returns is presented here. The PHD filter along with a peak extraction technique determines the most probable number of landmark generated measurements and their states Y_k. This output is used to define the validation gate in the measurement space. The measurements that fall outside the gate are discarded as clutter and the result is a new "clutter reduced (CR) measurement" set, $\widetilde{Z}_k$ consisting of mainly landmark generated measurements. The dynamic model describes the evolution of

[1] $y_{k,i}$ and $z_{k,i}$ are denoted as y_k and z_k for notational simplicity

the state of the landmarks induced by the motion of the underwater vehicle. The measurement model relates the measurements obtained by the blazed array sonar to the state of the landmarks. The PHD clutter filter recursion is a two step process:

- **PHD time update:** Given the motion model, the predicted PHD

$$D_{k|k-1}(y_k|Z^{(k-1)}) = \overbrace{\gamma_k(y_k)}^{\text{new landmarks}} + \int \overbrace{p_S(y_{k-1}).f_{k|k-1}(y_k|y_{k-1})}^{\text{existing landmarks}} .D_{k-1|k-1}(y_{k-1}|Z^{(k-1)})dy_{k-1} \tag{4.8}$$

  and,

  - $\gamma_k(y_k)$: PHD of the new incoming landmarks within the FOV
  - $p_S(y_{k-1})$: Probability of landmark being re-observed

- **PHD data update:** Given a new set of measurements Z_k , the updated PHD

$$D_{k|k}(y_k|Z^{(k)}) = (1-p_D)D_{k|k-1}(y_k|Z^{(k-1)}) + \sum_{z_k \in Z_k} \frac{p_D D_k(z_k)}{\lambda_c c_k(z_k) + p_D D_k(z_k)} D_k(y_k|z_k) \tag{4.9}$$

  where,

$$D_k(z_k|y_k) = \int f_k(z_k|y_k) D_{k|k-1}(y_k|Z^{(k-1)})dy_k \tag{4.10}$$

$$D_k(y_k|z) = \frac{f_k(z_k|y_k) D_{k|k-1}(y_k|Z^{(k-1)})}{D_k(z_k)} \tag{4.11}$$

  and,

  - $f_k(z_k|y_k)$: is the sensor likelihood function $L_z(y_k)$

- λ_c: average number of false alarms per scan, which is assumed to be Poisson distributed
- $c_k(z_k)$: distribution of each of the false alarms

At each time step, the PHD clutter filter not only propagates the PHD, but also the expected number of landmarks, given by

$$\tilde{N}_{k|k} = \int D_{k|k}(y_k|Z^{(k)})dy_k \tag{4.12}$$

Consequently the state of landmark oriented measurements is acquired by searching for $N_{k|k}$ highest peaks of $D_{k|k}(y_k|Z^{(k)})$, where $N_{k|k}$, represents the nearest integer to $\tilde{N}_{k|k}$. The PHD clutter filter recursion, involving eqns. 4.8 and 4.9 have multiple integrals, that have no closed form solutions, in general. There are various approaches for the practical implementation of the PHD recursion with sequential Monte Carlo (SMC) approximation [97], [110], [65] being the most prominent approach. We have used the SMC approximation for the practical implementation of the PHD clutter filter.

4.4.2 SMC Implementation of PHD Clutter Filter

This section describes SMC implementation of the PHD clutter filter. For any $k > 0$, let $\{w_k^{[j]}, y_k^{[j]}\}_{j=1}^{L_k}$ describe the particle approximation of the PHD. We start with L_{k-1} particles that are propagated to time k. At time k, we also generate additional J_k particles to explore for the additional landmarks that may appear in the FOV of the sonar. Thus time-update step of the PHD clutter filter has a particle representation given by $\{w_{k|k-1}^{[j]}, y_k^{[j]}\}_{j=1}^{L_{k-1}+J_k}$. Based on the new measurements the data-update step

maps the particle representation from the time-update step into $\{w_k^{[j]}, y_k^{[j]}\}_{j=1}^{L_{k-1}+J_k}$ by modifying the weights of the particles. The particle PHD clutter filter recursion can be summarized as follows:

Algorithm 1 SMC implementation of the PHD clutter filter recursion.

- ***Initialize***

At the start, L_{k-1} unweighted particles are distributed uniformly across the sensor's FOV.

- ***PHD time update step***

PHD Prediction of existing landmarks

for $j = 1, \ldots, L_{k-1}$ ***do***

Sample $\tilde{y}_k^{[j]} \sim \pi_k(\cdot|y_k^{[j]}, Z_k)$ *and compute its associated weight as*

$$\tilde{w}_{k|k-1}^{[j]} = \frac{\tau_{k|k-1}(\tilde{y}_k^{[j]}, y_{k-1}^{[j]})}{\pi_k(\tilde{y}_k^{[j]}|y_{k-1}^{[j]}, Z_k)} w_{k-1}^{[j]}$$

where $\tau(\cdot,\cdot) = f_{k|k-1}(\cdot|\cdot)$, *with* $f_{k|k-1}(\cdot|\cdot)$ *denoting the same meaning as in eqn. 4.8*

end for

PHD Prediction of new landmarks

for $j = L_{k-1} + 1, \ldots, L_{k-1} + J_k$ ***do***

Sample $\tilde{y}_k^{[j]} \sim \xi_k(\cdot|, Z_k)$ *and compute the weight of new landmarks as*

$$\tilde{w}_{k|k-1}^{[j]} = \frac{1}{J_k} \frac{\gamma_k(\tilde{y}_k^{[j]})}{\xi_k(\tilde{y}_k^{[j]}|, Z_k)}$$

end for

- ***PHD data update step***

for $j = 1, \ldots, L_{k-1} + J_k$ ***do***

Update weights

$$\tilde{w}_k^{[j]} = \left[1 - p_D + \sum_{z \in Z_k} \frac{p_D f_k(z|, y_k^{[j]})}{\lambda_c c_k(z) + \sum_{l=1}^{L_{k-1}+J_k} p_D f_k(z|y_k^{[l]}) \tilde{w}_{k|k-1}^{[l]}}\right] \tilde{w}_{k|k-1}^{[j]}$$

end for

- ***Estimate the number of landmarks from PHD***

Compute the total number of landmarks $\tilde{N}_{k|k} = \sum_{j=1}^{L_{k-1}+J_k} \tilde{w}_k^{[j]}$

$N_{k|k} = round(\tilde{N}_{k|k})$

- ***Resample the weights*** $\left\{\frac{\tilde{w}_k^{[j]}}{\tilde{N}_{k|k}}, \tilde{y}_k^{[j]}\right\}_{j=1}^{L_{k-1}+J_k}$ *to get* $\left\{\frac{w_k^{[j]}}{\tilde{N}_{k|k}}, y_k^{[j]}\right\}_{j=1}^{L_k}$

Rescale the weights to get $\left\{w_k^{[j]}, y_k^{[j]}\right\}_{j=1}^{L_k}$

It is important to note that during the resampling stage, the new weights $\{w_k^{[j]}\}_{j=1}^{L_k}$ are not normalized to one but to sum to $\tilde{N}_{k|k}$.

The SMC approximation of the PHD recursion provides an estimate of the number of landmarks as seen by the sensor at each step. It has been shown in [69] that the local maxima (peaks) of the PHD correspond to most probable locations or state of these $N_{k|k}$ landmarks.

A simple k-means clustering algorithm is used to partition the data and determine the peaks of the PHD. Given a set of the points, $\{\mathrm{y}_k^{[j]}\}_{j=1}^{L_{k-1}+J_k}$, where each point (in this case particles approximating the PHD) is a real vector, the k-means clustering algorithm aims to partition the $L_{k-1}+J_k$ points into $N_{k|k}$ sets, $S_k = \{S_{1_k}, S_{2_k}, \ldots, S_{N_{k|k}}\}$ so as to minimize the intra cluster distance measure.

$$\arg\min_{S_k} \sum_{i=1}^{N_{k|k}} \sum_{\mathrm{y}_k^{[j]} \in S_{k,i}} ||\mathrm{y}_k^{[j]} - \mu_{k,i}||^2 \tag{4.13}$$

where $\mu_{k,i}$ is the mean of $S_{k,i}$. The set of estimates of the most probable landmark locations at time step k is given by $\mathrm{Y}_k = \{\mu_{k,i}\}_{i=1}^{N_{k|k}}$.

4.4.3 Gating for Clutter Rejection

The PHD filter output, Y_k is used to define the validation gates in the measurement space. The measurements that fall outside the validation gate ν_g are discarded as clutter, there by resulting in the most probable landmark originated measurements. The implementation can be summarized as follows:

Algorithm 2 Implementation of gating function for clutter rejection.

- *Inputs to the gating function*

$Y_k = \{y_{k,1}, \ldots, y_{k,j}, \ldots, y_{k,N_k}\}$
$Z_k = \{z_{k,1}, \ldots, z_{k,i}, \ldots, z_{k,l_k}\}$
for $i = 1, \ldots, l_k$ ***do***
  for $j = 1, \ldots, N_k$ ***do***
    if then $f_k(z_{k,i}|y_{k,j}) > v_g$
      include $z_{k,i}$ *in* $\bar{Z}_k$
    end if
  end for
end for

4.5 Performance Validation in a Simulated Environment using Blazed Array Sonar

This section evaluates the performance of the clutter filtering algorithm using synthetically generated data. The advantage of generating synthetic data and simulating the underwater environment is that it allows realistic scenarios with exact locations of the landmarks and vehicle known at each time instant. Thus the performance of the clutter rejection filter is validated in terms of error in the estimated number of landmark generated measurements. A simple simulation environment is set up in which an underwater vehicle mounted with a blazed array sonar is moving in a straight line observing the landmarks in the field of view. The vehicle is moving at a nominal speed of $2m/s$ in an environment consisting of stationary features (landmarks). The number of landmarks at any time in the FOV of the sonar is unknown and time varying. The sonar is range limited to 40m with forward looking 45^o field of view. The synthetic sonar data was generated with seventeen landmarks in the environment immersed in clutter that was Poisson distributed with a density $\lambda_c = 5$. That is, clutter is uniformly distributed over the sensor FOV $[0, 40] \times [-3\pi/24, 3\pi/24]$ with an average rate of $\lambda_c = 5$ per scan. The other param-

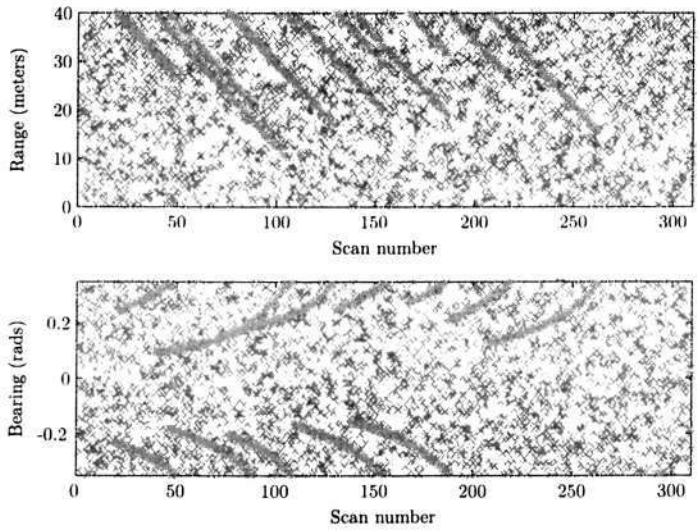
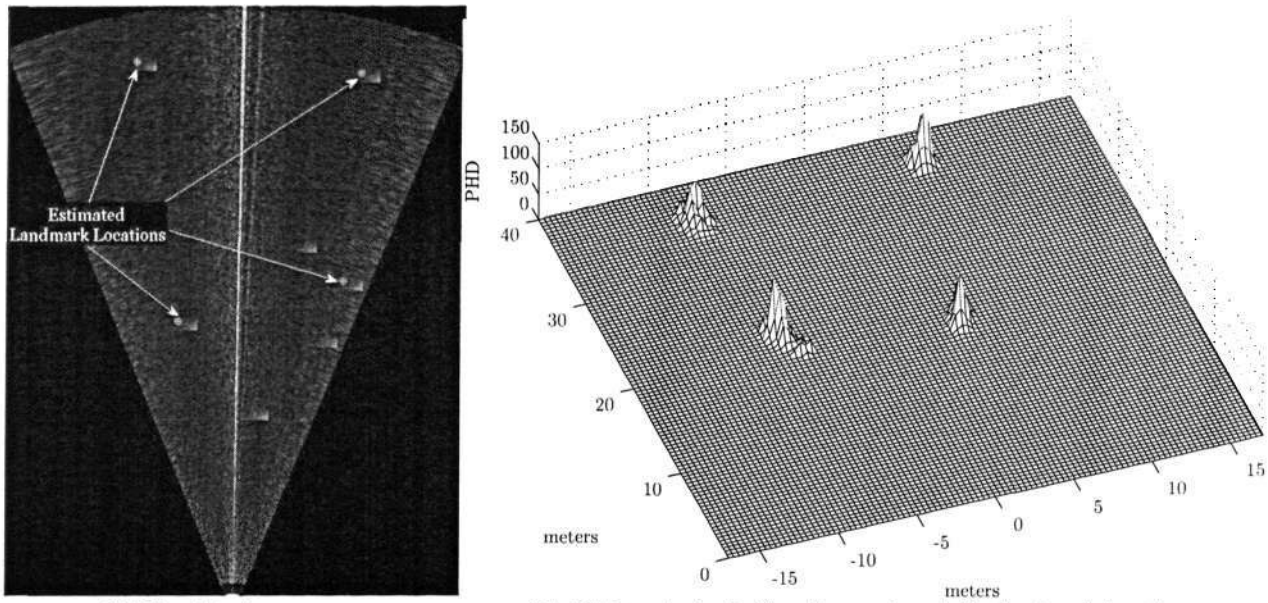


Figure 4.5 *Range and Bearing measurements containing actual landmark generated measurements (in green asterisk) immersed in a Poisson clutter (in gray crosses) with a density of* $\lambda_c = 5$.

eters used for the presented simulation were the zero mean velocity and steering control Gaussian process noise with standard deviations $\sigma_v = 0.3m/s$ and $\sigma_\gamma = 0.5^o$ respectively, the zero mean range and bearing Gaussian measurement noise with standard deviations of $\sigma_r = 0.3m$ and $\sigma_\theta = 0.5^o$ respectively. Due to the motion of the vehicle, the new landmarks would most likely appear in the FOV of the sonar at perimeter of the scanning sector. The PHD birth model will incorporate these new incoming landmarks by uniformly distributing the particles along the perimeter. Without loss of generality, the probability of detection is set to $p_D = 1$. It is important to note that due to explicit incorporation of detection uncertainty p_D in the measurement likelihood equation (eqn. 4.9), the PHD filter is able to handle the sensor's detection uncertainty (i.e., $p_D \neq 1$), while vector based frameworks always assume a unity probability of detection, irrespective of the sensors detection uncertainty, resulting in false detections, especially in the presence of clutter.

Eight hundred particles are generated for every landmark in the particle PHD filter implementation, and hence, the number of particles varies at each iteration in the simulation. Figure 4.5 shows the simulated range-bearing measurements at each time step, that encapsulates both landmark generated measurements and clutter generated measurements with an average clutter density of $\lambda_c = 5$. At each time instant, the true number of features, and their corresponding ground truth locations, within sensor FOV can be determined in a simulated environment and hence is used as a metric to determine clutter filtering performance. As outlined earlier

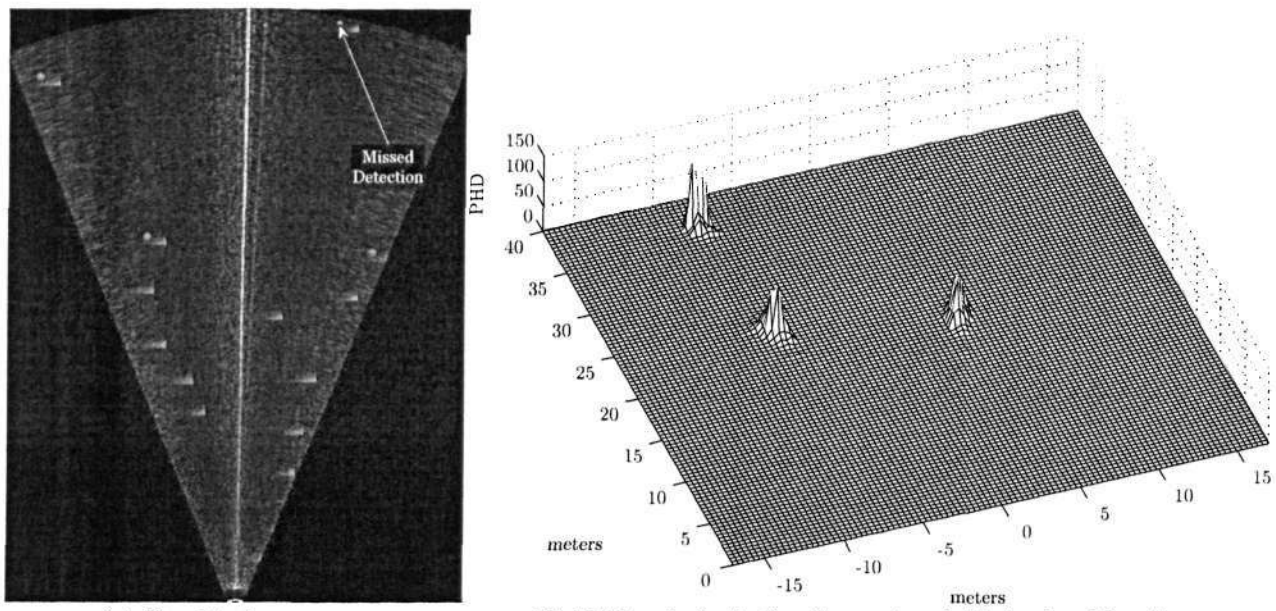


(a) Simulated sonar scan *(b) PHD peaks indicating the most probable landmark locations*

Figure 4.6 *Illustration of simulated sonar scan image with the actual locations of the landmarks (green asterisk) and the estimated landmarks (pink circles) along with its corresponding PHD local maxima from the Particle-PHD clutter filter.*

in Section 4.4.1, the integral of the probability hypothesis density over the measurement space at any time instant k, gives an estimate of the number of landmarks N_k in the FOV. Extracting the local maxima from these PHD based on the estimates of

the number of landmarks, provides us with the most probable locations of these landmarks as illustrated in fig. 4.6. As is evident from the illustration, there are four landmarks present in the FOV (as seen from the simulated sonar scan, on the right), and these correspond to four local maxima of the PHD. However, there are instances when the number of landmarks N_k estimated by the PHD clutter filter differs from the actual number of landmarks. When N_k is less than the actual number of landmarks, (as in the case of scan illustrated in fig. 4.7b, it results in miss detection. The complete dynamic behavior of the clutter filter in estimating landmark



(a) Simulated sonar scan *(b) PHD peaks indicating the most probable landmark locations*

Figure 4.7 *Illustration of simulated sonar scan image with the actual locations of the landmarks (green asterisk) and the estimated landmarks (pink circles) along with its corresponding PHD local maxima from the Particle-PHD clutter filter. Illustrates miss detection due to error in estimation of number of landmarks by the PHD clutter filter.*

generated measurements is shown in fig. 4.8. Figure 4.9 shows the range-bearing estimates to the landmarks obtained by the PHD clutter filter superimposed with

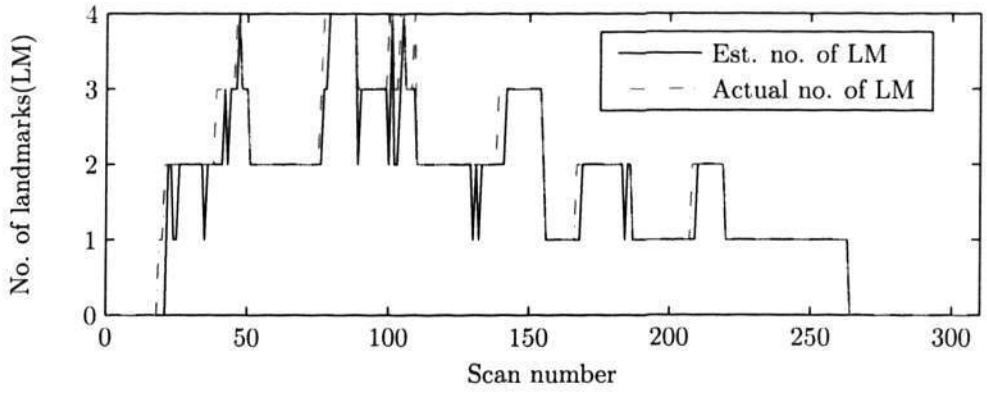


Figure 4.8 *Performance of the clutter filter based on the number of landmarks.*

the actual landmark generated measurements over time. Note that the filter has managed to effectively remove most of the clutter generated measurements.

4.6 Integrating PHD Clutter Filter with a Navigational (SLAM) Filter

In this section, a novel clutter rejection algorithm is presented based on the PHD clutter filter (discussed in the earlier section) for autonomous vehicle navigation. The novelty lies in the integration of the random finite set based PHD filter framework with the conventional vector based SLAM framework. In this implementation for the purpose of demonstration, we have chosen to integrate the PHD clutter filter with the conventional naive extended Kalman filter (EKF), though, it can be easily extended to more complex navigational frameworks. The integration of the PHD clutter filter with the conventional navigation filter is illustrated with the flowchart in fig. 4.10. In the proposed technique, the PHD clutter filter is used effectively

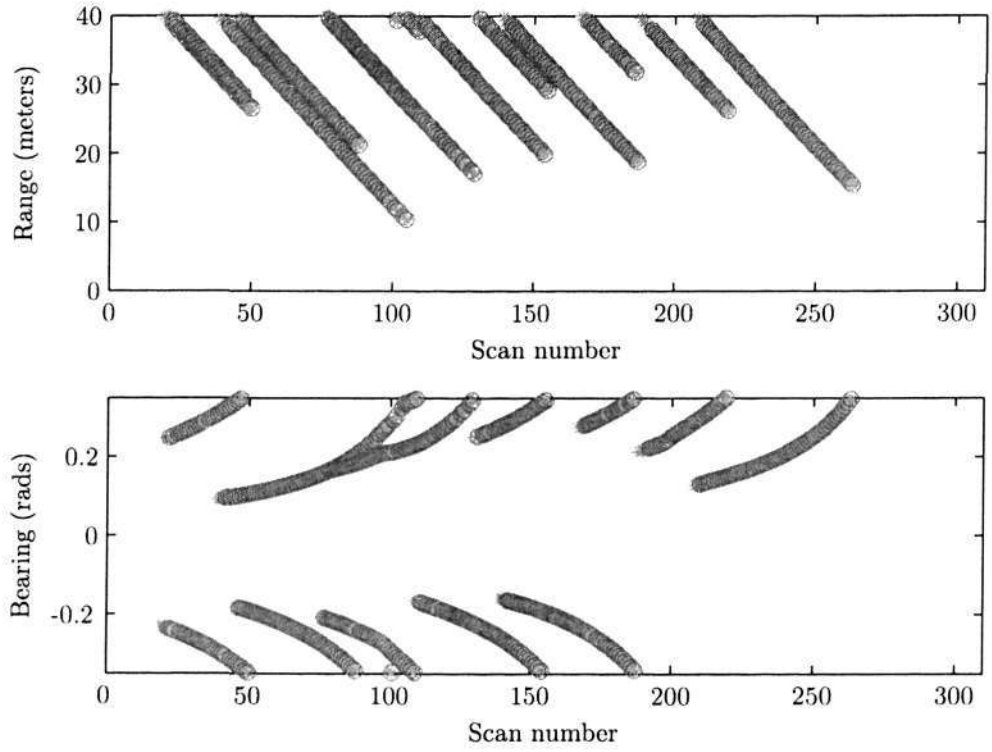


Figure 4.9 *Estimates of the range and bearing measurements from the PHD clutter filter (pink circles) superimposed with the actual landmark generated measurements (green asterisk).*

to reduce false measurements, thereby feeding mainly landmark originated measurements to the navigational filter. This effectively simplifies the data association technique needed within the navigation filter framework and completely obviates the need for external map/ feature management techniques. It is interesting to note that the proposed framework can be easily applied to any generic range-bearing-intensity (GRBI) sensor, even though our illustrations here are mainly limited to a blazed array sonar. An extended Kalman filter SLAM (EKF-SLAM) integrates the vehicle's navigation sensors to give an estimate of its own pose and at the same time retain the estimates of the previously observed landmarks to build a map as explained in Chapter 2. We have chosen the EKF framework for the SLAM imple-

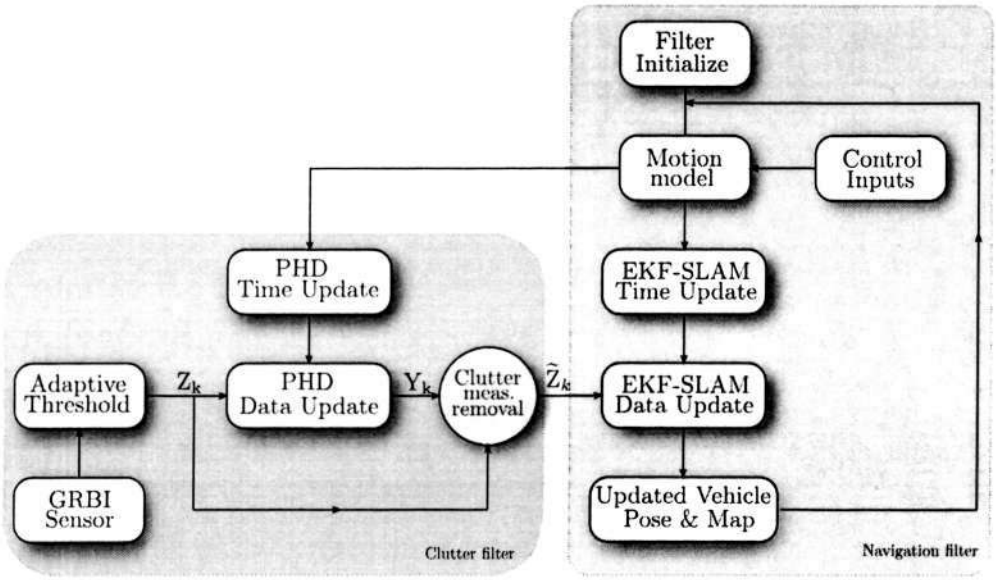


Figure 4.10 *Integration of the PHD clutter filter with the conventional EKF navigation filter.*

mentation because of the availability of the confidence bound (covariance) in the estimates without any additional computational requirements. It is a direct implementation of the seminal paper on SLAM [102], where the pose of the vehicle and the state of the landmarks are represented by a combined state vector x_{π_k} given by

$$x_{\pi_k} = \begin{bmatrix} x_k \\ M_k \end{bmatrix} \tag{4.14}$$

We explain the proposed technique detailing the integration of the two blocks, viz. clutter filter block and the navigation filter block, as illustrated in fig. 4.10 with the assistance of a case study with DVL and an angular rate sensor being used as navigation sensors providing the velocity and orientation of the underwater vehicle respectively with blazed array sonar as an imaging sensor mounted on the

underwater vehicle.

- **Vehicle Initialization**

  The initial vehicle location is assumed to known with certainty and is used to set the global reference frame G as shown in fig. 4.11. This frame is used in building the stochastic feature map. The state of the vehicle is given by

$$\mathbf{x}_k = [x_k \quad y_k \quad \psi_k]^T \tag{4.15}$$

  representing the position and orientation of the vehicle in the global reference frame.

- **EKF-SLAM time update**

  The motion model employed here uses the Doppler velocity log measure-

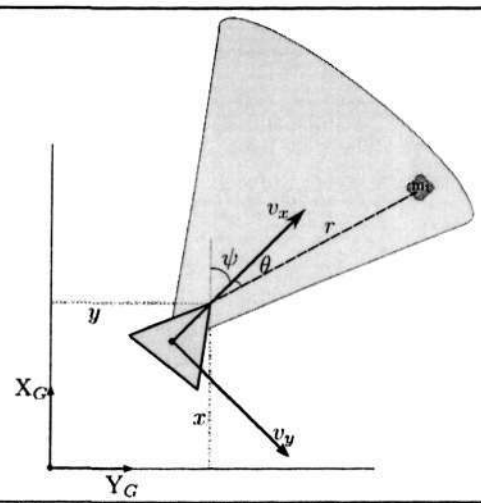


Figure 4.11 *Representation of the global and vehicle coordinate system. Vehicle is equipped with a blazed array sonar that provides a range r and a bearing θ measurement to a detected landmark.*

ments instead of controls to estimate posteriors overs poses. Practical experience suggests that DVL, while still erroneous, is usually more accurate than

the velocity motion models. Both methods suffer from drift, however velocity motion models additionally suffer from the mismatch between the actual motion controllers and its (crude) mathematical model. However with this model, the velocity information is available only after the vehicle has moved. Hence this kind of model is unusable for path planning and control, while it does not pose a problem for motion estimation, as in our case. Thus the motion model to predict the state of the vehicle is given by

$$\mathrm{x}_k = \begin{bmatrix} x_{k-1} + \Delta t(v_{x_{k-1}} cos(\psi_{k-1}) - v_{y_{k-1}} sin(\psi_{k-1})) \\ y_{k-1} + \Delta t(v_{x_{k-1}} sin(\psi_{k-1}) + v_{y_{k-1}} cos(\psi_{k-1})) \\ \psi_{k-1} + \Delta t v_{\psi_{k-1}} \end{bmatrix} \tag{4.16}$$

where, v_x, v_y and v_ψ are the linear and angular velocity measurements respectively in the vehicle coordinate frame obtained from the DVL and angular rate sensor respectively. The motion model can be written in short as a nonlinear equation,

$$\mathrm{x}_k = f_k(\mathrm{x}_{k-1}, \mathrm{u}_k) + \mathcal{N}(0, Q_k) \tag{4.17}$$

On the other hand, as landmarks correspond to fixed objects in the underwater environment, we assume that they are stationary and hence the predicted joint estimate is

$$\hat{\mathrm{x}}_{\pi_{k|k-1}} = [f_{k|k-1}(\hat{\mathrm{x}}_{k-1}, \mathrm{u}_k) \quad \hat{m}_{k-1,1} \ldots \hat{m}_{k-1,N_{k-1}}]^T \tag{4.18}$$

and its predicted covariance as

$$\mathrm{P}_{k|k-1} = \mathrm{F}_k \mathrm{P}_{k-1|k-1} \mathrm{F}_k^T + Q_k \tag{4.19}$$

where F_k is a Jacobian matrix as a result of linearization of the nonlinear function f.

- **Clutter filter**

  Point features are identified from the sonar scans returned by the blazed array sonar are fed as an input to the PHD clutter filter. The extraction of point features from the sonar data is essentially a two step process. The first step involves the detection of signals in presence of background noise. Adaptive threshold technique using OS-CFAR processor (discussed in Section 4.3) is used in the detection process to mainly identify the principal returns (landmark generated measurements). For each sonar scan the principal returns are then grouped together into clusters based on the regions of constant depth [62] and the cluster centroids are identified as point features. These point features which are the range-bearing measurements that encapsulate both landmark generated measurements and clutter generated measurements are fed as input to the PHD clutter filter. The details of the PHD clutter filter are explained in Section 4.4. The output of the filter is a set of clutter reduced (CR) measurements $\widetilde{\mathrm{Z}}_k$.

- **EKF-SLAM measurement update**

  The CR-measurements $\widetilde{\mathrm{Z}}_k$ are transformed from the sonar sensor coordinate frame to the vehicle coordinate frame, though not explicitly represented here for the sake of simplicity. Given the current vehicle pose x_k, the i^{th} feature at time k corresponds to a j^{th} landmark, $\mathrm{m}_{k,j} = [m_{k,j_x} \quad m_{k,j_y}]^T$ in the map. Hence

the resulting new measurement model is given by,

$$\begin{aligned}\tilde{\mathrm{z}}_{k,i} &= h_j(\mathrm{x}_k, \mathrm{m}_{k,j}) + \mathcal{N}(0, R_k) \\ &= \begin{bmatrix} r_{k,i} \\ \theta_{k,i} \end{bmatrix} = \begin{bmatrix} \sqrt{(x_k - m_{k,j_x})^2 + (y_k - m_{k,j_y})^2} \\ \tan^{-1}\left(\frac{(y_k - m_{k,j_y})}{(x_k - m_{k,j_x})}\right) - \psi_k + \frac{\pi}{2} \end{bmatrix} + \begin{bmatrix} w_{r_k} \\ w_{\theta_k} \end{bmatrix}\end{aligned} \tag{4.20}$$

The next step is to determine whether the new CR-measurement $\tilde{\mathrm{z}}_{k,i}$ corresponds to any landmarks $\mathrm{m}_{k,j}$, $j = 1, \ldots, N_{k-1}$ already existing in the map and should be used to update the state vectors or, on the contrary a new landmark and has to be incorporated into the map. Finding this association hypothesis is a process involving analysis of the discrepancy between the CR-measurements and its prediction. This is obtained by calculating the innovation and its associated covariance as

$$\upsilon_k^{ij} = \tilde{\mathrm{z}}_{k,i} - h_{k,j}(\hat{\mathrm{x}}_{\pi_{k|k-1}}) \tag{4.21}$$

$$\mathrm{S}_k^{ij} = \mathrm{H}_{k,j} P_{k|k-1} \mathrm{H}_{k,j}^T + R_k \tag{4.22}$$

where $\mathrm{H}_{k,j}$ is a Jacobian matrix as a result of linearization of the nonlinear function $h_{k,j}$. A simplest Nearest Neighbor association is taken in this case to be the closest association in a statistical sense. A common statistical discriminator is based on the normalized innovation squared between two estimates. The normalized innovation squared between the pseudo measurement and the estimated landmark location (referred to as Mahanabolis distance) is compared against a validation gate, d_{min}, for the association being considered. The normalized innovation squared forms a χ^2-distribution that can be used

to accept or reject a particular association with a given confidence level by the appropriate selection of d_{min}.

$$\varepsilon_k = {v_k^{ij}}^T {S_k^{ij}}^{-1} v_k^{ij} < d_{min} \tag{4.23}$$

Once the association is made, it is used to update the state estimates using standard EKF update equations.

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k^{ij} v_k^{ij} \tag{4.24}$$

$$P_k = [I - K_k^{ij} H_{k,j}] P_{k|k-1} \tag{4.25}$$

where,

$$K_k^{ij} = P_{k|k-1} {H_{k,j}}^T {S_k^{ij}}^{-1} \tag{4.26}$$

4.7 Algorithm Performance Validation in Offshore Underwater Environments

In this section, the proposed navigational framework wherein, a particle PHD clutter rejection filter is integrated with the EKF is implemented in an offshore underwater environment using a blazed array sonar as the primary imaging sensor. The data was acquired from a research testbed AUV, during a survey conducted along the coastlines of Singapore. The details of the system hardware and software architecture of the navigational modules along with specifications of the sensors are provided in the Appendix B. The aim of these tests was to validate the pro-

posed technique for its clutter rejection performance and benchmark it against the EKF-SLAM filter with an external map management filter.

4.7.1 Experimental Setup

In these experiments, the sonar targets (see fig. 4.12) were initially deployed in the underwater environment in which the vehicle was proposed to operate. An array consisting of five octahedron sonar reflectors were held together with strings and moored to the bottom of the sea bed using weights as shown in fig. 4.12, for maximizing the returns from the blazed array sonar. A marker buoy was attached to the array, which was suspended about 2m above the reflectors, holding the array in the water column. Four such arrays were deployed with the first three arrays about 18m apart and the fourth array about 30m apart and about $\sim$ 3m above the sea bed. These acted as identifiable and stable features. Naturally occurring objects such as coral reefs or rocky seabed if present, were also classified as a point features. If these features were stable they are also incorporated into the map. The AUVs mission was pre-planned and the way points were selected in such a manner, so that it traverses in a path that maximizes the visibility of the landmarks (reflectors) in the FOV of the blazed array sonar. The AUV was initialized using GPS information while at the surface and traversed along the way points maintaining a height of 5m from the seabed in a water column averaging 10m. The way points given to the AUV are listed in table 4.1 and the intended locations of the reflector arrays are tabulated in table 4.2. The vehicle orientation estimates from the fiber optic gyroscope during the experimental run is shown in fig. 4.13.

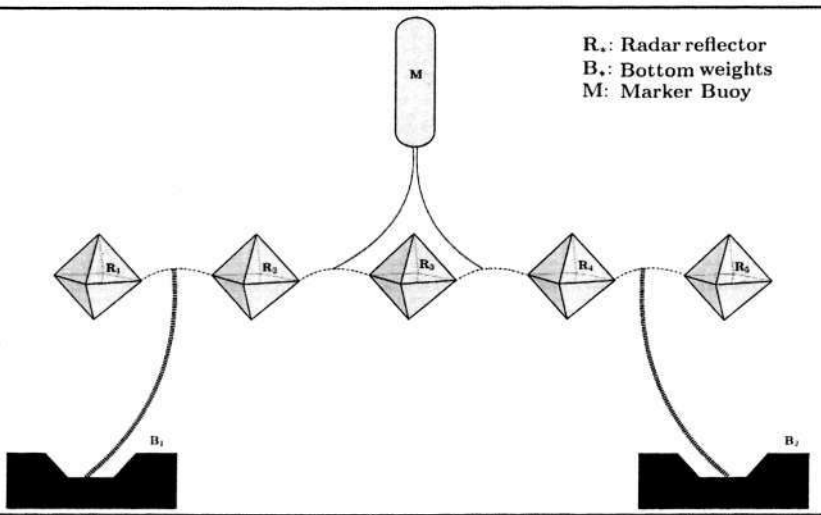


Figure 4.12 *Target setup: Five sonar reflectors tied in-line are suspended from the sea bed*

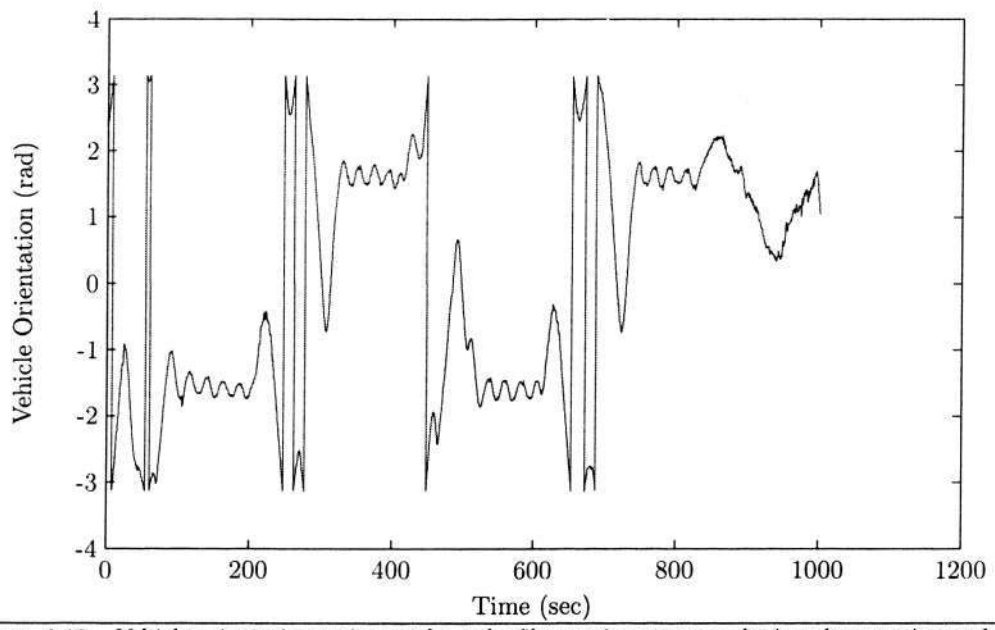


Figure 4.13 *Vehicle orientation estimates from the fiber optic gyroscope during the experimental run*

GPS co-ordinates		Global Map co-ordinates	
Latitude (deg)	Longitude (deg)	East (m)	North (m)
1.315810N	104.01672E	0	0
1.315240N	104.017416E	91.703550	-78.65044
1.315052N	104.017566E	108.45001	-99.64391
1.315053N	104.018979E	266.20163	-99.53214
1.315199N	104.019099E	279.59878	-83.22868
1.314864N	104.019095E	279.15225	-120.63728
1.315053N	104.018987E	267.09478	-99.53214
1.315046N	104.017484E	99.295290	-100.31392
1.314898N	104.017188E	66.248950	-116.84071
1.314686N	104.017326E	81.655698	-140.51421
1.315046N	104.017484E	99.295290	-100.31392
1.315053N	104.018991E	267.54135	-99.532148
1.315203N	104.019155E	285.85079	-82.782012
1.314872N	104.019150E	285.29261	-119.74394
1.315053N	104.018991E	267.54135	-99.532148
1.315046N	104.017484E	99.295290	-100.31392

Table 4.1 *Way points given to the AUV: GPS co-ordinates and its corresponding global map co-ordinates*

Landmark Id	East (m)	North (m)
Reflector Array 1	153.8581	-103.7717
Reflector Array 2	171.4977	-103.7717
Reflector Array 3	189.5839	-101.7616
Reflector Array 4	222.5186	-104.2183

Table 4.2 *Global map co-ordinates of the deployed reflector arrays*

4.7.2 Experimental Results

The results are benchmarked against EKF-SLAM algorithm with an external map management filter. The map management filter is based on the landmark quality algorithm [32] which rejects outliers by maintaining a tentative landmark list. Thus, instead of augmenting the map based on the measurement indicating the existence of a new landmark, it is first added to a tentative list of landmarks. Once a landmark has been consistently observed, it is transferred from the tentative list to the confirmed list to be estimated as a part of the EKF-SLAM algorithm. This technique tends to reduce the number of landmarks in the map by a significant factor as no-

ticed in the fig. 4.14. However, it is important to emphasize that the tentative list to be maintained, (i.e., the number of times (in our case 4) the landmarks have to be observed before being incorporated into the map) is an *ad hoc*, problem dependent parameter, that needs to be adjusted based on environmental conditions.

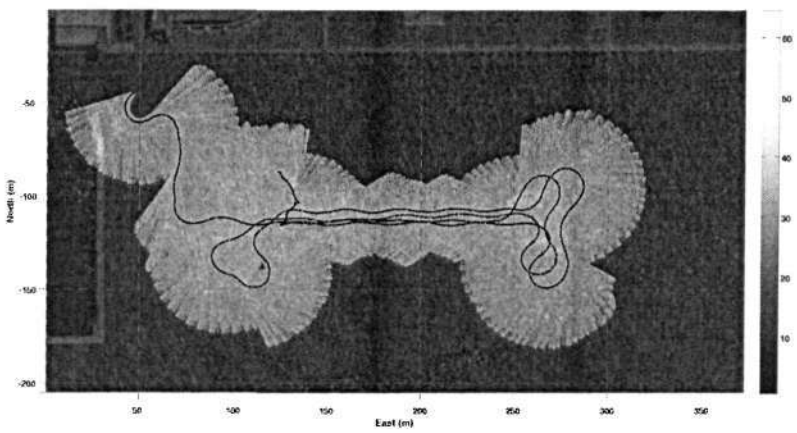


Figure 4.14 *Feature returns and the vehicle trajectory obtained from EKF-SLAM filter superimposed on satellite images of the experimental site*

Figure 4.14 shows a plot of the final map obtained by the EKF-SLAM algorithm with an external map management filter. It has been overlaid on the satellite image for a better understanding of the environment. The absolute location of all the potential point landmarks identified based on the sonar returns are also overlaid on the blazed array sonar scan images. These locations were computed using the estimated vehicle location at the instant of the corresponding sonar return. The lighter sections in the sonar scan images indicate stronger intensity returns indicated by the color bar in fig. 4.14. The sonar returns seen near the top left corner

of the fig. 4.14, illustrates the stronger returns obtained from the breakwaters that were present at the experimental site. These were characterized as point features and were also incorporated into the map. However it can be clearly observed that apart from landmark originated measurements, there are a lot of features that are incorporated into the map which are mainly clutter. While further fine tuning of the map management filter (varying the number of times the landmarks have to be observed before being incorporated into the map) reduces the clutter originated measurement however, it also tends to remove the landmark originated measurements. Moreover, this fine tuning has to be performed for each and every experiment. The proposed integration particle PHD clutter filter with the EKF-SLAM navigation filter not only rejects the clutter efficiently it also obviates the need for *ad-hoc* map management methods. Moreover, the PHD clutter filter implicitly has also performed two difficult tasks. First, it has intrinsically handled the calculation of the expected number of landmarks in the region under observation, and second it has located the most probable landmark locations without performing any explicit data-to-track association. The output of the PHD clutter filter has been used to define the validation gate in the measurement space. The measurements that fall outside the validation gate are discarded as clutter, thereby eliminating most of the clutter originated measurements. Moreover, since the CR measurements are fed into the EKF-SLAM navigation filter, the computational costs (quadratic with the number of landmarks) associated with filter are also reduced. Figure 4.15 shows the result of integrating the particle PHD filter with the EKF-SLAM algorithm.

Careful observation of fig. 4.15 reveals that the proposed technique has efficiently managed to reject the clutter and propagate the landmark based measurements in the map. Since there is no absolute position sensor on the underwater vehi-

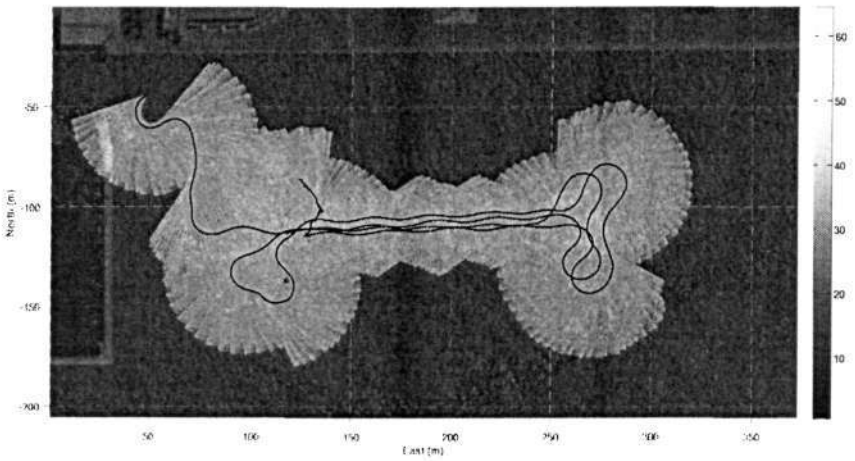
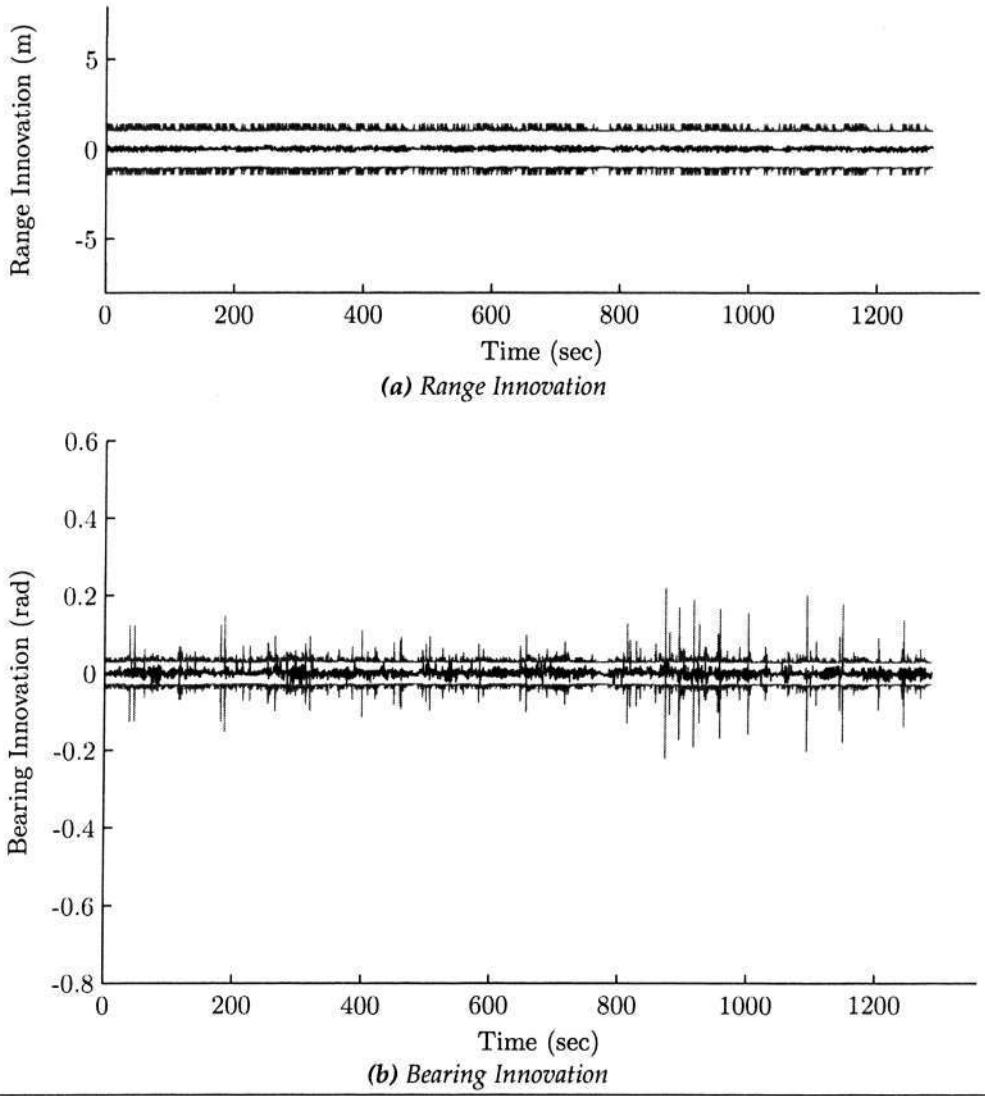


Figure 4.15 *Feature returns and the vehicle trajectory obtained from EKF-SLAM filter integrated with the particle PHD clutter filter superimposed on satellite images of the experimental site*

cle, the performance of the vehicle positioning cannot be measured against ground truth. To verify the performance of the proposed filter, the innovation sequence can be monitored to check the consistency of the estimates. The innovation sequence for the proposed algorithm is within the $2-\sigma$ covariance bounds, as observed in fig. 4.16.

The vehicle pose covariance estimates indicate that the proposed filter yields tighter covariance bounds over to the map managed version of the EKF-SLAM algorithm as observed in fig. 4.17. This is mainly attributed to clutter originated unstable measurements that have been incorporated into the state vectors, which result in data association errors. However, using more sophisticated data association strategies such as joint compatibility branch and bound (JCBB) [86], multiple frame data association (MFDA) [114] would lead to higher localization accuracy,



(a) Range Innovation

(b) Bearing Innovation

Figure 4.16 *The range and bearing innovation sequences plotted against their $2-\sigma$ confidence bounds. The innovation is plotted as a blue line while the confidence bounds are the red lines.*

while, using particle PHD filter as a clutter rejection filter and integrating it with the standard EKF-SLAM filter alleviates the need for such sophisticated data association strategies.

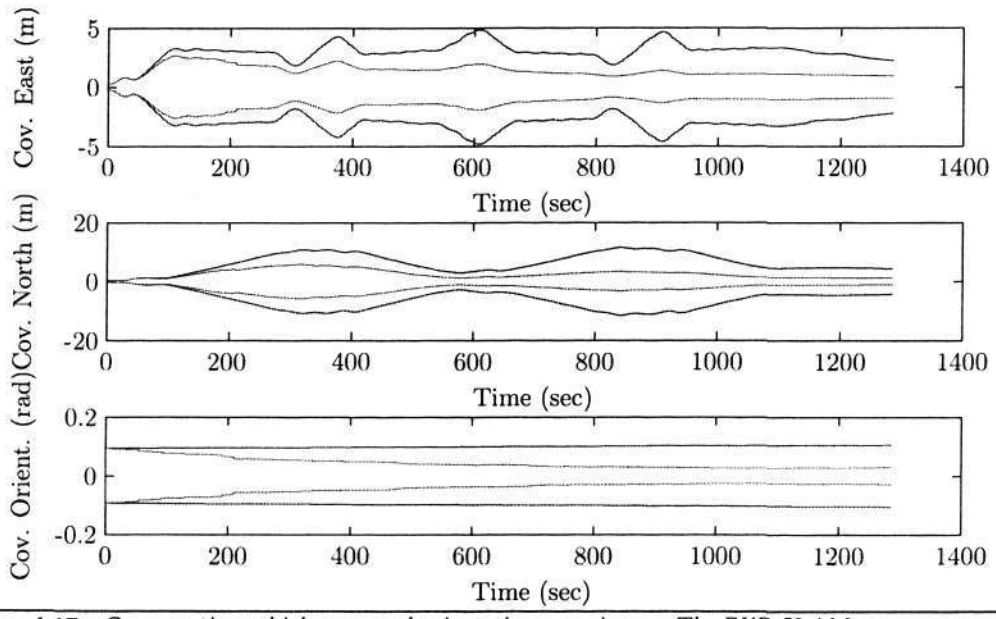


Figure 4.17 *Comparative vehicle pose and orientation covariances. The EKF-SLAM map management filter (blue) yields a poorer covariance estimate than the integrated particle PHD EKF-SLAM filter (red). This is mainly attributed to the data association mismatch occurring due to unstable features.*

4.8 Chapter Summary

This chapter addressed the issues of landmark detection from noisy sensor measurements and its impact on autonomous vehicle navigation applications. The fixed threshold and the adaptive threshold method with CFAR using order statistics were introduced for performing landmark detection with a blazed array sonar. The PHD clutter rejection filter was introduced along with its implementation details. The performance of the particle PHD clutter rejection filter was validated using synthetic sonar data. A novel autonomous navigation framework was introduced where a particle PHD clutter rejection filter was integrated with an EKF-SLAM algorithm. This proposed navigation framework was applied then applied in offshore underwater field trials and benchmarked against the standard EKF-SLAM

filter with a feature map management strategy. It presented the first instance of a deployable underwater implementation of the proposed algorithm. The proposed algorithm has the capability to be applied to any environment (land based indoor, outdoor environment and airborne) and incorporate measurements obtained from other sensing equipments such laser, radar, optical cameras although this implementation here is restricted to underwater environments and blazed array sonar as the sensing equipment. The next chapter presents a completely unified finite set statistical framework for autonomous navigation that explicitly accounts for measurement detection uncertainty, data association uncertainty, false alarms and map management.

5

A Unified Approach to Localization and Mapping using Finite Set Statistics

The previous chapter addressed the landmark detection problem in the presence of clutter using random finite sets and its incorporation within the autonomous vehicle navigation framework. This chapter presents a completely unified random finite set theoretic approach to the localization and the mapping problem. It shows the pitfalls of the conventional SLAM solutions that are founded on vector based stochastic framework, where the sensor models and the vehicle models are represented in state space form and the joint posterior or its statistics are obtained based on recursive Bayesian estimation. All of these SLAM solutions leading from the stochastic vector state-space approach require that we solve certain parallel problems in each recursion. These include effective solutions to the problems of data association, feature extraction, clutter filtering and landmark or map management.

The approach presented in this chapter is inspired by the recent developments in the target-tracking community where random finite sets have been perceived as a more natural representation of the multisource-multitarget problems [70]. Based on the finite set statistics (FISST) developed by Mahler, we presented the random finite set models within the autonomous vehicle navigation framework in Chapter 3. We use these models in this chapter and propose a novel unified random finite set theoretic framework, where the SLAM problem is reformulated so that the landmark map and the measurements are represented using random finite sets and the landmark map is jointly estimated with the vehicle state vector, whilst explicitly accounting for measurement detection uncertainty, data association uncertainty, false alarms and map management in the SLAM filter framework. It is henceforth referred to as FISST-SLAM. In this chapter, we also present and use a miss-distance metric based on Wasserstein distance measure, to quantify for the errors in landmark estimation within the autonomous vehicle navigation framework.

This chapter is organized as follows. Section 5.1 provides a background on the applications of RFS in various fields. In Section 5.2, we begin with the formulation of the SLAM problem from a random finite set theoretic framework. Thereafter, a solution to the joint vehicle and map posterior is obtained by decomposing it to a vehicle component and a map component. Section 5.3 describes the vehicle path posterior estimation using a particle filter approach. In Section 5.4 the map posterior estimation using an RFS optimal Bayes filter is explained. Thereafter an approximation of the RFS optimal Bayes filter by a first order statistical moment filter is described in Section 5.5. Section 5.6 and 5.7 explain the two different implementation methods of the proposed FISST-SLAM algorithm based on sequential Monte-Carlo methods and Gaussian Mixture model methods respectively. The implementational

issues with the proposed FISST-SLAM algorithm along with their computational complexities are discussed in Section 5.8. Sections 5.9 and 5.10 benchmark the proposed algorithm against FastSLAM using simulations and offshore underwater experiments respectively. In Section 5.11 we summarize and also discuss the failure modes of the proposed FISST-SLAM algorithm and techniques that can be used to address these issues.

5.1 Background

As presented in Chapter 2, the most common formulation of the SLAM problem is founded on a vector based stochastic framework, where the sensor models and the vehicle models are represented in state space form and the estimation of joint posterior or its statistics based on Bayes theorem [102], [80]. Rooted in this paradigm, the naive extended Kalman filter SLAM is the very first SLAM solution and since then the research community has made significant progress in this area. Amongst these the most notable are SLAM algorithms catering to non - linear process and measurement models and non - Gaussian pose distributions [76], computationally and memory efficient SLAM solutions, large scale SLAM [45], analysis of convergence, consistency [31] and observability of SLAM [59]. All of these SLAM solutions leading from the stochastic vector state space approach require that we solve certain parallel problems outside the SLAM Bayes recursion. These include effective solutions to the problems of data association [114], [86], feature extraction [95], clutter filtering [64] and landmark or map management [32]. A careful examination reveals that the need to solve these parallel problems outside the Bayes recursion stems primarily from the limitations of the representation of the map model, and

map measurement model used. More specifically, in the selection of the sensor model there is an implicit assumption that the number of landmarks are known a priori, the measurements are free from clutter, and there is neither detection nor data association uncertainty.

The approach presented in this chapter is inspired by the recent developments in the target-tracking community where random finite sets have been perceived as a more natural representation of the multisource-multitarget problems [69], [70]. Due to its versatility, ability to handle non-linear and variable number of targets and easy implementations of the first order statistical moment filter, it has been applied in various fields ranging from tracking multiple moving targets in terrain [98], using 3D-LIDAR (Velodyne) [60], [55], to detecting and tracking underwater objects [54], [27]. Other notable applications of the random finite set theory are in passive coherent location of targets observed from multiple bi-static radars [106], tracking corner features in optical image sequences [50], and tracking human figures in digital video [112] tracking an unknown time-varying number of speakers [67]. More recently, Mullane et. al [81] have presented an RFS approach to robotic pose estimation and mapping problems. Here the collection of vehicle state, the number of landmarks and their states are modeled as one finite random set and a PHD filter implementation is used to estimate the first order statistic of their joint posterior. As such it does not exploit the structure of the SLAM problem in the formulation as the one presented in this chapter. SLAM structure provides for the variable element map set (landmarks and their states) to be estimated separately conditioned on the vehicle path, thereby yielding an elegant and simpler realization of a random finite set theoretic SLAM filter.

5.2 Formulation of Bayesian Random Finite Set Mapping and Localization problem

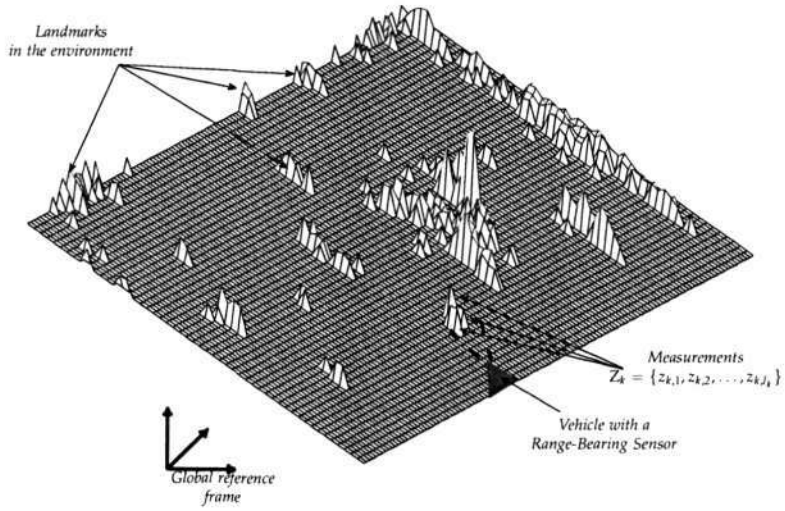


Figure 5.1 *An Illustration of Mapping and Localization Problem in random finite set framework. The vehicle is equipped with a sensor with a limited field of view obtaining measurements, which is a collection of landmark generated measurements and false alarms. These measurements are represented by an RFS Z_k.*

Let $U^{k-1} = [u_0, u_1, ..., u_{k-1}]^T$, denote the time sequence of controls applied to the vehicle up to time $k-1$, where u_k is the control input at time k. Let the map containing N_k static landmarks be represented by a set $M_k = \{m_{k,1}, \ldots, m_{k,N_k}\}$ at time k. Let the time sequence of the sets of measurements taken by a generic range-bearing-intensity sensor on-board the vehicle be denoted by $Z^{(k)} = Z_1, Z_2, \ldots, Z_k$, where Z_k is the set of measurements collected at time k , as shown in fig. 5.1. It is important to emphasize that the measurements are now represented by random finite sets and not in the classical vector state space representation as in Chapter 2. Let $X^k = [x_1, \ldots, x_k]$ denote the vehicle's complete pose history, where x_k represents vehicle's pose at time k. Now the SLAM problem requires that the probability

distribution

$$f_{k|k}(\mathrm{X}^k, \mathrm{M}_k | \mathrm{Z}^{(k)}, \mathrm{U}^k, \mathrm{x}_0) \tag{5.1}$$

be computed for all times k with x_0 being the initial state of the vehicle. This probability distribution describes the joint posterior density of the map state and vehicle path at time k given all the past observations and control inputs along with the initial vehicle state. The initial state of the vehicle x_0 is assumed to be known and hence will not be indicated. The problem is therefore to recursively calculate the probability of the vehicle path X^k and the map state M_k at time k, given the state of the vehicle path, X^{k-1} and the map state M_{k-1} at time $k-1$, and the set of all past control inputs U^k and the set of all past observations taken up to time k, $\mathrm{Z}^{(k)}$. The solution to this problem requires appropriate probabilistic models of the vehicle motion, map evolution and the observation process.

Using the chain rule of probability we decompose the joint SLAM posterior into a product of a vehicle component and a map state component as follows,

$$f_{k|k}(\mathrm{X}^k, \mathrm{M}_k | \mathrm{Z}^{(k)}, \mathrm{U}^k) = f_{k|k}(\mathrm{X}^k | \mathrm{Z}^{(k)}, \mathrm{U}^k) f_k(\mathrm{M}_k | \mathrm{X}^k, \mathrm{Z}^{(k)}, \mathrm{U}^k) \tag{5.2}$$

An important consequence of this decomposition is that it simplifies the joint posterior estimation into two separate estimation problems. That is one involving estimating the vehicle posterior $f_{k|k}(\mathrm{X}^k | \mathrm{Z}^{(k)}, \mathrm{U}^k)$ and the other estimating the map posterior conditioned on the vehicle path $f_{k|k}(\mathrm{M}_k | \mathrm{X}^k, \mathrm{Z}^{(k)}, \mathrm{U}^k)$. We use a particle filter to compute the vehicle path posterior $f_{k|k}(\mathrm{X}^k | \mathrm{Z}^{(k)}, \mathrm{U}^k)$, where the posterior distribution is represented by set of N_s particles, $\{\mathrm{X}^{k,[i]}\}_{i=1}^{N_s}$. The map posterior $f_{k|k}(\mathrm{M}_k | \mathrm{X}^k, \mathrm{Z}^{(k)}, \mathrm{U}^k)$, represented by the set of maps $\{\mathrm{M}_k^{[i]}\}_{i=1}^{N_s}$ is estimated using a random finite set theoretic approach. Note that each of the elements, $\mathrm{M}_k^{[i]}$ of this set

of maps represents the map conditioned on the path of the i^{th} vehicle trajectory posterior $X^{k,[i]}$. Further, the fact that the map posterior estimation is now conditioned on the vehicle path, it follows that the measurements of the individual landmarks that comprise the map are independent of one another. This fact is exploited in constructing the multi-landmark measurement likelihood functions from individual landmark likelihood functions.

5.3 Path Posterior Estimation

FISST-SLAM employs a particle filter for estimating the path posterior $f_{k|k}(X^k|Z^{(k)}, U^k)$. At each time step k, a set of particles $\mathcal{P}_k$ representing the posterior is maintained. Each particle $X^{k,[i]}$ represents a "guess" of the vehicle's path

$$\mathcal{P}_k = \{X^{k,[i]}\}_{i=1}^{N_s} = \{x_1^{[i]}, x_2^{[i]}, \ldots, x_k^{[i]}\}_{i=1}^{N_s} \tag{5.3}$$

The superscript $^{[i]}$ denotes the i-th particle in the set $\mathcal{P}_k$ and is calculated from the set $\mathcal{P}_{k-1}$ at time $k-1$, using the control input u_k and the measurements Z_k. For each particle, a *proposal distribution* is computed, conditioned on the specific particle history and a sample is drawn from it,

$$x_k^{[i]} \sim \pi(x_k|X^{k-1,[i]}, Z^{(k)}, u_k) \tag{5.4}$$

More specifically, as in FastSLAM 1.0, we sample the pose x_k in accordance with the $i-$th particle, by drawing a sample according to the motion posterior

$$x_k^{[i]} \sim f_{k|k-1}(x_k|x_{k-1}^{[i]}, u_k) \tag{5.5}$$

Here, $\mathrm{x}_{k-1}^{[i]}$ is the posterior estimate of the vehicle location at time $k-1$, residing in the i-th particle. The resulting sample $\mathrm{x}_{k}^{[i]}$ is then added to a transient set of particles, along with the pose history up to time $k-1$, $\mathrm{X}^{k-1,[i]}$. We then draw N_s particles according to a yet to be defined importance weight. The resulting set of N_s forms the new and final particle set $\mathcal{P}_k$.

This particle set $\mathcal{P}_k$ that accounts for the new measurements is obtained by sampling importance re-sampling, i.e., each particle from the temporary set is assigned a weight in accordance with an *importance function* which is calculated as follows,

$$\begin{aligned} v_k^{[i]} = \frac{\text{target distribution}}{\text{proposal distribution}} &= \frac{f_{k|k}(\mathrm{Z}_k|\mathrm{X}^{k,[i]},\mathrm{Z}^{(k-1)})f_{k|k-1}(\mathrm{x}_k^{[i]}|\mathrm{x}_{k-1}^{[i]},\mathrm{u}_k)}{f_{k|k-1}(\mathrm{x}_k^{[i]}|\mathrm{x}_{k-1}^{[i]},\mathrm{u}_k)} v_{k-1}^{[i]} \\ &= f_{k|k}(\mathrm{Z}_k|\mathrm{X}^{k,[i]},\mathrm{Z}^{(k-1)})v_{k-1}^{[i]} \end{aligned} \tag{5.6}$$

5.4 RFS Map Posterior Estimation

Due to the limited field of view (FOV) of the sensor and occlusions, the landmark map will not remain static over time. The motion of the vehicle causes new landmarks to appear and some landmarks to disappear in the limited field of view of sensor, resulting in the number of landmarks in the map to change randomly over time. A natural and more effective form of representation of the landmark map is by a random finite set (RFS), M_k,

$$\mathrm{M}_k = \{\mathrm{m}_{k,1},\ldots,\mathrm{m}_{k,i},\ldots,\mathrm{m}_{k,N_k}\} \in \mathcal{F}(\mathcal{M}), \tag{5.7}$$

where $\mathcal{F}(\mathcal{M})$ is a collection of all finite subsets of the space of maps $\mathcal{M}$ and $N_k = |M_k|$ is the number of elements or the cardinality of M_k representing the number of landmarks detected and observed at time k and $\mathrm{m}_{k,i}$ representing the state of the i^{th} landmark at time k. The measurements collected from the sensor at time k is also represented by a random finite set Z_k,

$$Z_k = \{z_{k,1}, \ldots, z_{k,i}, \ldots, z_{k,l_k}\} \in \mathcal{F}(\mathcal{Z}) \tag{5.8}$$

where $l_k = |Z_k|$ is the number of measurements (function of time k) collected from the sensor, at time k and $\mathcal{F}(\mathcal{Z})$ denotes all the subsets of the space of measurements $\mathcal{Z}$. We denote the sequence of the sets of such observations up to time k by $Z^{(k)} = \{Z_1, \ldots, Z_k\}$. Therefore the problem now becomes one of estimating (or updating) the multi-landmark map posterior conditioned on the vehicle path[1], for each particle $i \in N_s$, i.e.,

$$f_{k|k}(\mathrm{M}_k^{[i]} | \mathrm{X}^{k,[i]}, Z^{(k)}, \mathrm{U}^k) \tag{5.9}$$

To estimate this posterior of maps conditioned on the vehicle path we use a finite random set theoretic approach. We use the map evolution model and the measurement model that have been derived in Chapter 3 using FISST [69].

5.4.1 RFS Multi-landmark Map Evolution Model

Let $\zeta_{k-1} \in \mathrm{M}_{k-1}$[2] represent an individual landmark state at time $k-1$. Now this landmark may or may not appear in the sensor's view in the next time instant k, due to the limited FOV of the sensor and vehicle motion. The landmark $\zeta_{k-1} \in \mathrm{M}_{k-1}$

[1] Henceforth in this section, for clarity we will drop the suffix i denoting the particle

[2] $\zeta_{k-1} = \mathrm{m}_{k-1,i}$, where $i = 1, \ldots, N_{k-1}$

will appear in the next time instant k with a probability $p_S = p_{S,k}(\zeta_{k-1}, \mathrm{x}_k)$ or disappear with a probability $1 - p_{S,k}(\zeta_{k-1}, \mathrm{x}_k)$. For a landmark that is currently (time $k-1$) in view this probability represents its probability of persistence or survival, and for a landmark not in the current field of view, it represents the probability of reappearance. Conditional on the persistence of the landmarks at time k, the probability density of a transition from state ζ_{k-1} to ζ_k is given by a Markov process, $f_{k|k-1}(\zeta_k|\zeta_{k-1}, \mathrm{x}_k)$ (single landmark motion model). For a landmark $\zeta_{k-1} \in \mathrm{M}_{k-1}$ at time $k-1$, its state at the next time instant k is modeled as a RFS $S_{k|k-1}(\zeta_{k-1})$. The limited FOV sensor coupled with vehicle motion also results in new landmarks coming into view at time k. Let us denote these new appearances by an RFS Γ_k. Thus, the evolution of the map of landmarks is given by,

$$\mathrm{M}_k = \left[\bigcup_{\zeta_{k-1} \in \mathrm{M}_{k-1}} S_{k|k-1}(\zeta_{k-1}) \right] \cup \Gamma_{\mathbf{k}} \tag{5.10}$$

whose probability density function, $f_{k|k-1}(\mathrm{M}_k|\mathrm{M}_{k-1}, \mathrm{X}^k)$ denotes the RFS multi-landmark dynamic map evolution model as derived in the Section 3.2.

5.4.2 RFS Multi-landmark Map Measurement Model

In the classical stochastic vector based approach to SLAM, the measurement model is represented as a non-linear function mapping g_k of the vehicle pose vector x_k and the map state vector M_k into the measurement vector Z_k

$$\mathrm{Z}_k = g_k(\mathrm{x}_k, \mathrm{M}_k) + w(k) \tag{5.11}$$

where $w(k)$ is the additive measurement noise.

It is evident from the classical vector based multi-landmark map measurement model (eqn. 5.11) used in conventional SLAM that the effects due to detection uncertainty of sensor, filtering of the raw sensor measurements to remove clutter, and association of the filtered measurements to landmarks in the map need to be handled before performing a SLAM measurement update. As a consequence, in the conventional approach it is necessary to address the issues arising from detection uncertainty (missed detections), clutter or false alarms (measurements not due to landmarks) and data association (landmark to measurement association) by other explicit means before a measurement update is executed. By adopting a finite random set approach, we can construct a more natural multi-landmark (map) measurement model that explicitly accounts for detection uncertainty, clutter and unknown data association for map update.

Suppose the particular sensor mounted on the vehicle has a probability of detection p_D. It is obvious that if this sensor has a limited field of view (FOV), then p_D will be a function of both the landmark state $\zeta_k \in \mathrm{M}_k$ and the vehicle state x_k, i.e. $p_D = p_D(\zeta_k, \mathrm{x}_k)$. We represent the measurement generated by the sensor of a single landmark $\zeta_k \in \mathrm{M}_k$ at time k by an RFS $\Theta_k(\zeta_k)$. Now $\Theta_k(\zeta_k)$ will be a singleton set $\{z\}$ with probability $p_D(\zeta_k, \mathrm{x}_k)$ and the null set $\emptyset$ with probability $1 - p_D(\zeta_k, \mathrm{x}_k)$. In addition to the landmark originated measurements, there can be false detections due to false alarms and clutter. We represent these false detections by the RFS Ω_k. Given the predicted map state M_k, and the vehicle state x_k, the random observation set Z_k, consisting of landmark generated measurements along with missed detections, false alarms and/or clutter will have the form

$$Z_k = \Omega_k \cup \left[\bigcup_{\zeta_k \in M_k} \Theta_k(\zeta_k) \right] \tag{5.12}$$

whose probability density function $f_k(Z_k|M_k, x_k)$, also called as the sensor likelihood function is as derived in the Section 3.3.

5.4.3 Optimal Bayes RFS Multi-landmark Map Estimation Filter

Now that we have established the RFS multi-landmark map evolution and measurement models, we use the optimal RFS Bayes recursion [69] to estimate the map posterior conditioned on the vehicle path posterior. Thus given the map evolution model $f_{k|k-1}(M_k|M_{k-1}, x_k)$ and measurement model $f_k(Z_k|M_k, x_k)$, the multi-landmark Bayes recursion is as follows:

$$f_{k|k-1}(M_k|Z^{(k-1)}, X^k, U^k) = \int f_{k|k-1}(M_k|M_{k-1}, X^k).f_{k-1|k-1}(M_{k-1}|Z^{(k-1)}, X^{k-1}, U^{k-1})\delta M_{k-1} \tag{5.13}$$

$$f_{k|k}(M_k|X^k, Z^{(k)}, U^k) = \frac{f_k(Z_k|M_k, x_k) f_{k|k-1}(M_k|X^k, Z^{(k-1)}, U^k)}{f_k(Z_k|X^k, Z^{(k-1)}, U^k)} \tag{5.14}$$

where, $f_k(Z_k|X^k, Z^{(k-1)}, U^k)$ is the Bayes normalization factor and $\int f(\cdot)\delta(\cdot)$ denotes the set integral. The multi-landmark map evolution function (Markov density) and the multi-landmark likelihood function are constructed from the belief-mass function as explained in Sections 3.2 and 3.3 respectively.

As in the classical vector based optimal Bayes recursion discussed in Chapter 2, the optimal Bayes RFS recursion for map posterior estimation (as in eqns. 5.13 -

5.14) involves multiple integrals, which are computationally intractable. Moreover, the complexity of computing this recursion using numerical methods grows exponentially with the number of landmarks and so the optimal Bayes RFS filter must be approximated. To alleviate the complexity of computing the multi-landmark map posterior, a recursion was derived for the first order statistical moment of the multi-landmark posterior distribution, known as the Probability Hypothesis Density (PHD) filter [69], [110], which is incorporated in the FISST-SLAM framework as explained in the following section.

5.5 PHD Multi-landmark Map Estimation Filter

The probability hypothesis density (PHD) filter is an approximation to alleviate the computational intractability of the optimal Bayes RFS multi-landmark map estimation filter, due to [69]. Thus, instead of propagating the entire multi-landmark posterior density, PHD propagates only the first order statistical moment of the state as follows,

$$D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k)}) = \int f_{k|k}(\mathrm{M}_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k)})\delta\mathrm{M}_k \tag{5.15}$$

Let $D_{k|k-1}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k-1)})$ be the PHD associated with the predicted multi-landmark state and γ_k be the intensity function of the RFS Γ_k, that accounts for new landmarks appearing in the FOV. Let the false alarm by the sensor be Poisson distributed with average number λ_c and the spatial distribution governed by the probability density $c_k(z)$, then the multi-landmark map estimation filter using PHD is propagated in

time as follows[1]:

$$D_{k|k-1}(\zeta_k|\mathrm{x}_k^{[i]}, Z^{(k-1)}) = \gamma_k(\zeta_k) + \int p_S(\zeta_{k-1}) f_{k|k-1}(\zeta_k|\zeta_{k-1}).D_{k-1|k-1}(\zeta_{k-1}|\mathrm{x}_k^{[i]}, Z^{(k-1)})d\zeta_{k-1} \tag{5.16}$$

$$D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, Z^{(k)}) = (1-p_D)D_{k|k-1}(\zeta_k|\mathrm{x}_k^{[i]}, Z^{(k-1)}) + \sum_{z\in Z_k} \frac{p_D D_k(\mathrm{z}|\mathrm{x}_k^{[i]}, \zeta_k)}{\lambda_c c_k(\mathrm{z}) + p_D D_k(\mathrm{z}|\mathrm{x}_k^{[i]}, \zeta_k)} D_k(\zeta_k|\mathrm{z}, \mathrm{x}_k^{[i]}) \tag{5.17}$$

where,

$$D_k(\mathrm{z}|\mathrm{x}_k^{[i]}, \zeta_k) = \int L_z(\mathrm{x}_k^{[i]}) D_{k|k-1}(\zeta_k|Z^{(k-1)}, \mathrm{x}_k^{[i]})d\zeta_k \tag{5.18}$$

with $L_z(\mathrm{x}_k^{[i]}) = f_k(\mathrm{z}|\mathrm{x}_k^{[i]}, \zeta_k)$ is the sensor likelihood function and,

$$D_k(\zeta_k|\mathrm{z}, \mathrm{x}_k^{[i]}) = \frac{f_k(\mathrm{z}|\mathrm{x}_k^{[i]}, \zeta_k) D_{k|k-1}(\zeta_k|Z^{(k-1)}, \mathrm{x}_k^{[i]})}{D_k(\mathrm{z}|\mathrm{x}_k^{[i]}, \zeta_k)} \tag{5.19}$$

Based on the counting property of the PHD (refer Appendix A for more details), the total expected number of landmarks is,

$$N_{k|k} = \int D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, Z^{(k)})d\zeta_k \tag{5.20}$$

However, it is to be noted that the PHD recursion involving eqns. 5.16 - 5.17 have multiple integrals, that have no closed form solutions in general. Few approaches have been proposed to mitigate this problem based on numerical approximations [98], [110], [65] and Gaussian mixture models [109]. Based on these techniques, we present two alternate implementations to the proposed FISST-SLAM algorithm in

[1]Provided that all the assumptions stated in Section 3.5 are satisfied.

the following sections.

5.6 SMC Implementation of the FISST-SLAM Filter

The sequential Monte-Carlo implementation of the proposed FISST-SLAM algorithm is discussed here and henceforth would be referred to as SMC-FISST-SLAM filter. The path posterior estimation is implemented using a standard particle filter as described in Section 5.3. The PHD filter for map posterior estimation is implemented using an SMC approximation, where the estimation of landmark locations being made using a modified version of expectation-maximization (EM) that accounts for PHD particle weights as in [106], [37].

- **Step 1:**

  At time $k = 0$, initialize,

  Initialize with N_s number of vehicle pose particles $\{v_k^{[i]}, \mathrm{x}_k^{[i]}\}_{i=1}^{N_s}$.

  For every $i-$th particle representing the vehicle pose, PHD $D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}_k)$ is initialized with a set of L_k particles $\{w_k^{[j]}, \zeta_k^{[j]}\}_{j=1}^{L_k}$.

  For every particle representing the vehicles pose $\mathrm{x}_0^{[i]}$, the landmarks are yet to be estimated and hence the number of landmarks is $N_k = 0$, and their locations (mean) $\mu_{N_k,k}$, and their uncertainties (covariance) $\Sigma_{N_k,k}$ are represented as null sets i.e., $\left\langle \mathrm{x}_0^{[i]}, 0, \{\emptyset;\emptyset\} \right\rangle_{i=1}^{N_s}$

  At time $k \geq 1$, retrieve,

$$\left\langle \mathrm{x}_{k-1}^{[i]}, N_{k-1}^{[i]}, \{\mu_{1,k-1}^{[i]}; \Sigma_{1,k-1}^{[i]}\}, \ldots, \{\mu_{N_{k-1}^{[i]},k-1}^{[i]}; \Sigma_{N_{k-1}^{[i]},k-1}^{[i]}\} \right\rangle_{i=1}^{N_s}$$

- **Step 2:**

  Calculate weights for particles approximating vehicle pose

  The vehicle pose is sampled by drawing a sample according to the motion model,

  $$\mathrm{x}_k^{[i]} \sim f_{k|k-1}(\mathrm{x}_k|\mathrm{x}_{k-1}^{[i]}, \mathrm{u}_k)$$

  for $l = 1, \ldots, N_{k-1}^{[i]}$, update the path posterior particle weight based on the importance function

  $$v_k^{[i]} = |2\pi Q_l|^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\mathrm{z} - \hat{\mathrm{z}}_l)Q_l^{-1}(\mathrm{z} - \hat{\mathrm{z}}_l)^T\right] \quad \forall \mathrm{z} \in \mathrm{Z}_k$$

  end for

  where, $Q_l = H_l^T \Sigma_{l,k-1}^{[i]} H_l + R_l$ and $\hat{\mathrm{z}}_l = h(\mu_{l,k-1}^{[i]}, \mathrm{x}_k^{[i]})$

  and where, $H_l = h'(\mu_{l,k-1}^{[i]}, \mathrm{x}_k^{[i]})$ and R_l is zero mean Gaussian noise.

- **Step 3:**

  The following is recursed for every $i-$th particle representing the vehicle pose $\mathrm{x}_k^{[i]}$,

  - Retrieve the PHD

    Retrieve the set of L_{k-1} particles $\{w_{k-1}^{[j]}, \zeta_{k-1}^{[j]}\}_{j=1}^{L_{k-1}}$ representing the PHD $D_{k-1|k-1}(\zeta_{k-1}|\mathrm{x}_{k-1}^{[i]}, \mathrm{Z}_{k-1})$.

  - Perform PHD landmark prediction

    for $j = 1, \ldots, L_{k-1}$, sample according to the proposal distribution $\pi_k(\cdot|\zeta_k^{[j]}, \mathrm{x}_k^{[i]}, \mathrm{Z}_k)$,

$\tilde{\zeta}_k^{[j]} \sim \pi_k(\cdot|\zeta_k^{[j]}, \mathrm{x}_k^{[i]}, \mathrm{Z}_k)$ and compute the predicted weights using

$$\tilde{w}_{k|k-1}^{[j]} = \frac{\tau_{k|k-1}(\tilde{\zeta}_k^{[j]}, \zeta_{k-1}^{[j]})}{\pi_k(\tilde{\zeta}_k^{[j]}|\zeta_{k-1}^{[j]}, \mathrm{x}_k^{[i]}, \mathrm{Z}_k)} w_{k-1}^{[j]}$$

end for

where $\tau(\cdot,\cdot) = p_S(\cdot) f_{k|k-1}(\cdot|\cdot)$, with $f_{k|k-1}(\cdot|\cdot)$ denoting the same meaning as in eqn. 5.16.

Additional J_k particles are generated to explore the new incoming landmarks.

for $j = L_{k-1}+1, \ldots, L_{k-1}+J_k$, sample according to the proposal distribution $\xi_k(\cdot|\mathrm{x}_k^{[i]}, \mathrm{Z}_k)$, $\tilde{\zeta}_k^{[j]} \sim \xi_k(\cdot|\mathrm{x}_k^{[i]}, \mathrm{Z}_k)$ and compute the weights of new incoming landmarks into the FOV using

$$\tilde{w}_{k|k-1}^{[j]} = \frac{1}{J_k} \frac{\gamma_k(\tilde{\zeta}_k^{[j]})}{\xi_k(\tilde{\zeta}_k^{[j]}|\mathrm{x}_k^{[i]}, \mathrm{Z}_k)}$$

end for

where $\gamma_k(\cdot)$ denotes the PHD of the new incoming landmarks as in eqn. 5.16. Thus we have, $\{w_{k|k-1}^{[j]}, \zeta_k^{[j]}\}_{j=1}^{L_{k-1}+J_k}$ particle representation of the predicted PHD.

- Perform the PHD measurement update

  The PHD measurement update modifies the weights of the particle PHD representation from $\{w_{k|k-1}^{[j]}, \zeta_k^{[j]}\}_{j=1}^{L_{k-1}+J_k}$ to $\{w_k^{[j]}, \zeta_k^{[j]}\}_{j=1}^{L_{k-1}+J_k}$ based on the new measurements.

for $j = 1, \ldots, L_{k-1} + J_k$, update the PHD weights using

$$\tilde{w}_k^{[j]} = \left[1 - p_D + \sum_{z \in Z_k} \frac{p_D f_k(z|x_k^{[i]}, \zeta_k^{[j]})}{\lambda_c c_k(z) + \sum_{l=1}^{L_{k-1}+J_k} p_D f_k(z|x_k^{[i]}, \zeta_k^{[i][l]}) \tilde{w}_{k|k-1}^{[i][l]}} \right] \tilde{w}_{k|k-1}^{[j]} \qquad \forall z \in Z_k$$

end for

where, p_D, $f_k(\cdot|\cdot)$, λ_c and $c_k(\cdot)$ have the same meaning as in eqn. 5.17

- The number of landmarks are estimated from the PHD by computing

  $\hat{N}_{k|k} = \sum_{j=1}^{L_{k-1}+J_k} \tilde{w}_k^{[j]}$

  $N_{k|k} = [\hat{N}_{k|k}]_{\text{Nearest Integer}}$

- Determine the location (or states) of the landmarks

  The locations of the landmarks are determined by extracting $N_{k|k}$ local maxima from the PHD. This is based on the assumption that the local maxima (peaks) correspond to the most probable landmark locations [69]. Further, it is also assumed that the PHD in the region of these local maxima can be approximated by Gaussian distributions. We then, attempt to fit the Gaussian mixtures to the PHD using expectation-maximization (EM) that accounts for PHD particle weights as in [106], [37]. Landmark states are estimated by fitting $G = |Z_k| + N_{k-1|k-1}$ Gaussian mixtures to the updated PHD. The EM algorithm starts with the initial value $\theta_{g,k} = (\alpha_{g,k}, \mu_{g,k}; \Sigma_{g,k})$, where $\theta_{g,k}$ is a vector comprising the initial weight, mean, and the covariance of the g-th Gaussian mixture. The initial values for the Gaussian mixtures are chosen as the measurement locations Z_k and the local maxima at previous time instant $k-1$.

for $g = 1, \ldots, G$

– **E-step:**

$$p(g|\zeta_k^{[j]}) = \frac{p(\zeta_k^{[j]}|\theta_{g,k}^{[j]})\alpha_{g,k}}{\sum_{l=1}^{G} p(\zeta_k^{[j]}|\theta_{l,k}^{[j]})\alpha_{l,k}^{[j]}}$$

where,

$$p(\zeta_k^{[j]}|\theta_{g,k}^{[j]}) = |2\pi\Sigma_{g,k}|^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\zeta_k^{[j]} - \mu_{g,k})\Sigma_{g,k}{}^{-1}(\zeta_k^{[j]} - \mu_{g,k})^T\right]$$

– **M-step:**

$$\alpha_{g,k} = \sum_{j=1}^{L_{k\ 1}+J_k} w_k^{[j]} p(g|\zeta_k^{[j]})$$

$$\mu_{g,k} = \frac{1}{\alpha_{g,k}} \sum_{j=1}^{L_{k-1}+J_k} \zeta_k^{[j]} p(g|\zeta_k^{[j]})$$

$$\Sigma_{g,k} = \frac{1}{\mu_{g,k}} \sum_{j=1}^{L_{k-1}+J_k} p(g|\zeta_k^{[j]})(\zeta_k^{[j]} - \mu_{g,k})(\zeta_k^{[j]} - \mu_{g,k})^T$$

end for

The above recursion is repeated until EM algorithm converges and $N_{k|k}$ heaviest Gaussian mixtures are picked based on weights $\alpha_{g,k}$. The mean and covariance of these Gaussian mixtures $\left\langle \{\mu_{1,k}; \Sigma_{1,k}\}, \ldots, \left\{\mu_{N_{k|k},k}; \Sigma_{N_{k|k},k}\right\}\right\rangle$ represent the set of landmark position estimates and their uncertainty.

- Resample and rescale particle PHD weights $\left\{\frac{\tilde{w}_k^{[j]}}{N_{k|k}}, \tilde{\zeta}_k^{[j]}\right\}_{j=1}^{L_{k\ 1}+J_k}$ to get $\left\{\frac{w_k^{[j]}}{N_{k|k}}, \zeta_k^{[j]}\right\}_{j=1}^{L_k}$

Multiply the weights by $\hat{N}_{k|k}$ to get $\left\{w_k^{[j]}, \zeta_k^{[j]}\right\}_{j=1}^{L_k}$

Steps 1 to 3 are recursed at each time step (with step 3 recursed for every particle representing the vehicle path posterior), until the end of the run.

In stochastic vector based SLAM framework, one of the most popular analytical approximations of the nonlinear Bayes filter is the extended Kalman filter (EKF), where the posterior density is approximated by a Gaussian distribution and propagated in time by applying the Kalman recursions to local linearizations [102]. In [109], it has been shown that the posterior intensity of the multi-landmark state propagated by the PHD recursions is a weighted sum of various functions, many of which are non-Gaussian. In the same vein as the EKF, these non-Gaussian constituent functions are approximated by a Gaussian. An approximation of the posterior intensity at each time step is obtained by applying the Gaussian mixture PHD recursions to a locally linearized model [109]. We apply this closed form solution based on Gaussian mixture models for implementing the proposed FISST-SLAM algorithm in the following section.

5.7 Gaussian Mixture Implementation of the FISST-SLAM Filter

The Gaussian mixture implementation of the proposed FISST-SLAM algorithm is discussed here and henceforth would be referred to as GM-FISST-SLAM filter. In order to apply the GM techniques for the proposed FISST-SLAM algorithm, the following assumptions have to be satisfied.

- The probability of persistence p_S and the probability of detection p_D are not time-varying and are constants.
- The PHDs of new incoming landmarks into the FOV $\gamma_k(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k-1)})$ is assumed to be Gaussian mixture of the form

$$\gamma_k(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k-1)}) = \sum_{j=1}^{J_{\gamma,k}} w_{\gamma,k}^{[j]} \mathcal{N}(\zeta_{k-1}; \mu_{\gamma,k}^{[j]}, \Sigma_{\gamma,k}^{[j]}) \tag{5.21}$$

- The single landmark Markov transition density $f_{k|k-1}(\zeta_k|\zeta_{k-1})$ and the likelihood function $f_k(\mathrm{z}|\mathrm{x}_k^{[i]}, \zeta_k)$ are assumed to be Gaussian in nature.
- The predicted and updated PHDs are also Gaussian mixtures.

We now outline Gaussian mixture implementation of the FISST-SLAM algorithm.

- **Step 1:**

  At time $k = 0$, initialize,

  Initialize with N_s number of vehicle pose particles $\{v_k^{[i]}, \mathrm{x}_k^{[i]}\}_{i=1}^{N_s}$.

  For every $i-$th particle representing the vehicle pose, PHD $D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}_k)$ is initialized with a weighted sum of J_k Gaussians

$$D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}_k) = \sum_{j=1}^{J_k} w_k^{[j]} \mathcal{N}(\zeta_k; \mu_k^{[j]}, \Sigma_k^{[j]}) \tag{5.22}$$

  For every particles representing the vehicles pose $\mathrm{x}_0^{[i]}$, the landmarks are yet to be estimated and hence the number of landmarks is $N_k = 0$, and their locations (mean) μ_k, and their uncertainties (covariance) Σ_k are represented as null sets i.e., $\left\langle \mathrm{x}_0^{[i]}, 0, \{\emptyset; \emptyset\} \right\rangle_{i=1}^{N_s}$

At time $k \geq 1$, retrieve,

$$\left\langle \mathrm{x}_{k-1}^{[i]}, N_{k-1}^{[i]}, \{\mu_{k-1}^{[i][j]}; \Sigma_{k-1}^{[i][j]}\}_{j=1}^{N_{k-1}} \right\rangle_{i=1}^{N_s}$$

- **Step 2:**

  Calculate weights for particles approximating vehicle pose

  The vehicle pose is sampled by drawing a sample according to the motion model

  $$\mathrm{x}_k^{[i]} \sim f_{k|k-1}(\mathrm{x}_k | \mathrm{x}_{k-1}^{[i]}, \mathrm{u}_k)$$

  for $l = 1, \ldots, N_{k-1}^{[i]}$, update the path posterior particle weight based on the importance function

  $$v_k^{[i]} = |2\pi Q_l|^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\mathrm{z} - \hat{\mathrm{z}}_l) Q_l^{-1} (\mathrm{z} - \hat{\mathrm{z}}_l)^T\right] \quad \forall \mathrm{z} \in Z_k$$

  end for

  where, $Q_l = H_l^T \Sigma_{k-1}^{[i][l]} H_l + R_l$ and $\hat{\mathrm{z}}_l = h(\mu_{k-1}^{[i][l]}, \mathrm{x}_k^{[i]})$

  and where, $H_l = h'(\mu_{k-1}^{[i][l]}, \mathrm{x}_k^{[i]})$ and R_l is zero mean Gaussian noise

- **Step 3:**

  The following is recursed for every $i-$th particle representing the vehicle pose $\mathrm{x}_k^{[i]}$,

  - Retrieve the PHD

    Suppose the assumptions mentioned above hold, then the PHD at time $k-1$, $D_{k-1|k-1}(\zeta_{k-1} | \mathrm{x}_{k\ 1}^{[i]}, Z^{(k-1)})$ is a Gaussian mixture of the form

    $$D_{k-1|k-1}(\zeta_{k-1} | \mathrm{x}_{k-1}^{[i]}, Z^{(k-1)}) = \sum_{j=1}^{J_{k-1}} w_{k-1}^{[j]} \mathcal{N}(\zeta_{k-1}; \mu_{k-1}^{[j]}, \Sigma_{k-1}^{[j]}) \tag{5.23}$$

consisting of J_{k-1} Gaussians, with $w_{k-1}^{[j]}$, $\mu_{k-1}^{[j]}$ and $\Sigma_{k-1}^{[j]}$ being their weights, mean and covariances respectively.

- Perform PHD landmark prediction

  For the existing landmarks in the FOV (under static assumption), If the PHD of the existing landmarks the FOV at time k is a Gaussian mixture of the form

$$D_{k|k-1}(\zeta_{k-1}|\mathrm{x}_{k-1}^{[i]}, \mathrm{Z}^{(k-1)}) = p_S \sum_{j=1}^{J_{k-1}} w_{k-1}^{[j]} \mathcal{N}(\zeta; \mu_{k-1}^{[j]}, \Sigma_{k-1}^{[j]}) \tag{5.24}$$

  then,

  for $j = 1, \ldots, J_{k-1}$,

$$\begin{aligned} w_{k|k-1}^{[j]} &= p_S . w_{k-1}^{[j]} \\ \mu_{k|k-1}^{[j]} &= \zeta_{k-1}^{[j]} \\ \Sigma_{k|k-1}^{[j]} &= \Sigma_{k-1}^{[j]} \end{aligned}$$

  endfor

  For new incoming landmarks

  If the PHD of the new incoming landmarks represented by a RFS $\gamma_k(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k-1)})$ at time k is a Gaussian mixture of the form

$$\gamma_k(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k-1)}) = \sum_{j=1}^{J_{\gamma,k}} w_{\gamma,k}^{[j]} \mathcal{N}(\zeta_{k-1}; \mu_{\gamma,k}^{[j]}, \Sigma_{\gamma,k}^{[j]}) \tag{5.25}$$

  then,

for $j = J_{k-1} + 1, \ldots, J_{k-1} + J_{\gamma,k}$,

$$\begin{aligned} w_{k|k-1}^{[j]} &= w_{\gamma,k}^{[j]} \\ \mu_{k|k-1}^{[j]} &= \mu_{\gamma,k}^{[j]} \\ \Sigma_{k|k-1}^{[j]} &= \Sigma_{\gamma,k}^{[j]} \end{aligned}$$

endfor

Thus predicted PHD at time k is a Gaussian mixture of the form,

$$D_{k|k-1}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k-1)}) = \sum_{j=1}^{J_{k-1}+J_{\gamma,k}} w_{k|k-1}^{[j]} \mathcal{N}(\zeta_k; \mu_{k|k-1}^{[j]}, \Sigma_{k|k-1}^{[j]}) \tag{5.26}$$

and the predicted number of landmarks is given by

$$\hat{N}_{k|k-1} = p_S.\hat{N}_{k-1|k-1} + \sum_{j=1}^{J_{\gamma,k}} w_{\gamma,k}^{[j]} \tag{5.27}$$

- Perform the PHD measurement update

  The PHD measurement update, as in eqn. 5.17, is also a Gaussian mixture given by

$$D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k)}) = D_{k|k-1}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k-1)})[(1 - p_D) + \sum_{z \in \mathrm{Z}_k} D_{L,k}(z|\zeta_k, \mathrm{x}_k^{[i]})] \tag{5.28}$$

  where for every $z \in \mathrm{Z}_k$

$$D_{L,k}(z|\zeta_k, \mathrm{x}_k^{[i]}) = \sum_{j=1}^{J_{k-1}+J_{\gamma,k}} w_{k|k}^{[j]} \mathcal{N}(z; \mu_{k|k}^{[j]}, \Sigma_{k|k}^{[j]}) \tag{5.29}$$

for $j = 1, \ldots, J_{k-1} + J_{\gamma,k}$,

$$w_{k|k}^{[j]} = \frac{p_D w_{k|k-1}^{[j]} f_k^{[j]}(z|\zeta_k, \mathrm{x}_k^{[i]})}{\lambda_c c_k(z) + \sum_{l=1}^{J_{k-1}+J_{\gamma,k}} w_{k|k-1}^{[l]} f_k^{[l]}(z|\zeta_k, \mathrm{x}_k^{[i]})} \tag{5.30}$$

$$f_k^{[j]}(z|\zeta_k, \mathrm{x}_k^{[i]}) = \mathcal{N}(z; H_k \mu_{k|k-1}^{[j]}, S_k^{[j]}) \tag{5.31}$$

$$\mu_{k|k}^{[j]} = \mu_{k|k-1}^{[j]} + K_k^{[j]}(z - H_k \mu_{k|k-1}^{[j]}) \tag{5.32}$$

$$\Sigma_{k|k}^{[j]} = [I - K_k^{[j]} \nabla H_k] \Sigma_{k|k-1}^{[j]} \tag{5.33}$$

$$K_k^{[j]} = \Sigma_{k|k-1}^{[j]} \nabla H_k^T [S_k^{[i]}]^{-1} \tag{5.34}$$

$$S_k^{[j]} = R_k + \nabla H_k \Sigma_{k|k-1}^{[j]} \nabla H_k^T \tag{5.35}$$

where,

$$\nabla H_k = \left. \frac{\partial f_k(\zeta_k, 0)}{\partial \zeta_k} \right|_{\zeta_k = \mu_{k|k}^{[j]}} \tag{5.36}$$

endfor

Thus, there are $J_k = (1 + |Z_k|)(J_{k-1} + J_{\gamma,k})$ Gaussian components in the updated PHD with $(1 + |Z_k|)$ components for each prediction term at time k and the Gaussian mixture is of the form,

$$D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, Z^{(k)}) = \sum_{j=1}^{J_k} w_{k|k}^{[j]} \mathcal{N}(\zeta_k; \mu_k^{[j]}, \Sigma_k^{[j]}) \tag{5.37}$$

- The number of landmarks are estimated from the updated PHD by computing

$$\hat{N}_{k|k} = (1 - p_D)\hat{N}_{k|k-1} + \sum_{z \in Z_k} \sum_{j=1}^{J_{k-1}+J_{\gamma,k}} w_{k|k}^{[j]}(z) \tag{5.38}$$

$N_{k|k} = [\hat{N}_{k|k}]_{\text{Nearest Integer}}$

- Determine the location (or states) of the landmarks

  $N_{k|k}$ heaviest Gaussian mixtures are picked based on weights. The mean and covariance of these Gaussian mixtures $\left\{\mu_{k|k}^{[j]}; \Sigma_{k|k}^{[j]}\right\}_{j=1}^{N_{k|k}}$ represent the set of landmark position estimates and their uncertainty.

Steps 1 to 3 are recursed at each time step (with step 3 recursed for every particle representing the vehicle path posterior), until the end of the run.

5.8 Implementational Issues with the FISST-SLAM Filter

The two implementations of the FISST-SLAM algorithm discussed above have their own merits and demerits. The underlying assumptions of the Gaussian mixture implementation as discussed in the earlier section are far more restrictive than the SMC implementation of the FISST-SLAM algorithm. For mildly non-linear problems, it has been shown that the GM implementation of the PHD filter [25] will maintain an approximation error that converges to zero as the number of Gaussian mixtures tends to infinity. However, in a high noise scenario (process and measurement noise), it may be more appropriate to use the SMC-FISST-SLAM filter, which can adequately incorporate non-linear dynamic models and non-Gaussian state and measurement noises. While employing the SMC implementation can adequately model for the system non linearities, it comes at a cost of inculcating large number of particles and thereby increasing the computational demand in each FISST-SLAM recursion. Moreover, SMC-FISST-SLAM filter has an additional cost of extracting multi-landmark states using EM-algorithm. The SMC implementation works well

when the updated PHD $D_{k|k}(\cdot)$ has an estimated $N_{k|k}$ clusters. However, if the number of landmarks differs from the number of clusters, then the probable location of these landmarks estimated can become unreliable. While the GM-FISST-SLAM filter allows the landmark location estimates to be directly extracted from the PHD posterior intensity $D_{k|k}(\cdot)$, thereby completely avoiding the clustering techniques that are employed in the SMC-FISST-SLAM filter.

The time complexity of FISST-SLAM filter is $\mathcal{O}(NK)$, where N is number of particles representing the vehicle path posterior and K is either the number of particles (in the case of SMC-FISST-SLAM filter) or the number of Gaussian mixtures (in the case of GM-FISST-SLAM filter) approximating the PHD representing the landmarks. Thus, computational complexity of the FISST-SLAM algorithm is linear in terms of the map size. However, since each update of PHD multi-landmark map estimation filter is limited to the FOV, the computational complexity is effectively constant in time.

The next section discusses the simulation results for the proposed FISST-SLAM algorithm based on SMC implementation (SMC-FISST-SLAM filter) and Gaussian mixture implementation (GM-FISST-SLAM filter) described above. The results are benchmarked against FastSLAM 1.0 with an efficient map or feature management strategy.

5.9 Simulations & Results

Simulation tests are run using the parameters summarized in table 5.1. A total of 72 landmarks are placed in random locations in a 600 m x 500 m rectangular area. The vehicle's motion is simulated based on the motion model given by eqn. 5.5.

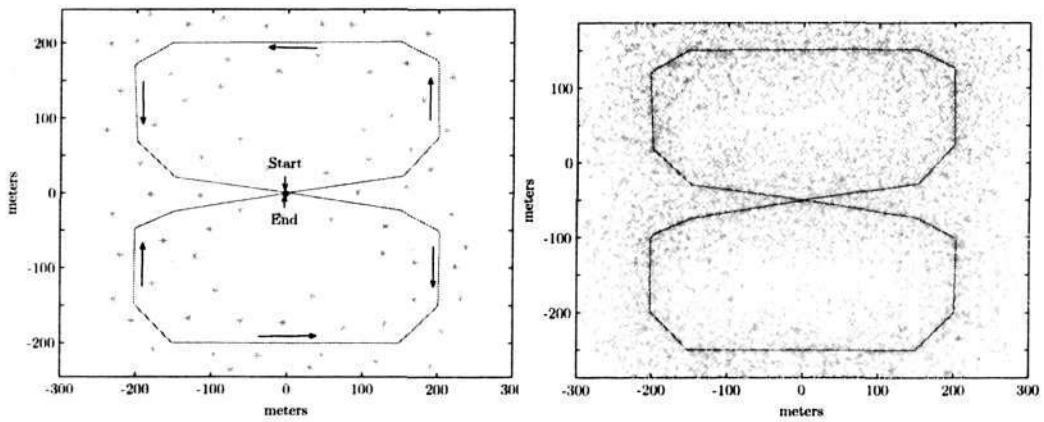
Parameters		Values
Velocity	V	4 m/s
Sensor Field of View	Range	$0 - 100$ m
	Bearing	180°
Control Noise	Velocity $[\sigma_v]$	0.3 m/s
	Steer $[\sigma_a]$	2°
Measurement Noise	Range $[\sigma_r]$	0.3 m
	Bearing $[\sigma_\theta]$	0.5°

Table 5.1 *Parameters used in the simulation*

The vehicle is equipped with a range-bearing sensor with a maximum observable range of $100m$ and a FOV of 180^0. The range and bearing measurements for a landmark are only generated when it falls within the defined sensor FOV. Clutter is assumed to be Poisson distributed and is generated uniformly over the observation space $[-\pi/2, \pi/2] \times [0, 100]$ (FOV) with an average rate of λ_c (range and bearing) per scan. The sensor data and control parameters representative of a real vehicle equipped with a generic range-bearing-intensity sensor is tabulated in table 5.1. The control inputs (speed and yaw) are updated at 20 Hz and observations (range and bearing) were obtained every 1 Hz. The simulated true vehicle path and the true locations of the landmarks immersed in clutter with densities $\lambda_c = 0$ and $\lambda_c = 10$ are shown in fig. 5.2. In the simulation, the vehicle starts at the origin and returns to the origin tracing a figure of '8'.

5.9.1 Map Error Metric

The feature based SLAM algorithms produce estimates of the location of the vehicle (localization) and the location of the landmarks (mapping) in the environment. However, literature on classical vector based SLAM mainly focuses on the algorithm's performance based on vehicle localization error, with very little emphasis



(a) with $\lambda_c = 0$ *(b) with $\lambda_c = 10$*

Figure 5.2 *Ground truth of the vehicle path (in red) and landmark locations (green asterisk) with clutter (grey cross) densities set at $\lambda_c = 0$ and $\lambda_c = 10$. A total of 72 landmarks are deployed in the simulated environment. The vehicle starts at the origin and returns to the origin tracing a figure of '8'. Total distance traversed is approximately 2.1 km.*

on the overall mapping error. Most of the error metrics are based on Euclidean distance or Mahalanobis distance measures, which are only valid for individual features, under the correct data association assumption. However such metrics do not provide any information on the quality of the entire multi-landmark map state. Further, such metrics do not account for missed detections or false alarms.

Thus in order to quantify the overall quality of the multi-landmark map built, we use the Wasserstein miss-distance metric based on the developments in the target tracking community [49]. Here, we compute the Wasserstein "miss-distance" between the landmark set $\widetilde{\mathrm{M}}_k$ at time k corresponding to the ground truth and the other set representing the estimated landmark set M_k. The ground truth or the actual location of the set of landmarks $\widetilde{\mathrm{M}}_k$ at time k is obtained by traversing the vehicle in the simulated path with absolute position certainty. The Wasserstein miss-distance measure between the actual landmarks $\widetilde{\mathrm{M}}_k$ and their estimates M_k at time k is then given by,

$$\begin{aligned}\mathbf{W}_p(\mathrm{M}_k, \widetilde{\mathrm{M}}_k) &= \inf_{\mathrm{M}_k, \widetilde{\mathrm{M}}_k} \{\mathbf{E}[|\mathbf{M_k} - \widetilde{\mathbf{M}}_\mathbf{k}|^\mathbf{p}]\}^{\frac{1}{p}} \\ &= \inf_{\mathrm{C}} \sqrt[p]{\sum_{i=1}^{|\mathrm{M}_k|} \sum_{j=1}^{|\widetilde{\mathrm{M}}_k|} \mathrm{C}_{ij}.\mathrm{d}_p(\mathrm{m}_{k,i}, \widetilde{\mathrm{m}}_{k,j})} \\ &= \inf_{\mathrm{C}} \sqrt[p]{\sum_{i=1}^{|\mathrm{M}_k|} \sum_{j=1}^{|\widetilde{\mathrm{M}}_k|} \mathrm{C}_{ij}. \parallel \mathrm{m}_{k,i}, \widetilde{\mathrm{m}}_{k,j} \parallel^p}\end{aligned} \tag{5.39}$$

where the infimum is taken over all $|\mathrm{M}_k| \times |\widetilde{\mathrm{M}}_k|$ "transportation matrices" C (where a transportation matrix is one whose entries C_{ij} satisfy $\mathrm{C}_{ij} > 0$, $\sum_{j=1}^{|\widetilde{\mathrm{M}}_k|} \mathrm{C}_{ij} = \frac{1}{|\mathrm{M}_k|}$ and $\sum_{i=1}^{|\mathrm{M}_k|} \mathrm{C}_{ij} = \frac{1}{|\widetilde{\mathrm{M}}_k|}$). In our case p is set to 2.

5.9.2 Performance Evaluation of FISST-SLAM and FastSLAM

The results are benchmarked against FastSLAM 1.0 along with an efficient map management strategy, as in [76], which we from now on refer to as landmark quality (LQ) FastSLAM. This is done by maintaining a landmark existence probability based on log odds [36]. Whenever, the $i-$th landmark m_i is observed, the log odds for its existence o_i is incremented by a fixed value or else its value is decremented. If the log odds is above a predefined threshold τ_{lo}, the landmark is incorporated into the map. In our simulation the log odds value was incremented in steps of 0.01 and the threshold was set at $\tau_{lo} = 0.08$. This was determined after some trial simulation runs that yielded good results.

Both LQ-FastSLAM and the proposed FISST-SLAM algorithm (with both SMC and GM implementations) were executed for the scenario described above. Throughout the simulation the number of particles used to estimate the vehicle trajectory component was set to 80 for both the algorithms. Five hundred particles are used for

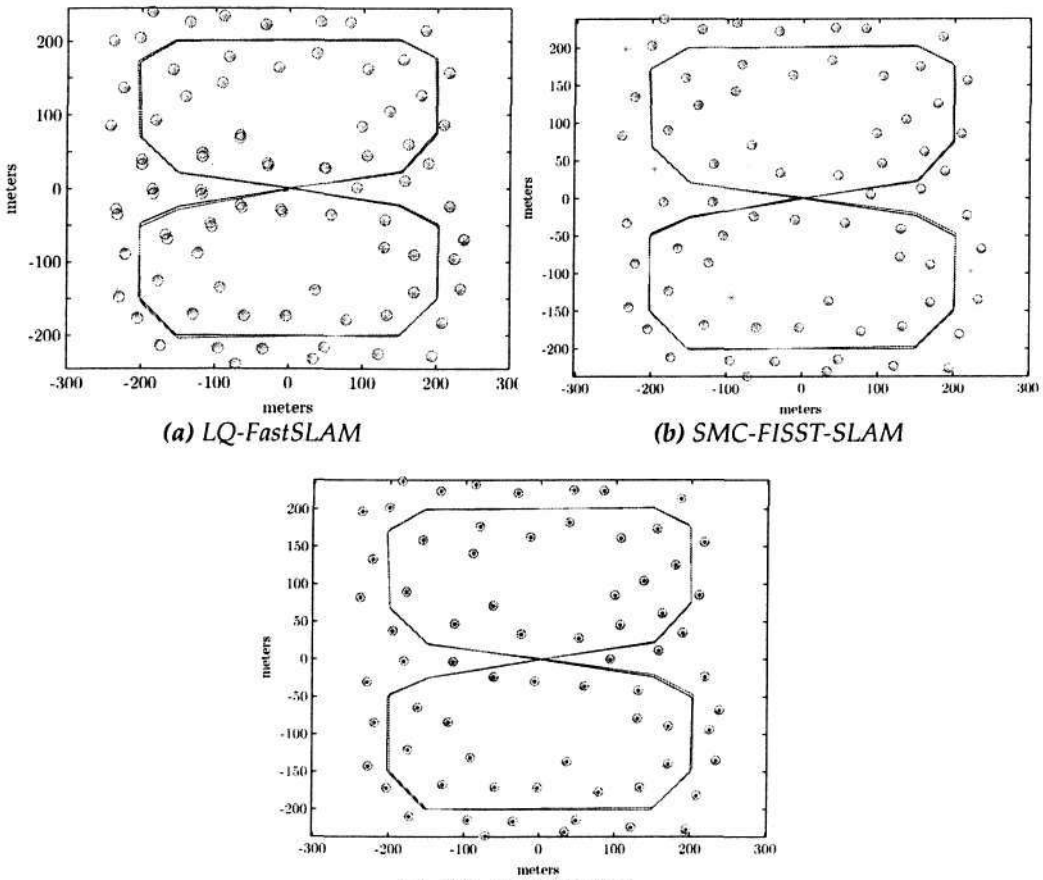
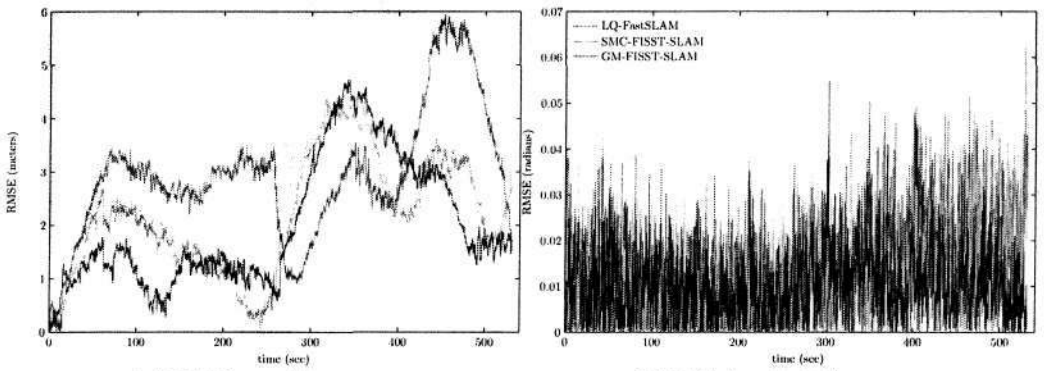


Figure 5.3 *Estimated (in blue) and true (in red) trajectories superimposed with estimated (red circles) and actual landmark locations (green asterisk) with* $\lambda_c = 1$.

each expected landmark in the SMC implementation of the PHD multi-landmark estimation filter. Thus the number of particles used per iteration vary throughout the simulation. In case of the GM implementation of the PHD multi-landmark estimation filter, the maximum allowable number of Gaussian terms was set to one hundred. Without loss of generality the sensor probability of detection p_D was set to 1. We examine the performance of the algorithms under varying clutter conditions and benchmark them in terms of vehicle pose errors and the number of landmarks

and their states correctly incorporated into the map at each time step.



(a) Vehicle position error *(b) Vehicle orientation error*

Figure 5.4 *A comparison of vehicle pose estimation error between FISST-SLAM and LQ-FastSLAM for clutter rate $\lambda_c = 1$*

The comparisons of the estimated vehicle trajectory and final map for LQ-FastSLAM and FISST-SLAM implementations with low clutter density ($\lambda_c = 1$) are shown in fig. 5.3. At low clutter rate, with a reduced data association ambiguity, all the filters perform consistently, maintaining low vehicle pose errors as shown in fig. 5.4. Figure 5.5 shows the time evolution of the number of landmarks incorporated into the map and their corresponding multi-landmark map error metric under mild clutter scenarios for each approach. Thus it can be clearly observed that under mild clutter conditions, all the approaches perform equally well, with relatively low miss-distance values, even though, LQ-FastSLAM has incorporated a few false detections into the map.

The comparisons of the estimated vehicle trajectory and final map for LQ-FastSLAM and FISST-SLAM implementations with dense clutter scenario ($\lambda_c = 10$) are shown in fig. 5.6. Although, FastSLAM with a feature management strategy does equally well as FISST-SLAM under low clutter, it's performance under high clutter is much

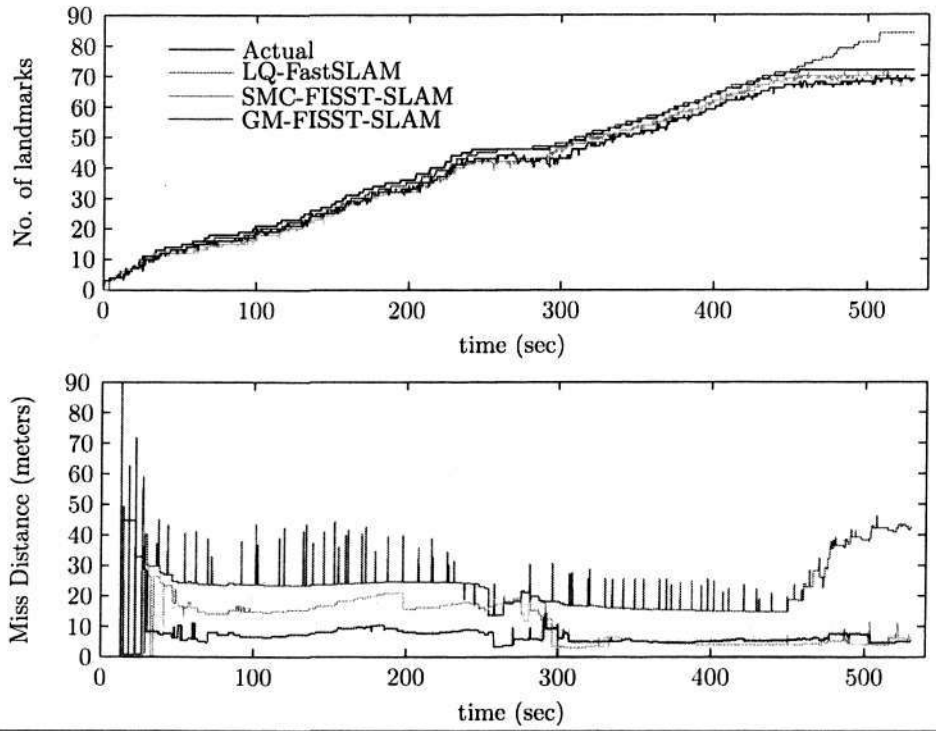
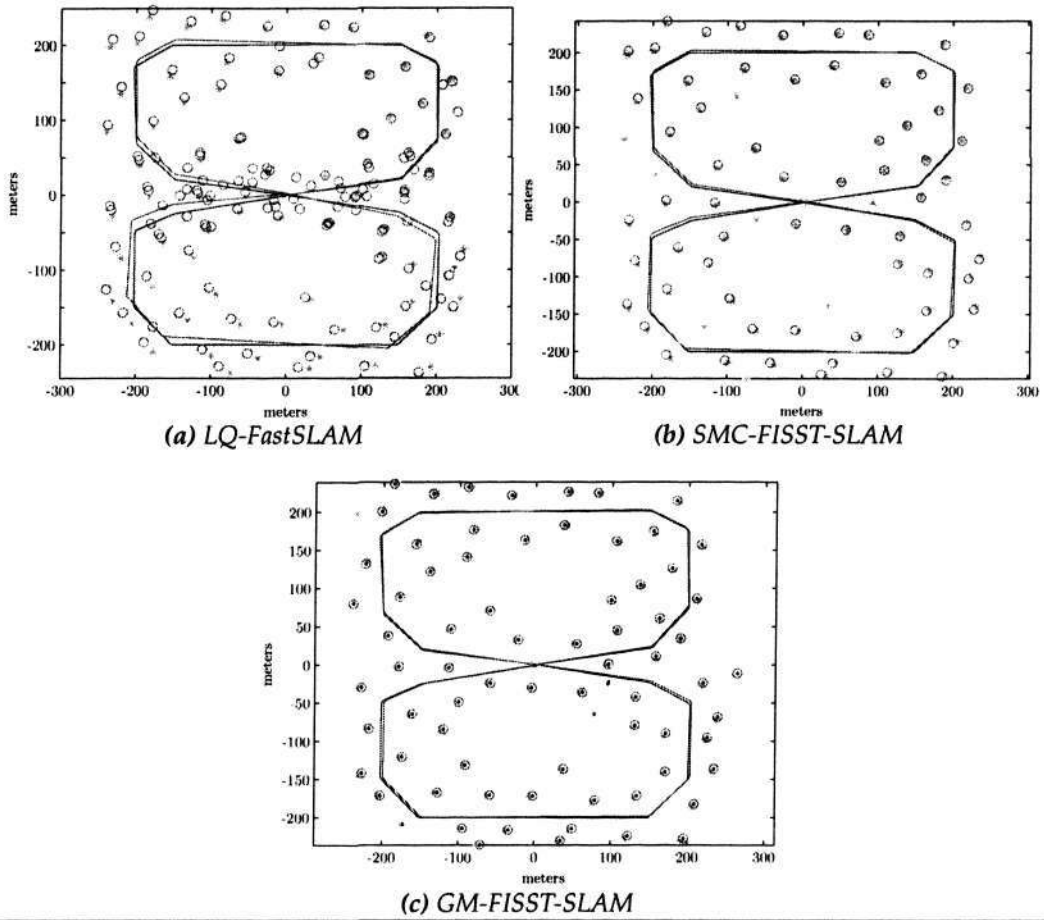


Figure 5.5 *No of landmarks incorporated into map and a Wasserstein distance map metric at each time for mild clutter scenarios. A total of 72 landmarks are deployed in the simulated environment.*

less to be desired.

This is clearly evident from the large number of clutter returns (as seen in fig. 5.6a) incorrectly estimated as actual landmarks (in this case 136, fig. 5.8) after the complete run with a large RMS error of 9.45m and a loop closing error of 3.14 m. The relatively large errors, as observed in fig. 5.7 are mainly due to the clutter that is incorporated into the map as valid landmarks, resulting in miss-associations. However, increasing the log-odds predefined threshold τ_{lo}, the amount of clutter incorporated into the map is reduced, although, this is at the expense of missed detections of valid landmarks and hence the quality of the landmark map built.

As for the proposed FISST-SLAM algorithm (for both the implementations), it's

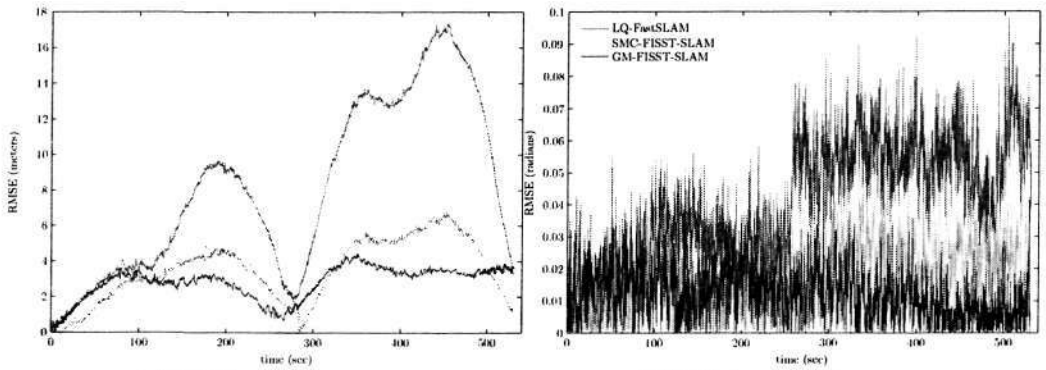


(a) LQ-FastSLAM *(b) SMC-FISST-SLAM*

(c) GM-FISST-SLAM

Figure 5.6 *Estimated (in blue) and true (in red) trajectories superimposed with estimated (red circles) and actual landmark locations (green asterisk) with* $\lambda_c = 10$.

performance is much less affected by the increase of clutter as compared to LQ-FastSLAM as is clearly seen in fig. 5.7 and fig. 5.8. For the complete closed-loop run the estimated number of landmarks in the dense clutter scenario is 71 (GM-FISST-SLAM filter) which is 1 short of the actual number and the landmark positions have been estimated to a very high accuracy. It is also important to note that RMS error is about half of LQ-FastSLAM in dense clutter environment with a loop closing error of 1.21m (GM-FISST-SLAM filter), as observed in fig. 5.7. Figure 5.8 shows



(a) Vehicle position error *(b) Vehicle orientation error*

Figure 5.7 *A comparison of Vehicle pose estimation error between FISST-SLAM and LQ-FastSLAM for clutter rate* $\lambda_c = 10$

the time evolution of the number of landmarks incorporated into the map and their corresponding multi-landmark map error metric under dense clutter scenarios for each approach. Unlike the mild clutter scenario where the multi-landmark map error error metric was relatively lower for all the methods, for the dense clutter scenario, the map error error metric is consistently high for LQ-FastSLAM as compared to FISST-SLAM demonstrating the superiority of FISST-SLAM. Figures 5.5 and 5.8 shows that the multi-landmark map error metric remains consistently high for SMC-FISST-SLAM filter in comparison with GM-FISST-SLAM filter. This is mainly attributed to the landmark state estimation using clustering techniques in SMC implementation. When the estimated number of landmarks from the PHD is incorrect, the clustering output that determines the landmark states become unreliable. Moreover, the map error metric tends to penalize sets with different cardinalities. This results in higher peaks in miss-distance metric at time instances where the estimated number of landmarks are incorrect as shown in figs. 5.5 and 5.8. The consequence of this is higher vehicle localization errors as observed in figs. 5.4 and

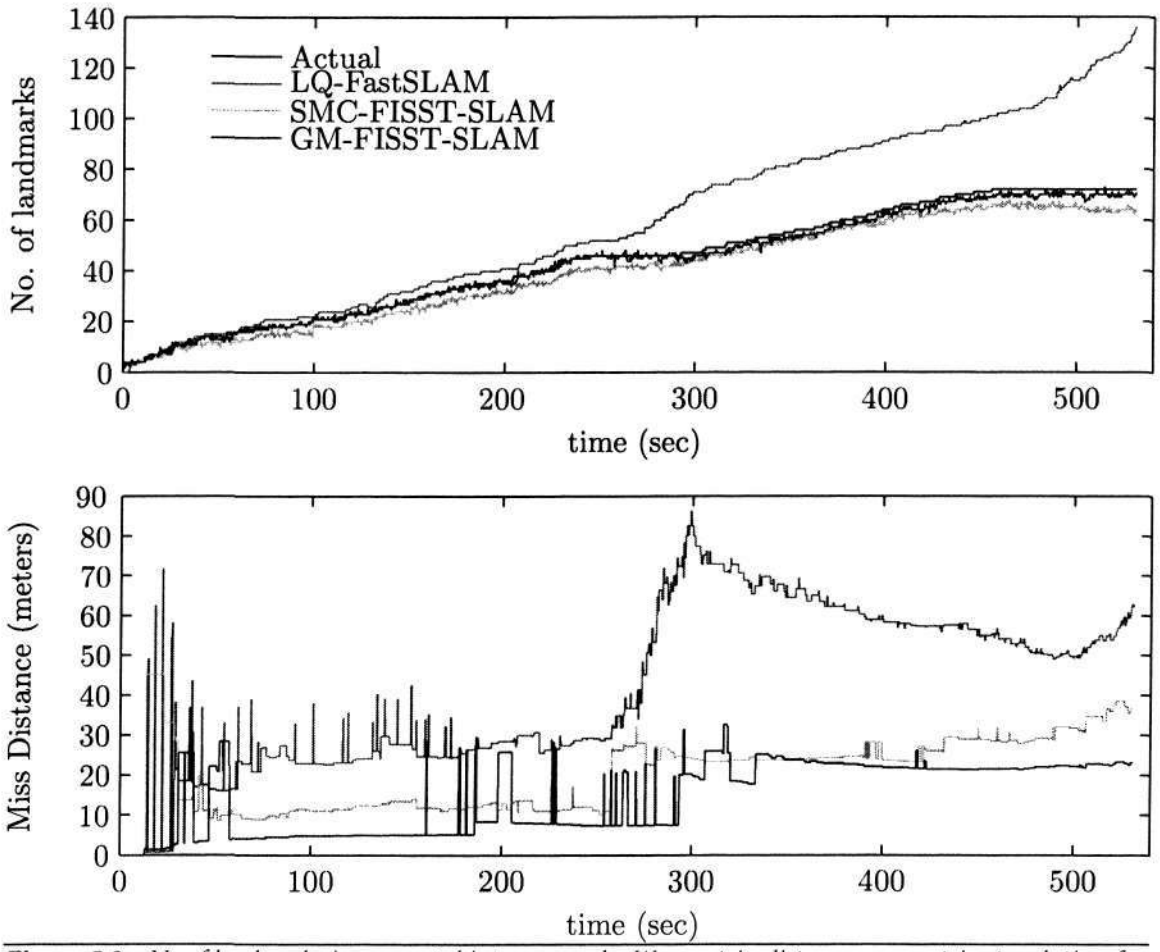


Figure 5.8 *No of landmarks incorporated into map and a Wasserstein distance map metric at each time for dense clutter scenarios. A total of 72 landmarks are deployed in the simulated environment.*

5.7.

These simulations demonstrate that for FastSLAM to be effective it is essential that an even more elaborate feature map management strategy with more sophisticated data association algorithms such as MFDA [114], JCBB [86], MHT [21] is needed alongside the core SLAM algorithm. The proposed FISST-SLAM algorithm explicitly accounts for measurement detection uncertainty, data association uncertainty, false alarms and feature map management in the SLAM filter framework,

and thereby obviating the need for external heuristic feature map management strategies. It is also important to note that FISST-SLAM jointly estimates the vehicle pose, number of landmarks and their states without any explicit clutter detection or feature management strategies, unlike FastSLAM. All of these results demonstrate the effectiveness of the proposed FISST-SLAM algorithm, particularly in high clutter scenarios.

5.9.3 Performance Evaluation in Dynamic Environments

Real world environments are not static. They contain moving objects, such as people, temporary objects that appear static for a while such as parked cars, or for instance in underwater, marine life such as fishes and other sea-creatures. In dynamic environments, any SLAM algorithm (either vector based or the proposed approach) must somehow manage moving objects. It can detect them and ignore them; it can track them as moving landmarks, however should never add a moving landmark into the map and assume it as stationary. Thus, the main problem associated with dynamic environments is that of feature (landmark) selection and map management. To explicitly handle moving objects, in [48] a dynamic object detection routine was setup which eliminated the moving objects before being handled by the SLAM algorithm. In [111], moving objects were added to the estimated state, where the implemented solution involved a stationary SLAM update followed by a separate tracker for tracking moving objects.

In the conventional vector state space based approach (like FastSLAM) for dynamic environments, the following aspects of feature management are critical and need to be explicitly handled.

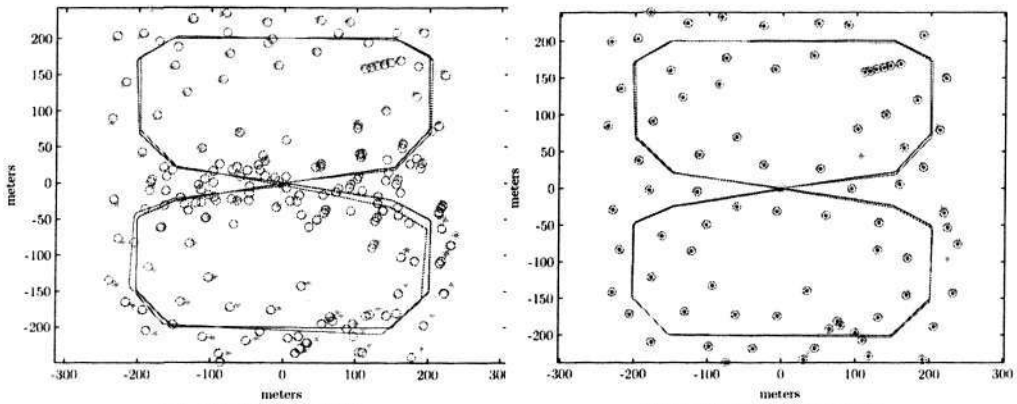
- Addition of features into the map. Dynamic or unstable features are avoided from being added into the map by deferred decision making techniques where in a provisional list of landmarks is maintained until a satisfactory evidence of their stability is determined.

- Removal of redundant or obsolete features from the map. For whatever reason maybe the dynamic nature of the environment or mobile nature of certain objects, some landmarks may cease to exist and no longer provide useful information. These obsolete features should be deleted from the map to maintain an uncluttered and contemporary representation of the environment.

This is commonly implemented in the vector state space based approaches by maintaining a landmark existence probability based on log odds [36]. Whenever, the $i-$th landmark m_i is observed, the log odds for its existence o_i is incremented by a fixed value or else its value is decremented. If the log odds is above a predefined threshold τ_{lo}, the landmark is incorporated into the map. However, in the case of the proposed FISST-SLAM algorithm, the dynamic and unstable landmarks are automatically discarded within the SLAM framework.

The performance of the proposed algorithms are further evaluated in dynamic environments. Simulation tests are run using the same parameters summarized in table 5.1. However, apart from having 78 static landmarks, three more time-varying landmarks are introduced into the environment. Each time-varying landmark is assumed to moving in a two-dimensional plane according to constant acceleration linear Gaussian dynamics with zero-mean Gaussian white noise with respective standard deviations of $\sigma_l = 1$ and $\sigma_{\dot{l}} = 0.1$. For simplicity during the simulations, without any loss of generalization, there is no distinction made between the static

and dynamic objects. The landmarks that are deemed as stable by the simulator are automatically added into the map and the unstable or dynamic landmarks are eliminated. Both LQ-FastSLAM and the proposed GM-FISST-SLAM algorithm were executed for the scenario described above. The comparisons of the estimated vehicle

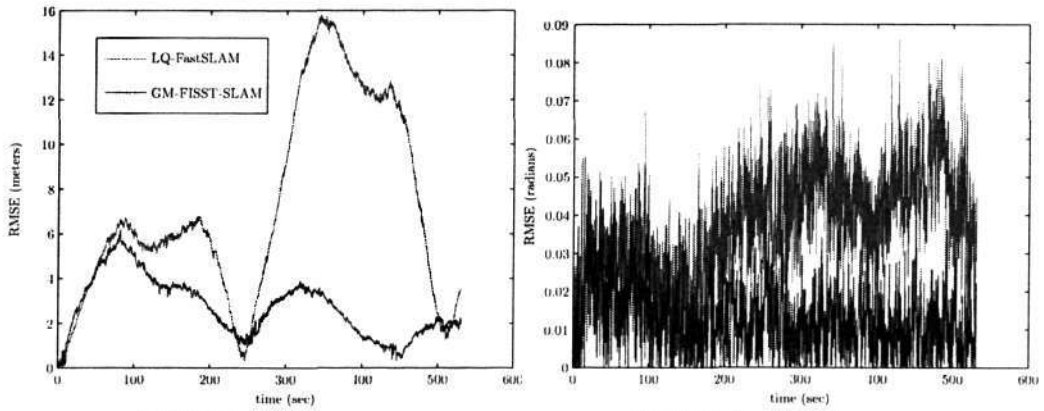


(a) LQ-FastSLAM *(b) GM-FISST-SLAM*

Figure 5.9 *Estimated (in blue) and true (in red) trajectories superimposed with estimated (red circles) and actual landmark locations (green asterisk) with three time-varying and 78 static landmarks.*

trajectory and the final map for LQ-FastSLAM and FISST-SLAM implementations with the clutter density λ_c set to 10 in dynamic environments are shown in fig. 5.9.

The performance of FastSLAM with the feature management strategy deteriorates under high clutter coupled with time-varying landmarks. This is clearly evident from the large number of clutter returns (as seen in fig. 5.11) incorrectly estimated as actual landmarks (in this case 188) after the complete run with a large RMS error of 15.8 m and a loop closing error of 3.6 m. The relatively large errors, as observed in fig. 5.10a are mainly due to the clutter and dynamic landmarks that is incorporated into the map as valid static landmarks, resulting in miss-associations. As for the proposed FISST-SLAM algorithm, it's performance also affected due to the presence of time-varying landmarks, however, it manages to filter the dynamic



(a) Vehicle position error *(b) Vehicle orientation error*

Figure 5.10 *A comparison of Vehicle pose estimation error between FISST-SLAM and LQ-FastSLAM for clutter rate $\lambda_c = 10$ in dynamic environment*

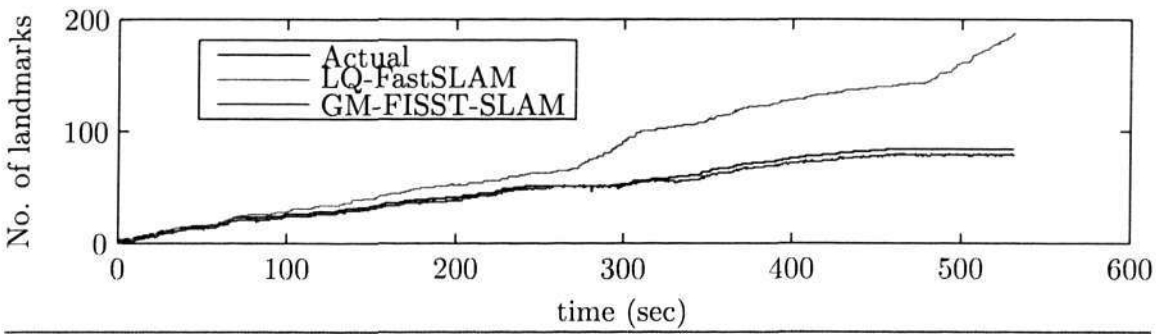


Figure 5.11 *No of landmarks incorporated into map as a function of time densely cluttered dynamic environment.*

landmarks as clutter and retains only the static landmarks within the map. For the complete closed-loop run the estimated number of landmarks in the dense clutter scenario in dynamic environment is 79, which is one excess of the actual number of static landmarks. It is also important to note that RMS error is about one-fourth of LQ-FastSLAM in dynamic environment with a loop closing error of 1.94 m as observed in fig. 5.10a.

However, if one needs to explicitly track dynamic objects in environments re-

quires what is known as dynamic-SLAM, where the dynamic objects need to be carefully modeled, which can then be introduced into RFS Multi-Landmark Map Evolution model (Section 5.4.1).

5.10 Algorithm Performance Evaluation in Offshore Underwater Environments

This section analyzes the performance of the proposed FISST-SLAM filter (SMC implementation) in the same offshore environment as discussed in Section 4.7.1. It also benchmarks it against the classical vector based FastSLAM 1.0 which is implemented with an efficient log-odds feature map management strategy (LQ-FastSLAM). The blazed array sonar is used as primary imaging sensor containing noisy range bearing intensity measurements, as was presented in Chapter 4. In this experiment, point features from the sonar signal returns were used as inputs to the filter, although the filters can easily be modified to accommodate other features such as lines, gradients etc. The extraction of point features from the sonar data is essentially a two step process. The first step involves in the detection of signals in presence of background noise. Adaptive threshold technique using OS-CFAR processor (discussed in Section 4.3) is used in the detection process to mainly identify the principal returns (landmark generated measurements). For each sonar scan the principal returns are then grouped together into clusters based on the regions of constant depth [62] and the cluster centroids are identified as point features. These point features which are the range-bearing measurements that encapsulate both landmark generated measurements and clutter generated measurements are fed as

an input to the filters. Since there is no absolute position sensor on the AUV, and GPS does not work underwater, a qualitative analysis for the vehicle pose estimates is performed as and when the AUV surfaces to obtain a GPS fix.

Both the LQ-FastSLAM and FISST-SLAM methods adopt particle based estimates for the vehicle pose and use 80 particles to sample the vehicle trajectories. The estimated vehicle trajectory and the corresponding multi-landmark map for FISST-SLAM and LQ-FastSLAM are shown in fig. 5.12. The noisy blazed array sonar measurements have increased the data association ambiguity and hence making the landmark map management task more challenging. Under such noisy conditions, in the presence of large amount of clutter, the merits of the proposed FISST-SLAM filter framework are clearly noticed and can be seen to outperform the LQ-FastSLAM (refer fig. 5.12). It is also important to note that the GPS fixes (shown using green asterisk) which are assumed to be the ground truth in these experiments are obtained only when the AUV ascends to the surface. The vehicle trajectory from the proposed approach closely tracks the ground truth and is almost similar to that of LQ-FastSLAM algorithm due to the particle representation of the vehicle trajectory. However, the landmark map estimate obtained from LQ-FastSLAM hardly compares with that of the ground truth, while FISST-SLAM manages to consistently detect the landmarks. As explained in Section 4.7.1, the three sonar reflector arrays are placed approximately 18m apart and fourth array 30m apart are consistently detected along with the breakwaters of the marina by the proposed FISST-SLAM filter, rejecting clutter generated measurements in the process as shown in fig. 5.12b. In comparison, LQ-FastSLAM also detects these landmarks, however, along with a large number of spurious measurements which are clearly evident in fig. 5.12a.

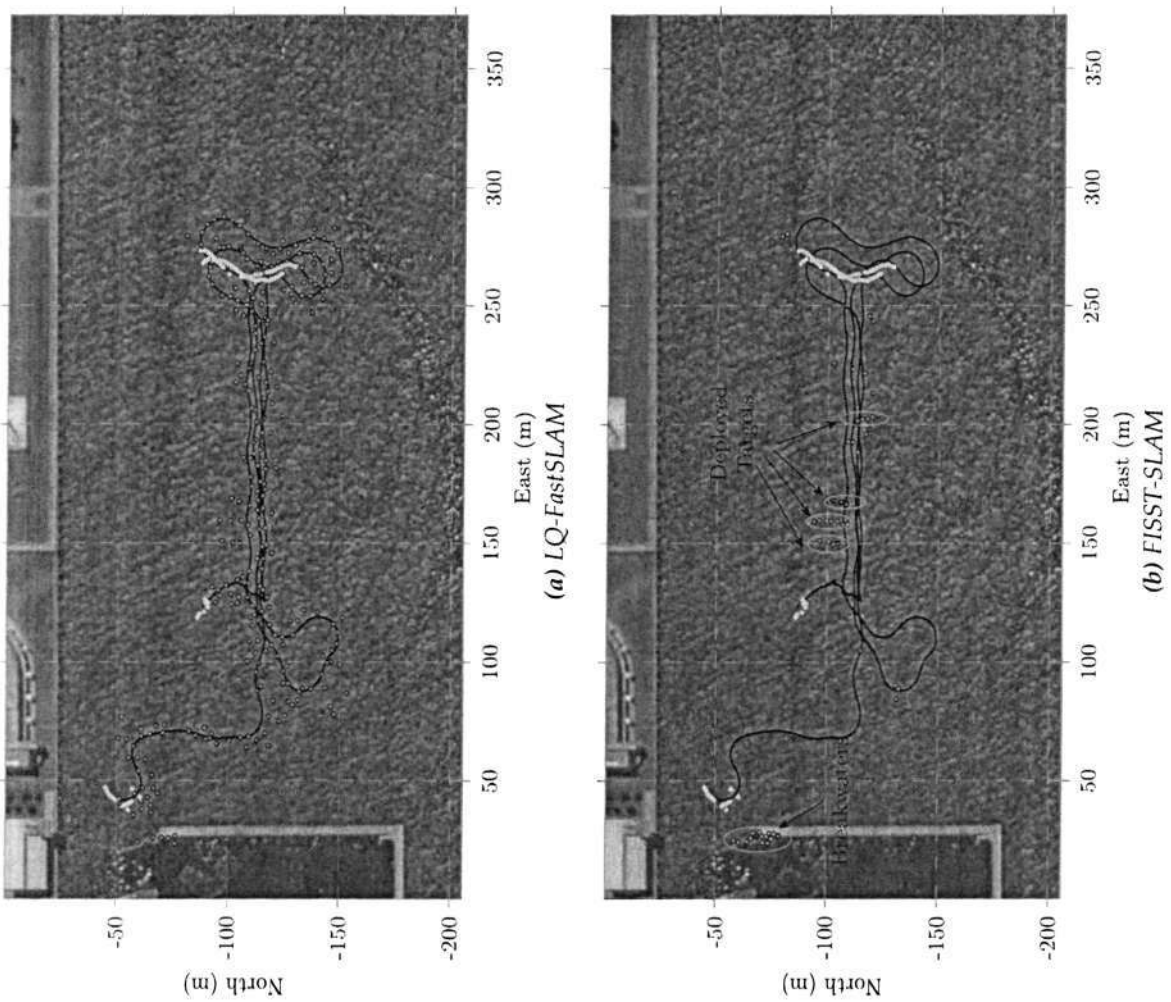
Additional results validating the proposed algorithm are shown in fig. 5.13

which reproduces an acoustic multi-landmark map mosaic generated by placing the blazed array sonar scans according to the LQ-FastSLAM and the proposed FISST-SLAM estimated AUV trajectories. The resulting acoustic map matches almost perfectly with the real position of the breakwaters of the marina as seen from the high intensity returns at the upper left corner of the figure. These figures also highlight the landmark detection and their extraction process based on principal sonar signal returns from individual sonar scans and their subsequent inclusion into the multi-landmark map based on SLAM filter. It is evident from fig. 5.13 that the proposed FISST-SLAM filter outperforms LQ-FastSLAM filter in rejecting the noisy clutter measurements and consistently incorporating landmark generated measurements into the multi-landmark map.

5.11 Chapter Summary

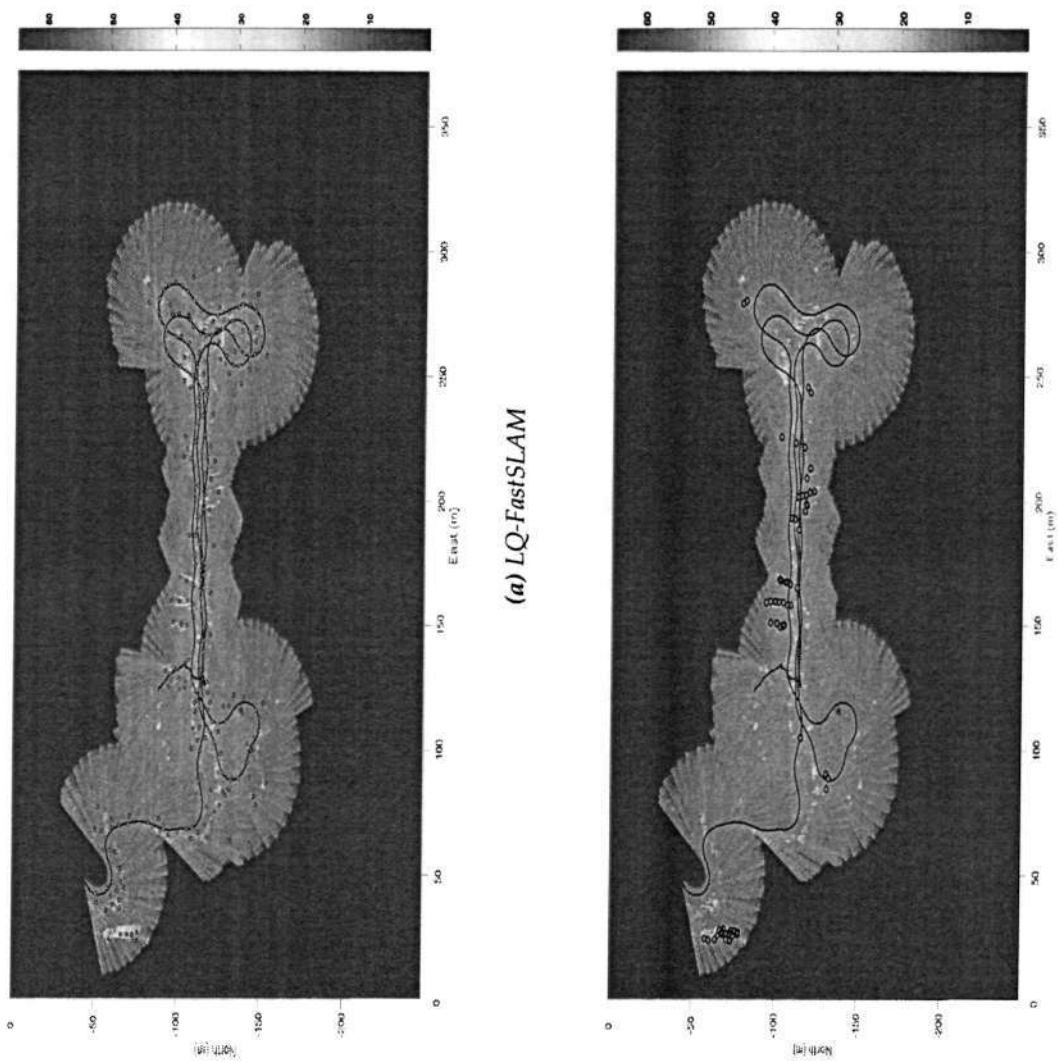
In this chapter we presented a random finite set theoretic formulation and a solution to the SLAM problem. As compared to the classical vector based formulation of SLAM, the major attribute of this FISST-SLAM formulation is that it provides for a more accurate representation and processing of key elements that affect SLAM performance, in particular, the ever evolving map of landmarks, and the multiple-landmark measurement model. Missed detections, false alarms, association of landmarks to observations and new landmarks are accounted for directly and automatically with in the SLAM filter framework, unlike in stochastic vector realizations of SLAM algorithms, where these issues have to be dealt with explicitly with added heuristics outside the Bayes optimal filter. In this chapter we also presented the incorporation of the first order statistical moment (PHD) filter into the

FISST-SLAM framework to alleviate the computational complexity of the optimal Bayes RFS multi-landmark map estimation filter. Two alternative implementations of the FISST-SLAM algorithm using sequential Monte Carlo and Gaussian mixture approximation methods were also developed. In this chapter we emphasized on importance of the overall multi-landmark map error metric, that is often neglected in the classical vector based SLAM literature. Based on the developments in the target tracking community, a Wasserstein miss-distance metric was introduced to quantify for the overall multi-landmark mapping error. To our knowledge, this quantitative metric is used for the first time in the field of autonomous vehicle navigation. Extensive experiments in form of rigorous simulation studies and real offshore underwater sea trials were conducted to validate the proposed filter framework. The results from all the experiments, high clutter scenarios in particular, comprehensively demonstrate the superiority of the FISST-SLAM filter in comparison with the classical stochastic vector based FastSLAM filter.



(a) LQ-FastSLAM (b) FISST-SLAM

Figure 5.12 *The estimated AUV trajectory (black line) and corresponding multi-landmark map (pink circles) superimposed on satellite images for all the filters based on blazed array sonar measurements. The GPS ground truth (green asterisk) is overlaid indicating locations when the AUV surfaced.*



(a) LQ-FastSLAM

(b) FISST-SLAM

Figure 5.13 *Acoustic mosaic built overlapping blazed array sonar scans along the FastSLAM estimated trajectory (top) and the FISST-SLAM estimated trajectory (bottom). The multi-landmark map (pink circles) estimated by both the algorithm are also overlaid. The sonar reflectors placed and breakwaters present in the test environment are clearly picked up the FISST-SLAM algorithm.*

6

Conclusions and Recommendations

This chapter summarizes the work presented in this thesis and discusses the contributions made to the field of autonomous underwater vehicle navigation. Some research directions opened as a result of the contributions are then discussed. This chapter is organized as follows. While Section 6.1 presents, a chapter-by-chapter summary of the research undertaken, Section 6.2, presents the overall conclusions of this thesis. Section 6.3 presents the possible failure modes of the proposed framework while Section 6.4 briefly discusses the scope for future research.

6.1 Summary

In **Chapter 1** we presented the main objectives of this thesis. Navigation, which is the combination of precise positioning and guidance, was put forth as being the fundamental challenge to the mobile robotics community and more specifically to the underwater mobile robotics community. This chapter described a number of existing navigation methods currently used by underwater vehicles, highlighting

their limitations. Special importance was provided to onboard sensing systems with emphasis on feature based simultaneous localization and mapping (SLAM) algorithms.

In **Chapter 2** we presented the classical vector based state of the art SLAM algorithms. A probabilistic approach to sensor and vehicle modeling, localization and mapping was presented. It was shown that the EKF-SLAM filter is a form of the optimal Bayesian filter constrained to estimating only up to the second order statistical moments, namely, the mean and variance of a distribution. These constraints have made the EKF an efficient and widely used navigation algorithm. However, a number of issues associated with problems with the application of the EKF-SLAM filter were also discussed along with methods in the literature to mitigate these problems. Numerical methods in form of sequential Monte-Carlo techniques or particle filter techniques applied to the SLAM problem in the classical vector based SLAM literature were also presented. The latter half of the Chapter 2 discussed the limitations of the classical vector approaches that are due to the representation of the map evolution and measurement models and proposed an alternative framework using random finite sets.

By adopting a finite random set theoretic framework, a multi-landmark map evolution model that accounts for the new landmarks appearing in the field of view of sensor and a natural multi-landmark (map) measurement model that explicitly accounts for detection uncertainty, clutter and unknown data association for autonomous vehicle navigation was constructed in **Chapter 3**. An optimal Bayesian random finite set multi-landmark map estimation filter was then presented. First order statistical moments filter, also referred to as PHD filter was then discussed to overcome the computational complexity of the optimal Bayes filter.

In **Chapter 4**, we addressed the issues of landmark detection from noisy sensor measurements and its impact on autonomous navigation applications. The fixed detection and the adaptive detection methods were introduced for performing landmark detection with a blazed array sonar. The PHD clutter rejection filter was introduced along with its implementational details. The performance of the particle PHD clutter rejection filter was validated using synthetic sonar data. A novel clutter rejection algorithm was presented based on the PHD clutter rejection filter for autonomous vehicle navigation. The novelty of the approach lies in the formulation of an integrated framework, where a random finite set based PHD filter was used as a clutter filter and integrated with the conventional vector based EKF-SLAM filter. This proposed navigation framework was applied then applied in offshore underwater field trials and benchmarked against the standard EKF-SLAM filter with an advanced map management strategy.

In **Chapter 5**, we presented a random finite set theoretic formulation and a solution to the SLAM problem. As compared to the classical vector based formulation of the SLAM problem, the major attribute of this FISST-SLAM formulation is that, it provides for a more accurate representation and processing of key elements that affect SLAM performance. Missed detections, false alarms, association of landmarks to observations and new landmarks are accounted for directly and automatically with in the SLAM filter framework, unlike in classical vector realizations of SLAM algorithms, where these issues have to be dealt with explicitly with added heuristics outside the Bayes optimal filter. In this chapter we also presented the incorporation of the first order statistical moment (PHD) filter into the FISST-SLAM framework to alleviate the computational complexity of the optimal Bayes RFS multi-landmark map estimation filter. Two alternative implementations of the FISST-SLAM algo-

rithm using sequential Monte Carlo and Gaussian mixture approximation methods were also developed. In this chapter we emphasized on importance of the overall multi-landmark map error metric, that is often neglected in the classical vector based SLAM literature. The Wasserstein metric was defined, and subsequently shown to be an easily computable and a meaningful miss-distance metric for the purpose of multi-landmark mapping performance evaluation. Extensive experiments in form of rigorous simulation studies and real offshore underwater sea trials were conducted to validate the proposed filter framework.

6.2 Conclusions

The successful development and application of key random finite set theoretic based algorithms and performance metrics in this thesis have demonstrated that, random finite set based methods are not only theoretically sound, but are also viable and effective approaches to the simultaneous localization and mapping problem in practice. In particular, the proposed FISST-SLAM algorithm, offer principled, reliable and efficient solutions for the joint detection and estimation of an unknown number of landmarks observed with unreliable sensors in the presence of clutter. Moreover, the proposed approaches were seen to outperform the conventional vector state space based stochastic approaches in certain circumstances in practical scenarios. The key conclusions of this dissertation are given as follows.

Adopting a finite random set theoretic framework enabled us to construct, a multi-landmark map evolution model that accounts for the new landmarks appearing in the field of view of sensor and a natural multi-landmark (map) measurement model that explicitly accounts for detection uncertainty, clutter and unknown data

association for autonomous vehicle navigation. These models were then used develop a clutter rejection algorithm based on moment approximations to the posterior density of the random finite set state for autonomous vehicle navigation systems and integrated in a novel fashion with the conventional vector state space based SLAM framework. Using the same models, a mathematically principled unifying random finite set theoretic SLAM framework, referred to as FISST-SLAM was developed, where the SLAM problem was reformulated so that the landmark map and the measurements were represented using random finite sets and the landmark map was jointly estimated with the vehicle state vector. Subsequently, two alternative implementations of the FISST-SLAM algorithm were also developed using sequential Monte Carlo and Gaussian mixture approximation methods. Additionally, using an easily computable and a meaningful miss-distance performance metric called the Wasserstein metric, the multi-landmark mapping performance was evaluated. Finally, the feasibility of the proposed algorithms were demonstrated through the use of extensive simulations and experiments using AUVs in real offshore underwater environments.

6.3 Failure Modes

While within the context of this thesis we have demonstrated that the unified random finite set theoretic framework has significant advantages over the classical vector based solutions for the SLAM problems, we are compelled to also point out what we think are the weak points of this framework. In particular, we believe the main weakness stems from the usage of the first order statistical moment within the FISST-SLAM filter framework. In the FISST-SLAM filter, in order to

alleviate the computational intractability of the RFS multi-landmark map distribution $f_{k|k}(\mathrm{M}_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k)})$, we have used its first order statistical moment, the PHD $D_{k|k}(\zeta_k|\mathrm{x}_k^{[i]}, \mathrm{Z}^{(k)})$ to estimate the landmark states. This first order approximation along with the assumption of Poisson distribution needed to arrive at closed form formulation for the PHD multi-landmark filter results in great loss of information which affects the FISST-SLAM filter performance. This loss in higher order cardinalities have significant impact in low signal to noise ratio scenarios in particular, resulting in unreliable estimates of the number of landmarks [38]. The PHD filter used for multi-landmark map estimation approximates the unknown number of landmarks (cardinality distribution) by a Poisson distribution with matching mean. Since the mean and variance of the Poisson distribution are the same, when the number of landmarks present is very high, the cardinalities based on the PHD filter tend to have a correspondingly high variance. In order to overcome this problem of erratic estimation of the number of landmarks in extremely low SNR scenarios, a modified version of the PHD filter with a relaxed Poisson assumption has been proposed in form of cardinalized PHD (CPHD) filter [71]. Thus incorporating CPHD filter for multi-landmark map estimation within the FISST-SLAM filter framework should further improve its performance.

6.4 Recommendations

This thesis has laid the foundation for a promising onboard sensing framework which is free of additional external deployments for autonomous underwater vehicle navigation. However, as always, there is room for additional improvement. Further areas of research that we have identified as warranting future investigation

are:

- Due to the erratic estimation of the number of landmarks in extremely low SNR scenarios, a modified version of the PHD filter with a relaxed Poisson assumption has been proposed in form of cardinalized PHD (CPHD) filter [71]. Incorporating CPHD filter for multi-landmark map estimation within the FISST-SLAM filter framework should further improve its performance. However, this comes at a cost of significant increase in the computational complexity.

- In this thesis, we have considered only the static landmarks in the environment. However, given the dynamic nature of the underwater environment, it would be extremely useful to incorporate moving features into the map. The map evolution presented in FISST-SLAM framework can be extended to include the dynamic features.

- In this thesis, we have assumed that the underwater vehicle has only three degrees of freedom and thus work under the constraint that pitch, roll and heave effects are assumed negligible. However, it will be desirable to extend the scope to six degree of freedom, so that the vehicle can navigate without any constraints in a three dimensional space. The framework put forth in this thesis should not theoretically pose a significant problem for this extension. However, the determining a point in a 3D space using a range bearing intensity sensor such as a blazed array sonar may pose a challenge. Perhaps, an alternative could be using two blazed array sonars mounted perpendicular to one another or 3D imaging sonars.

- The problem of cooperative mapping and localization by multiple AUVs is

particularly useful and extremely fascinating. The data fusion requirements for such systems can be very challenging due to inconsistencies involved in combining information from multiple sources. The random finite set framework developed in this thesis can be extended to perform cooperative mapping and localization of multiple AUVs.

6.5 A Final Word

For robotics researchers with an interest in autonomous vehicle navigation, the work in this thesis represents a departure from the classical vector based SLAM filter framework and a recognition that more complex, less structured environments must be addressed with the more elegant random finite set based solutions. The work described in this thesis sets the stage for the future of autonomous underwater vehicle navigation systems.

Appendix A

Finite Set Statistics

A.1 Random Finite Sets

This chapter introduces some of the key concepts of point process theory and its relation to Finite Set Statistics (FISST). Point process theory is a theoretical framework to analyze the occurrences of certain events in time or of features in a region [29]. Consider a multiple landmark scenario where at one time instant there are l measurements, given by $Z = \{z_1, z_2, \ldots, z_l\}$. The collection of all the subsets of Z is denoted by $\mathcal{F}(\mathcal{Z})$. The set $\mathcal{F}(\mathcal{Z})$ contains 2^l elements. The triple $(\Omega, \mathcal{F}(\mathcal{Z}), P)$ is called the *probability space*. The cardinality of a subset A of Z is denoted by $|A|$.

Definition 1. *A random finite set Ξ with values in $\mathcal{F}(\mathcal{Z})$ is a map $\Xi : \Omega \to \mathcal{F}(\mathcal{Z})$ such that $\Xi^{-1}(\{A\}) = \{\omega \in \Omega : \Xi(\omega) = A\} \in \mathcal{F}(\mathcal{Z})$, for any $A \subseteq Z$ [29], [87].*

To explain the concept of RFS, let us begin with hypothetical example as illustrated in fig. A.1.

Example 1. *Single Landmark Case:*

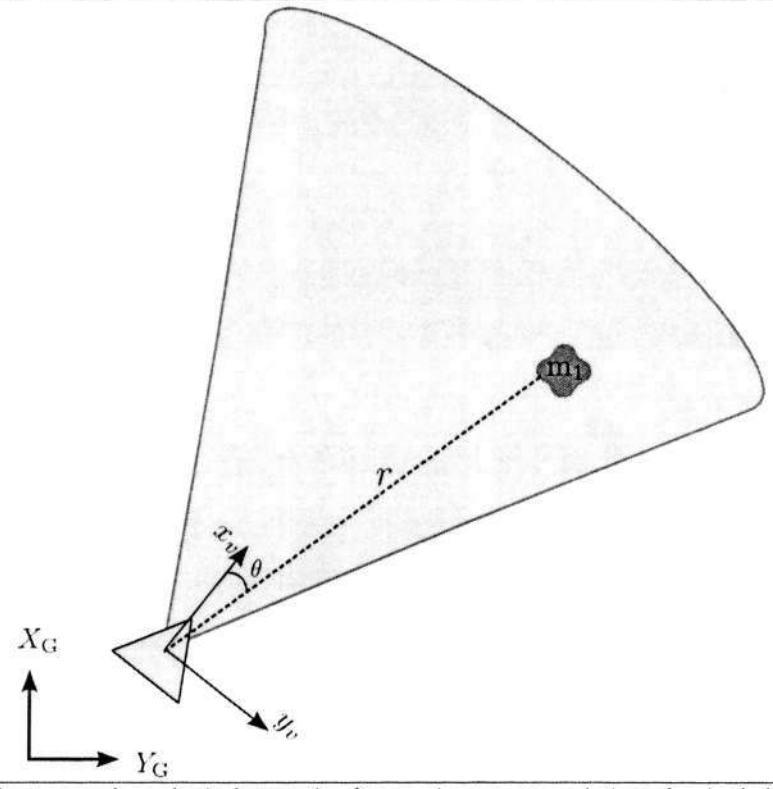


Figure A.1 *Illustrates a hypothetical scenario of an environment consisting of a single landmark (in red) observed by a vehicle (yellow) equipped with a GRBI sensor*

Consider a single landmark present in the environment. This is being observed by a generic range-bearing-intensity (GRBI) sensor (e.g., sonar, laser, radar, etc.). The measurement received by the sensor can be modeled as a vector $\mathrm{z} = (r, \theta, \mathrm{I})$ where r is the range, θ is the bearing, and I is the intensity. When the intensity falls below a particular threshold, the landmark is no longer detected by the sensor. This phenomenon can be mathematically modeled as a single randomly varying vector of the form $\Xi = \{\mathrm{Z}\}$, where Z is a three dimensional vector. The random finite set Ξ contains one element, and its instantiations is a singleton $\{\mathrm{z}\}$. Let S be the entire field of view of the sensor.

A.2 Single Landmark Tracking

We would want to detect this landmark using a sensor that provides an observation z_k at time k. At any time step k, we have a set consisting of the current observations and all the past observations, represented by $Z^k = [z_1, \ldots, z_k]^T$. The aim is to recursively estimate the probability density of the state given all the observations up to time k or commonly referred to as *posterior density* i.e.,

$$f_{k|k}(\mathrm{m}|Z^k) \tag{A.1}$$

The state m can encapsulate information such as pose, velocity or acceleration and can be estimated from the posterior density by using a variety of estimators. The most popular ones being

- Expected *a posteriori* (EAP) estimator

$$\hat{m}^{EAP} = \int \mathrm{m} f_{k|k}(\mathrm{m}|Z^k) d\mathrm{m} \tag{A.2}$$

- Maximum *a posteriori* (MAP) estimator

$$\hat{m}^{MAP} = \arg\max_{m} f_{k|k}(\mathrm{m}|Z^k) \tag{A.3}$$

A.2.1 Bayesian Filter

Given the state at time step $k-1$, Bayes theorem is used to determine the posterior density at time k. It is achieved in two steps:

- **Time-update**

Given a state m' at time $k-1$, the probability density function (pdf) of a transition to the state m at time k is

$$f_{k|k-1}(\mathrm{m}|\mathrm{m}') \tag{A.4}$$

This is commonly referred to as the *Markov transition density*. Given the motion model and the Bayesian posterior density $f_{k-1|k-1}(\mathrm{m}'|\mathrm{Z}^{k-1})$ at time $k-1$, a time-updated density is obtained using the Chapman-Kolmogorov equation[19]:

$$f_{k|k-1}(\mathrm{m}|\mathrm{Z}^{k-1}) = \int f_{k|k-1}(\mathrm{m}|\mathrm{m}')f_{k-1|k-1}(\mathrm{m}'|\mathrm{Z}^{k-1})d\mathrm{m}' \tag{A.5}$$

- **Data-update**

  This step incorporates the measurement available at time k via the sensor likelihood function given by

$$L_{k,\mathrm{z}}(\mathrm{m}) \triangleq f_k(\mathrm{z}|\mathrm{m}) \tag{A.6}$$

  The observation z_k is used to update (weight) the density produced by the time-update step to determine the final posterior density at time k:

$$f_{k|k}(\mathrm{m}|\mathrm{Z}^k) = \frac{f_k(\mathrm{z}_k|\mathrm{m})f_{k|k-1}(\mathrm{m}|\mathrm{Z}^{k-1})}{\int f_k(\mathrm{z}_k|\mathrm{m})f_{k|k-1}(\mathrm{m}|\mathrm{Z}^{k-1})d\mathrm{m}} \tag{A.7}$$

The posterior density function $f_{k|k}(\mathrm{m}|\mathrm{Z}^k)$ encapsulates everything about the state, based on the current observations and a priori information.

A.3 Belief Mass Functions

Belief mass functions (BMF) $\beta_{\Xi}(S)$ are the main building blocks for the formal multi-landmark RFS Bayes modeling. Specifically, they are central to the process of constructing true RFS map evolution model (Section 3.2) and RFS measurement model (Section 3.3). The concept of BMFs are explained further using a single landmark case, as illustrated in example 1. Consider a single landmark present in the environment. This is being observed by a generic range-bearing-intensity (GRBI) sensor. The measurement received by the sensor can be modeled as a vector $\mathrm{z} = (r, \theta, \mathrm{I})$ where r is the range, θ is the bearing, and I is the intensity. When the intensity falls below a particular threshold, the landmark is no longer detected by the sensor. Let $1 - p$ be the probability that the landmark is no longer detected by the sensor. This phenomenon can be mathematically modeled as a randomly varying finite set Ξ such that:

- $\Xi = 0$ with probability $1 - p$.
- $|\Xi| = 1$ with probability p and $\Xi = \{\mathrm{Z}\}$, where Z is a two dimensional vector conditioned on the fact that $\Xi \neq 0$.

The probability that Ξ will be contained within the sensor FOV S, is given by

$$\begin{aligned}
\beta_{\Xi}(S) &= P(\Xi \subseteq S) \\
&= P(\Xi \subseteq S, \Xi = \emptyset) + P(\Xi \subseteq S, \Xi \neq \emptyset) \\
&= P(\Xi = \emptyset) + P(\Xi \neq \emptyset).P(\Xi \subseteq S | \Xi \neq \emptyset) \\
&= 1 - p + p.P(\mathrm{Z} \in S | \Xi \neq \emptyset) \\
&= 1 - p + p \int_S f_{\mathrm{Z}}(\mathrm{z}) d\mathrm{z}
\end{aligned} \tag{A.8}$$

where, $f_Z(z)$ is the probability density function of Z. Thus $\beta_\Xi(S)$ is completely characterized by the function defined by

$$f_\Xi(Z) \triangleq \begin{cases} 1-p & \text{if } Z = \emptyset \\ p.f_Z(z) & \text{if } Z = \{z\} \\ 0 & \text{if } Z = \{z_1, z_2\}, z_1 \neq z_2 \\ 0 & \text{if otherwise} \end{cases} \tag{A.9}$$

Now, the belief mass function can be expressed as set integral of $f_\Xi(Z)$ as follows:

$$\begin{aligned} \beta_\Xi(S) &= 1 - p + p \int_S f_Z(z)dz \\ &= 1 - p + \int_S p.f_Z(z)dz + \frac{1}{2}\int_{S\times S} 0 dz_1 dz_2 + 0 + \dots \\ &= f_\Xi(\emptyset) + \int_S f_\Xi(z)dz + \frac{1}{2}\int_{S\times S} f_\Xi(\{z_1, z_2\})dz_1 dz_2 + \dots \\ &= \int_S f_\Xi(Z)\delta Z \end{aligned} \tag{A.10}$$

In eqn. A.10, the first term is the probability of having no landmarks in S, the second term is the probability of having one landmark in S, and the next term is the probability of having two landmarks in S, and so on. If the set integral is taken over entire space, and not confined to any subset of S, then it should integrate to one, like the probability density function.

A.3.1 Set Integrals

The belief mass function $\beta_{\Xi}(S)$ discussed above, can be expressed as a set integral over a region S (note the term δZ) can be written as

$$\begin{aligned}\int_S f_{\Xi}(Z)\delta Z &\triangleq \sum_{l=0}^{\infty}\frac{1}{l!}\int_{\underbrace{S\times\ldots\times S}_{l}} f_{\Xi}(\{z_1,\ldots,z_l\})dz_1\ldots dz_l \\ &= f(\emptyset)+\int_S f(\{z\})dz+\frac{1}{2}\int_{S\times S} f(\{z_1,z_2\})dz_1dz_2+\ldots\end{aligned} \tag{A.11}$$

For all $l \geq 2$, define the function $f_{\Xi_l}(z_1,\ldots,z_l)$ in l vector variables $z_1,\ldots,z_l$ given by,

$$f_{\Xi_l}(z_1,\ldots,z_l) \triangleq \begin{cases} \frac{1}{l!}f_{\Xi}(\{z_1,\ldots,z_l\}) & \text{if } z_1,\ldots,z_l \text{ are distinct} \\ 0 & \text{if otherwise} \end{cases} \tag{A.12}$$

Then each term in eqn. A.11 is given by,

$$\int_{\underbrace{S\times\ldots\times S}_{l}} f_{\Xi}(\{z_1,\ldots,z_l\}) = l!\int_{\underbrace{S\times\ldots\times S}_{l}} f_{\Xi_l}(z_1,\ldots,z_l) \tag{A.13}$$

A.4 Probability-generating Functionals

Probability-generating functionals(p.g.fl) are often used to transform difficult mathematical problems into simpler ones. Given the probability density function $f_{\Xi}(Z)$ of the RFS Ξ, the p.g.fl is

$$G_{\Xi}[h] \triangleq \int h^{Z}.f_{\Xi}(Z)\delta Z \tag{A.14}$$

where h^Z for a test function $h(z)$ is given by

$$h^Z \triangleq \begin{cases} 1 & \text{if } Z = \emptyset \\ \prod_{z \in Z} h(z) & \text{if otherwise} \end{cases} \tag{A.15}$$

under the assumption that all random variables are independent. Substituting eqn. A.15 in eqn. A.14, we get

$$G_\Xi[h] \triangleq f_\Xi(\{\emptyset\}) + \int h(z).f_\Xi(\{z\})dz + \frac{1}{2}\int h(z_1).h(z_2).f_\Xi(\{z_1,z_2\})dz_1dz_2 + \ldots \tag{A.16}$$

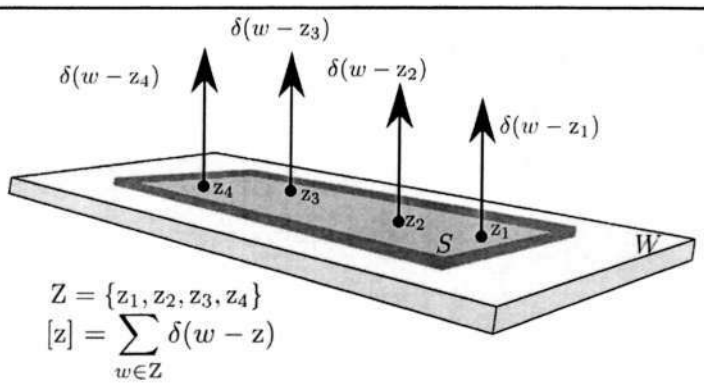


Figure A.2 *Point process theory*

If the test function $h(\cdot)$ is the Dirac probability measure $\mathbf{1}_S(\cdot)$, of the elements of Z existing in the region S of the state space W, i.e., $S \subseteq W$ (as shown in fig. A.2) and

$$h(z_i) = \mathbf{1}_S(z_i) = \begin{cases} 1 & z_i \in S \\ 0 & z_i \notin S \end{cases} \tag{A.17}$$

then, the belief mass function $\beta_\Xi(S)$ of Ξ from eqn. A.9 can be rewritten as the

expected value of $\mathbf{1}_S$

$$\beta_\Xi(S) = \int \mathbf{1}_S f_\Xi(Z)\delta Z \tag{A.18}$$

A.5 Set Density and Set Derivatives

In ordinary calculus, it is well known that the probability density function of a random variable is the Radon-Nikodym derivative of the induced measure with respect to some base measure, i.e., for any random vector Z, there is an essentially unique function $f_Z(z)$ such that, $\forall S$,

$$\int_S f_Z(z)dz = P(Z \in S) \tag{A.19}$$

The density function $f_Z(z)$ is the probability density function of Z. Along the same veins, using the first Radon-Nikodym theorem [44], it has been shown that the set integral (as in eqn. A.11) has an inverse operator as follows,

$$\int_S \frac{\delta\beta_\Xi}{\delta Z}(\emptyset)\delta Z \triangleq P(\Xi \subseteq S) = \beta_\Xi(S) \tag{A.20}$$

The probability density function for a RFS Ξ can be constructed by taking set derivatives of the belief mass function of Ξ:

$$\beta_\Xi(S) = \int_S f_\Xi(Z)\delta Z \tag{A.21}$$

where,

$$f_\Xi(Z) = \left[\frac{\delta\beta_\Xi}{\delta Z}(S)\right]_{S=\emptyset} \tag{A.22}$$

A.6 Set Calculus

If $\beta_1, \beta_2, \ldots, \beta_n$ are the belief measures, and $a_1, a_2, \ldots, a_n$ are real numbers, then

- **Constant Rule**

$$\frac{\delta}{\delta Z}(a_1) = 0 \qquad \text{if} \qquad Z = \emptyset \tag{A.23}$$

- **Sum Rule**

$$\frac{\delta}{\delta Z}(a_1\beta_1(S) + a_2\beta_2(S)) = a_1\frac{\delta}{\delta Z}\beta_1(S) + a_1\frac{\delta}{\delta Z}\beta_1(S) \tag{A.24}$$

- **Product Rule**

$$\frac{\delta}{\delta Z}(\beta_1(S)\beta_2(S)) = \sum_{W \subseteq Z}\left[\frac{\delta}{\delta W}\beta_1(S)\frac{\delta}{\delta(Z-W)}\beta_2(S)\right] \tag{A.25}$$

- **Power Rule**

$$\frac{\delta}{\delta Z}p(S)^n = \begin{cases} \dfrac{n!}{(n-|z|)!}p(S)^{n-|z|}\displaystyle\prod_{z\in Z} f_p(z) & \text{if } |Z| \leq n \\ 0 & \text{if } |Z| > n \end{cases} \tag{A.26}$$

- **Convolution formula**

  If $\Xi = \Xi_1 \cup \ldots \cup \Xi_n$ and $\beta_\Xi(S) = \beta_{\Xi_1}(S) \cup \ldots \cup \beta_{\Xi_n}(S)$ then, probability density function of Ξ is

$$f_\Xi(Z) = \sum_{W_1 \uplus \ldots \uplus = Z} f_{\Xi_1}(W_1)\ldots f_{\Xi_n}(W_n) \tag{A.27}$$

  where the summation is taken over all mutually disjoint subsets $W_1, \ldots, W_n$ of

Z, such that $W_1 \cup \ldots \cup W_n = Z$

- **Special case: n=2**

$$f_{\Xi}(Z) = \sum_{W \subseteq Z} f_{\Xi_1}(W) f_{\Xi_2}(Z - W) \tag{A.28}$$

A.7 Optimal Bayes RFS Multi-landmark Filter

The concepts of FISST calculus is used to construct the multi-landmark version of the Bayes filter, in which full FISST Bayes posterior is propagated in time as a two-step process as follows [68]:

- **Time-Update Step**

  The predictor equation is analogous to single target Bayesian prediction A.5 with set integral instead of the vector integral given by

$$f_{k|k-1}(\mathrm{M}|\mathrm{Z}^{(k-1)}) = \int f_{k|k-1}(\mathrm{M}|\mathrm{M}') f_{k-1|k-1}(\mathrm{M}'|\mathrm{Z}^{(k-1)}) \delta \mathrm{M}' \tag{A.29}$$

  where, $\mathrm{Z}^{(k)} : \mathrm{Z}_1, \ldots, \mathrm{Z}_k$ is the time sequence of measurement sets.

- **Data-Update Step**

  The corrector equation is analogous to single target Bayesian update A.7 given by

$$f_{k|k}(\mathrm{M}|\mathrm{Z}^{(k)}) = \frac{f_k(\mathrm{Z}_k|\mathrm{M}) f_{k|k-1}(\mathrm{M}|\mathrm{Z}^{(k-1)})}{\int f_k(\mathrm{Z}_k|\mathrm{M}) f_{k|k-1}(\mathrm{M}|\mathrm{Z}^{(k-1)}) d\mathrm{M}} \tag{A.30}$$

A.8 First order Statistical Moment Filter

As an analogous statistic to the mean of the constant gain Kalman filter, Mahler proposed the usage of probability hypothesis density (PHD) as the first order statistical moment to approximate the multi-landmark posterior of the Bayes filter [68].

A.8.1 The Probability Hypothesis Density

Let $f_{\Xi}(\mathrm{M})$ be the multi-landmark probability distribution function of a random finite set Ξ.

Definition 2. *The integral of PHD $D_{\Xi}(\mathrm{m})$ in any region S provides the expected number of landmarks in that region.*

$$N_{\Xi}(S) = \int_{S} D_{\Xi}(\mathrm{m}) = E[|S \cap \Xi|] \tag{A.31}$$

A.8.2 Differentiation Formula for PHD

Using the concepts of p.g.fl and its properties, one can construct the PHD [69]. Taking a gradient derivative of p.g.fl. eq A.14, we get:

$$\frac{\delta G_{\Xi}[h]}{\delta \mathrm{m}} = \int \left(\frac{\delta}{\delta \mathrm{m}} h^{\mathrm{M}} \right) f_{\Xi}(\mathrm{M}) \delta(\mathrm{M}) \tag{A.32}$$

Explicitly expanding the derivative term, and using set derivative properties [69]

and eqn. A.15,

$$\begin{aligned}\frac{\delta}{\delta \mathrm{m}} h^{\mathrm{M}} &= \frac{\delta}{\delta \mathrm{m}} \prod_i h(\mathrm{m})\Big|_{\mathrm{m}=\mathrm{m}_i} \\ &= \frac{\delta}{\delta \mathrm{m}} [h(\mathrm{m})|_{\mathrm{m}=\mathrm{m}_1} \dots h(\mathrm{m})|_{\mathrm{m}=\mathrm{m}_n}] \\ &= \left(\frac{\delta}{\delta \mathrm{m}} h(\mathrm{m})\Big|_{\mathrm{m}=\mathrm{m}_1}\right) h(\mathrm{m}_2) \dots h(\mathrm{m}_n) + \dots + h(\mathrm{m}_1) \dots h(\mathrm{m}_{n-1}) \left(\frac{\delta}{\delta \mathrm{m}} h(\mathrm{m})\Big|_{\mathrm{m}=\mathrm{m}_n}\right) \\ &= \sum_{i=1}^{n} h(\mathrm{m}_1) \dots \left(\frac{\delta}{\delta \mathrm{m}} h(\mathrm{m})\Big|_{\mathrm{m}=\mathrm{m}_i}\right) \dots h(\mathrm{m}_n)\end{aligned} \tag{A.33}$$

If the test function $h(\cdot)$ is a Dirac probability measure as in eqn. A.17, then derivative

$$\frac{\delta}{\delta \mathrm{m}} h(\mathrm{m})\Big|_{\mathrm{m}=\mathrm{m}_i} = \delta(\mathrm{m} - \mathrm{m}_i) \tag{A.34}$$

becomes a Dirac-delta function. Hence eqn. A.33 can be simplified as,

$$\begin{aligned}\frac{\delta}{\delta \mathrm{m}} h^{\mathrm{M}} &= \sum_{i=1}^{n} (1) \dots \delta(\mathrm{m} - \mathrm{m}_i) \dots (1) \\ &= \sum_{w \in \mathrm{M}} \delta(\mathrm{m} - w)\end{aligned} \tag{A.35}$$

So the gradient derivative of p.g.fl. in eqn. A.32 can be re-written as,

$$\frac{\delta G_{\Xi}[h]}{\delta \mathrm{m}} = \int \left(\sum_{w \in \mathrm{M}} \delta(\mathrm{m} - w)\right) f_{\Xi}(\mathrm{M}) \delta(\mathrm{M}) \tag{A.36}$$

Thus, the gradient derivative of p.g.fl., when $h(\cdot)$ is $\mathbf{1}_S$ is the first order moment

density given by,

$$\begin{aligned} D_{\Xi}(\mathrm{m}) = D_{\Xi}(\{\mathrm{m}\}) &= \left.\frac{\delta G_{\Xi}[h]}{\delta \mathrm{m}}\right|_{h=\mathbf{1}_S} \\ &= \int \left(\sum_{w \in \mathrm{M}} \delta(\mathrm{m} - w) \right) f_{\Xi}(\mathrm{M}) \delta(\mathrm{M}) \\ &= \int [\mathrm{m}].f_{\Xi}(\mathrm{M}) \delta(\mathrm{M}) \end{aligned} \tag{A.37}$$

It is called the probability hypothesis density (PHD) of $f_{\Xi}(\mathrm{M})$.

Definition 3. ***Multi-landmark moment Density:*** *The m-th order moment density is given by the m-th order derivative of p.g.fl.*

$$D_{\Xi}(\{\mathrm{m}_1, \ldots, \mathrm{m}_n\}) = \left.\frac{\delta^m G_{\Xi}[h]}{\delta \mathrm{m}_1, \ldots, \delta \mathrm{m}_n}\right|_{h=\mathbf{1}_S} \tag{A.38}$$

So, the ***first order moment density (i.e., the PHD)*** *is given by*

$$D_{\Xi}(\{\mathrm{m}\}) = \left.\frac{\delta G_{\Xi}[h]}{\delta \mathrm{m}}\right|_{h=\mathbf{1}_S} \tag{A.39}$$

A.8.3 Integration Formula for PHD

We know that from eqn. A.37 that,

$$\begin{aligned} D_{\Xi}(\mathrm{m}) &= \int [\mathrm{m}].f_{\Xi}(\mathrm{M})\delta(\mathrm{M}) \\ &= \int \left(\sum_{w \in \mathrm{M}} \delta(\mathrm{m} - w) \right) f_{\Xi}(\mathrm{M})\delta(\mathrm{M}) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int [\delta(\mathrm{m} - \mathrm{m}_1) + \ldots + \delta(\mathrm{m} - \mathrm{m}_n)] f_{\Xi}(\{\mathrm{m}_1, \ldots, \mathrm{m}_n\}) d\mathrm{m}_1 \ldots d\mathrm{m}_n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int f_{\Xi}(\{\mathrm{m}, \mathrm{m}_1, \ldots, \mathrm{m}_n\}) d\mathrm{m}_1 \ldots d\mathrm{m}_n \\ &= \int f_{\Xi}(\{\mathrm{m}\} \cup \mathrm{W})\delta W \end{aligned} \tag{A.40}$$

Definition 4. ***Probability Hypothesis Density:*** *Given the* $\Xi = \Xi_{k|k}$ *and* $f_{\Xi}(\mathrm{M}) = f_{k|k}(\mathrm{M}|Z^{(k)})$ *then its PHD is given by,*

$$D_{k|k}(\mathrm{M}|Z^{(k)}) = \int f_{k|k}(\mathrm{M} \cup W|Z^{(k)})\delta W \tag{A.41}$$

Lemma 1. ***PHD of a Poisson process:*** *If the multi-landmark distribution is a Poisson process*

$$f(\mathrm{M}) = e^{-\mu} \prod_{\mathrm{m} \in \mathrm{M}} \mu s(\mathrm{m}) \tag{A.42}$$

then its PHD is given by

$$D(\mathrm{m}) = \mu.s(\mathrm{m}) \tag{A.43}$$

Proof:

$$
\begin{aligned}
D(\mathrm{m}) &= \int f(\{\mathrm{m}\} \cup \mathrm{M}) \delta \mathrm{M} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{\mathrm{m}, \mathrm{m}_1, \ldots, \mathrm{m}_n\}) d\mathrm{m}_1 \ldots d\mathrm{m}_n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int e^{-\mu} \mu^{m+1} . (s(\mathrm{m}).s(\mathrm{m}_1) \ldots s(\mathrm{m}_n))\, d\mathrm{m}_1 \ldots d\mathrm{m}_n \qquad \text{(A.44)} \\
&= e^{-\mu} \mu . s(\mathrm{m}) \sum_{n=0}^{\infty} \frac{\mu^n}{n!} \\
&= \mu . s(\mathrm{m})
\end{aligned}
$$

Appendix B

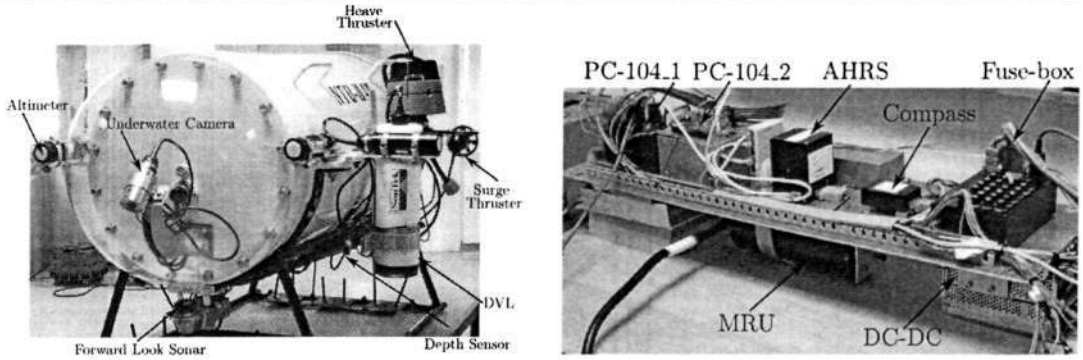
Research Test-bed Vehicles, Sensors and Specifications

B.1 The Research Test-bed Vehicles

B.1.1 The Autonomous Underwater Vehicle: NTU_UAV

The low-cost Nanyang Technological University Underwater Autonomous Vehicle(NTU_UAV) has 4 Degrees-of-freedom (DOF) using four-propeller thrusters (two for surge and two for heave) as shown in Fig. B.1. It carries on-board sensors, which can receive the inertial information, bottom and water tracking speeds of the vehicle, optical images, sonar images, range to obstacles and to the bottom, and the diving depth.

The internal sensors provide the motion parameters of the vehicle with respect to vehicle's coordinate system. The Motion Reference Unit (Seatex MRU-6 [3]) and Attitude Heading Reference Unit (Crossbow [2]) measure the accelerations and the



(a) NTU_UAV *(b) Internal layout*

Figure B.1 *Autonomous Underwater Vehicle: NTU_UAV with its onboard sensors*

Autonomous Underwater Vehicle	
Maximum Speed	*4 knots*
Cruise Speed	*2 knots*
Thrusters	*2xTecnadyne 250 (Surge), 2xTecnadyne 520 (Heave)*
Gross Vehicle Weight	*100 kg (Positively buoyant)*
Maximum Operating Depth	*100 m*
Hull	*Aluminum alloy*
Overall length	*130 cm*
Diameter	*80 cm*
Battery	*Silver-Zinc alkaline 24v, 135 AH*

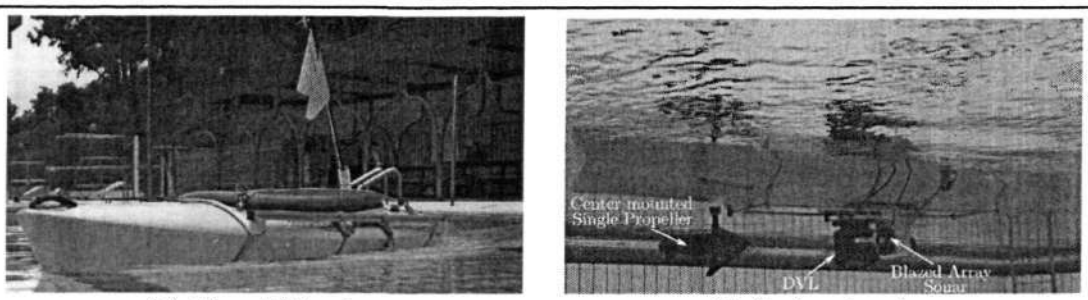
Table B.1 *Specifications of the NTU_UAV*

angular variations of the AUV with the assistance of a triple axis accelerometers, angular rate sensors and a magnetometer. TCM2 [4] combines three-axis magnetometer and a two-axis tilt sensor and acts as an integrated compass module. The three acoustic range finders or altimeters (Tritech PA-600 [8]) detect the distance to the objects in the underwater environment and are used as sensory inputs to an obstacle avoidance system. The monocular CCD camera (Tritech [8]) mounted in front of the AUV provides information for inspection, short-range target identification/tracking and for docking purposes. The pressure sensor (Druck PTX 1830 [5]) measures the external pressure experienced by the vehicle and is in turn translates it to measure the depth.

The speed of the vehicle with respect to the bottom or water can be measured using the Doppler Velocity Log (Sontek Argonaut DVL [7]). This sensor also provides the range to the bottom.

B.1.2 The Autonomous Surface Craft

A commercial off the shelf autonomous surface craft (ASC) from Robotics marine system was purchased and modified to suit our requirements. This ASC is equipped with a main vehicle computer and two PC-104's, a battery system, center aligned propulsion and steering systems along with Wi-fi communications system, a GPS, Compass and the vehicle specifications are listed in table. B.2. The detailed specifications of the sensors used a provided in the next section.



(a) Above Water view *(b) Underwater view*

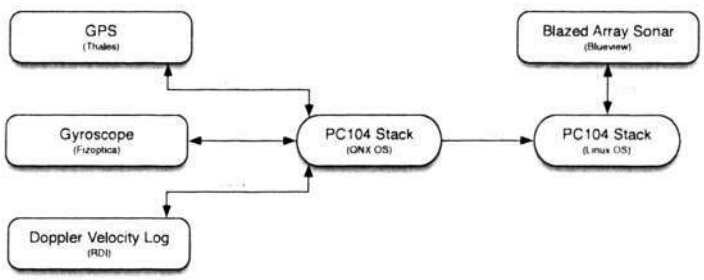
Figure B.2 *Autonomous Surface Craft in Swimming pool*

Autonomous Surface Craft	
Maximum Speed	*5 knots*
Cruise Speed	*3 knots*
Gross Vehicle Weight	*90 kg*
Hull	*High-density polyethylene*
Overall length	*304.8 cm*
Beam	*76.2 cm*
Battery	*12 v, 100 AH*

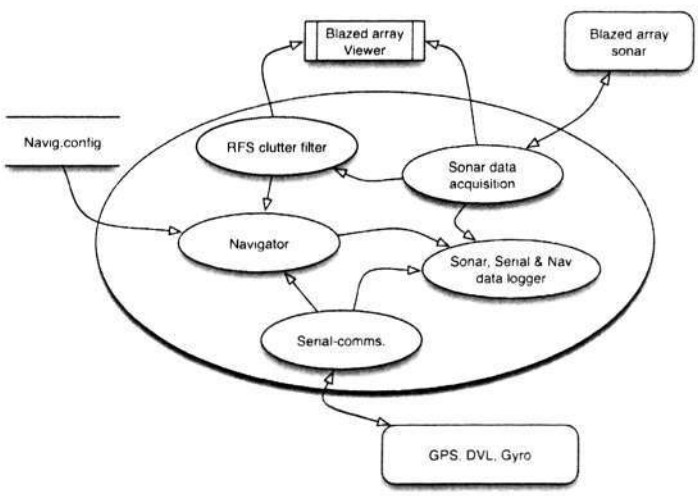
Table B.2 *Specifications of the ASC*

B.1.3 Navigational Software Architecture of Research Test-beds

An ASC and an AUV have been used interchangeably for algorithmic verification. The surface craft was mostly used in the shallow swimming pools for initial algorithmic validation and the underwater vehicle was used along the coastal waters of Singapore. Since its initial development stage both these systems have been developed with the same system architecture for ease of operation.



(a) Hardware Architecture



(b) Software Architecture

Figure B.3 *System Architecture of the Research Testbeds*

The systems are equipped with two PC104 stacks for online data capturing and

processing (as in fig. B.3a). The navigation sensor suite mainly comprises of a Doppler velocity log, a fiber optic gyroscope, compass, GPS and a depth sensor. The system is also equipped with a multi-beam sonar from Blueview that is interfaced to with PC104 Linux OS stack through a TCP/IP interface. The PC104 QNX OS stack captures data from the navigational sensory suite and transmits them to the PC104 Linux OS stack at 115k. This stack processes the incoming data along with the data obtained from the sonar to provide the navigational information. The software architecture encompasses a spectrum of functionality which ranges from multi-sensor communications, dynamic control, high precision navigation and path planning, concurrent mission task arbitration and execution, mission logging and playback. However, only navigational aspects of the software architecture are detailed as illustrated in fig. B.3b.

B.2 Sensor Specifications

This section details the main sensors used on-board the research test bed platforms during the course of this thesis.

B.2.1 Blazed Array Sonar

The blazed array sonars are acoustic analogous equivalent of diffraction grating in optics [104]. The signal flow in a blazed array sonar is illustrated in fig. B.4. A broadband pulse is transmitted through a transmitting array that creates a frequency dispersed sound field. It interacts with the underwater environment and the sound field is received by a receiver array which is then combined to form a single signal. An acoustic image is generated through time scale processing. Through out

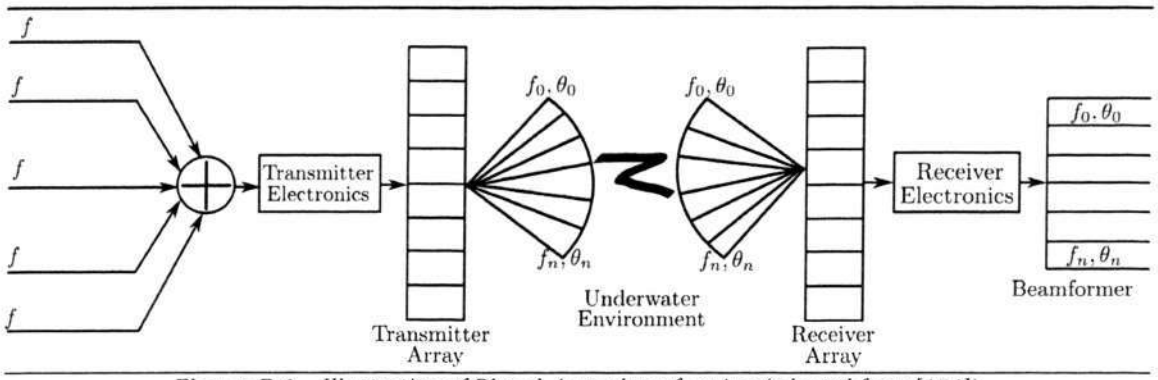


Figure B.4 *Illustration of Blazed Array beamforming (adapted from [104])*

the course of this thesis, we have used a 900kHz blazed array sonar from Blueview technologies (fig. B.5) whose specifications are listed in table B.3.

Figure B.5 *Blueview P900E-20*

B.2.2 Doppler Velocity Log

The Doppler Velocity Log (DVL) calculates the vehicle velocities, relative to the ocean floor or the water column based on the principle of Doppler shift [108]. Most DVLs have four transducers, each aligned at an angle of 30^0 from the vertical and arranged in a Janus configuration (facing opposite directions) as shown in the fig. B.6[6]. Each transducer transmits sound waves (referred to as ping) in a narrow beam at a known frequency (usually high frequency in the order of $300 - 1200kHz$).

Blazed Array Sonar	
Model	Blueview P900E-20
Field of View	$45^o \times 20^o$
Max Range	55 m
Beam Width	$1^o \times 20^o$
Number of Beams	256
Beam Spacing	0.18^o
Range Resolution	2.54 cm
Update Rate	Up to 10 Hz
Frequency	900 kHz
Communication	Ethernet
Voltage	12-48 VDC
Power	10 Watts
Weight in Air	1.86 kg
Weight in Water	0.45 kg
Depth Rating	300 m
Size	7.0in × 4.0in OD

Table B.3 *Technical Specifications of the Blazed Array Sonar*

The pings forming an acoustic beam strikes particulate matter suspended in the water column or the seabed. When the pings strike scattering centers, some of the sound energy is reflected along the acoustic beam to the transducer. The returned sound has a frequency (Doppler) shift proportional to the velocities of the scattering centers and water they are traveling in along the acoustic beam. In our experiments during the course of this thesis we have used DVLs from Sontek (Argonaut) and RD instruments (Workhorse navigator), with the latter as shown in fig. B.6 being used rather extensively.

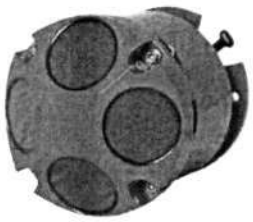


Figure B.6 *RDI DVL Workhorse Navigator*

The detailed specifications of the RD instruments Workhorse Navigator WHN

600 are provided in table B.4.

Doppler Velocity Log	
Model	*RDI Workhorse navigator WHN 600*
Bottom Velocity Single-ping precision S.D. @ 1 m/s	*±0.3 cm/s*
Long-term accuracy	*±0.2% ± 0.1cm/s*
Minimum altitude	*0.7 m*
Maximum altitude	*90 m*
Velocity range	*±10 m/s*
Velocity resolution	*0.1 cm/s*
Ping rate	*7 Hz max*
Water Reference Velocity Accuracy	*±0.3% ±0.2 cm/s*
Operating temperature	*-5 to 45° C*
Storage temperature	*-30 to 75° C*
Depth rating	*3000 m*
Weight in air	*15.8 kg*
Weight in water	*8.8 kg*
DC input	*20-50 VDC*
Current	*0.4 A*
Peak power @ 24 VDC	*21 w*
Average power (typical)	*3 w*
Integrated Sensor1:Compass	*±2° @ 60° dip, 0.5 g*
Integrated Sensor2:Tilt sensor	*±0.5° up to ±15°*
Integrated Sensor3:Temperature sensor	*−5° to 45° C*

Table B.4 *Technical Specifications of the Doppler Velocity Log*

B.2.3 Fiber Optic Gyroscope

A fiber optic gyroscope (FOG) provides precise rotational rate information based on the principle of interference of light to detect mechanical rotation. There are two light beams that travel along the fiber in opposite directions. The beam traveling against the rotation experiences a slightly shorter path than the other beam. The resulting phase shift affects the beams interference when they are combined. The intensity of the combined beam then depends on the rotation rate of the vehicle. A FOG from Fizoptika (fig. B.7) is used in our experiments whose specifications are listed in table B.5.

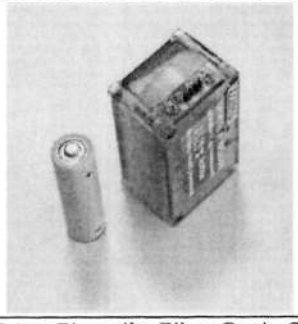


Figure B.7 *Fizoptika Fiber Optic Gyroscope*

Fiber Optic Gyroscope	
Model	*Fizoptika VG095M*
Rate range	*300 deg/s*
Scale Factor (SF)	*10 mV/deg/s*
Frequency range	*0 to 0.45 kHz*
Angle random walk	*0.1 deg /h*
Bias stability (steady state)	*15 deg / h (RMS)*
SF variation (steady state)	*0.1% (RMS)*
Readiness time	*0.1 s*
Nominal Operating Temperature	*−30° C to +70° C*
Maximum Endurance Temperature	*−55° C to +85° C*
Vibration (operating)	*2 g (RMS), 20 Hz to 500 Hz*
Vibration (endurance)	*6 g (RMS), 20 Hz to 2000 Hz*
Shocks (endurance)	*90 g, 1 ms*
Acceleration (operating)	*5 g*
Acceleration (endurance)	*20 g, 5 s*
Power dissipation	*1 W*
Total weight	*80 grams*
Housing material	*Aluminum alloy*
MTBF	*20000 hours (20° C, predicted)*
Lifetime (predicted)	*15 years*

Table B.5 *Technical Specifications of the Fiber Optic Gyroscope*

B.2.4 Global Positioning System

The Global Positioning Systems (GPS) system of satellites is extensively used to provide precise positioning data for land-based and air-based systems. Unfortunately, electromagnetic signals attenuate rapidly in water, and hence the use of GPS in the underwater environment is severely limited. However, in our experiments, we have used GPS (fig. B.8) for the purpose of validation by bringing the AUV to the surface at regular intervals. The detailed specifications of the GPS used in our experiments

is listed in table B.6.

Figure B.8 *Thales AC12 GPS receiver unit*

Global Positioning System	
Model	Thales A12/AC12
Real-time Accuracy	1 m
Single Frequency	Yes
Dual Frequency	No
Static Mode	Yes
Kinematic Mode	Yes
Real-time position accuracy- autonomous	3.0 m (CEP),5.0 m (95%)
Real-time position accuracy- differential	0.8 m (CEP), 1.5 m (95%)
Raw data output (code & carrier)	Yes
Velocity Accuracy	0.1 knots (95%)
Time to First Fix: re-acquisition	1 s
Time to First Fix: Cold	150 s
Time to First Fix: Warm	45 s
Time to First Fix: Hot	10 s
Channels	12
Channels: L1 GPS code and carrier	10
Channels: SBAS (WAAS/EGNOS/MSAS)	2
Raw Data Rates	1 Hz
Position data Output	Yes
Position Data Rate	1 Hz
Ouput: NMEA-0183 version	3
1PPS timing signal (5V TTL)	Yes
1PPS timing signal (5V TTL): Precision - stand-alone	250 ns
User Selectable standard datums	Yes
User-definable datums	Yes
RTCM Mode	Remote
RTCM Message Types	1, 3 and 9
Data Link: minimum	1200 bps
Power Consumption receiver	1 W
Input voltage	10 to 18 Vdc
Unit Height	29.5 mm
Unit Width	104.6 mm
Unit Length	104.6 mm
Weight	240 grams
Connector	SMA
Speed (max)	514 m/s
Altitude (max)	18,288 m
Raw data output (code & carrier)	Yes
Time to First Fix: re-acquisition	1 s

Table B.6 *Technical Specifications of the GPS*

Appendix C

Author's Publications

C.1 Journal Papers

- Bharath Kalyan, Kwang Wee Lee, and Sardha Wijesoma. FISST-SLAM: A finite set statistical approach to simultaneous localization and mapping. *International Journal of Robotic Research*, pp. 1251-1262, Vol. 29, No. 10, September 2010.

- Balasuriya Arjuna, Wijesoma Sardha, Bharath Kalyan, and Thomas Lim. An overview of NTU UAV: A research testbed. *Robotics Research Center Journal*, 3-1, June 2004.

- Sardha Wijesoma, Bharath Kalyan and Kwang Wee Lee. RBRFS-SLAM: Rao-Blackwellized Random Finite Set Theoretic Approach to SLAM. *Elsevier-RAS* (under review).

C.2 Technical Reports

- Sardha Wijesoma, Kwang Wee Lee, Bharath Kalyan, Dong Jun Feng and M.D.P. Moratuwage. Technical report on smart navigation and mapping system for the next generation of AUVs and ROVs in marine operations. *Technical Report TR-MPA-D2*, July 2009.

- Sardha Wijesoma, Bharath Kalyan, and Kwang Wee Lee. Technical report on simultaneous localization and map building for underwater systems. *Technical Report TR-DSO-D5*, September 2008.

C.3 Conference Papers

- Bharath Kalyan, K.W. Lee, W.S. Wijesoma, D Moratuwage, and Nicholas M. Patrikalakis. A random finite set based detection and tracking using 3D LIDAR in dynamic environments. Istanbul, Turkey, October 2010. IEEE Systems, Man & Cybernetics.

- Kwang Wee Lee, Bharath Kalyan, Sardha Wijesoma, Martin Adams, Franz S. Hover, and Nicholas M. Patrikalakis. Tracking random finite objects using 3D-LIDAR in marine environments. In Proceedings of the 2010 ACM Symposium on Applied Computing, pages 1282âĂŞ1287, Sierre, Switzerland, 2010. ACM.

- Parijath Deshpande, Mehul Sangekar, Bharath Kalyan, Mandar Chitre, Shiraz Shahabudeen, Venugopalan Pallayil, and Koay Teong Beng. Design and development of AUVs for cooperative missions. In *Defense Technology Asia*, Singapore, March 2007.

- Bharath Kalyan, Arjuna Balasuriya, and SardhaWijesoma. Multiple target tracking in underwater sonar images using Particle-PHD filter. In *OCEANS 2006 - Asia Pacific*, pages 1âĂŞ5, Singapore, 2006.

- Bharath Kalyan, Arjuna Balasuriya, Hayato Kondo, Toshihiro Maki, and Tamaki Ura. Motion estimation and mapping by autonomous underwater vehicles in sea environments. In *IEEE Oceans*, Brest, France, June 2005.

- Bharath Kalyan, Arjuna Balasuriya, Toshihiro Maki, Tamaki Ura, and Hayato Kondo. Embedding vision, Sonar INS based motion estimation algorithms for AUV navigation. In *2nd Workshop on Integration of Vision and Inertial Sensors, International Conference on Robotics and Automation*, Palau de Congresses de Catalunya, Barcelona, Spain, April 2005.

- Bharath Kalyan, Arjuna Balasuriya, Multi-Sensor Data fusion Approach for Terrain aided Navigation of Autonomous Underwater Vehicles, *IEEE Oceans*, November 9-12, 2004, Kobe, Japan

- Bharath Kalyan and Balasuriya.A. Sonar based automatic target detection scheme for underwater environments using CFAR techniques: A comparative study. In *IEEE Underwater technology*, Taiwan, 2004.

- Bharath Kalyan and Balasuriya.A. Multiple sensors based navigation scheme for AUV position estimation. In *IEEE Underwater technology*, Taiwan, 2004.

- Bharath Kalyan and Balasuriya.A. Sonar and vision based navigation schemes for autonomous underwater vehicles. In *International Conference on Control, Automation, Robotics and Vision*, Kunming, China, 2004.

- Balasuriya.A, Wijesoma.S, Bharath Kalyan, and Thomas Lim. Development of a test-bed AUV: the NTU-UAV. In *Second International Conference on Computational Intelligence Robotics and Autonomous Systems (CIRAS)*, Singapore, 2003.

References

[1] *Blueview Technologies - Blazed Array Sonar.* Available from: http://www.blueview.com/p900-20. x, 16

[2] *Crossbow Technology.* Available from: http://www.xbow.com/. 173

[3] *Motion Reference Unit - Kongsberg Maritime.* Available from: http://www.km.kongsberg.com/. 173

[4] *PNI Sensor Corporation - Sensors and Modules.* Available from: http://www.pnicorp.com/products/. 174

[5] *Pressure Transducer, Transmitters and Differential Pressure Sensors - GE Sensing.* Available from: http://www.gesensing.com/productservices/tech_pressure.htm. 174

[6] *RD Instruments.* Available from: http://www.rdinstruments.com/navigator.html. 6, 178

[7] *SonTek DVL water velocity sensors.* Available from: http://www.sontek.com/. 175

[8] *Tritech International Limited.* Available from: http://www.tritech.co.uk/. x, 14, 15, 174

[9] P.E. An, A.J. Healey, S.M. Smith, and S.E. Dunn. New experimental results on GPS/INS integration using ocean voyager II AUV. In *IEEE AUV*, pages 249–255, Monterey, CA, June 1996. 9

[10] T. Aoki, T. Murashima, S. Tsukioka, H. Nakajyoh, and M. Ida. Development of deep sea free swimming ROV "UROV7K". In *MTS/IEEE Oceans*, **3**, pages 1307–1311, 1999. 3

[11] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian baysian tracking. *IEEE Transactions on Signal Processing*, **50**(2), February 2002. 33

[12] T.C. Austin. The application of spread spectrum signaling techniques to underwater acoustic navigation. *IEEE AUV Technology*, pages 443–449, 1994. 7, 8

[13] N. Ayache and O. Faugeras. Maintaining a representation of the environment of a mobile robot. *IEEE Transactions on Robotics and Automation*, **5**:804–819, December 1989. 66

[14] T. Bailey. *Mobile Robot Localisation and Mapping in Extensive Outdoor Environments*. PhD thesis, The University of Sydney, 2002. 37

[15] T. Bailey and H.F. Durrant-Whyte. Simultaneous localization and mapping (SLAM): part II state of the art. *Robotics & Automation Magazine, IEEE*, **13**(3):108–117, 2006. 23

[16] B.A.A.P. Balasuriya. *Computer Vision for Autonomous Underwater Vehicle Navigation*. PhD thesis, Dept. of Naval Architecture and Oceanic Enginerring, 1998. 12

[17] B.A.A.P. Balasuriya and T. Ura. Underwater cable following by Twin-Burger 2. In *IEEE International Conference on Robotics and Automation*, **1**, pages 920–925, 2001. 2

[18] Y. Bar-Shalom and T.E. Fortman. *Tracking and Data Association*. Academic Press, 1988. 36, 42

[19] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley and Sons, 2001. 160

[20] M.R. Black and B. Butler. Arctic ocean trials of trackpoint ultrashort baseline acoustic positioning system. *IEEE AUV Technology*, pages 297–302, 1994. 2, 7, 8

[21] S.S. Blackman. Multiple hypothesis tracking for multiple target tracking. *Aerospace and Electronic Systems Magazine, IEEE*, **19**(1):5–18, 2004. 138

[22] S.S. Blackman and R.F. Popoli. *Design and Analysis of Modern Tracking Systems*. Artech House, Norwood, MA, 1999. 42

[23] R.N. Carpenter. Concurrent mapping and localization with FLS. In *Workshop on Autonomous Underwater Vehicles*, pages 133–148, Cambridge,MA, USA, 1998. 17, 67

[24] S. Challa, R.J. Evans, and D. Musicki. Target tracking - a bayesian perspective. In *Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on*, **1**, pages 437–440 vol.1, 2002. 43

[25] D. Clark and B.N. Vo. Convergence analysis of the gaussian mixture PHD filter. *IEEE Transactions on Signal Processing*, **55**(4):1204, 2007. 62, 128

[26] D.E. Clark and J. Bell. Convergence results for the particle PHD filter. *Signal Processing, IEEE Transactions on*, **54**(7):2652–2661, 2006. 62

[27] D.E. Clark and J. Bell. Multi-target state estimation and track continuity for the particle PHD filter. *IEEE Transactions on Aerospace and Electronic Systems*, **43**(3), July 2007. 106

[28] T.R. Clem, J.T. Bono, P.S. Davis, D.J. Overway, L. Vaizer, A. Torres, T. Austin, R.P. Stokey, and G. PackarD. Initial buried minehunting demonstration of the laser scalar gradiometer operating onboard REMUS 600. In *IEEE Oceans*, September 2006. 2, 3

[29] D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes, Volume 1.* Springer, 2nd edition, November 2003. 157

[30] D.E. di Massa and W.K. Stewart. Terrain-relative navigation for autonomous underwater vehicles. In *OCEANS '97. MTS/IEEE Conference Proceedings*, **1**, pages 541–546 vol.1, 1997. 10

[31] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. M.Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, **17**(3):229–241, June 2001. 105

[32] M.W.M.G. Dissanayake, S.B. Williams, H.F. Durrant-Whyte, and T. Bailey. Map management for efficient simultaneous localization and mapping (SLAM). *Autonomous Robots*, **12**(3):267–286, 2002. Available from: `http://portal.acm.org/citation.cfm?id=591919`. 38, 68, 96, 105

[33] A. Doucet, N. de Freitas, and N.J. Gordon. *Sequential Montecarlo Methods in Practice.* Springer, 2001. 32, 33

[34] H.F. Durrant-Whyte. Uncertain geometry in robotics. *Robotics and Automation, IEEE Journal of*, **4**(1):23–31, 1988. 5

[35] H.F. Durrant-Whyte and T. Bailey. Simultaneous localization and mapping: part i- the essential algorithms. *Robotics & Automation Magazine, IEEE*, **13**(2):99–110, 2006. 10, 23, 28

[36] A. Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, **22**(6):46–57, 1989. 38, 66, 132, 140

[37] O. Erdinc, P. Willett, and Y. Bar-Shalom. Probability hypothesis density filter for multitarget multisensor tracking. In *Information Fusion, 2005 8th International Conference on Information Fusion (FUSION)*, **1**, page 8 pp., 2005. 117, 120

[38] O. Erdinc, P. Willett, and Y. Bar-Shalom. A physical-space approach for the probability hypothesis density and cardinalized probability hypothesis density filters. In *Signal and Data Processing of Small Targets 2006, 18 April 2006*, **6236** of *Proc. SPIE - Int. Soc. Opt. Eng. (USA)*, pages 623619.1–623619.12, USA, 2006. SPIE - The International Society for Optical Engineering. Copyright 2008, The Institution of Engineering and Technology. Available from: http://dx.doi.org/10.1117/12.673194. 154

[39] R.M. Eustice, H. Singh, J.J. Leonard, M. Walter, and R. Ballard. Visually navigating the RMS titanic with SLAM information filters. In *Robotics: Science and Systems*, pages 57–64, Cambridge, MA, 2005. MIT Press. 12

[40] S.D. Fleischer. *Bounded-Error Vision-Based Navigation of Autonomous Underwater Vehicles*. PhD thesis, Stanford University, 2000. 11

[41] P.P. Gandhi and S.A. Kassam. Analysis of CFAR processors in homogeneous background. *IEEE Transactions on Aerospace and Electronic Systems*, **24**(4):427–445, July 1988. 74

[42] R. Garcia, J. Batlle, and X. Cufi. Positioning an underwater vehicle through image mosaicking. In *IEEE International Conference on Robotics and Automation*, 2001. 12, 18

[43] B.C. Garner. Gravitational field maps and navigational errors. *IEEE Journal of Oceanic Engineering*, **27**(3), July 2002. 10

[44] I. R. Goodman, R.P.S. Mahler, and H.T. Nguyen. *Mathematics of Data Fusion*. Kluwer Academic Publishers, 1997. Available from: http://portal.acm.org/citation.cfm?id=549931. 57, 165

[45] J.E. Guivant and E.M. Nebot. Optimization of the simultaneous localization and map-building algorithm for real-time implementation. *Robotics and Automation, IEEE Transactions on*, **17**(3):242–257, 2001. 23, 31, 105

[46] J.S. GUTMANN AND K. KONOLIGE. Incremental mapping of large cyclic environments. In *Computational Intelligence in Robotics and Automation, 1999. CIRA '99. Proceedings. 1999 IEEE International Symposium on*, pages 318–325, 1999. 65

[47] P.E. HAGEN, N. STORKERSEN, B.E. MARTHINSEN, G. STEN, AND K. VESTGARD. Military operations with HUGIN AUVs: lessons learned and the way ahead. In *Oceans*, **2**, pages 810–813, Europe, June 2005. 2

[48] D. HAHNEL, R.TRIEBEL, W.BURGARD, AND S.THRUN. Map building with mobile robots in dynamic environments. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2003. Available from: `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.4177`. 139

[49] J.R. HOFFMAN AND R.P.S. MAHLER. Multitarget miss distance via optimal assignment. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, **34**(3):327–336, 2004. 131

[50] N. IKOMA, T. UCHINO, AND H. MAEDA. Tracking of feature points in image sequence by SMC implementation of PHD filter. In *Proc. Soc. of Instrument & Contr. Engineers (SICE)*, pages 1696–1071, Hokkaido, Japan, August 2004. 106

[51] J.S. JAFFE. Computer modeling and the design of optimal underwater imaging systems. *IEEE Journal of Oceanic Engineering*, **15**(2), April 1990. 11

[52] S. JULIER AND J. UHLMANN. A new extension of the kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, Orlando, FL*, 1997. Available from: `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.5.2891`. 30

[53] B. KALYAN AND B.A.A.P. BALASURIYA. Sonar and vision based navigation schemes for autonomous underwater vehicles. In *International Conference on Control, Automation, Robotics and Vision*, Kunming, China, 2004. 74

[54] B. KALYAN, B.A.A.P. BALASURIYA, AND W.S. WIJESOMA. Multiple target tracking in underwater sonar images using Particle-PHD filter. In *OCEANS 2006 - Asia Pacific*, pages 1–5, Singapore, 2006. 106

[55] Bharath Kalyan, K.W. Lee, W.S. Wijesoma, D Moratuwage, and Nicholas M. Patrikalakis. A random finite set based detection and tracking using 3D LIDAR in dynamic environments. Istanbul, Turkey, October 2010. IEEE Systems, Man & Cybernetics. 106

[56] B. Kamgar-Parsi. Registration algorithms for geophysical maps. In *OCEANS '97. MTS/IEEE Conference Proceedings*, **2**, pages 974–980 vol.2, 1997. 10

[57] J. C. Kinsey, R. M. Eustice, and L. L. Whitcomb. Underwater vehicle navigation: recent advances and new challenges. In *IFAC Conf. on Manoeuvring and Control of Marine Craft*, 2006. 4, 5, 11

[58] K. Konolige. Improved occupancy grids for map building. *Auton. Robots*, **4**(4):351–367, 1997. 66

[59] K.W. Lee, W.S. Wijesoma, and I.G. Javier. On the observability and observability analysis of SLAM. In *IEEE/RSJ Intenational Conference on Intelligent Robots and Systems (IROS)*, pages 3569–3574, Beijing, China, 2006. 105

[60] Kwang Wee Lee, Bharath Kalyan, Sardha Wijesoma, Martin Adams, Franz S. Hover, and Nicholas M. Patrikalakis. Tracking random finite objects using 3D-LIDAR in marine environments. In *Proceedings of the 2010 ACM Symposium on Applied Computing*, pages 1282–1287, Sierre, Switzerland, 2010. ACM. Available from: http://portal.acm.org/citation.cfm?id=1774088.1774362. 106

[61] J.J. Leonard, A. Bennett, C.M. Smith, and H.J.S. Feder. Autonomous underwater vehicle navigation. Technical report, MIT Marine Robotics Laboratory Technical Memorandum, 1998. 4

[62] J.J. Leonard and H.F. Durrant-Whyte. *Directed Sonar Sensing for Mobile Robot Navigation*. Kluwer Academic Publishers, 1992. Available from: http://portal.acm.org/citation.cfm?id=573569. 67, 91, 143

[63] J.J. Leonard and H.J.S. Feder. Decoupled stochastic mapping. *IEEE Journal of Ocean Engineering*, **26**(4):561–571, October 2001. 23, 31

[64] X.R. Li. Tracking in clutter with strongest neighbor measurements. i. theoretical analysis. *Automatic Control, IEEE Transactions on*, **43**(11):1560–1578, 1998. 105

[65] L. Lin, Y. Bar-Shalom, and T. Kirubarajan. Track labeling and PHD filter for multitarget tracking. *Aerospace and Electronic Systems, IEEE Transactions on*, **42**(3):778–795, 2006. 62, 79, 116

[66] F. Lu and E. Milos. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, **4**:333–349, 1997. 65

[67] W. K. Ma, B. N. Vo, S. S. Singh, and A. Baddeley. Nonlinear signal Processing-Tracking an unknown Time-Varying number of speakers using TDOA measurements: A random finite set approach. *IEEE Transactions on Signal Processing*, **54**(9):3291–3304, 2006. 106

[68] R.P.S. Mahler. An introduction to Multisource-Multitarget statistics and its applications. Technical report, Lockheed Martin Technical Monograph, 2000. 41, 43, 57, 167, 168

[69] R.P.S. Mahler. Multi-target bayes filtering via first-order multi-target moments. *IEEE Transactions on Aerospace and Electronic Systems*, **39**(4):1152–1178, 2003. 39, 41, 58, 59, 60, 61, 81, 106, 111, 114, 115, 120, 168

[70] R.P.S. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Artech House Publishers, February 2007. 39, 61, 104, 106

[71] R.P.S. Mahler. Unified sensor management using CPHD filters. In *Information Fusion, 2007 10th International Conference on*, pages 1–7, 2007. 154, 155

[72] S. Majumder. *Sensor fusion and feature based navigation for sub sea robots*. PhD thesis, Australian centre for field robotics, Department of aerospace, mechanical and mechatronic engineering, University of Sydney, August 2002. 17, 23, 67

[73] R.L. Marks, S.M. Rock, and M.J. Lee. Real-time video mosaicking of the ocean floor. *IEEE Journal of Oceanic Engineering*, 1995. 12

[74] H. Medwin and J.E. Blue. *Sounds in the sea*. Cambridge University Press, 2005. 70

[75] P.H. Milne. *Underwater Acoustic Positioning Systems*. Gulf Publishing, 1983. 8

[76] M. Montemerlo. *FastSLAM: a factored solution to the simultaneous localization and mapping problem with unknown data association*. PhD thesis, School of Computer Science, Carnegie Melon University, June 2003. 105, 132

[77] M. Montemerlo and S. Thrun. Simultaneous localization and mapping with unknown data association using FastSLAM. In *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, **2**, pages 1985–1991 vol.2, 2003. 23, 34, 35

[78] B. Moran. *Underwater shape reconstruction in two dimensions*. Ph.D thesis, Massachusetts Institute of Technology, Dept. of Ocean Engineering, 1994. 67

[79] H. Moravec. Sensor fusion in certainty grids for mobile robots. *AI Mag.*, **9**(2):61–74, 1988. Available from: http://portal.acm.org/citation.cfm?id=46187. 66

[80] P. Moutarlier and R. Chatila. An experimental system for incremental environment modelling by an autonomous mobile robot. In *The First International Symposium on Experimental Robotics I*, pages 327–346. Springer-Verlag, 1990. Available from: http://portal.acm.org/citation.cfm?id=661279. 5, 105

[81] J. Mullane, B.N. Vo, M.D. Adams, and W.S. Wijesoma. A random set formulation for bayesian SLAM. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 1043–1049, 2008. 106

[82] K.P. Murphy. Bayesian map learning in dynamic environments. *In Neural Info. Proc. Systems (NIPS*, pages 1015–1021, 1999. Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.6616. 34

[83] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *Automatic Control, IEEE Transactions on*, **39**(6):1237–1241, 1994. 43

[84] K. Nasahashi, T. Ura, A. Asada, T. Obar, T. Sakamaki, K. Kim, and K. Okamura. Underwater volcano observation by autonomous underwater vehicle âĂIJr2D4âĂİ. In *Oceans-Europe*, **1**, pages 557–562, Europe, June 2005. 2

[85] S. Negahdaripour and X. Xu. Mosaic based positioning and improved motion estimation methods for automatic navigation of submersible vehicles. *IEEE Journal of Oceanic Engineering*, **27**(1), 2002. 12, 18

[86] J. Neira and J.D. Tardos. Data association in stochastic mapping using the joint compatibility test. *IEEE Transactions on Robotics and Automation*, **17**(6):890–897, December 2001. 37, 99, 105, 138

[87] H.T. NGUYEN. *An Introduction to Random Sets*. Chapman & Hall/CRC, 1 edition, March 2006. 157

[88] I. NYGREN AND M. JANSSON. Terrain navigation for underwater vehicles using the correlator method. *IEEE Journal of Oceanic Engineering*, **29**(3):906–915, July 2004. 10

[89] K. PANTA, B. N. VO, AND S. SINGH. Novel data association schemes for the probability hypothesis density filter. *Aerospace and Electronic Systems, IEEE Transactions on*, **43**(2):556–570, 2007. 77

[90] G.W. PULFORD. Taxonomy of multiple target tracking methods. *Radar, Sonar and Navigation, IEE Proceedings -*, **152**(5):291–304, 2005. 42

[91] D. REID. An algorithm for tracking multiple targets. *Automatic Control, IEEE Transactions on*, **24**(6):843–854, 1979. Available from: `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1102177`. 37

[92] D. RIBAS, P. RIDAO, J.D. TARDÃŞS, AND J. NEIRA. Underwater SLAM in man-made structured environments. *J. Field Robot.*, **25**(11-12):898–921, 2008. Available from: `http://portal.acm.org/citation.cfm?id=1464493.1464494&coll=GUIDE&dl=GUIDE`. 17, 67

[93] H. ROHLING. Radar CFAR thresholding in clutter and multiple target situations. *Aerospace and Electronic Systems, IEEE Transactions on*, **AES-19**(4):608–621, 1983. 74

[94] I. T. RUIZ, Y. PETILLOT, D.M. LANE, AND C. SALSON. Feature extraction and data association for AUV concurrent mapping and localisation. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, **3**, pages 2785–2790 vol.3, 2001. 16, 67

[95] S. SE, D.G. LOWE, AND J. LITTLE. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *International Journal of Robotics Research*, **21**(8):735–758, August 2002. 105

[96] M. SHOR AND L. NADAV. Performance of order statistics CFAR. *IEEE Transactions on Aerospace and Electronic Systems*, **27**(2):241–224, March 1991. 74

[97] H. SIDENBLADH. Multi-target particle filtering for the probability hypothesis density. In *Proc. Intl Conf. on Information Fusion*, pages 800–806, Cairns, Australia, 2003. 62, 79

[98] H. Sidenbladh and S.L. Wirkander. Tracking random sets of vehicles in terrain. In *Proc. 2003 IEEE Workshop on Multi-Object Tracking*, Madison WI, June 2003. 58, 106, 116

[99] H. Singh, C.N. Roman, L. Whitcomb, and D. Yoerger. Advances in fusion of high resolution underwater optical and acoustic data. In *International Symposium on Underwater Technology*, **3**, pages 206–211, May 2000. 3

[100] H. Singh, L.L. Whitcomb, D.R. Yoerger, and O. Pizarro. Microbathymetric mapping from underwater vehicles in the deep ocean. *Computer Vision and Image Understanding*, **79**(1):143–161, July 2000. 2

[101] R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. In I.J. Cox and G.T. Wilfong, editors, *Autonomous Robot Vehicles*, pages 167–193. Springer-Verlag, 1990. 23, 28

[102] R.C. Smith and P. Cheeseman. On the representation and estimation of spatial uncertainly. *Int. J. Rob. Res.*, **5**(4):56–68, 1987. Available from: `http://portal.acm.org/citation.cfm?id=33842`. 5, 23, 88, 105, 122

[103] H. Thomas and E. Petit. From autonomous underwater vehicles (AUVs) to supervised underwatervehicles (SUVs). In *MTS/IEEE OCEANS*, **2**, pages 875–887, October 1997. 9

[104] R.L. Thompson, J. Seawall, and T. Josserand. Two dimensional and three dimensional imaging results using blazed arrays. In *OCEANS, 2001. MTS/IEEE Conference and Exhibition*, **2**, pages 985–988 vol.2, 2001. xiii, 177, 178

[105] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005. 25, 31

[106] M. Tobias and A. D. Lanterman. Probability hypothesis density-based multitarget tracking with bistatic range and doppler observations. *IEE Proc. Radar, Sonar, and Navigation*, **152**(3), 2005. 106, 117, 120

[107] T. Ura and K. Kim. On-site INS update of an AUV "r2d4" by SSBL based position estimation. In *OCEANS '04. MTTS/IEEE TECHNO-OCEAN '04*, **3**, pages 1606–1611, 2004. 3

[108] R.J. Urick. *Principles of Underwater Sound*. McGraw-Hill Company, 1975. 6, 13, 70, 71, 72, 73, 178

[109] B.N. Vo AND W.K. Ma. The gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, **54**(11):4091–4104, November 2006. 62, 116, 122

[110] B.N. Vo, S. Singh, AND A. Doucet. Sequential monte carlo methods for multi-target filtering with random finite sets. *IEEE Transactions on Aerospace and Electronic Systems*, **41**(4):1224–1245, 2005. 62, 79, 115, 116

[111] C.C. Wang. *Simultaneous localization, mapping and moving object tracking*. PhD thesis, Carnegie Mellon University, 2004. Available from: http://portal.acm.org/citation.cfm?id=1023485. 139

[112] Y.D. Wang, J.K. Wu, A.A. Kassim, AND W.M. Huang. Tracking a variable number of human groups in video using probability hypothesis density. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, **3**, pages 1127–1130, 2006. 106

[113] L. L. Whitcomb, D.R. Yoerger, AND H. Singh. Advances in doppler-based navigation of underwater robotic vehicles. In *IEEE International Conference on Robotics and Automation*, pages 399–406, 1999. 4, 8

[114] W.S. Wijesoma, L.D.L. Perera, AND M.D. Adams. Toward multidimensional assignment data association in robot localization and mapping. *IEEE Transactions on Robotics*, **22**(2):350–365, April 2006. 99, 105, 138

[115] S. Williams, M.W.M.G. Dissanayake, AND H.F. Durrant-Whyte. Towards Terrain-Aided navigation for underwater robotics. *Advanced Robotics*, **15**(5):533–550, May 2001. 16, 67

[116] S. Williams AND I. Mahon. Simultaneous localisation and mapping on the great barrier reef. In *IEEE International Conference on Robotics and Automation*, New Orleans, USA, May 2004. 12

[117] S.B. Williams, M.W.M.G. Dissanayake, AND H.F. Durrant-Whyte. An efficient approach to the simultaneous localisation and mapping problem. *IN PROC. IEEE INT. CONF. ROBOTICS AND AUTOMATION*, **1**:406–411, 2002. Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.1721. 23, 31