Cooperative 3D Mapping and Localisation of Multiple Mobile Robots

Julian Ryde

A thesis submitted for the degree of Doctor of Philosophy

Department of Computer Science

University of Essex

January 2008

Abstract

Developing a coherent and unified approach to localisation and mapping is one of the prerequisites for fully establishing the mobile robotics era in the 21st century. Therefore, this research attempts to address this simultaneous localisation and mapping problem (SLAM) for both a single robot and cooperating multiple robots equipped with 3D range sensors.

To begin with various methods and hardware set-ups for allowing robots to reliably detect each other are explored and tested. Once robots have detected others the question of how their relative pose can be determined most accurately is investigated, especially in the case when two corresponding robots detect each other simultaneously. This situation, referred to as mutual localisation, results in reliable and accurate relative pose determination.

The second aspect of this work develops hardware mechanisms for mobile robots to obtain 3D information about the environment around them. For robots to rapidly acquire 3D information of their surrounding spatial structure a conventional 2D laser range scanner is augmented with a rotating mirror to build a 3D laser range scanner. A data structure, termed an occupancy list, is devised for efficient probabilistic storing of the 3D map. Algorithms for coherently combining multiple 3D scans from different view points are developed, thus allowing mobile robots to generate an internal representation of their environs.

These contributions are necessary to explore the thesis, that multiple cooperating mobile robots equipped with 3D sensors not only vastly improve but are also essential for navigation and mapping within indoor environments.

Combining these contributions produces a system that is capable of mapping indoor environments in full 3D in real-time with one or more robots.

Acknowledgements

First and foremost I would like to thank my supervisor Professor Huosheng Hu for giving me the benefit of his substantial experience and expertise. His patience and tireless dedication to supervision greatly enhanced my PhD experience.

Acknowledgement must also be given to Dr. Ulrich Nehmzow and Dr. Dongbing Gu for their valuable insights throughout this research.

I would like to thank my friends and family for their support and encouragement, especially my parents for their proof reading. Finally I would like to thank Tricia Tate for her companionship and patience.

Funding was provided by the University of Essex, EPSRC and Professor Huosheng Hu. The dissertation is dedicated to the memory of Mary Rydzewska.

Contents

List of Figures

Chapter 1

Introduction

Mobile robots must be able to sense their surroundings if they are to fulfil a purpose outside of controlled static environments. They should combine sensory readings taken over time into a coherent internal representation. Ideally, they ought to be able to do this alone if necessary as well as take advantage of other nearby robots, sharing information and working cooperatively to localise themselves and map their environment. The culmination of this research is the production of 3D representations of indoor environments by both a single robot and a team of robots.

1.1 Background

There are various definitions for robotics. Indeed Brady (1985) defined it as "the intelligent connection of perception to action". For mobile robotics one aspect of the action that must be present is the ability to move. This ability eventually introduces the robot to uncontrolled environments. Previously fixed robots operated in controlled environments such as assembly lines in automated manufacturing. Researchers achieved remarkable success in situations where the environment and robot could be accurately modelled and analytic techniques prevailed. Unfortunately, the paradigms that emerged from these successful implementations are poorly suited for the dynamic, inaccurately modelled and chaotic world that mobile robots often find themselves outside of the laboratory.

It was these limitations that lead to the ascension of probability in robotics, especially mobile robotics. Furthermore mobile robotics highlights the need for improved sensors to help deliver reliable operation in unknown environments. Improvements to robot sensors were clearly necessary starting with the limited success of ultrasound, the beguiling hopes for camera based vision, the accuracy of laser ranging and data fusion work with combinations of these.

Mobile robots are expected to operate in a variety of locales, both indoors and outdoors, as well as, structured and unstructured. These environments are usually dynamic and vary enormously. Outdoors, global localisation is made considerably easier through the use of GPS, however for indoor robots there is no such global location signal infrastructure. Indoors it is possible to use the wireless network signal strength as a global signal locator (Serrano et al., 2004), however location accuracy is around 2m and orientation determination appears all but impossible with standard wireless network cards. Another disadvantage is that the environment has to contain such a wireless infrastructure and wireless networks are far from ubiquitous.

Concentrating on indoor environments gives mobile robot application areas such as homes, offices, factories and hazardous environments including nuclear facilities and urban military reconnaissance.

1.1.1 Sensors

Sensory perception is an important aspect for robotics, especially so for mobile robotics. There has been a large volume of work dedicated to improving the interpretation of sensory data but perhaps not as much towards enhancing the sensor hardware. Many of the first mobile robots were equipped with sonar sensors. Sonar sensors use the speed of sound, usually ultrasound, to actively determine the range to obstacles. As technology improved more advanced sensors such as cameras and laser range finders were adopted. Laser range finders use the speed of light and a scanning mechanism to measure ranges at a number of angles around the robot. The high velocity of light means they require very precise timing over short intervals and are consequently more

expensive than sonar ranging systems.

The introduction of cameras as robotic sensors spurred an enormous research effort into computer vision. Computer vision is the process by which the image data from one or more cameras is analysed and interpreted to generate or extract useful information. The literature is detailed, deep and expansive and a full description is outside the scope of this discussion. A good introduction to the literature available is Dudek and Jenkin (2000) with more recent literature referenced in Lewis (2006).

1.2 Motivation and Objectives

The motivation for this research arises from a need. The need is nuclear power station decommissioning. This environment would appear to be ideal proving grounds for mobile robots because it is very difficult, time consuming and expensive for humans to work safely within these environments. Nuclear decommissioning may be broken down into two phases, the first is the surveying and mapping phase. In this phase the focus is on determining the levels of radiation and the structure of the contaminated environment. The second phase is the process of dismantling, removing and preparing the radioactive material for safe storage. Of the two phases of nuclear decommissioning this research focuses on part of the first phase, that of determining the structure of the environment. The type of environments to be surveyed are cluttered and complicated with non-horizontal floor sections, ramps and walkways. Thus conventional approaches, although useful, may be improved.

The following is the thesis of this research.

> Multiple cooperating mobile robots equipped with 3D sensors not only vastly improve but are also essential for navigation and mapping within indoor environments.

This is the driving force and motivation for all the experiments and developments throughout the research. The thesis consists of two aspects. The first is that cooperation between robots is beneficial to localisation and mapping in particular. Single robot

solutions are possible, indeed one is described in Chapter 5, however multiple robot mapping is more reliable as demonstrated in Chapter 6. The second aspect is that for robots to perform adequately it is imperative that they are equipped with sensors that return range information in 3D. This 3D range information greatly improves obstacle detection over 2D planar scanners. Mapping the environment in 3D improves the localisation and map building by providing more information thus reducing the probability of different locations providing similar sensor readings. Even 2D maps benefit from the extra information provided by 3D sensors (Wulf et al., 2004). Extension of the sensors into three dimensions allows the full pose to be calculated enabling the robots to build maps when operating on non-planar surfaces. The research is constrained to indoor mapping because the challenges faced are significantly different. Outdoor mobile robots have access to global positioning satellite information and therefore can determine their global position with high certainty.

This research is not confined in relevance to nuclear decommissioning, indeed this is merely a good example because it is almost impossible to accomplish without mobile robots. Other problems may be tackled without mobile robots but it is only a matter of time before it becomes preferable to address them with mobile robots. One such situation is urban search and rescue, the 3D nature of the work is especially relevant for this as damaged buildings are often non-planar and very cluttered. These buildings are usually structurally unstable and risking further human life is undesirable. Military application is inevitable and with an indoor focus for this research allows it to partially address the difficulty of fighting asymmetric urban warfare. Soldiers need no longer enter buildings without knowledge of the internal structure and location of potential threats. Other applications of this research might include automatic surveying of buildings for a variety of scenarios where indoor manual surveying would need to be done. In the entertainment field, virtual tours of museums could be made available and landmark buildings may be digitised and imported into computer games.

The aim of this work is to enable mobile robots to work in concert, detect other robots, perceive their surroundings and develop internal representations of the world

and their position within it all in three dimensions. Without being able to accurately sense their surroundings mobile robots have no hope of reliable operation. Once these capabilities have been established the aim is to answer the question of whether cooperating robots equipped with 3D sensors map significantly better than a single robot acting in isolation.

1.3 Research Methodology

Throughout this research several overarching principles guide the process of inquiry.

An emphasis is placed on producing experimental results with physical hardware rather than relying on results generated by experiments run in simulation. Quantitative as opposed to qualitative measurements of performance are preferred when possible. A number of measures are taken to assess map quality including repeatability, similarity of various dimensions and localisation performance.

Algorithms and analysis employ probability and sample based methods with a focus on not just returning an answer but also an estimate of error and if possible a full distribution of results. This prepares the robot well for an uncertain world. In this vein errors are discussed throughout the dissertation. Areas where errors are analysed include triangulation, inter-robot detection, mutual localisation, robot odometry, 3D scans, maps and localisation.

Wherever possible assumptions have been minimised. Many mapping algorithms rely on feature extraction for which assumptions have to be made about the environment and what features it may contain. While this may be appealing for indoor environments with many straight lines and planes relying on the presence of these limits the applicability of the research. One of the advantages of multiple robot mapping is that certain assumptions are allowed for instance features on the robots themselves are likely to be present in observations made by other robots.

Needless complexity is avoided wherever possible. Some manifestations of this are in hardware development of the 3D laser scanner and optimisation of the computer algorithms. It is believed that simplicity begets reliability and additional complexity is

only introduced if it is required and returns a significant increase in performance. For instance algorithm optimisations going from exhaustive search localisation to sample based searches, although increasing complexity, are necessary so that the localisation computation runs in under a second rather than hours.

1.4 Dissertation Contributions

The work described in this dissertation has been published in five international peer reviewed conference proceedings and one book chapter. The details of the publications are listed in Section 7.2.1. The following are the principle contributions to the field of mobile robotics.

- Methods for locating cooperating robots

  Fast extraction of circular features from range data

  Mutual localisation

- Development of 3D laser scanner

- 3D mapping

  3D environment representation mechanism dubbed occupancy lists

  Iterative multi-resolution sample based scan alignment and mapping algorithm

- Cooperative 3D mapping of indoor environments

  Comparison of cooperative mapping performance in 3D to robots mapping individually

In order to effectively cooperate, robots need to be aware of other robots. This requires reliable inter-robot detection methods. The two main methods explored involve mounting cylindrical targets and the detection of their circular signatures in the range

scans. Covering the cylinders in retroreflecting tape makes them more distinguishable as well as reducing the detection effort because few naturally occurring objects in indoor environments retroreflect.

Much research has been done localising robots with respect to landmarks and this work is employed in this research however what is unique is that other robots, rather than merely acting as passive landmarks, also observe the observing robot. Effective combination of these two simultaneous observations is the crux of mutual localisation.

It is found that standard mobile robot sensors are lacking for the purposes of this work and so a custom sensor is built by extending a 2D laser range finder so that it acts as a 3D laser range finder.

Many researchers use occupancy maps for 2D mapping however these do not scale well to 3D maps and so a new environment representation is devised that maintains a hashed list of occupied cells only.

Multiple scans from different poses have to be combined to build maps. The alignment of 2D scans has been done many times by previous researchers however the alignment of full 3D scans is still the subject of ongoing research. The main method is the iterative closest points algorithm (Besl and McKay, 1992; Nüchter et al., 2005) usually referred to as ICP registration. This work however uses an iterative multi-resolution sampling process where the scans are initially downsampled by a factor of 16. Exhaustive search is performed on the downsampled scans identifying prospective poses. Then the downsampling factor is reduced and samples taken around those prospective poses. The process is repeated each time reducing the downsampling factor until sufficient accuracy is obtained.

The contributions in each chapter are steps required to achieve and test the system described in Chapter 6 which is the successful integration of the contributions to produce a multi-robot 3D mapping system. It should be noted however that planning the movement of the robots is outside the scope of this research. Consequently, for all experiments the motion of the robots is controlled by teleoperation.

Almost all tasks that could be conceivably executed by mobile robots require the robots to both sense their immediate surroundings and have an awareness of where they are. Many research labs have robots with 2D laser range scanners and so the low cost method described for turning these into 3D range scanners would be relevant to these research groups. The contributions described in this research also address how data from 3D sensors in general can be stored and combined into global maps. These ideas are relevant to many medium sized (1m) indoor mobile robots and the capabilities they would be expected to have. Finally the research described here on multiple mobile robots is the first to consider how multiple robots with 3D sensors can generate maps with the aid of various combinations of inter-robot sightings.

These contributions represent a step towards the goal of useful indoor autonomous mobile systems.

1.5 Dissertation Outline

The overview of the dissertation structure is presented in Figure 1.1 which also indicates chapter dependencies. An exploration of the work relating to and the foundation of this research is given at the beginning of each chapter. In this manner each chapter is relatively self-contained. There are, however, dependencies between the chapters with the work divided into two parallel streams prompted by the thesis stated in Section 1.2.

One stream is the progress towards a cooperating multi-mobile robot system. This starts with addressing the problem of reliable inter-robot detection. This is solved by a combination of circular feature extraction and retroreflective marking. Once this inter-robot sensing is established, methods for including this extra information into the localisation and mapping process are refined.

The second stream is working towards enabling 3D sensory perception for mobile robots. This is solved by extending 2D laser range scanner technology to produce a 3D laser range scanner.

Finally the results of the two streams are combined to produce the cooperative 3D

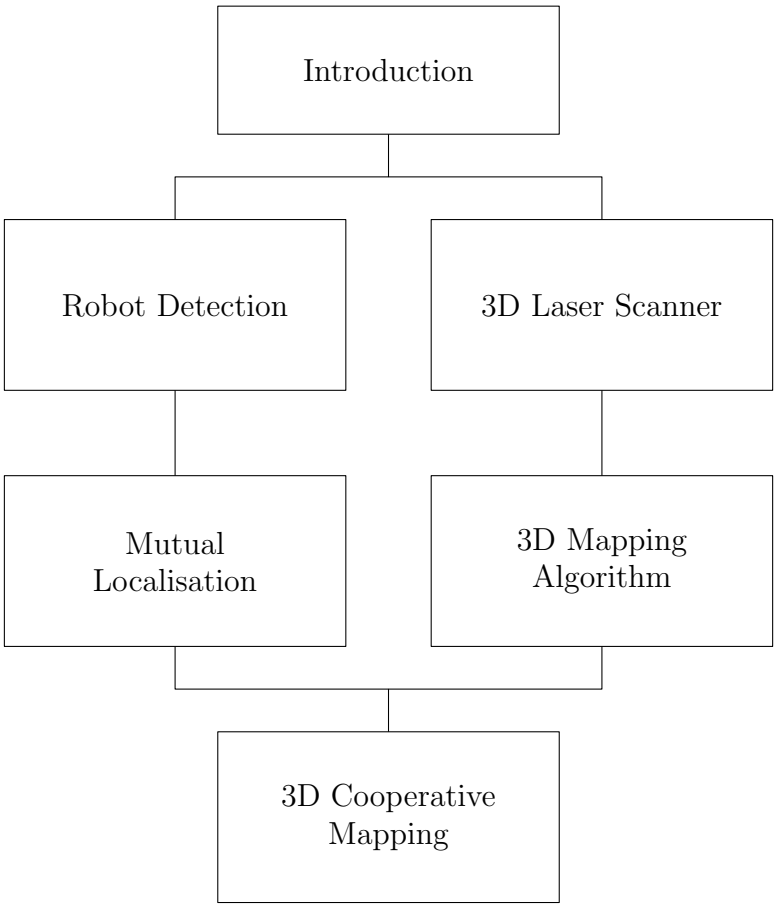

Figure 1.1: Overview of the relationships and dependencies between chapters.

mapping multi-robot solution explained in Chapter 6. A summary of the results in the context of real world deployment is discussed in Chapter 7. The development of the mobile robot 3D scanner opens up a large area of research, the implications of which are explored in Chapter 7.

In Chapter 2 the problem of successfully detecting other robots is addressed, as it is an important pre-requisite for cooperative localisation and mapping. This is tackled with feature extraction algorithms to detect the presence of circular targets in laser range data and is based on 2D data. The conventional Hough transform is compared to novel least squares based fitting algorithm. The latter is found to be superior in both computational speed, accuracy and reliability.

Chapter 3 introduces the idea of mutual localisation which is made possible by a robotic hardware set-up that enables robots to be detected and detect others simultaneously. It is then shown how this mutual localisation gives a better relative pose constraint between two robots than standard cooperative localisation employed in much of the related literature.

Chapter 4 explains the novel 3D laser scanner which enabled the collection of 3D data. Two scanner designs are described, the standard and the enhanced scanner. The standard scanner is easier to deploy and so is deployed on multiple robots. The enhanced scanner delivers more accurate data but is harder to physically deploy and is therefore tested on a single robotic platform. The enhanced scanner consists of an augmented 2D scanner facing upwards with a rotating mirror driven by a stepper motor above it.

Chapter 5 presents occupancy lists as a structure for expressing probabilistic maps in 3D as an alternative to occupancy grids that scale poorly to three dimensions. Rather than storing the probability of occupancy of every voxel in the environment in a three dimensional array, only those voxels with high probability of occupancy are stored in a hashed list structure. To produce maps multiple occupancy lists have to be aligned and combined and these processes are also addressed in this chapter. An algorithm similar to multi-resolution particle filters which is specifically tailored and

optimised for occupancy lists turns multiple scans into one coherent map. The probability of localisation error is derived for the scan alignment and localisation process. The experimental results testing this whole process are presented and show maps with a resolution of 0.03m over different indoor environments.

Chapter 6 provides a theoretical framework tying together the work from previous chapters. This theoretical framework includes pose constraints brought about by various robot detections. Mutual localisation, occupancy lists and the multi-resolution particle filter are combined to yield a full 3D mapping solution that can be deployed on many robots. The chapter concludes with experimental evidence supporting the thesis by showing the multi-robot mapping is comparable in accuracy but more reliable. A team of four robots give rise to a threefold reduction in mean position error compared to 3D mapping by a single robot.

Chapter 2

Robot Detection

It is greatly beneficial for cooperating robots to reliably and accurately detect one another. For localisation and mapping purposes when these detections occur they must be able to infer their relative poses so that they may amalgamate their mapping information. This chapter explores various methods for inter-robot detection and tests them. A novel fast circular detection algorithm tuned for range data is developed and its performance compared to that of a conventional Range Weighted Hough Transform. Cooperative localisation achieves both continuous localisation and map building. Robust continuous localisation is achieved at robot speeds of 0.2m/s with successful detection in 98% of laser scans.

2.1 Introduction

The ability of a mobile robot to locate itself within the environment is one of the most fundamental problems currently thwarting mobile autonomous operation. Most current localisation methods require a prior map or attempt to build one. Therefore, equally important is the ability of a mobile robot or robots to build accurate environment representations (maps). However, building a map without knowledge of robot pose (position and orientation) is difficult. The problem of locating a robot given a prior map has been solved in many fashions, however there is not yet a universally adopted method. Building the map in the first place is still an active research area, especially

as it is a prerequisite to most localisation algorithms.

Some tasks are best tackled by a team of robots and accurate mapping is such a task. For teams of robots to take advantage of the multi-robot approach they must be able to locate themselves relative to one another in a fast, accurate and reliable manner.

This chapter lays some of the experimental and theoretical ground work for enabling robots to work as a team and detect one another. Related work in the fields of feature extraction and robot detection is explored in Section 2.2.

The accuracy of the two sensors used for map building, the odometry sensor and the laser scanner, is assessed. The experimental platform is described in Section 2.3 which also contains the method for transforming the laser scans according to the robot pose to the global map.

Two methods for extracting circular features of other robots from laser scan data are compared in Section 2.4, the Hough Transform and a custom least squares fitting approach. It is found that the least squares fitting approach is more accurate, more reliable and faster. The hardware set-up used to enable the mobile robots to detect one another is explained and tested in Section 2.5.

2.2 Related Work

Researchers have investigated many methods that allow robots to detect each other. One such way involves robots observing targets, with rotational symmetry of order one, mounted on each robot. This allows determination of the relative pose of the observing robot. With this approach pose errors increase rapidly with range as it is very sensitive to the ability of the observer to calculate the orientation of the observed target. This requires high resolution data of the target which diminishes linearly with increasing separation. The use of multiple landmarks (Ryde and Hu, 2005a) increases the accuracy of the pose estimate by effectively deploying a large target consisting of two separated landmarks.

Cooperative localisation sometimes referred to as co-location is assumed by many

for instance (Betke and Gurvits, 1997), but actual implementations are not discussed as often. Initially the problem of co-location might be perceived as straightforward but robust, accurate and fast co-location is not readily available.

If robots may mutually detect each other as explained in Chapter 3 then only a team of two is required. For this chapter the detection of other cooperating robots is done using geometric targets and a laser range finder, for later work the geometric targets are enhanced with retroreflective tape. Laser range finders are expensive and so there is motivation to only equip one robot with such a sensor to reduce the cost of the robot team. If only one robot may reliably detect the positions of others then a team of three robots is required, one superior robot and two robots acting as mobile landmarks. The first step in the process of cooperative localisation is the ability for robots to detect each other. A number of methods are considered for this. The main contenders are an infra-red beacon approach, omni-directional vision or using distinct shapes (geometric beacons) and a laser scanner. The infra-red beacon approach suffers from two main drawbacks, significant hardware deployment and the return of target angles only. The former is surmountable however the latter is a significant limitation. Localisation given separation angles of indistinguishable targets is possible (Åström, 1993), however the problem is highly non-linear and difficult to implement. Successful implementations involving robots with omni-directional vision are possible however whilst the bearing accuracy is good they suffer from large range errors. In these the apparent size of the coloured targets mounted on other robots indicates their distance from the observing robot.

A laser scanner and geometric beacons provides both the range and angle to targets vastly simplifying the process of relative localisation, the only difficulty remaining is reliable geometric target extraction from the laser range data. The angle information alone is readily detected with an infrared beacon and sensor. Unfortunately employing this results in rather unwieldy geometry. The error estimates in position vary enormously with the geometric configuration of the landmarks. It is important to appreciate that triangulating when the orientation of the robot is not known against

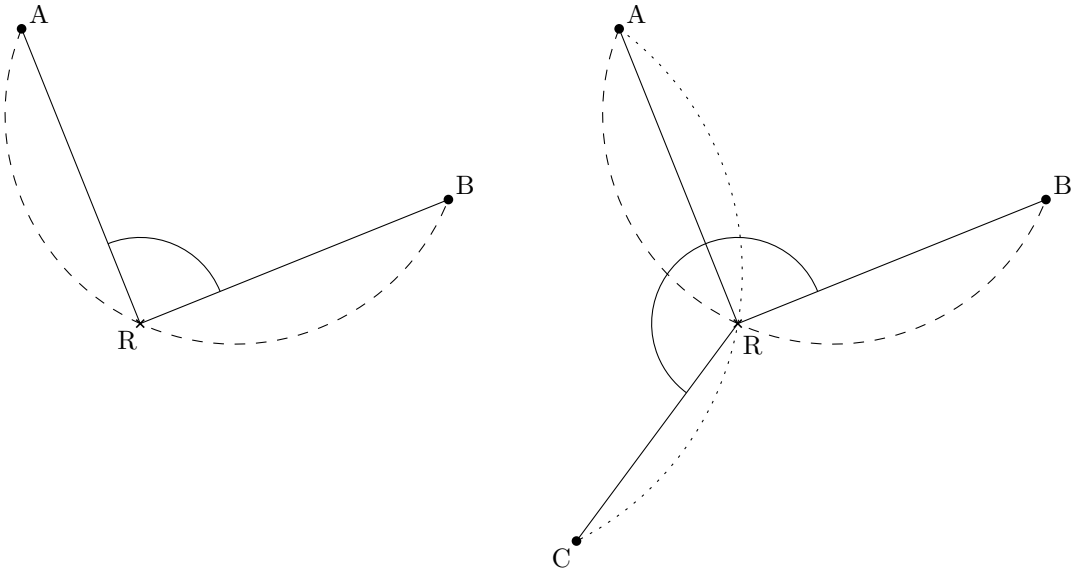

Figure 2.1: Locus of points satisfying constant angles between landmarks (A, B and C) and the position of the robot (R). The dashed line indicates a particular locus of points for which the angle ∠ARB is constant. The dotted line is the locus for a constant ∠ARC.

some external coordinate frame is more difficult than when magnetic compasses are used.

A robot with a global orientation sensor requires only two landmarks with their associated compass bearings and known positions in order to locate itself. A magnetic compass is a global orientation sensor that works well outdoors but tends to be unreliable indoors. Without such a sensor two landmarks constrain the position of the robot to a circle passing through them, Figure 2.1. This arises from a chord theorem in geometry which states that angles along the circumference of a circle subtended by the same arc or chord are equal. One more landmark is required to produce another circle and the intersection of these is the position of the robot, as illustrated in 2.1. A thorough derivation using the complex plane is undertaken in Betke and Gurvits (1997) and a good summary of beacon localisation systems is covered by Borenstein et al. (1996). It is important to note that error in the robot position increases rapidly as it approaches the circle passing through all three landmarks. All positions on this circle are undetermined. Alternatively, if both range and bearing information are available from each landmark then the position distributions arising from many landmarks may be combined easily because the position estimates are independent.

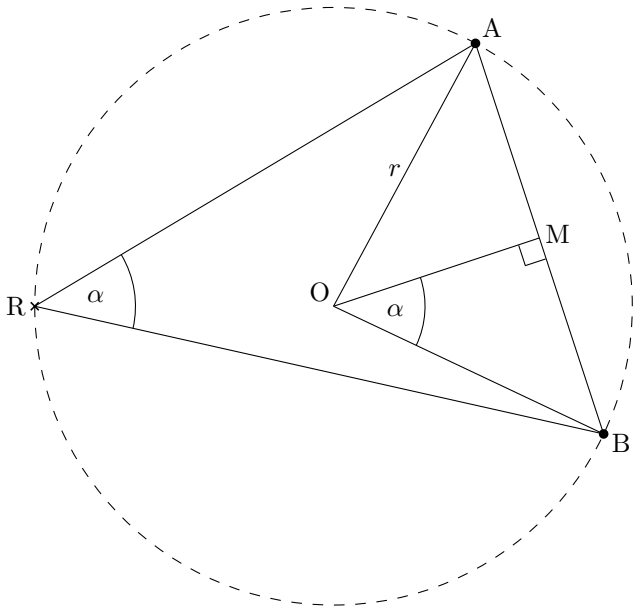

Figure 2.2: Relationship between the radius (r) of the landmark coincident circle and the angular separation (α) of the landmarks A and B.

The relationship between the landmark separation angle and the radius of the circle passing through the landmarks is obtained using the chord theorem which states that the angle subtended by a chord at the centre of a circle is twice the angle subtended at the circumference. In Figure 2.2 the angle at the circle centre ($\angle$AOB) is bisected by the perpendicular bisector of the line AB. From the right angled triangle ($\triangle$MOB) produced,

$$r = \frac{AB}{2\sin\alpha}. \tag{2.1}$$

With (2.1) the circles may be represented analytically and the robot position computed from their intersection. This intersection is unique as the remaining intersection is one of the landmarks. It can be seen from this short treatise that it would be preferable to extract both the range and angular separation of the landmarks in order to yield a robust landmark localisation system. This may be accomplished with the angle from the infrared beacons as a search starting point for finding the robot from the laser range finder data and for uniquely identifying it. A thorough theoretical analysis of the performance and ultimate limits of cooperative localisation is expounded in Mourikis and Roumeliotis (2006).

Multiple robots are able to localise and map by placing cylindrical beacons on

top of the mobile beacon robots and extracting the position of this circular feature from the laser range data. Some success in this feature extraction has been achieved using a circular Hough transform (Hough, 1962) algorithm to find and track cylinders as the robot moves. The Hough transform is a useful tool for extracting geometric primitives from image data. Jensfelt and Christensen (2001) use a version called the Range Weighted Hough Transform (RWHT) to reliably detect walls (straight lines) in cluttered environments. The RWHT takes into account that smaller range readings are more accurate and so weights the contribution of these range readings accordingly. Ahman et al. (1993) establish the RWHT which is subsequently used by Schiele and Crowley (1994).

2.3 Experimental Method

Any approach to map building requires the ability to accurately estimate position. This may be achieved to some degree through the effective use of dead reckoning also known as odometry. Most localisation methods require the error characteristics of both the odometry and exteroceptive sensors to be a priori knowledge.

Kelly (2004) undertakes a rigorous analysis of error propagation in odometry. In terms of calibration trajectories Kelly (2004) makes the important point that closed or symmetric trajectories may not achieve satisfactory calibration parameters. An important question for calibrating the odometry is to ascertain the path dependence of the errors.

Quite possibly the most straightforward calibration is that required to ensure the robot travels in a straight line. This rarely happens in practice without calibration. Calibration is both a time consuming and repetitive process which makes it suitable for automation. The robot's odometry sensor (wheel encoders) is calibrated when it can detect enough static landmarks and therefore track its pose accurately. During this process the error profile of the odometry sensor is recorded. When the robot can no longer track its position with external sensors it can then maintain a more reliable estimate of its pose from the calibrated odometry sensor.

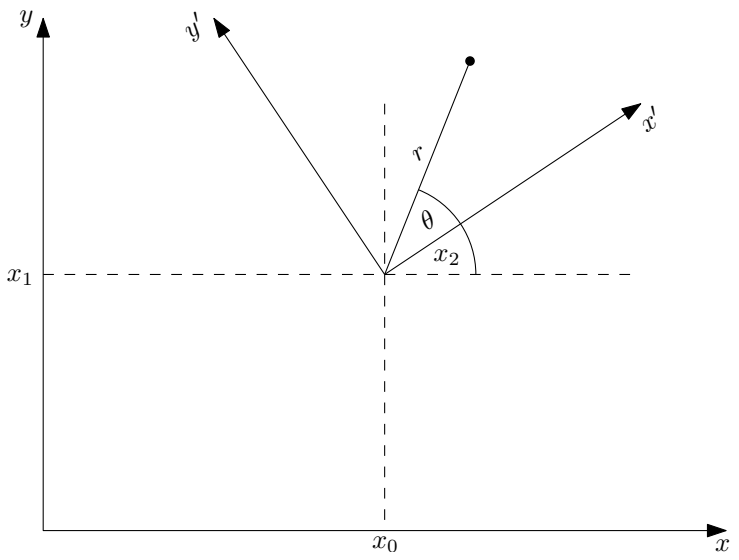

Figure 2.3: The geometry of the robot and coordinate system used to describe pose. The coordinate system in the robot frame is x', y' and a laser return is (r, θ).

The experimental platform is a Pioneer 2 robot, which is holonomic and equipped with a SICK LMS 200 laser range finder. The range finder has a scanning angle width of 180° and a resolution of 0.5°. The laser scanner is an almost ideal sensor with unrivalled accuracy, acquisition time and range. The main disadvantages with this sensor are cost (£4000), mass (4.5kg) and power consumption (17.5W). The performance and error characteristics of this LMS are detailed in Aboshosha and Zell (2003).

The Pioneer 2 robot is made by the iRobot company for research purposes. It is controlled by an on-board computer running Red Hat Linux. Communication with the robot is accomplished via wireless LAN. Its sensor payload includes 16 sonar sensors, 16 bump switches and internal odometry wheel encoders.

The robot records its change in pose from the wheel encoders and converts these to a global pose consisting of the three independent degrees of freedom x_0, x_1 and x_2, the heading. These can be represented by the vector 2.2. The corresponding coordinate system is illustrated in Figure 2.3. The position coordinates of the robot are x_0 and x_1 measured in metres and x_2 is the heading of the robot measured in radians from the

x-axis.

$$\vec{x} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \tag{2.2}$$

Determining the error in odometry can be separated into the error in each of the coordinates x_0, x_1 and heading, x_2. These coordinates are independent as a result of the holonomicity of the robot. The heading odometry is tested by rotating the robot physically and recording the odometry reading. The translational component of the pose is analysed in a similar way. The robot is physically propelled along a known straight path on the floor and the odometry recorded.

The range data returned from the sonar and laser range sensors is converted into the relative Cartesian coordinate system by the coordinate transformation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} \tag{2.3}$$

where r is the range and θ, the heading, is the angle from straight ahead. This is the standard two dimensional polar to Cartesian coordinate transformation.

The following transforms the local coordinates of a landmark that is at position (r, θ), relative to the robot with pose (x_0, x_1, x_2), to the corresponding position in the global frame.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \begin{bmatrix} r\cos(\theta + x_2) \\ r\sin(\theta + x_2) \end{bmatrix} \tag{2.4}$$

Because the laser range data is naturally polar it is more efficient to perform the rotation about the local origin in polar coordinates and then translate the result. This consideration lies behind the structure of (2.4). If the local coordinates are Cartesian then the standard Cartesian two-dimensional rotation matrix is applied,

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{2.5}$$

When the position vector is pre-multiplied by $R(\theta)$ the resulting position is rotated by θ radians about the origin. To complete the coordinate transform the position of the robot is then added as in (2.4). These coordinate transforms enable the scans taken at multiple poses to be combined into a single map. The poses may be acquired from odometry after calibration however odometry is not accurate enough to develop high quality maps. One solution to this problem is to track the robot's pose with respect to easily identifiable landmarks in the environment. These landmarks may be previously deployed in the environment, which is usually considered less than desirable, as every environment the robot operates in will have to be prepared in such a manner. Alternatively, other robots may have landmarks mounted on them and in this manner provide landmarks that can be deployed where needed. The laser scanner is the sensor of choice for mapping and so the reminder of the chapter discusses the geometric feature extraction methods necessary to identify and locate such geometric landmarks.

2.4 Hough Transform versus Least Squares Fitting

The SICK LMS 200 although not specifically designed for mobile robotics is an outstanding sensor capable of 0.25° angular resolution and millimetre distance resolution at ranges of one to eight metres. Ye and Borenstein (2002) contains a detailed discussion of its technical merits. The reliability under varying ambient illumination conditions, good accuracy and reliability are among the many reasons for selecting this sensor as the focus for research.

Geometric primitives (shapes) extraction from noisy data is a large area of research within the computer vision community. The two main approaches considered are the Hough Transform and least squares fitting. These approaches are then optimised for laser range data and their performance tested in the following sections.

2.4.1 Hough Transform

The Hough transform (Hough, 1962) has been hugely successful in the vision community thanks to its tolerance of noise and shape edge discontinuity. The most popular use for the Hough transform is for the detection of straight lines in images. Each point in the image votes for all the possible lines that go through that point. Straight lines are expressed by the parameters indicated in Figure 2.4. These parameters are chosen rather than the usual $y = mx + c$ to avoid problems inherent in the gradient of vertical lines. This vote is recorded in an accumulator array which is a digitised parameter space (Hough space). For example straight lines have position and gradient so the Hough space could be position versus gradient. Figure 2.6 shows an example accumulator grid used to extract straight lines from the typical laser scan in Figure 2.5. Points correspond to sinusoidal curves in the parameter space of perpendicular distance versus angle of perpendicular H_θ. The surface plot in Figure 2.6 shows two dominant peaks around $H_\theta = 0$ and perpendicular distances of one and two metres. These peaks correspond to the two lines of points running up the laser scan in Figure 2.5. These peaks are very prominent illustrating the efficacy of extracting straight lines from laser scanner data using the Hough Transform.

When all the points in the image have contributed their votes the most likely parameters appear as peaks in Hough space and correspond to the most prevalent straight lines. The Hough transform may be generalised to any geometric primitive however the introduction of each new parameter adds another dimension to the Hough space. With each new parameter the geometric increase in storage and processing required for the accumulator grids have clear repercussions on performance.

The standard Hough Transform is particularly ineffective for circle extraction from laser range data because of the uneven distribution of points in Cartesian space. The laser scanner samples at regular intervals of θ which results in an increasing density of points on nearer objects. A nearby straight edge obstacle is detected in preference to the actual circles as demonstrated in Figure 2.7 due to the higher sample density along its edge.

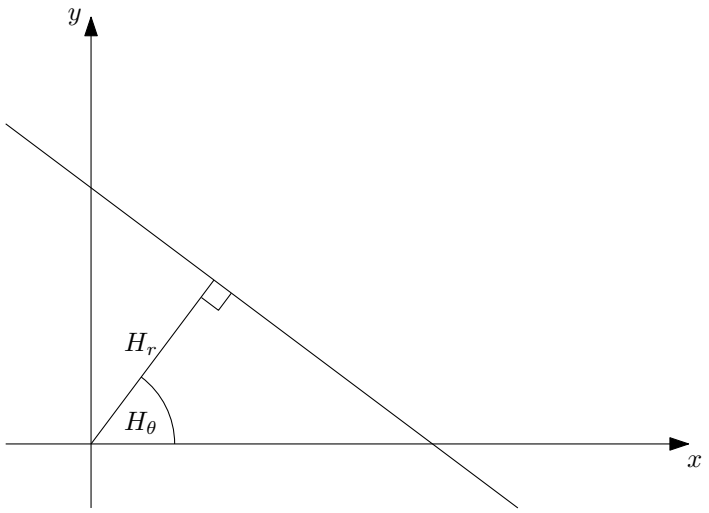

Figure 2.4: The parameters used to describe a straight line, the perpendicular distance (H_r in metres) and the angle of the perpendicular (H_θ in radians)

.

This is rectified by a Range Weighted Hough Transform(RWHT) as discussed in Forsberg et al. (1995). The weight function applied is a simple linear increase from the origin of the scan. This linear increase negates the effect of the $1/r$ fall in point density. The improvement is immediately apparent in Figure 2.8 where the peaks of the circle centres may now be distinguished from nearby walls.

Only a two-dimensional Hough parameter space is required for the circle search because the radius of the circles is known so the two parameters are the coordinates of the centre of the candidate circle. Comparing the accumulator grids for the line and circle searches (Figures 2.6 and 2.8) it is clear that the Hough Transform is much better at extracting straight lines from laser range scans. The peak in the straight line accumulator array are much more dominant than in the circle array where the transform potentially confuses short sections of straight lines with circles. This confusion of straight lines with circles is a serious problem that refuses to be resolved despite careful tuning of the algorithm. A possible solution is the initial removal of straight lines and then perform the circular Hough transform on the remaining data points. This results in a much larger execution time and reduces the generality of the algorithm to environments with flat structures. In order to avoid confusion as to which points contributed to which peaks when the first straight line has been found all points sufficiently near

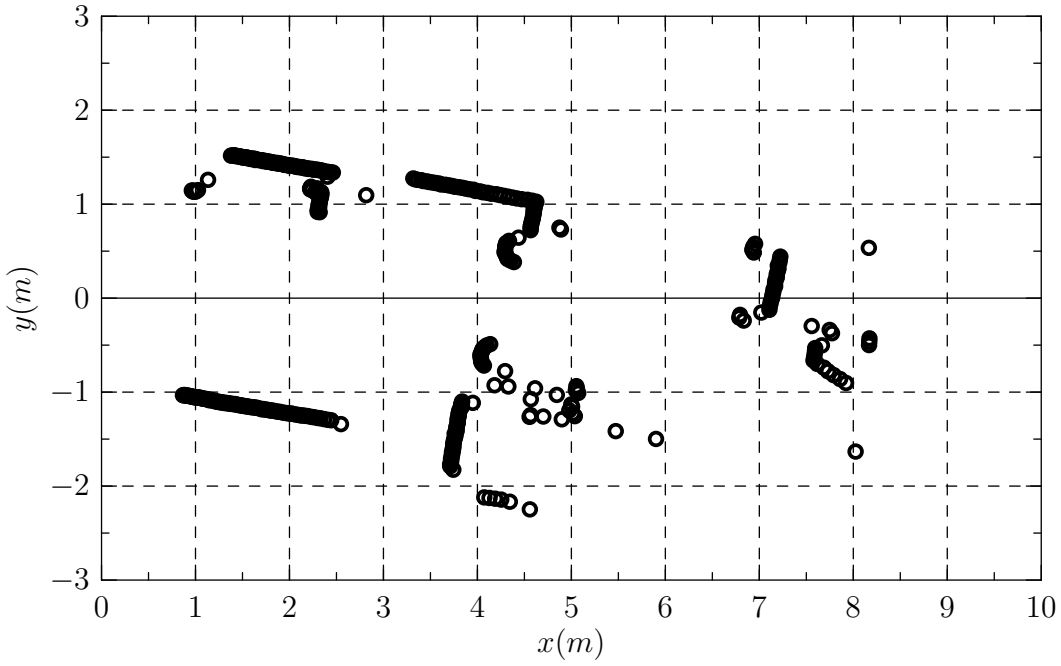

Figure 2.5: Typical laser scan at high resolution (0.25°) and 100° scan angle

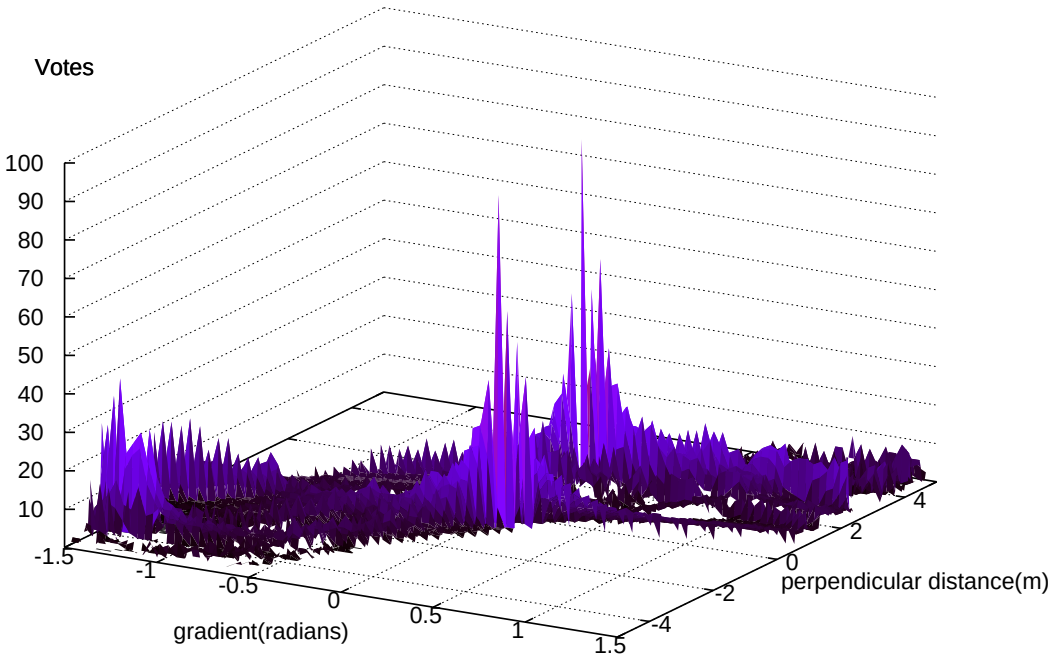

Figure 2.6: Plot of the Hough transform accumulator array resulting from searching for straight lines within the data in Figure 2.5.

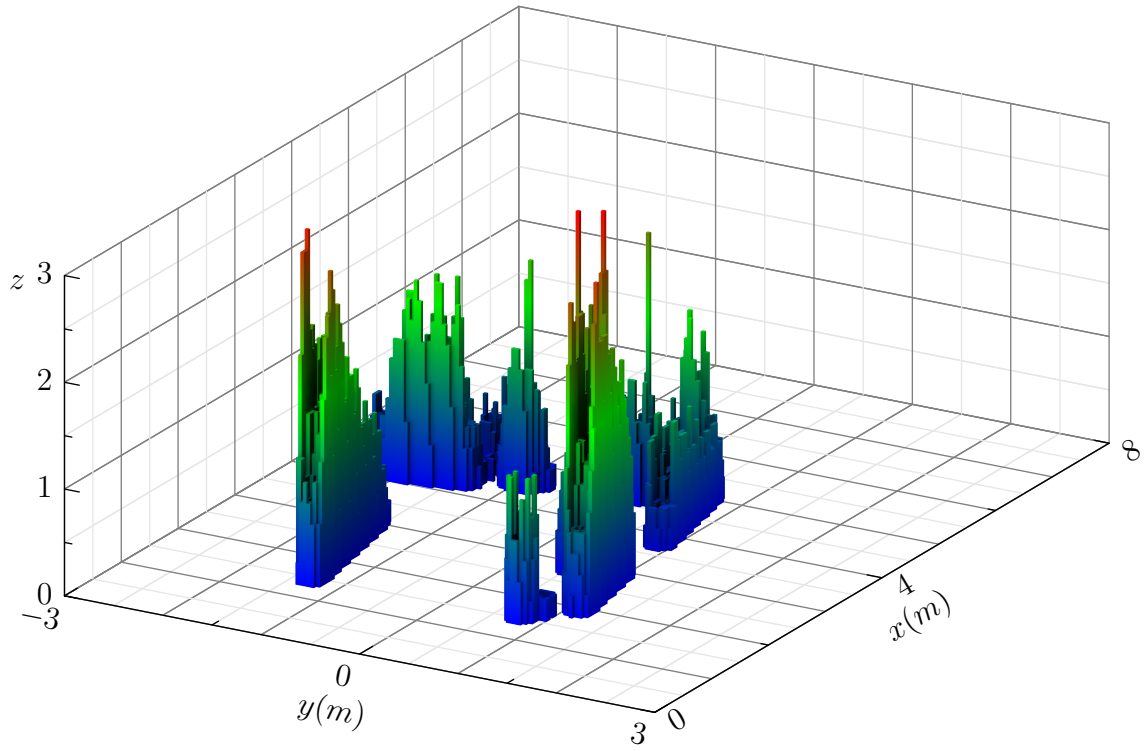

Figure 2.7: Plot of the Hough transform accumulator array resulting from searching for circles within the data in Figure 2.5.

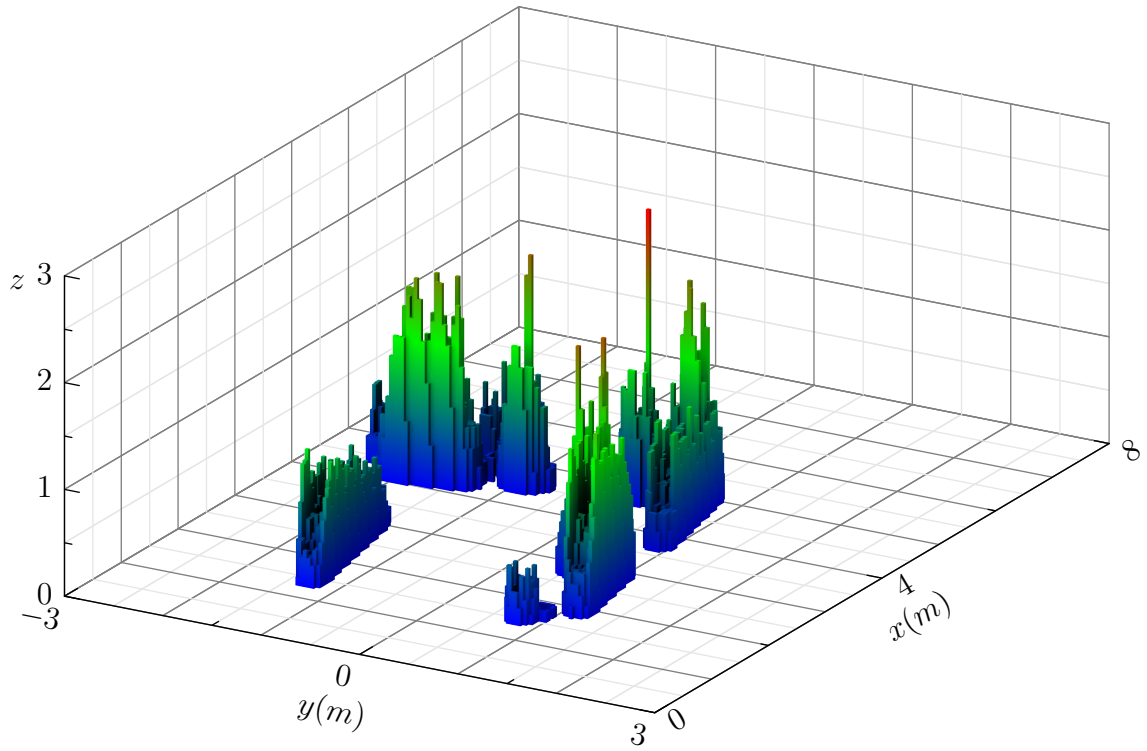

Figure 2.8: Plot of the accumulator array resulting from searching for circles using a $1/r$ Range Weighted Hough Transform within the data in Figure 2.5.

that line in the laser scan have to be removed and the Hough transform repeated. This is done until all significant edges have been removed then the search for circles is executed.

There are a number of reasons why the circular Hough transform is not particularly suited to this application. The Hough Transform always returns an answer even if the geometric primitive is not present in the data and determination of the peak significance, by comparison with others, and the kind of data expected, requires a relatively complicated statistical analysis. In robotics as in any other science returning an answer to a question is not good enough; the confidence of the answer must also be available. Ideally the entire probability distribution function for the answer should be returned.

2.4.2 Least Squares Circle Fitting

Poor performance of the Hough transform has triggered research into least squares curve fitting methods. It is evident from Chernov and Lesort (2003) that fitting circles to points is a non-trivial process because the resulting equations are highly non-linear and circles cannot be elegantly expressed in the Cartesian coordinate system. No least squares algorithm suitable for range data is available and so one is devised.

One of the problems with the circular Hough transform is that there is much information that is specific to laser range scans that is not included in the search for circles. One important property of circles is that they are symmetric when viewed from any angle. Also within the range data there is an inherent sequence of the data that is not clear in Cartesian coordinates. If a circle is to be detected a sequence of points from the range scan must lie close to the circumference of that circle. Relaxing the requirement for the detection of occluded targets allows the algorithm illustrated in Figure 2.9.

The algorithm operates by assuming a circular target is on the end of each scan angle and calculates the mean of the least squares difference. The scan angle with the lowest of these is the most likely contender for the bearing to the circle centre. Figure 2.9 illustrates the geometry involved with laser scan points depicted by crosses. Point

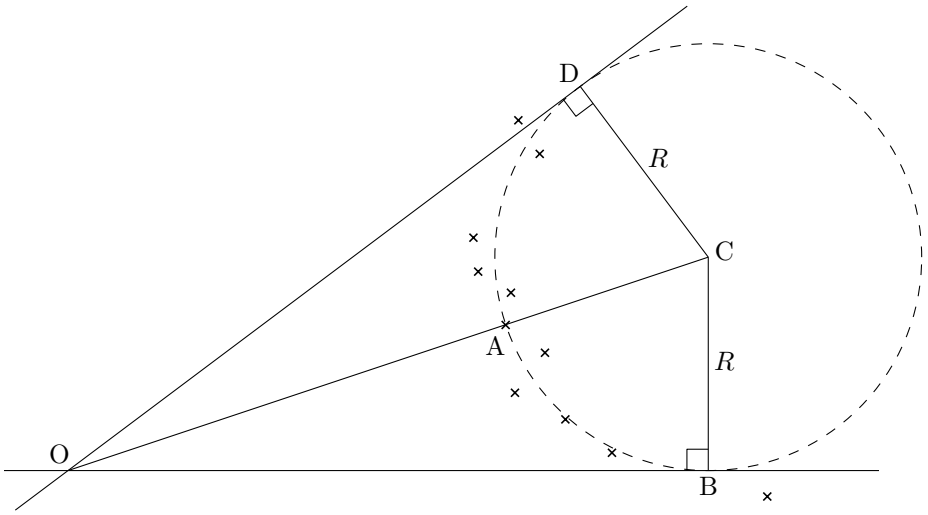

Figure 2.9: Geometric construction illustrating least squares method for circle location. Crosses indicate laser returns and the dashed circle is the current circle hypothesis that is being tested.

A is the current scan point being evaluated and the circle represents the search target. The candidate circle for A is assumed to be positioned with centre C as shown on the line OA where O is the origin of the laser scan. Assuming the laser scan returns points evenly distributed over θ then the number of nearest neighbours to be assessed is calculated. Thus points that lie within $\angle$AOB from A are candidate points where

$$\angle AOB = \sin^{-1} \frac{R}{R + |OA|}. \tag{2.6}$$

Care has to be taken regarding scan points lying near D and B which are subject to glancing edge effects, for instance reflection of the laser from the surface of the target onto another obstruction and back. Knowing the position of the hypothesis circle and the points to be tested against it, the mean least squares difference may be found. This gives an indication of how far on average the points are from the circumference of the hypothesis circle consequently giving a likelihood for the hypothesis circle. This is repeated for each point in the scan and the points that exceed a threshold probability imply successful circle detection at that position. Figure 2.10 plots the reciprocal root mean least square differences for the example laser scan in Figure 2.5.

What is apparent from an inspection of Figure 2.10 is the accurate detection and localisation of the two circular targets with the smaller of the two circle peaks being

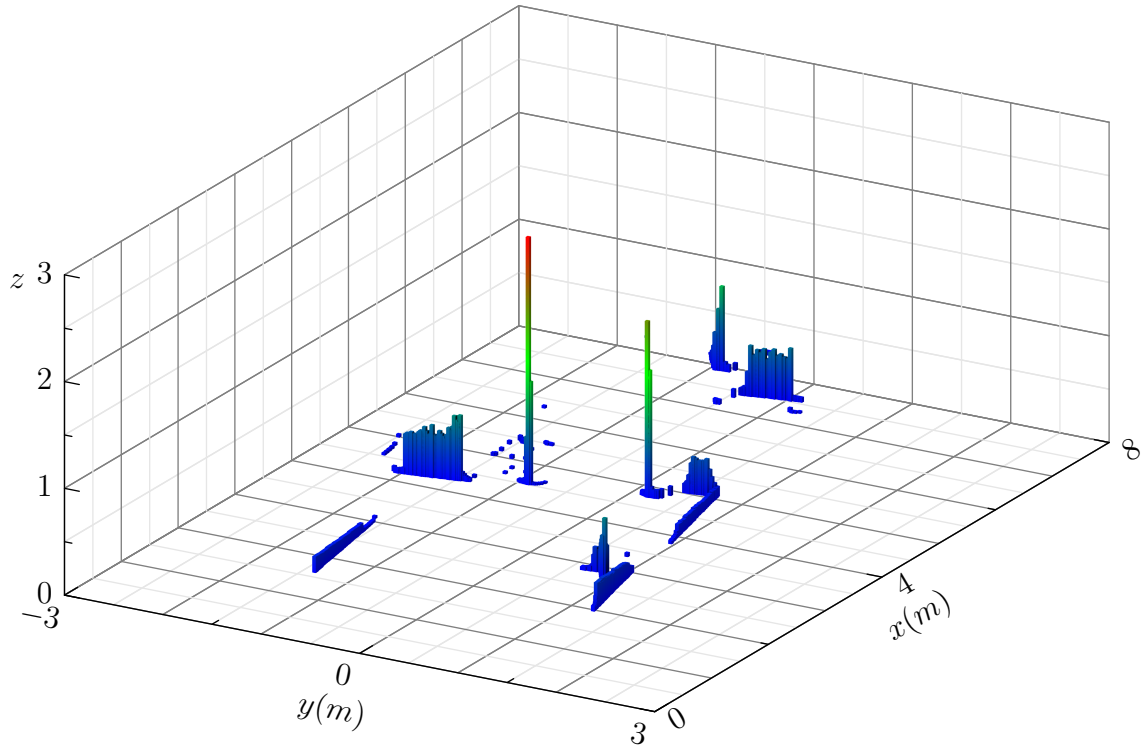

Figure 2.10: Reciprocal root mean least squares difference of the laser scan in Figure 2.5.

nearly twice as big as the largest background peak. This ensures a superior performance with circle detection reliability of 98% versus 50% for the range weighted Hough transform. Comparison of Figure 2.10 and Figure 2.8 emphasizes the effectiveness of the least squares algorithm over the Hough transform for reliable circular target extraction from laser range data. Because the least squares algorithm takes advantage of range data specific characteristics like data sequence and a single observation point, which the more generic Hough transform cannot incorporate, the least squares is not only 25 times more accurate but much faster to compute and requires less memory. The accumulator array for the Hough transform is CPU cycle and memory intensive, especially with increasing parameter space resolution.

2.4.3 Detecting Circles in Range Data

Fast circle detection from laser range data is discussed in the previous section and in Ryde and Hu (2005a). However, this detection process may be dramatically improved by adding retrorelective tape to targets. This approach determines the position of

the centre of the circle without fully utilising all the circumference data points. The algorithm essentially supplies the points that are on the circumference of the target circle and belongs to a class of detection algorithms. Given the relevant points, finding the centre of that circle is the job of fitting algorithms. It is perhaps a surprisingly complex feat to achieve due to the non-linearity and poor expression of circles in Cartesian coordinates. Because it is easier for humans to visualise in such a coordinate system it is tempting to convert range readings from the laser scanner into Cartesian coordinates. For fitting circular data this turns out to be a mistake.

Leaving the range readings in polar coordinates has a number of advantages. The points have a definite sequence and the function passing through the points is guaranteed to be single valued for all θ; the same cannot be said for points on a circle in Cartesian coordinates. In sequence means that continuous convex surfaces, such as cylinders, in the environment will produce adjacent points in polar coordinates. The equation of a circle at position (a, b) with radius R in Cartesian coordinates is

$$(x-a)^2 + (y-b)^2 = R^2. \tag{2.7}$$

Using standard Cartesian to polar substitutions and application of some trigonometric identities gives the following relationship between r and θ for the circumference of a circle with centre (c, α).

$$r = \frac{R^2 - c^2}{r} + 2c\cos(\alpha - \theta) \tag{2.8}$$

Solving this for r using the quadratic formula gives

$$r = c\cos(\theta - \alpha) \pm \sqrt{R^2 + c^2\cos^2(\theta - \alpha) - c^2}. \tag{2.9}$$

Now as the laser scanner detects surfaces nearer to the origin only the solution involving the negative square root need be considered. The Taylor expansion about α illustrates the suitability of using a quadratic fit in polar coordinates to find the centre of the

circle.

$$r = c - R - \frac{(cR - c^2)(\theta - \alpha)^2}{2R} + \frac{(cR^3 - 4c^2R^2 + 3c^4)(\theta - \alpha)^4}{24R^3} + \cdots \tag{2.10}$$

A good description of polynomial fitting is given in Weisstein (2005). In general it is possible to algebraically minimise the vertical offsets for the general kth degree polynomial

$$y = a_0 + a_1 x + \ldots + a_k x^k. \tag{2.11}$$

Fitting in the Cartesian coordinate system produces inaccurate results because the vertical offsets are an approximation to the true perpendicular offset. This approximation is better when the gradient is small, an assumption that is valid for circles in laser range data but most definitely invalid when the range data is converted to Cartesian coordinates. In Cartesian coordinates the circles could have large gradients depending on their location within the field of view of the laser scanner. For a second degree polynomial, $k = 2$, solving the matrix equation

$$\mathbf{r} = \mathbf{X}\mathbf{a} \tag{2.12}$$

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = \begin{bmatrix} 1 & \theta_1 & \theta_1^2 \\ 1 & \theta_2 & \theta_2^2 \\ \vdots & \vdots & \vdots \\ 1 & \theta_n & \theta_n^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \tag{2.13}$$

by inverting $\mathbf{X}$ allows the calculation of the polynomial coefficients as

$$\mathbf{a} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{r} \tag{2.14}$$

Once the quadratic coefficients have been found, determination of the minimum is necessary to find the angle along which the centre of the circle lies. Thus the polar

coordinates of the circle centre are

$$(c, \alpha) = \left(a_0 - \frac{a_1^2}{4a_2} + R, \frac{-a_1}{2a_2} \right). \tag{2.15}$$

The range of the circle centre is the range of the quadratic minimum plus the radius of the circle.

This only works if three or more points on the circumference of the circle are detected. For the cases of one and two points geometric approaches must be undertaken. In the following experiments if the number of retroreflective returns in a scan is less than three then the result is discarded. Ideally these results should be included because fewer sample points occur at large separations where better orientation accuracy is possible.

2.5 Cooperative Localisation and Mapping

Given that a robot can observe other stationary robots how may it determine changes in its pose? The constituents of pose, position and orientation, imply that changes in pose may be described as linear combinations of two geometric transforms, translation and rotation. An important consideration is whether the observed robots are distinguishable; if they are then the determination of unique pose change between landmark observations is possible. The rotation is calculated from the change in angle of the lines joining the landmarks and the translation is the average displacement of each landmark to its image landmark. If the landmarks are indistinguishable then alignment is ambiguous because each one cannot be associated with absolute certainty to the same landmark in the subsequent sensor updates.

Given the relative positional information of indistinguishable landmarks, three are sufficient to unambiguously determine position. Initially two would appear to be sufficient however the ambiguity of identity means that landmarks may be rotated 180° about the point midway between them.

Even though only three asymmetrically distributed indistinguishable landmarks are

needed for unambiguous pose determination the fewer landmarks required the better. Is it possible to have reliable pose updates using only the relative positions of two landmarks? There are a number of ways that this may be achieved; the simplest is to use distinguishable landmarks, for instance circles of sufficiently different radii. If indistinguishable landmarks have to be used then they may be placed in such a way that localisation is only required in one half plane; for instance if they are against a wall then the robot cannot be localised in the half plane behind the wall and still be able to detect the landmarks. Finally odometry, the limited speed of the robot and fast updates means that the 180° change required for ambiguity would never happen between updates without being detected by the odometry sensors.

From two observations of relative positions of two landmarks the change in pose is determined as follows. The change in angle of the line joining the landmarks is the rotation. The translation may be found by applying the rotation to align the two pairs of points and then calculating the translation of the average position of each pair. For observations of three or more indistinguishable landmarks the situation is a little more complicated when establishing the rotation involved because there are multiple lines joining the landmarks to consider. What is important to appreciate is that even supposedly indistinguishable landmarks observed from two positions may be differentiated by virtue of their separations; this simplifies the SLAM problem (Wang et al., 2007).

Essentially if the landmarks are arranged in such as way as the distances between them all are substantially different (larger than the error in the laser range finder) then these separations become easily identifiable unique vectors. This condition is more likely to be satisfied for fewer landmarks. This is not a particularly limiting requirement; in fact any lines of symmetry in the distribution of the landmarks create ambiguity in the pose of the robot. Once the lines joining landmarks are identified then their average change in orientation allows calculation of the robot rotation. Aligning the landmarks enables the translation to be found in a manner similar to the case of two landmarks. The displacement of the mean position or centroid of the landmarks

at each observation is equal to the negative change in position of the robot. Maps are then built with the pose changes and associated scan data.

Algorithm 1 summarises the processes involved in generating maps using cooperative localisation data. First the laser data is acquired from the SICK LMS which has a variety of angular resolution and range modes. The two main modes used are 180° at 0.5° resolution and 100° at 0.25°. Better angular resolution enables more accurate cooperative localisation over a larger area surrounding the reference targets. The increased data associated with higher resolutions results in reduced data acquisition rate and increased processing time. The most appropriate laser mode ultimately depends on the operational environment, small rooms (less than eight metres) need wide angle perception and larger rooms require finer angular resolution.

Algorithm 1 Map building

```
loop
    Acquire laser data 180° at 0.5° or 100° at 0.25°
    Detect circles with radius 0.125m
    if RMS < 0.01 then
        Calculate pose change
        Align scan with global map
        Append scan to global map
    end if
end loop
```

Once the laser range data has been acquired the circular geometric targets are located and extracted from the range data. The least squares fitting algorithm, locates circles of prior known radius from range data. If the best candidate has a root mean least squares exceeding 0.01m then the scan is rejected and the next one processed. Possible causes of rejection include situations when the targets cannot be detected because they are outside the angular or distance range.

Once the positions of the stationary targets have been extracted the change in pose is used to align and add the scan minus the targets to the global map. It is important to remove the targets once they have been detected as they are not part of the environment and so should not be added to the global map.

Rejection of scans may occur if the robot is moving too fast such that there is

enough distortion in the range scan to make reliable detection of the targets difficult. The upper limits on the speed and rotation of the robot may be calculated as follows.

It takes time to scan the target during which the robot will have moved if it is scanning in motion. First the time taken for the scan to span the target must be calculated. A target with radius Rm and range rm subtends an angle

$$\theta = 2\sin^{-1}\frac{R}{r+R} \tag{2.16}$$

For a scanning frequency f the time taken for the scan to span the target is

$$t = \frac{2}{\pi f}\theta \tag{2.17}$$

During the time taken to scan, a robot moving at velocity v will have moved a distance of vt. This introduces an aberration to the scan meaning that points on one side of the target will be displaced by this amount. If this velocity aberration exceeds the threshold used for geometric target detection then localisation will fail. This threshold should be set at around the accuracy of the laser scanner currently 0.01m. This gives the following upper bound on the velocity

$$v < 0.01\frac{\pi f}{4\sin^{-1}\frac{R}{r+R}}. \tag{2.18}$$

Now for typical values as observed from experiments f=10Hz, R=0.1225m, r=4m mean the velocity has to be less than 2.6m/s. Through experiments it is found that reliable target detection breaks down at velocities far lower than this. This is because the rotational speed of the robot gives the target an apparent velocity that increases with distance to the target. Thus the angular velocity of the robot ω is limited by

$$\omega < 0.01\frac{\pi f}{4r\sin^{-1}\frac{R}{r+R}}. \tag{2.19}$$

For the typical values above this insists the angular velocity remain below 0.7rad/s.

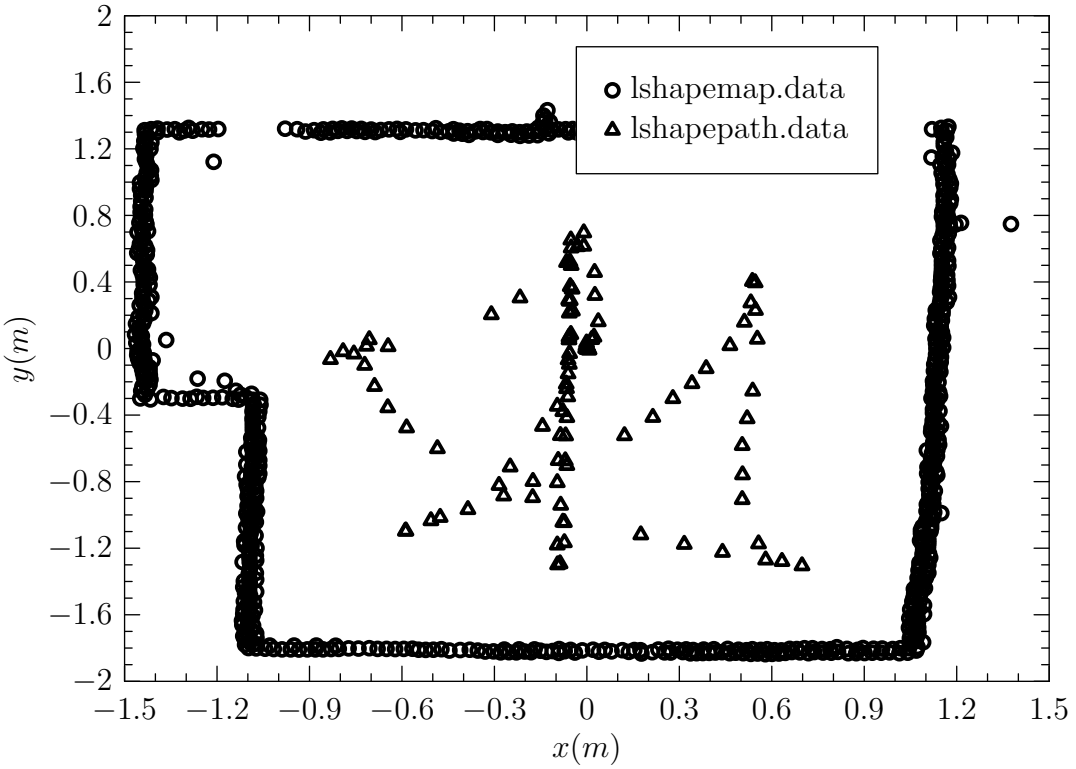

Figure 2.11: Global map along with the robot trajectory generated by localising with circular beacon detection. The two circular beacons are located at (-1.1, 1.3) and (0.9, 1.3).

This means that for accurate localisation the motion of the robot is limited much more by the angular velocity constraint than the linear velocity. Even when the robot is driving in a straight line it tends to twitch as it advances and these twitches exceed the angular velocity constraint at the relatively low linear velocities of 0.2m/s.

To conclude it is currently possible to perform localisation at speeds of around 0.2m/s however in order to build the most accurate maps it is preferable to move a short distance to a new pose then stop and scan. In this way continuous localisation can be provided however these scans will not be incorporated into the global map, only the ones taken when stationary are used.

2.5.1 Results

An example of the maps generated is shown in Figure 2.11. The map is the result of driving a robot around a pen of known dimensions in which two circular targets are located. When both targets are within the robots field of view it can localise itself

and then add its current scan to the global map. This gives an qualitative idea of the mapping accuracy and a more quantitative idea may be obtained by comparing dimensions extracted from this map to the actual measurements obtained from the ground truth. This analysis shows a deviation of around 4% between dimensions from the global map obtained and those from the ground truth.

Figure 2.11 highlights an interesting artefact in the region of the concave corner of the room. When this corner is observed from positions towards the bottom of the map pixel mixing produces spurious range returns erroneously indicating obstacles in free space. The mixed pixel problem is due to the laser spot size sampling from different surfaces, Ye and Borenstein (2002). These kind of sensor errors should be removed by a good global representation such as occupancy grids.

Experiments are performed to test the localisation accuracy delivered. These involved driving the robot along a straight line. The deviation of the positions from this straight line gives an indication of the localisation error in the direction perpendicular to the line. This error depends approximately linearly on the angular resolution of the laser scanner, the range and separation of the geometric targets. The localisation error is of the order of 0.02m at ranges of 0 to 8m at a scan resolution of 0.25°.

It is instructive to carry out a quick analysis of how the localisation error varies with the range, separation, size and position of the targets.

Error in the range to the targets introduces an error in the position estimation of the robot. Figure 2.12 illustrates the dependence of the error on range error. In Figure 2.12 the origin O is the true position of the robot and O' is its worst case perceived position if the range to the target A is over estimated and that to target B is underestimated. The error estimate is greatly simplified if a far field approximation is used which means

$$\overline{AB} \ll \overline{OM} \tag{2.20}$$

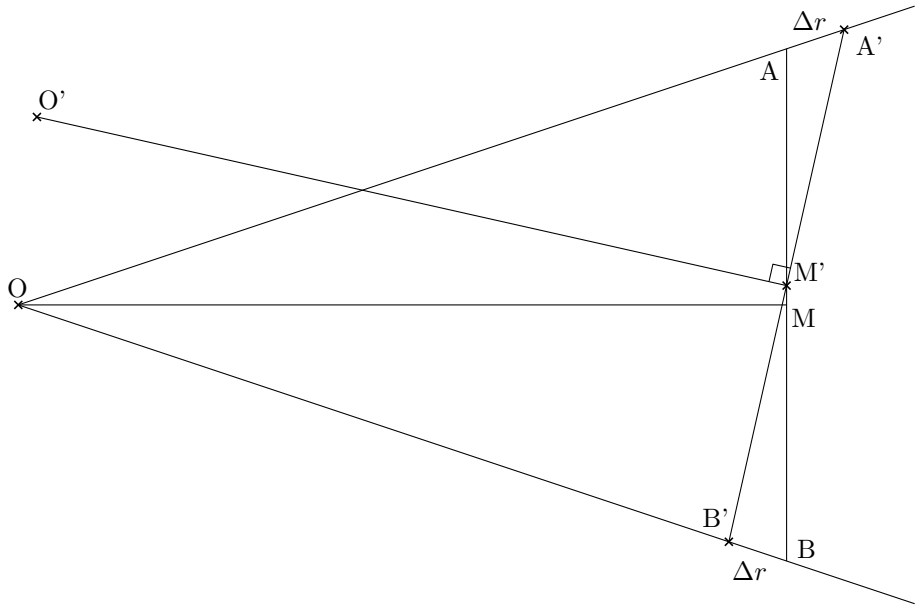

Figure 2.12: Geometric construction used to calculate localisation error

in this approximation the following similarities prove useful

$$\sin\theta \approx \tan\theta \approx \theta \tag{2.21}$$

for θ in radians. As

$$\overline{OM} = \overline{O'M} \tag{2.22}$$

then for the displacement of O' the angle of rotation is

$$\angle OMO' = \arctan \frac{2}{\sqrt{2}} \frac{\overline{AA'}}{\overline{AB}} \approx \sqrt{2} \frac{\overline{AA'}}{\overline{AB}}. \tag{2.23}$$

Note that the root two factor is due to the addition of the errors in quadrature. Finally the position error

$$\overline{OO'} \approx \sqrt{2} \frac{\overline{AA'}}{\overline{AB}} \overline{OM} \tag{2.24}$$

The far field approximation, expressed in (2.20), breaks down for large target separations or when the robot is close to the targets however this is immaterial because in these situations the error is at its least. It should now be clear from Figure 2.12 that the dependence of position error (σ_x) on angular error (σ_θ) for the laser scanner

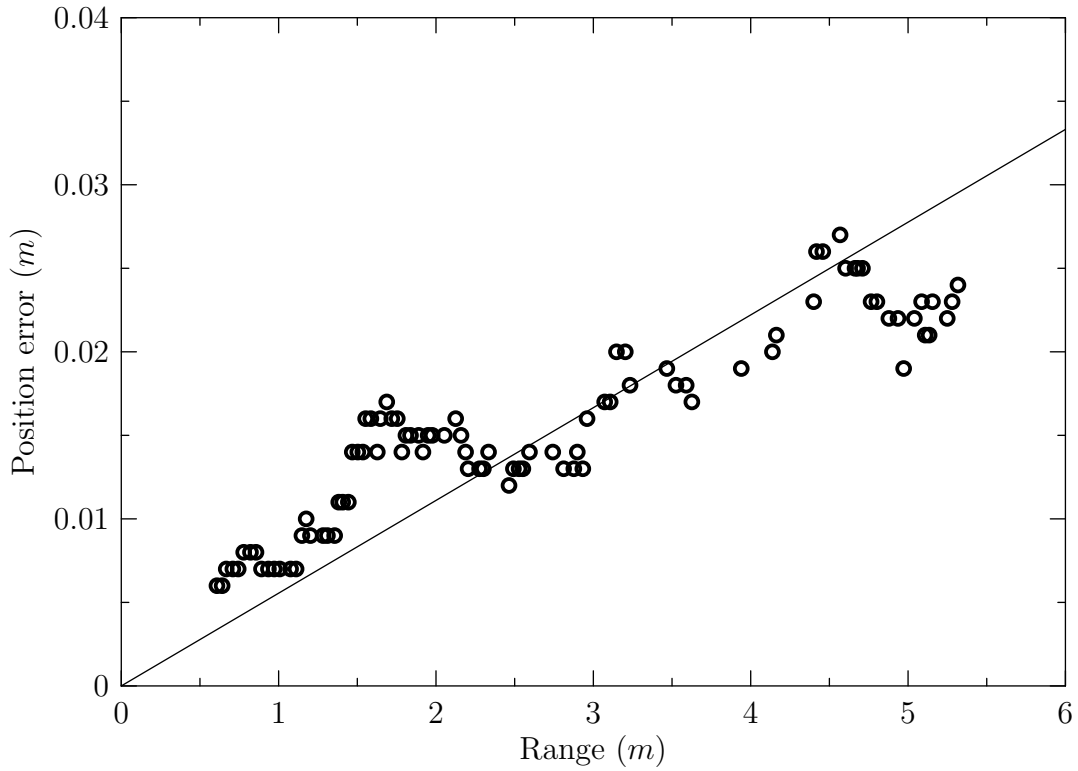

Figure 2.13: The increase of position error with range to beacons.

is approximated by

$$\sigma_x \approx \overline{OM}\sigma_\theta. \tag{2.25}$$

The angular error for the SICK LMS 200 is circa 0.009 radians so at a range of 4m the position error due to angular error is around 0.03m. Targets separated by 2m with radii 0.1225m at a range of 4m observed with a range error of 0.01m produced a position error of 0.03m. This prediction agrees with the error observed at this range in Figure 2.13.

2.6 Summary

The focus of this chapter is enabling multi-robot cooperative localisation and mapping through reliable robot detection. The chapter starts with an assessment of the work of others on inter-robot detection, more general landmark detection and the benefits of cooperative localisation and mapping. The robotic experimental platform is presented along with some of the mathematical tools necessary for localisation with respect to

landmarks mounted on cooperating robots.

A few cooperative localisation schemes are proposed and assessed. These include infrared or geometric reference targets and various extraction algorithms involving Hough transforms and least squares fitting.

The effectiveness of extracting circular landmarks from laser range data by two methods is compared in detail. These methods are the traditional Range Weighted Hough Transform (RWHT) and an optimised least squares method introduced by this research. The performance of the least squares method is decidedly superior to that of the RWHT in terms of accuracy, reliability and execution speed.

The work on circular feature detection is extended, improved and applied to multi-robot localisation and mapping. The best set-up, as judged by a combination of cost, deployment ease, reliability and accuracy, is mounting circular targets on robots and locating the circular features in the laser range data. Circular feature detection is performed by the novel least squares fitting approach which is optimised for the sequential nature of the range data and the highly symmetric aspect of the circular geometric targets.

Under this scheme both continuous localisation at 0.2m/s and map building are achieved. A global map geometrically accurate to 4% is produced. The localisation error is both theoretically analysed and experimentally measured. It is found that position error increases linearly with range to landmarks with a position error of 0.02m at a distance of 5m. In the next chapter this cooperative localisation error dependence on range is eliminated. The robust localisation scheme lays the foundation for mapping featureless or highly symmetric environments when multiple robots are available.

Chapter 3

Design of Mutual Localisation Algorithm

In this chapter the conventional premise of cooperative localisation is extended by the idea of simultaneous mutual localisation. In mutual localisation rather than robots observing other robots as conventional landmarks, pairs of robots observe one another simultaneously. The additional information through simultaneous mutual observation allows robot teams to determine their relative poses accurately and robustly. This novel localisation scheme coupled with the custom circular target fitting approach, described in Chapter 2, delivers good relative pose determination with average position and orientation errors of 0.01m and 1.4°. The experimental results show that the proposed scheme can be used to produce maps with a geometric error of about 2%.

3.1 Introduction

This chapter demonstrates mutual localisation which is more effective than standard cooperative localisation. The pivotal idea for mutual localisation is that rather than one robot acting as an observer and the other as a beacon, both robots act as observers and beacons to one another. Thus two or more robots operating as a team can use these simultaneous mutual observation events to precisely determine their relative poses regardless of range. This has a number of immediate significant advantages. The

system requires only two robots, is more resistant to environmental distractions and has increased relative pose accuracy. The pose error no longer increases with separation as is the case with standard cooperative localisation (as demonstrated in Section 3.4.2). Once reliable inter-robot localisation is established then these extra constraints when applied to the mapping process improve it significantly in terms of reliability as demonstrated in Chapter 6.

How may a robot best determine the relative pose of its fellow team members? Relative position can be determined given a mechanism for extracting sensor readings of other robots from background data. One possibility could involve each robot carrying a cylinder which may be detected and extracted from the laser range data using the Hough transform or the circle least squares fitting algorithm presented in Chapter 2. Finding the relative orientation using a single, stationary cylinder is impossible due to the circular symmetry of the target. To determine orientation, should the robot under observation be moved in some fashion or should a different identification geometry be used? What geometric shapes would allow fast algorithms to find them? The ability to distinguish robots would vastly simplify the process of determining relative locations allowing unimodal probability distributions to be used to model the layout of the robots.

Section 3.2 discusses the work of others in representing and updating robot pose. This section also covers methods other researchers have adopted to allow robots to measure their pose relative to other robots and landmarks.

Section 3.3 addresses the various physical configurations possible for enabling robots equipped with laser range finders to detect one another quickly. Also discussed are the mechanism for handling mutual localisation when the origin of the beacon is not coincident with that of the laser scanner. The results for both simulation and real experiments are presented in Section 3.4. Section 3.5 performs a theoretical analysis of the error dependence of mutual localisation upon the various experimental quantities particularly robot separation.

3.2 Related Work

Localisation can be subdivided into local and global. Local localisation methods involve tracking the progress of the robot from a known position and updating its current position along the trajectory. Global localisation can estimate the robot pose with no prior information. The former technique although tending to be fast to compute is ultimately doomed to fail after long periods of operation. Current methods involve using local techniques for fast updates to position, incorporating both proprioceptive and exteroceptive input. These pose estimates are augmented with less frequent global localisations when the pose error exceeds a threshold.

Robot pose has been tracked in increasingly complex ways. Initially Leonard and Durrant-Whyte (1991) represented pose as a Gaussian distribution, subsequently multiple Gaussian distributions are used by Jensfelt and Kristensen (2001) and finally arbitrary probability density functions, sampled from using a variety of numerical methods such as adaptive particle filters, (Fox, 2003). Position probability grids were first espoused by Burgard et al. (1996) and used successfully by Fox (1998). These approaches have, at their heart, a Bayesian update process. Bayesian filters address the problem of how best to update the probability representation of the underlying state of a system given noisy data reflecting that state.

Mobile robot localisation can be treated as a state estimation problem. The underlying state is the robot's pose and potentially the map. In the case of localisation with a prior map the observation model dictates the data expected from the sensors for a given the pose in the map, $p(z_t|x_t)$. The state-space model derived from the underlying physics predicts the future state of the system given its current state and control inputs, concisely expressed as $p(x_t|x_{t-1}, u_{t-1})$. For localisation this corresponds to where the robot is expected to be, x_t, given its current pose, x_{t-1}, and the drive commands that have been issued to it, u_{t-1}. Finally, the noise model, w_t, is used to quantify the error between the true state of the system and that which is observed.

One of the most basic forms of Bayes' Theorem states that the probability of the hypothesis given the data is equal to the product of the probability of the data given

the hypothesis and the probability of the hypothesis divided by the probability of the data. The latter is usually assigned the role of a normalising constant. This is concisely expressed as

$$p(\mathrm{H}|\mathrm{D}) = \frac{p(\mathrm{D}|\mathrm{H})p(\mathrm{H})}{p(\mathrm{D})}, \tag{3.1}$$

where H and D are the hypothesis and data respectively. Comparing (3.1) and (3.2) helps identify the meaning of the terms posterior, prior, likelihood and evidence.

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}} \tag{3.2}$$

For dynamic systems most literature expresses the posterior at time t as the belief $Bel(x_t)$ specifically

$$\mathrm{Bel}(x_t) = p(x_t|z_t, u_{t-1}, z_{t-1}, u_{t-2} \ldots, u_0, z_0) \tag{3.3}$$

where x is the state of the dynamic system under inspection. One of the most important simplifying assumptions is the Markov assumption; it is assumed that the data readings, z_t, and control commands, u_t, are independent of historical readings and commands. Asserting this assumption allows real time updates to maps but may fail when a robot closes an open loop, namely returning to a previously visited position. When a robot discovers a place in the environment it has previously explored, it has new information regarding its displacement between poses. This information may be used to help correct previous cumulative odometry and scan alignment errors that accrued whilst traversing the loop.

The Markov assumption requires the posterior to be updated in two ways. A predictive update is carried out when a new control input u_{t-1} is applied. The prediction of the new system state is determined thus

$$\mathrm{Bel}^{-}(x_t) \longleftarrow \int p(x_t|x_{t-1}, u_{t-1})\mathrm{Bel}(x_{t-1})dx_{t-1}. \tag{3.4}$$

The second belief update is a correction upon receiving new sensor data and is applied

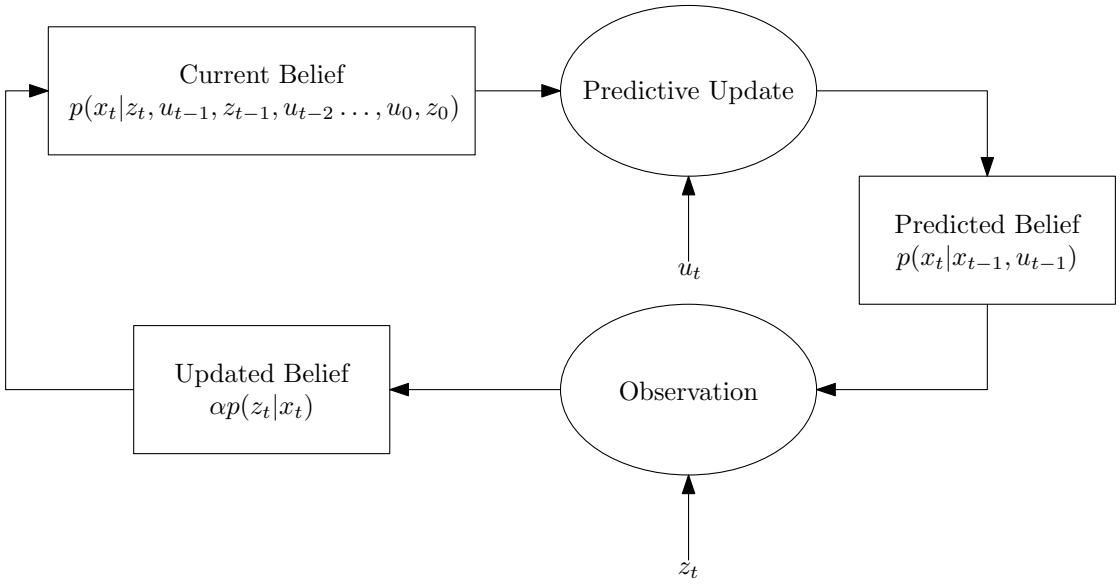

Figure 3.1: Flowchart illustrating the belief update cycle.

as follows

$$\text{Bel}(x_t) \longleftarrow \alpha p(z_t|x_t)\text{Bel}^{-}(x_t) \tag{3.5}$$

where α is the normalising constant ensuring that the sum of beliefs over all possible system states is one. The term $p(x_t|x_{t-1}, u_{t-1})$ describes how the system evolves under control, namely how the new state depends upon the previous state only and the previous control commands (states and controls prior to $t-1$) need not be considered as per the Markov assumption. The $p(z_t|x_t)$ describes the probability of receiving observations given a state often referred to as the perceptual or sensor model. This attempts to answer the question, if the robot were in a particular position, what would the output of its sensors be? This update process is summarised in Figure 3.1.

In order to use Bayesian update methods the sensor model, $p(z_t|x_t)$, and the robotic platform movement or odometry model, $p(x_t|x_{t-1}, u_{t-1})$, need to be established and effectively represented. The probability density function for the $Bel(x_t)$ should be adequately expressed and a brief discussion of the various mechanisms for describing this distribution ensues.

Kalman filters, introduced by Kalman (1960), model the belief as a Guassian distribution described by two parameters only, the mean and standard deviation. The Kalman filter places strict constraints through the assumptions of system linearity

and normally distributed errors. Non-linearity may be handled by approximating the system with a sum of linear terms via the Taylor series expansion. Despite these constraints the Kalman filter has enjoyed much success, (Leonard and Durrant-Whyte, 1991; Lu and Milios, 1997a; Arras and Vestli, 1998; Weigel et al., 1999). Some of the problems arising from the Kalman filter are highlighted by Fox et al. (1998b). However the Kalman filter is still popular even today, Cox and Leonard (1994); Durrant-Whyte et al. (2001); Leonard and Feder (1999); Neira et al. (1999); Deans (2002) who use Kalman filters for their ease of implementation and light computational burden.

Some of the main deficiencies of Kalman filters include the inability to localise without an initial accurate estimate and having no way of dealing with indistinguishable features or landmarks. Indistinguishable landmarks result in multi-modal pose probability distributions and hence are poorly handled by the uni-modal Kalman filter.

The weaknesses of the Kalman filter prompted the development of the multiple hypothesis approach which expresses belief state by a number of Gaussian distributions. This allows the robot to navigate using indistinguishable features and in highly symmetric environments. The problem of global localisation is handled by multiple hypothesis tracking; however it is difficult to represent an essentially uniform initial distribution with many Gaussian distributions. The assumptions and their problems underlying the Kalman filter are also present in this method. Successful implementations have been achieved by Austin and Jensfelt (2000); Roumeliotis and Bekey (2000); Jensfelt and Kristensen (2001); Castellanos et al. (2002).

A stumbling point with many applications of multi-hypothesis tracking and indeed particle filters, is when the belief consolidates with high certainty into one hypothesis it becomes difficult for the robot to entertain alternative hypotheses once again. Thus they become brittle in the face of the 'kidnapped robot problem' and may not recover from large localisation failures. This is addressed somewhat by ad hoc methods establishing when hypotheses should be abandoned and how new ones should be spawned, Cox (1993).

As is often the case in mobile robotics, hybrid approaches are adopted to capi-

talise on the different advantages of various methods without suffering the setbacks inherent to any particular one. The pose and map state space may be expressed using a combination of parametric functions or discrete sampled values. Rao-Blackwellised particle filters use a mixture of particle and Kalman filters thus providing the benefits of both. Doucet and Andrieu (2000); Bergman et al. (2002) use the Rao-Blackwell theorem to deliver an optimal estimator (given a starting, approximate one) in order to provide accurate state estimators for posteriors including robot pose and maps. This alleviates problems caused by the Markov assumption which requires locking down the global map after every update cycle. The question of how and when to incorporate new sensor data into the global map is important. If sensor data is amalgamated too soon then new geometric knowledge of the environment cannot be assimilated into the global map, for instance when the robot revisits locations (loop closing). Lu and Milios (1997b) essentially never consolidate the global map; instead each scan is stored with the current best guess for the translation and rotation relative to a global frame. In this manner, if the robot revisits a location, the information can be incorporated into the network of sensor scan poses; this idea is also adopted by Konolige and Gutmann (1999). Combining 3D range and image data is done in Newman et al. (2006) where the map is built from the point clouds and the camera images enable the detection of loop closures.

3.2.1 Inter-robot Localisation

The ability to measure the relative position between robots is one of the most important benefits arising from cooperating multiple robots. Kurazume and Nagata (1994) discuss the improvements over dead reckoning achieved using relative position determination between three robots. Navarro-Serment and Khosla (1999) propose using trilateration between four robots and achieve results an order of magnitude more accurate than standard dead reckoning techniques. It means that effectively there are active artificial mobile beacons in the environment that help resolve pose ambiguity. The use of multiple robots enables the application of methods normally reserved for

operational arenas containing artificial landmarks as exemplified by Hu and Gu (2000). Some research (Åström, 1993) shows the possibility of determining landmark locations by simply measuring separation angles between them at various unknown positions. This technique still works if the landmark and the measurement positions are all unknown at the beginning. The angle between the robot and the beacon is expressed in 3.6.

$$\beta\left(\vec{x},\vec{r}\right) = \text{atan2}\left(r_y - x_y, r_x - x_x\right) - x_\theta \tag{3.6}$$

where the beacon position is $r = (r_x, r_y) \in \Re^2$, the robot pose is $x = (x_x, x_y, x_\theta) \in \Re^3$. This is tackled by Åström (1993) as an optimisation problem; the function in (3.7) is minimised by adjusting the beacon positions against the constraints imposed by the landmark angles α_{ij}.

$$f = \sum_{i,j \in I} \left(\alpha_{ij} - \beta\left(\vec{x}_i, \vec{r}_j\right)\right)^2 \tag{3.7}$$

When the optimisation has converged the beacon positions are known and extracting the robot position at each measurement point is then possible. By measuring landmark separation angles alone the scale and orientation of the map of calculated landmark positions is undetermined. The scale of the map may easily be found in practice by measuring the distance separation between two landmarks; and the orientation of the map is often not required because absolute orientation indoors is rarely adopted due to the unreliability of magnetic compasses.

Cooperative localisation, sometimes referred to as co-location is useful despite its omission in some multiple mobile robot literature. For instance Yamauchi (1998) where robots detect each other as obstacles that should simply decay over time when the occupancy grid is updated. Co-location methods include various forms of vision such as an omni-directional camera (Omnicam) (Alur et al., 2002) or COPIS (Yagi and Yachida, 1996) and standard cameras (Schmitt et al., 2002; Jennings et al., 1999).

Research conducted by Khosla et al. (2004) looks at how to test whether robots can perceive each other given the current map and their positions. This is useful for

coordinating the actions of robot teams because maintaining a line of sight between team members is essential for some mapping algorithms. The research also introduces the idea of a virtual sensor which involves incorporating calculated data into sensor data as if it were perceived by the sensors. Localisation and mapping using heterogeneous robot teams with sonar sensors is examined extensively by Grabowski and Khosla (2001); Khosla et al. (2003).

Work on increasing the robustness of localisation amongst robot team members is carried out by Navarro-Serment et al. (2001) who tested various localisation fault tolerance algorithms in simulation. The main problems are due to colinearity and destructive interference. Colinearity is a problem when only bearing or range information is available; when both are present this situation is not as debilitating for the system. The latter is a problem inherent to the sensing mechanism deployed. Laser range finders cannot suffer from destructive interference as the phase of the laser changes so rapidly over time any phase coherence necessary for destructive interference will be so short as to be undetectable.

Quantitative measures of map quality and localisation performance are notably absent from much of the literature as observed by Lee (1996). Some of the best efforts (Elfes, 1991; Lim and Cho, 1992) assess the map quality using an average entropy measure.

3.3 System Design

Mutual localisation relies on the ability of each robot to both observe and be observed by other robots. This mechanism is achieved in Spletzer et al. (2001); Howard et al. (2003) amongst others. Omnidirectional vision is employed in Spletzer et al. (2001) which enables relative angles to be measured at all times. The range estimates are produced by measuring the apparent colour blob size, however they are inaccurate and susceptible to lighting conditions. Laser range finders with two retroreflective cylindrical targets are used in Howard et al. (2003) but relative pose is determined by observation of the target's orientation; consequently accuracy deteriorates quite rapidly

Figure 3.2: Pioneer 2 equipped with laser scanner and retroreflective beacon.

with range. The ideal set-up would use the laser scanner which benefits from requiring only one sensor and the inherent accuracy of the laser scanning returning both range and bearing as shown in Figure 3.2. The two-dimensional nature of the laser scanner and the 180° field of view are the main disadvantages. The two-dimensional scan means that it is difficult to mount an observable beacon in the plane without blocking some part of the scan.

A number of possible mechanisms for mutual observation are shown in Figure 3.3. The simplest, for theoretical analysis, are those where the position of the beacon on the robot coincides with the origin of the laser scanner. This may be accomplished by mounting the laser scanners at different heights, Figure 3.3(d), and angling the laser scanners, Figure 3.3(a). Alternatively the laser scanner may be encapsulated in a cylinder with a slot cut out allowing the laser rays to escape as shown in Figure 3.3(b). Adequate tilting ensures that the laser rays of other robots only perceive the outer cylinder. The main disadvantage of this approach is the variation in tilt forwards as opposed to that at the sides where the rays are horizontal; for fields of view less than 180° this may not be a problem. Also two targets could be located either side of the scanner, Figure 3.4(b), however detection of one of the targets gives rise to an

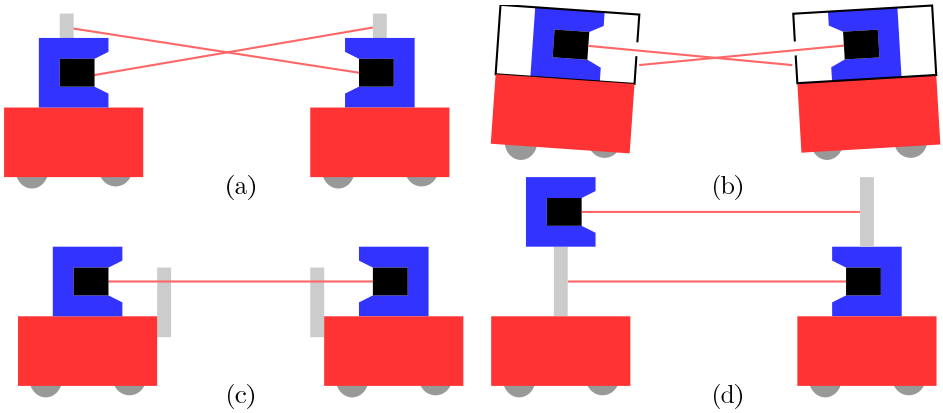

Figure 3.3: Mechanisms for mutual observation.

ambiguous pose.

One of the main drawbacks of retroreflective targets is the limited observation angle, however the detection process may be improved by circle detection. In experiments it was observed that retroreflectivity failed for angles of incidence exceeding 50°. This is an important consideration for straight targets; circular targets however always present an observable perpendicular surface. This restriction makes set-ups such as (a) and (d) in Figure 3.4 unsuitable. Another consideration when selecting appropriate targets is size. The SICK LMS 200 has a maximum resolution with a field of view of 180° of 0.5° which produces a ray separation of approximately 1cm for every metre of range. For example a cylinder of diameter 0.1m will always be visible up to 10m.

Having eliminated mutual observation strategies involving tilting or different heights leaves the strategies illustrated in Figure 3.4. The straight line target approach suffers from the limited observation angle of retroreflective tape but offers additional information on the orientation of the target and consequently the pose of the observer. Using cylinders means that there will always be a good retroreflective return, but when one cylinder is occluded the orientation of the target is unknown. Angling the straight 'wings' back 45° allows retroreflective returns at most positions but it is more complicated to extract target pose from observation and it admits pose ambiguity if only one is observed. In this chapter the single cylinder approach depicted in Figure 3.3(c) and Figure 3.4(c) is adopted because of the good performance of circle detection, Section 2.4.3.

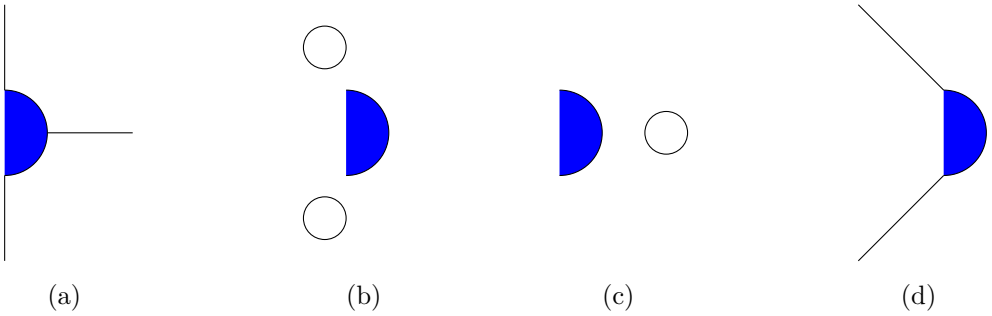

Figure 3.4: Overhead view of beacons allowing mutual observation.

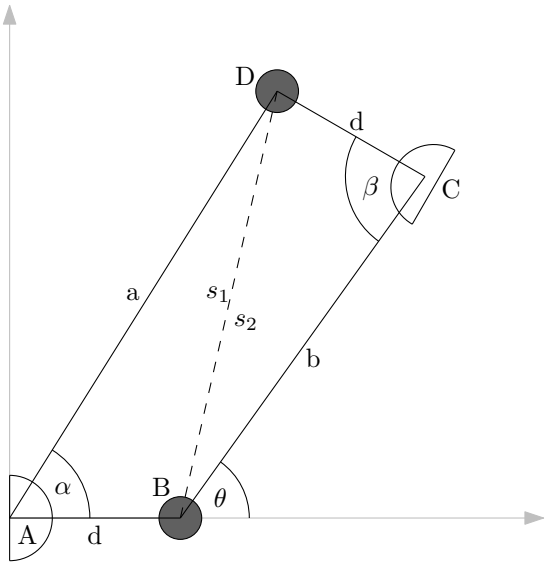

Figure 3.5: Beacon geometry used to calculate the relative pose. AB is the stationary robot and CD the mapping robot. D and B are the beacons with C and A the origins of the laser scanners.

3.3.1 Non-Coincident Beacons

It is usually not feasible, in 2D, to have the origins of the laser scanner and beacons coincident on a robot. If these origins are displaced then calculating the relative pose is made more complex. The best strategies, which are most suitable on balance from both a hardware and theoretical perspective, are thin targets mounted directly in front of the robot or larger targets but at different heights to the laser scanner. The disadvantages of the former are that the laser data from the central angles of the range scan is discarded and thin targets require the robots to be close for simultaneous mutual observations to occur. Displacement of the beacon from the origin of the laser scanner introduces the complications evident in Figure 3.5, where the laser scanners are represented as semi-circles with the forward part of the scanner corresponding to the curve of the semicircle and the beacons as solid circles.

In Figure 3.5 the robots are at A and C, with C mapping whilst A remains stationary. The pose of A may be constrained to the origin and x-axis without loss of generality. The beacons, B and D, are attached to the robots at A and C respectively. The distances AB and CD are both equal to d. The separation of the beacons is labelled as s_1 and s_2 because s_1 is the separation as observed by the robot at A and s_2 is the separation as observed by the robot at C, which in practice will not be identical. The robot at A observes the beacon D at range a and angle α whilst that at C observes the beacon B at range b and angle β. The beacon separation may be calculated from the robot observations by the cosine rule for triangles.

$$s_1 = \sqrt{d^2 + a^2 - 2ad\cos\alpha} \tag{3.8}$$

$$s_2 = \sqrt{d^2 + b^2 - 2bd\cos\beta} \tag{3.9}$$

Comparison of the two values s_1 and s_2 (which should be approximately equal) allows erroneous localisations to be detected and rejected. Typical large errors stem from distractors in the environment or error sensitive geometric configurations.

To acquire the position of the robot at C it is best to consider the geometry in the

complex plane rather than the more readily apparent geometric methods involving the sine and cosine rules. Using vectors in the complex plane gives two expressions for the position of C found by traversing the two available paths from the origin to C.

$$C = d + be^{i\theta} \tag{3.10}$$

$$C = ae^{i\alpha} + de^{i(\theta-\beta)} \tag{3.11}$$

Equating (3.10) and (3.11) gives the following expression for θ.

$$e^{i\theta} = \frac{ae^{i\alpha} - d}{b - de^{-i\beta}} \tag{3.12}$$

Eliminating $e^{i\theta}$ from (3.10) and simplifying gives C in terms of the observed angles, ranges and the parameter d.

$$C = \frac{abe^{i(\alpha+\beta)} - d^2}{be^{i\beta} - d} \tag{3.13}$$

The real and imaginary parts of C give the Cartesian coordinates of the second robot. The orientation is $\theta - \beta + \pi$ with θ determined from the argument of (3.12). For the implementation the complex number calculations are done in Java using the Java Math Expression Parser (JEP) because Java does not natively handle complex numbers.

Once the pose of robot C is acquired its scan data may be used to update the map as explained in Section 3.3.3.

3.3.2 Dealing with Distractors

Both robots search for regions of high intensity returns indicative of retroreflective objects in the environment. It is assumed that there are relatively few distractors as naturally occurring retroreflective materials are rare. The system however is robust to occasionally occurring retroreflective objects because of the mutual nature of the

observations. If one or both of the robots detect a spurious retroreflective distractor then the estimated separation of the robots as observed by each robot are likely to differ substantially. Typically it is found that if the difference between s_1 and s_2, Figure 3.5, exceeds a threshold of 0.02m then the occupancy grid should not be updated.

3.3.3 Updating the Map

The various map representations available to robotics researchers are explored more deeply in Section 5.2.1 however these experiments rely on standard occupancy grids for expressing the map. Occupancy grids are 2D arrays of cells each containing an occupancy probability. This is the probability that the cell contains an obstacle.

From Bayes' theorem for hypothesis evaluation (3.2) and assuming a binomial hypothesis that the cell can either be occupied or free then $p(H|D) + p(\neg H|D) = 1$ so $\frac{p(D|H)p(H)}{p(D)} + \frac{p(D|\neg H)p(\neg H)}{p(D)} = 1$ consequently

$$p(H|D) = \frac{p(D|H)p(H)}{p(D|H)p(H) + p(D|\neg H)p(\neg H)} \tag{3.14}$$

$$= \left(1 + \frac{p(D|\neg H)p(\neg H)}{p(D|H)p(H)}\right)^{-1} \tag{3.15}$$

For maps that are standard occupancy grids each cell holds the probability that the particular space it represents contains an obstacle. The grid is initialised so that all cells contain 0.5 representing complete uncertainty. New sensor data is incorporated via the Bayesian update process described in (3.22). Occupancy grids are an effective map representation as they are robust, handle uncertainty well and allow fusion of data from different sensors (Štěpán et al., 2005). Underlying the Baysian update for occupancy of a single cell is

$$P(\text{occupied}) = \left(1 + \frac{(1 - \text{likelihood})(1 - \text{prior})}{\text{likelihood} \times \text{prior}}\right)^{-1} \tag{3.16}$$

where the likelihood is provided by the data reading in conjunction with the sensor model and the prior is the current value of the occupancy for that cell. The result is

the new updated occupancy probability for this cell.

One of the severe limitations of the Bayesian update process stems from the independence assumption regarding successive scans. If the robot is stationary then successive scans are dependent and the robot updates the map with ever increasing certainty. Any errors that are a function of the position are almost permanently recorded and it becomes exceedingly unlikely that they will be corrected by conflicting observations from other positions.

A solution similar to Konolige (1997) alleviates this problem. The work in Konolige (1997) applies to sonar sensors however with suitable modification works well with laser scanners. In practice the laser scanner errors tend to depend only on position. Using pose buckets is not as helpful because there are still multiple updates to the map from the same position albeit at different orientations. Thus a position bucket system is used which only includes sensor data from new positions regardless of orientation. This is a cautious approach based on the premise that it is better to discard possibly useful data than risk irrevocably damaging the map. It may take slightly longer to map but is more reliable.

A notation and derivation similar to that of Thrun (2003) is adopted and is modified for the properties of the laser scanner rather than those of sonar sensors. For our purposes the poses and the measurements therewith are denoted $z_1, \ldots, z_T$ at various times up to the Tth time. This leads us to rephrase the problem of map building in the canonical form

$$p(M|z_1, \ldots, z_T) \tag{3.17}$$

where it must be emphasized that $z_1, \ldots, z_T$ is the set of measurements and corresponding poses. That is, the probability of the map, M, given such measurements should be determined.

Mapping and localisation may be modelled as a partially observable Markov decision process (POMDP). The state of the world can be only incompletely observed from any vantage point within the environment. Under a first-order Markov assumption the map

is assumed to contain the information conveyed by all observations up to that point and consequently can be recursively updated. The recursive Bayesian estimator means

$$p(m_k|m_0,\ldots,m_{k-1}) = p(m_k|m_{k-1}) \tag{3.18}$$

and also that

$$p(z_k|m_0,\ldots,m_k) = p(z_k|m_k). \tag{3.19}$$

Applying these assumptions gives rise to

$$p(m_0,\ldots,m_k,z_1,\ldots,z_k) = p(m_0)\prod_{i=1}^{k} p(z_i|m_i)p(m_i|m_{i-1}) \tag{3.20}$$

Finally, this produces

$$p(m_k|z_1,\ldots,z_k) = \frac{p(z_k|m_k)p(m_k|z_1,\ldots,z_{k-1})}{p(z_k|z_1,\ldots,z_{k-1})} \tag{3.21}$$

which only holds if the environment is time invariant.

For a particular map cell m and laser return cell l being either occupied in the map m or free in the map $\neg m$ then

$$p(m|l) = \left(1 + \frac{p(l|\neg m)p(\neg m)}{p(l|m)p(m)}\right)^{-1} \tag{3.22}$$

The log odds approach to probabilities simplifies the update step and enhances the expressive power of the map, which mostly contains very high probabilities (cells with probabilities under a threshold are removed from the map). The odds of an event A with probability $p(A)$ are

$$o(A) = \frac{p(A)}{p(\neg A)} = \frac{p(A)}{1-p(A)}, \tag{3.23}$$

and conversely probabilities can be calculated from the odds as

$$p(A) = \frac{o(A)}{1 + o(A)}. \tag{3.24}$$

Thus (3.22) can be expressed as

$$o(m|l) = \frac{p(l|m)p(m)}{p(l|\neg m)p(\neg m)} \tag{3.25}$$

$$o(m|l) = \frac{p(l|m)}{p(l|\neg m)} o(m) \tag{3.26}$$

Therefore in terms of log odds

$$\log o(m|l) = \log \frac{p(l|m)}{p(l|\neg m)} + \log o(m) \tag{3.27}$$

$$\log o(m|l) = \log p(l|m) - \log p(l|\neg m) + \log o(m) \tag{3.28}$$

For many of the experiments in this research the map resolution is better than 0.1m, the operating arena size is 30m by 30m and the origin is the initial position of the observing robot. The quantities $p(\neg m)$ and $p(m)$ can be approximated by assuming the map cells are independent. If this is the case then the coverage (percentage of occupied cells) c, means $p(m) = c$ and $p(\neg m) = 1 - c$. The term $p(l|m)$ represents the probability of observing the cell as occupied given that the corresponding one in the environment is occupied and $p(l|\neg m)$ is the probability that the cell is perceived as occupied given the probability that the environmental cell is free. The relationship between these is given as $p(\neg m) = 1 - p(m)$. These predictive probabilities are calculated from the sensor model which can be established by scanning a known environment. The most common situations producing the observation of occupied cells where none exist in the map are pixel mixing and observing previously unobserved regions of the map.

There are a number of glaring inaccuracies produced by the independent cell as-

sumption. Most notably if many intervening cells between the origin and the return point are occupied then the probability that this is an accurate reading is diminished, and so less weight should be assigned to that reading. Also the probability of occupancy for a cell is clearly affected by that of the surrounding cells as can be seen from the statistical analysis of a typical map. The problem with the map cell independence assumption is more evident in 3D maps.

Due to the fidelity and narrow beam width of the laser scanner straightforward ray-trace models can be applied for the sensor model. When updating the occupancy grid, all cells along the path of the laser ray have their probabilities decreased whilst the cell containing the end point of the ray has its probability increased. This method assumes that the occupancy of a cell is independent of the occupancy of others; this is mathematically convenient but physically inaccurate. This cell independence along the ray trace path is especially noticeable in situations where a laser return has been registered at a pose that has many intervening occupied cells according to the map. These intervening occupied cells, when considered together, make the probability of the data given the map and pose very unlikely. This implies the sensor reading, pose or map are incorrect.

3.4 Experimental Results

A number of experiments are undertaken to test the effectiveness of mutual localisation for mobile robot localisation and mapping. These experiments are conducted in both simulation and laboratory environments. For all experiments described in this chapter the map update is performed according to (3.22) and assumes a raytrace sensor model for the laser scanner.

3.4.1 Simulation Experiments

Preliminary work was undertaken in the simulation environment of Player/Stage (Gerkey et al., 2003). The simulation set-up consisted of the retroreflective cylindrical beacons

with radii 0.01m mounted 0.2m directly in front of each robot. The quality of a typical map generated is compared with the true map, Figure 3.6. Poses where the beacons are separated by less than two metres are discarded. The occupancy grid is a result of 54 updates. One of the advantages of simulation is the ready availability of the true map allowing an assessment of map fidelity, Figure 3.6. A crucial disadvantage of simulations is that they often use an overly simple model of the real world and so results in simulation tend to be better than they would be if the experiment were done for real. Without real world errors the map produced in Figure 3.6 is very good and although this shows the viability of the algorithm, it does not guarantee operation in real experiments. Algorithms are often tested in simulation because it is faster to experiment in simulation; indeed this testing may even be automated. It is interesting to compare Figure 3.6 and Figure 3.9 in which the experiment is done with real mobile robots. The shadows observed behind obstacles in both simulation and experiment result from occlusion and the probabilistic occupancy grid updates. These shadow cells have been observed as free less frequently due to occlusion by the obstacles. Consequently, the probability that they are free is lower and is represented as grey, rather than white, pixels in the occupancy grids.

3.4.2 Real Experiments

Localisation performance is assessed by manoeuvring the robot along a series of straight lines. For this experiment the laser scanners are mounted at different heights and retroreflective cylinders 0.04m in radius are mounted in front (0.12m) and above or below the scanner. Deviation of the pose from the straight line quantified both the position error perpendicular to the line and the orientation error. The position and orientation results from these experiments are summarised in Figure 3.7 and Figure 3.8. The average position error is small at around 0.01m and the orientation error is approximately 1.4°.

The orientation accuracy is harder to measure because the accuracy of mutual localisation exceeds the accuracy in aligning the robot. There is no dependence of the

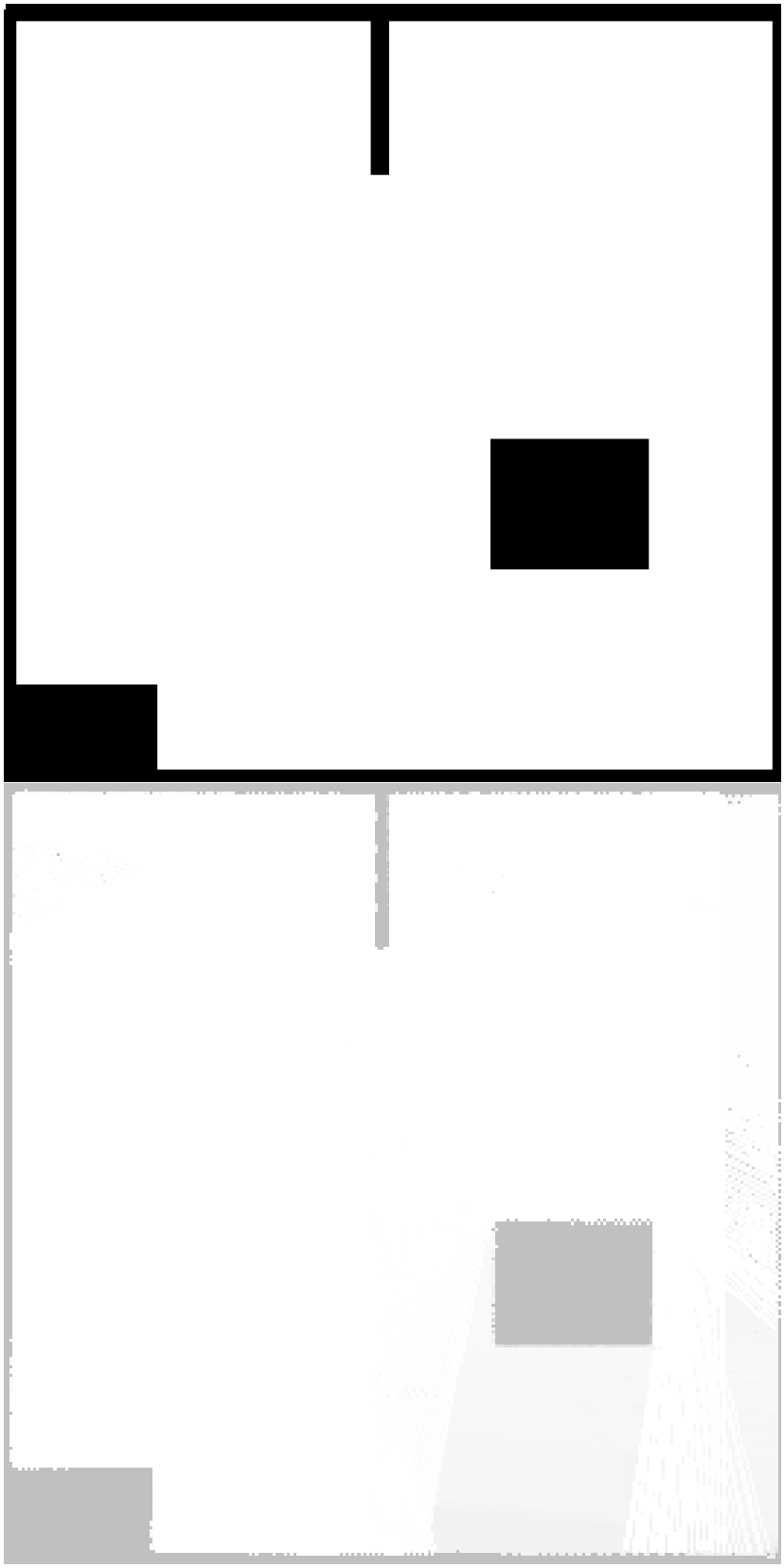

Figure 3.6: Plan of room used in the simulation with map produced.

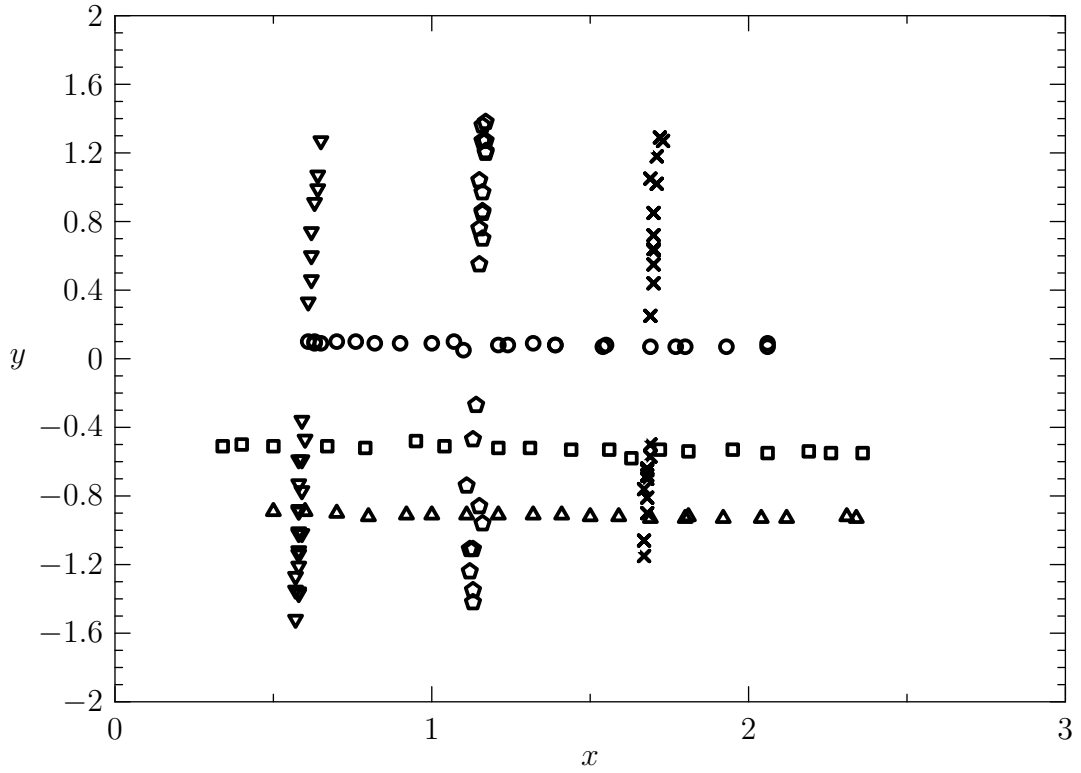

Figure 3.7: Mutual localisation for multiple straight trajectories.

orientation error on θ (Figure 3.5) and possibly a weak dependence on b (Figure 3.5) as shown in Figure 3.8.

Now that the localisation error has been established an experiment to assess the mapping accuracy is performed. In this experiment a square enclosure of internal dimensions 2.42m by 2.49m is mapped by two robots employing mutual localisation. Figure 3.9 shows the overhead view and corresponding occupancy grid that is produced. The cell size of the occupancy grid is 0.01m and the dimensions of the enclosure are found to be 2.42m by 2.44m. These results give a geometric error of around 2%.

The experimental set-up involved Pioneer 2 robotic platforms with SICK LMS 200 laser scanners depicted in Figure 3.2. Cylindrical beacons 1cm in diameter wrapped in retroreflective tape are mounted 0.07m directly in front of the laser scanner. The retroreflective nature of these beacons makes them easy to detect however they are too narrow for mutual localisation much over 2m. One robot remains stationary whilst the other is teleoperated to explore the enclosure. A ray trace sensor model is assumed when updating the map.

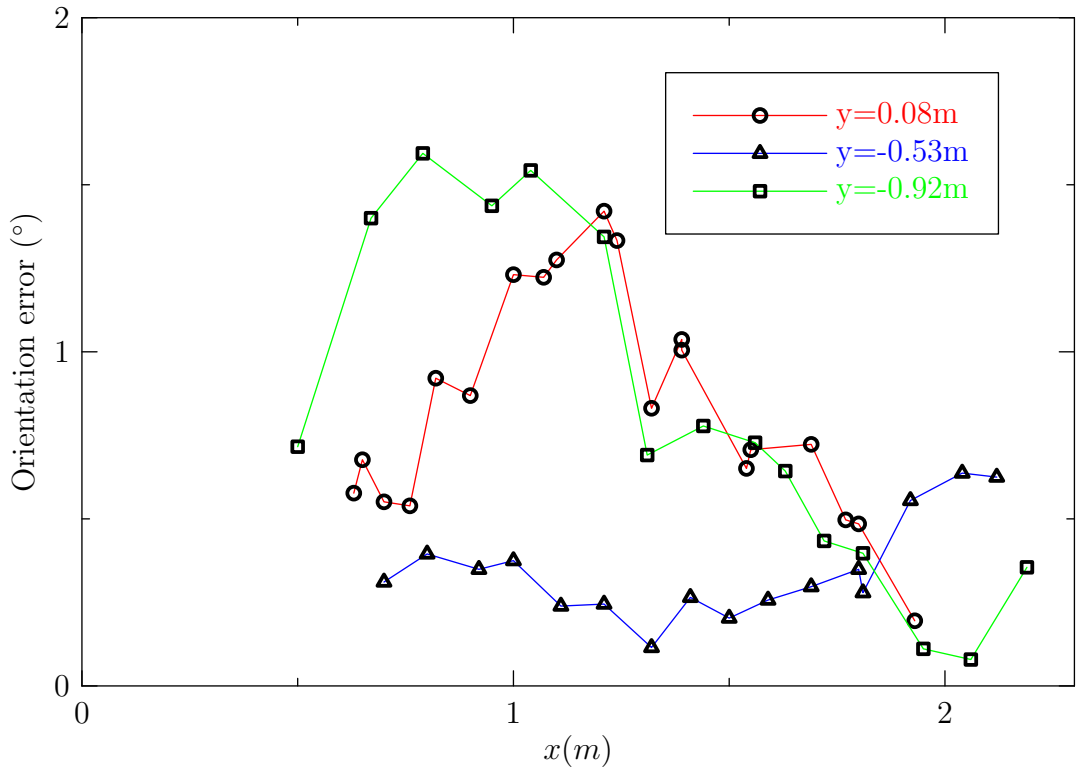

Figure 3.8: Orientation error for straight trajectories parallel to the x-axis.

Two categories of errors need to be considered for a useful occupancy grid, geometric errors and reliability errors. Geometric errors are determined by comparison of distances in the map with the real distances and are a result of systematic and random errors. The reliability error can be assessed by discrepancies between the occupancy of cells in the map with corresponding cells in the real environment.

3.4.3 Map Building Results

The results of a mutual localisation and mapping experiment are displayed in Figure 3.9. The map may be compared to an overhead view of the arena which is corrected for spherical aberration. The robots are equipped with thin retroreflective beacons as depicted in Figure 3.2. The robot on the right remains stationary throughout the experiment whilst that on the left explores. The left robot infers its position through mutual localisation with the stationary robot. Knowing its position the exploring robot can update the global map accordingly with each laser scan to produce the occupancy map in Figure 3.9. From the size of the arena and a comparison between the photograph

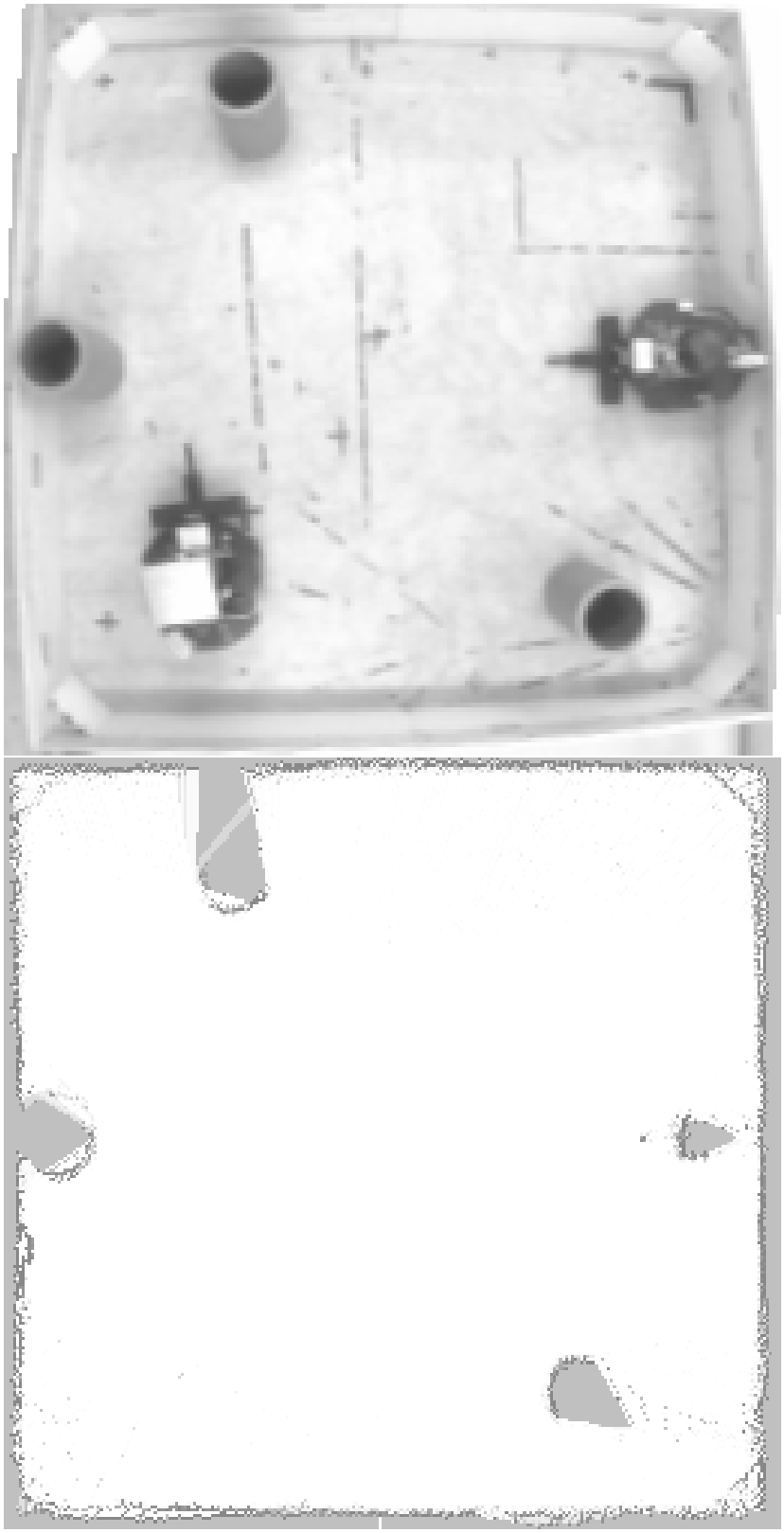

Figure 3.9: Overhead view of enclosure 2.49m by 2.42m and corresponding occupancy grid generated. The robot on the right remains stationary whilst that on the left is teleoperated.

and map in Figure 3.9, it may be concluded that the error is of the order of 0.02m.

3.5 Analysis

The expression for the position of the robot given in (3.13) depends upon the observed angles and ranges of the mounted beacons as well as the mount point of the beacons on the robot $(d, 0)$ relative to the origin of the laser scan. To properly judge the accuracy of this method an error analysis is conducted.

$$\Delta C(a, b, \alpha, \beta, d) = \frac{\partial C}{\partial a}\Delta a + \frac{\partial C}{\partial b}\Delta b + \frac{\partial C}{\partial \alpha}\Delta \alpha + \frac{\partial C}{\partial \beta}\Delta \beta + \frac{\partial C}{\partial d}\Delta d \tag{3.29}$$

shows the dependence of the error function for C (Figure 3.5) upon the errors in d and the observed ranges and angles. Partial differentiation of C with respect to a, b, α, β and d illustrates the functional dependence of the error upon these quantities. Finally assume d is small compared to the separation and the following limits ensue.

$$\lim_{d\to 0}\frac{\partial C}{\partial a} = \lim_{d\to 0}\left(\frac{be^{i(\beta+\alpha)}}{be^{i\beta}-d}\right) = e^{i\alpha} \tag{3.30}$$

$$\lim_{d\to 0}\frac{\partial C}{\partial b} = \lim_{d\to 0}\left(\frac{ae^{i(\beta+\alpha)}}{be^{i\beta}-d} - \frac{e^{i\beta}\left(abe^{i(\beta+\alpha)}-d^2\right)}{\left(be^{i\beta}-d\right)^2}\right) = 0 \tag{3.31}$$

$$\lim_{d\to 0}\frac{\partial C}{\partial \alpha} = \lim_{d\to 0}\left(\frac{abie^{i(\beta+\alpha)}}{be^{i\beta}-d}\right) = aie^{i\alpha} \tag{3.32}$$

$$\lim_{d\to 0}\frac{\partial C}{\partial \beta} = \lim_{d\to 0}\left(\frac{abie^{i(\beta+\alpha)}}{be^{i\beta}-d} - \frac{bie^{i\beta}\left(abe^{i(\beta+\alpha)}-d^2\right)}{\left(be^{i\beta}-d\right)^2}\right) = 0 \tag{3.33}$$

$$\lim_{d\to 0}\frac{\partial C}{\partial d} = \lim_{d\to 0}\left(\frac{abe^{i(\beta+\alpha)}-d^2}{\left(be^{i\beta}-d\right)^2} - \frac{2d}{be^{i\beta}-d}\right) = \frac{a}{b}e^{i(\alpha-\beta)} \tag{3.34}$$

The error in the position as expressed in (3.13) not only depends on the partial derivative but also on the errors in the quantities involved.

$$\Delta a \approx 0.02m, \Delta b \approx 0.02m, \Delta\alpha \approx \frac{w}{a}, \Delta\beta \approx \frac{w}{b} \tag{3.35}$$

where w is the apparent width of the mounted target at that observation angle. In the case of a circular target the observed width is always the same. For the far field approximation assuming, in the worst case, only one laser beam is reflected from the target then the error is given in (3.35). In this case the angular error is proportional to the size of the target over the range. Consequently the total error in C is independent of robot separation. Although the error is independent of robot separation, increased distance between the robots decreases the frequency of simultaneous mutual detection events. This in turn reduces the frequency of simultaneous observation pose updates. Asynchronous observations may still be used to track position and orientation separately.

Combining the partial differentials and (3.35) gives the overall error in the position expressed as

$$\Delta C = e^{i\alpha}\Delta a + aie^{\alpha i}\Delta\alpha + \frac{a}{b}e^{(\alpha-\beta)i}\Delta d. \tag{3.36}$$

In the far field approximation with small d when a and b are similar in size the quotient $\frac{a}{b}$ tends to one. The potentially dominating error term involving $\Delta\alpha$ is subdued by the $\frac{1}{a}$ dependence of $\Delta\alpha$. By taking the modulus of each term in (3.36) and substituting (3.35) the maximum error is given by

$$\Delta C_{max} \approx \Delta a + w + \Delta d \tag{3.37}$$

This result is important because it demonstrates that error in the far field approximation for large robot separations is not unduly sensitive to the quantities measured for any position or orientation combination for two robots.

For accurate mutual localisation it is vital that the mutual observation events are simultaneous. By simultaneous it is meant that there has been minimal pose change between the times of the observation events. This is especially important for the orientation part of the pose. If synchronising mutual observations proves difficult an alternative solution is to use the pose information whilst moving and only update the map when the robot is stationary or the rate of pose change sufficiently small. For the experiments described in this research the robots are all moved to their new positions whereupon they are issued the command to scan. As they are stationary whilst scanning then the resulting scans can be considered simultaneous.

3.6 Summary

This chapter starts with an exploration of related work involving 2D mapping and localisation. The theory behind probabilistic updates of occupancy grids is explained, particularly with reference to Bayesian probability. The necessity of the Markov assumption to simplify the map update process is introduced and some resulting maps from both simulation and experiment displayed.

Various hardware configurations that allow robots to both observe others and be themselves observed are assessed. The configurations are optimised for robots equipped with laser range detectors.

A new approach to cooperative localisation has been explored in this chapter. Traditionally cooperative localisation has been the natural extension of the corpus of work on artificial landmarks. Namely that the benefits of artificial landmarks can be enjoyed without the need to deploy them or otherwise alter the environment. Localisation and mapping was accomplished by mounting beacons on robots and using them as artificial landmarks. By leaving the beacon robots stationary an exploratory robot could navigate and map an area. When that area has been mapped the team may move on to another area. By mounting beacons on robots we have not only mobile landmarks but also observing landmarks.

Ensuring that robots may observe team-mates and be observed themselves means

that simultaneous mutual localisation events can occur. These events allow superior relative pose determination. Firstly, the mutual localisation is robust with respect to spurious readings because simple checks on the validity of the mutual pose are available; for instance the separation of the robots should be similar as measured by both observing robots. Secondly, the accuracy in the pose does not deteriorate with separation, a very useful property. Increasing separation merely decreases the frequency of mutual localisation events. Having to place the beacon at a different position to the origin of the laser sensor introduces some complications. These are handled and the error properties of mutual localisation are derived theoretically and validated experimentally.

Having demonstrated the effectiveness of cooperative localisation and mapping through mutual localisation in 2D the remaining chapters extend these ideas for 3D sensors and maps.

Chapter 4

Developing a Novel 3D Laser Scanner

Many real-world applications require mobile robots to have 3D perception and mapping capabilities. This chapter proposes a novel mechanism for augmenting a traditional 2D laser range finder to produce almost omni-directional 3D scans at degree resolution in 3 seconds. The 3D laser scanner is built by adding a rotating mirror to a conventional 2D scanner. The resulting sensor delivers 3D scans at a configurable frequency up to 1Hz, is accurate to 0.02m over a range of 8m and is easy to deploy on conventional mobile robots.

4.1 Introduction

The dearth of suitable sensors has hindered many real world applications of mobile robotics. Cameras are cheap low mass sensors with low power requirements and return images that our brains may easily interpret as full 3D scenes. That these images may be so effortlessly interpreted by the human brain has encouraged many researchers to rely on cameras as sensors for mobile robots. Unfortunately the interpretation, that of converting the raw data into information, has proved to be remarkably difficult. In hindsight perhaps this is to be expected as a large proportion of the human brain is dedicated to processing visual information. In the process of interpreting visual data

much background information is applied by the brain, for instance the expected shape and size of objects, shadows and occlusion. The field of image depth perception is extensive and complicated as is evidenced by the evolution of art and the vast array of techniques used by artists to bring still 2D images to life. In this way visual perception is not only a sensor problem but also an artificial intelligence one.

One of the most demanding aspects of computer vision is the conversion of the 2D image information into 3D. Converting from 3D to 2D has been long established, in the simplest case merely involving a projection onto a plane. The projection however discards information and so the inverse transformation of 2D to 3D is impossible to perform unambiguously from a single image. Unambiguous transformation requires extra information which is obtained in many ways. Multiple points of view are necessary and may be provided by multi-camera set-ups with the simplest being stereoscopic, alternatively on mobile platforms multiple points of view may be acquired over time by the motion of the platform. This is sometimes referred to by others in the field as structure from motion. In multi-camera set-ups the baseline separation of the cameras is known to high accuracy making depth perception easier than with multiple view points with unknown separation.

Establishing correspondence points within images taken from different positions and then applying triangulation enables the range to these points to be calculated. Finding corresponding points in successive images has proved difficult to do quickly and reliably in varying light conditions and dynamic environments.

Once correspondence points have been established the range to the point is determined by triangulation. Further problems include occlusion and the functional dependence of range error upon range (4.6). From (4.6) it is clear that the baseline separation s of the observation points may be adjusted to reduce triangulation errors. The separation required, $s = 2x$ for the minimum error, is large compared to the range. This larger viewpoint separation increases the likelihood of occlusion. Finally good correlation points are rare on featureless surfaces such as uniformly coloured and illuminated walls, a common characteristic of indoor environments.

Due to the current state of AI research it is necessary to have sensors that do more of the interpretation of data, converting data to information. For this reason research was aimed at developing an appropriate mobile robotic sensor.

Compared to image-based sensors laser range finders are heavier and more expensive. However the difficulties extracting geometric data about the environment from images versus the high quality geometric data provided by the laser scanners provided the motivation for extending the 2D laser range finders to produce 3D scanners. Although this takes some effort from a hardware perspective it has proved invaluable in circumventing the difficult 2D to 3D transformation problem. The lack of inexpensive commercially available 3D laser scanners suitable for mobile robots prompted the development of a custom built 3D laser scanner as described in this chapter.

Section 4.2 includes relevant experimental efforts in mobile robot 3D sensors. This section includes a derivation of the error functions for any system relying on triangulation for range finding. Triangulation underpins many camera based ranging techniques such as stereoscopic vision. The hardware and software elements of the new scanner are presented in Section 4.3. In this section an alternative laser scanner design is introduced and the efficacy compared to that of the first design. The sensor model appropriate for such a scanner is explored in Section 4.4. Section 4.5 contains 3D scans generated by the 3D scanners and assesses their accuracy. The effective visualisation of the 3D data is briefly addressed in this section.

4.2 Related Work

Extracting 3D information about the world surrounding a robot has proved difficult and the two main approaches, vision and laser range finding, have been dogged by problems. Vision is often computationally intensive and suffers from sensitivity to changes in illumination. Many of the difficulties stem from having to solve the correspondence problem which can be alleviated by structured light approaches, however the data spatial density does not come close to that provided by laser scanners. Non-visual localisation and mapping has taken place in 2D, mainly due to limitations of the

sensors in the case of laser range finders, or processor speed and algorithms for that of stereoscopic vision.

Recently in a drive to test the benefits of 3D sensors researchers have mounted 2D laser scanners on nodding or rotating mechanisms to obtain 3D scans (Lingemann et al., 2004; Nüchter et al., 2005; Pfaff et al., 2007). Alternatively, two laser scanners mounted with their scan planes orthogonal (Howard et al., 2004b) are also popular. It must be noted that the orthogonal mounting technique can produce 3D maps however it does not give the robot 3D sensory perception which is necessary for reliable obstacle avoidance. Significant attention is being focused on 3D laser mapping (Montemerlo and Thrun, 2004; Howard et al., 2004a; Hähnel et al., 2002).

A completely different approach is undertaken by Weingarten et al. (2004). This scanner has a modulated infra-red illumination and a CMOS/BCCD image sensor taking four samples of the reflected infra-red illumination intensity to determine the phase difference of the reflected data for each image pixel. From this phase difference and the frequency of illumination modulation the distance to image pixels can be ascertained unambiguously up to 7.5m. This sensor delivers 19480 range measurements at 30 FPS, with no moving parts, whilst comparing favourably in accuracy to the SICK LMS 200. Its main disadvantage is price, low availability and a relatively narrow field of view, although this can be altered by the lens arrangement.

There are a few commercial 3D laser range finders however these are often prohibitively expensive especially when required for multiple robots. Companies such as Riegl and Deltasphere sell scanners that could be used but they are designed for surveying and are typically tripod mounted, rather than for operation on mobile robots. As a result they have a narrow field of view sacrificing scan rate for high point density.

This multitude of custom 3D range scanners illustrates the rapidly increasing interest in full 3D robotic sensory perception. The benefits of full 3D mapping are abundant and so the rapid expansion of this field is inevitable. The detection of negative and over-hanging obstacles greatly enhances avoidance behaviour. Once 3D maps of environments have been built they can be customised for and distributed to different

robots. For instance various 2D occupancy grids may be built for robots of different sizes or with 2D sensors at different heights. Severely cluttered non-manifold environments such as search and rescue situations may be reliably mapped. Maps based upon the ceilings of rooms (Jeong and Lee, 2005) will remain accurate for longer and an unoccluded view of the ceiling is usually readily accessible to a robot even in crowded environments (Burgard et al., 1999). An alternative approach taking advantage of full 3D perception (Wulf et al., 2004) uses virtual 2D scans produced by collapsing the 3D data vertically into a plane and then taking the furthest point for each 2D scan angle. This produces good 2D representations in cluttered environments.

The disadvantages of 3D sensing technologies are slower scan acquisition time and the geometric increase in data that need be processed.

In summary the performance of many mobile robots has been hampered by the absence of accurate and reliable 3D spatial information. Lacking this information mobile robots are doomed to be ineffective and confined to research laboratories.

4.2.1 Triangulation Error

As mentioned in Section 4.2 stereoscopic vision and other approaches that rely on a triangulation for range finding are beset by problems. One of which is the dependency of accuracy on range. Figure 4.1 depicts the geometry of a multiple point of view triangulation range finding system. The separation of the points of view is $s = s_1 + s_2$. In order to determine the functional dependence of error upon range the equation for the range is first determined. By solving for the intersection of the observation lines

$$x = \frac{s}{\tan\beta - \tan\alpha} \tag{4.1}$$

consequently

$$y = x\tan\beta \tag{4.2}$$

Along the perpendicular bisector of the observation points ($y = s/2$) $-\alpha = \beta$ where

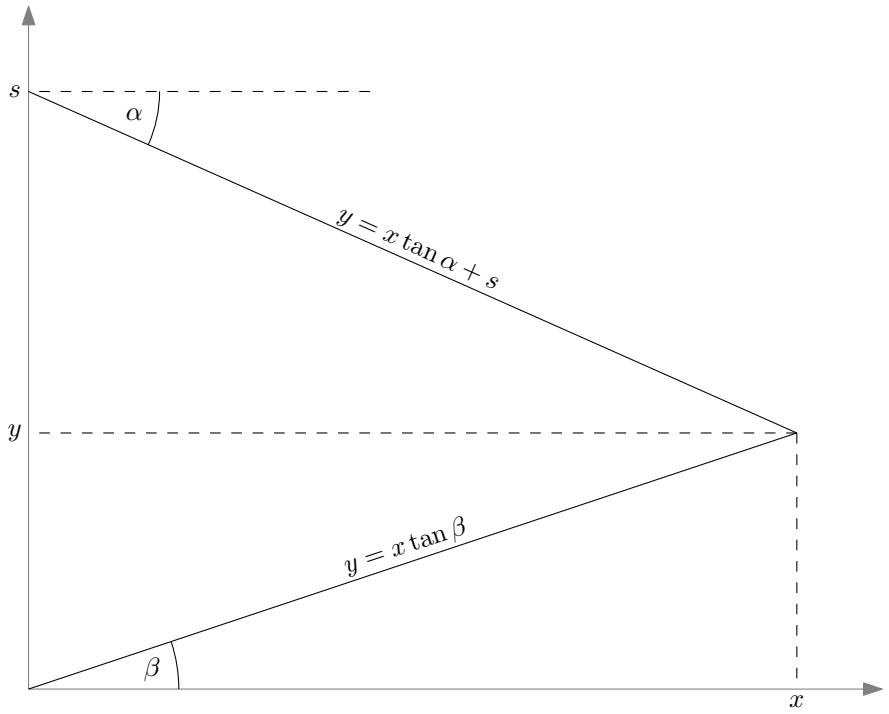

Figure 4.1: Triangulating the point (x, y) by observation from the origin and $(0, s)$ where s is the separation. It should be noted that in this example α is negative.

triangulation is most accurate

$$x = \frac{-s}{2\tan\alpha}. \tag{4.3}$$

Differentiating with respect to α

$$\frac{dx}{d\alpha} = \frac{s}{2\sin^2\alpha} \tag{4.4}$$

From the right-angled triangle

$$\sin\alpha = \frac{s/2}{\sqrt{\frac{s^2}{4} + x^2}} \rightarrow \frac{1}{\sin^2\alpha} = 1 + \frac{4x^2}{s^2} \tag{4.5}$$

Substituting (4.4) into (4.5) and simplifying leads to the following expression for the dependence of the range error (Δx) upon the angular error $(\Delta\alpha)$

$$\Delta x = \left(\frac{s}{2} + \frac{2x^2}{s}\right)\Delta\alpha. \tag{4.6}$$

From (4.6) it is clear that the error in range increases with the square of the range.

Increasing the baseline separation s reduces the error. In fact minimising (4.6) results in $s = 2x$ showing how wide a baseline is needed for the most accurate range determination by triangulation.

For the situations in which a stereoscopic system could be mounted on a medium sized robot, $s = 1$, using cameras with 640 horizontal resolution and FOV of 30° results in errors of 0.1m at a range of 8m. This error is the best case error when the pixel resolution is available on the centreline between the cameras. This compares poorly with a laser scanner having a typical range error under 0.02m. Optimal triangulation would warrant a separation of 4m which is unfeasible for a medium sized indoor robot. Multiple cooperating robots could establish correspondence points amongst themselves and thus employ an observation separation that is closer to optimal.

4.3 3D Scanner Design

As discussed in Section 4.2 there are a number of approaches for acquiring 3D sensor data. Some of the more successful involve two 2D scanners perpendicular to one another and rotation of one 2D scanner. With two perpendicular scanners one is responsible for providing data for localisation whilst the other records cross-sections of the environment. Merging these cross-sections produces 3D representations. These are however not true 3D sensors and are subject to a number of constraints. Motion is constrained to a plane and obstacle detection is not particularly enhanced over standard 2D scanners. Commercial 3D scanners do exist however they are very expensive and are not suitable for deployment on mobile robots as the scanners have narrow fields of view and long scan times. Instances of successful true 3D scanners often involve rotating or rocking a 2D laser scanner. Specific hardware configurations vary in the details of the 2D laser scanner motion however they tend to be quite complicated and require sophisticated precision engineered mechanical assemblies due to the weight and connections of the 2D laser scanner.

For these reasons the approach implemented in this work is to extend 2D scanners by placing a rotating mirror in front of or above (for the enhanced scanner) to divert the

scan planes vertically to produce 3D scans. As will be seen the hardware apparatus is relatively straightforward and quick to deploy allowing multiple conventional 2D robotics platforms to be equipped with 3D sensing capabilities.

4.3.1 Hardware Design

Two types of 3D laser scanner are employed in this research, Figure 4.2. The accuracy of the laser scanner is fundamentally dependent on precise knowledge of the mirror position. There are two main methods to achieve this rotational accuracy. The closed and open feedback control loop methods. With the closed feedback loop the angular velocity of the motor does not have to be as smooth and consistent as long as the angular sensor is accurate and reliable with a high update rate. This places less demanding requirements on the driving mechanism at the cost of a more complex feedback mechanism and the requirement for a rotation sensor. Conversely with the open feedback loop control approach the driving mechanism must be very reliable and consistent over several mirror revolutions. Thus the cost with this approach is in the drive mechanism but there is no need to measure the mirror angle frequently and return this information to the host robot. As explained in Section 4.3.3 this is accomplished to high accuracy with a quartz oscillator driving a high resolution stepper motor. The closed loop scanner is accomplished with a standard DC motor and a rotation sensor.

The main problems with nodding or rotating the laser scanner are difficulties in hardware implementation and high power consumption as the 2D SICK scanners are heavy. The mechanism used in this research is a rotating mirror scanner which produces 3D scans with a field of view up to 120° by 270° (4.11). It is light, has low power consumption and is easily deployed on conventional robotics platforms. As shown in Figure 4.2 the 3D laser range scanner consists of a SICK LMS 200, rotating reflecting surface prototype and software to convert the range scans into 3D data. The mirror rotation period is 40 seconds. The slow rotation speed, mass of the mirror compared to the laser scanner, continuous rather than reciprocating motion, all contribute to a very low power consumption of 0.01W. The angular position of the mirror, ϕ, is fed

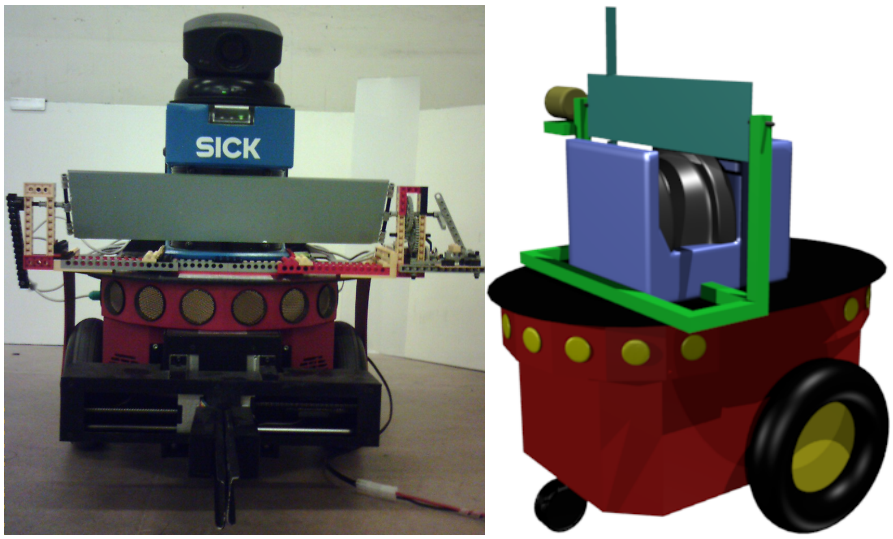

Figure 4.2: Pioneer 2 robot with rotating mirror mechanisms enabling 3D scans. The first generation of 3D laser scanner and the enhanced version.

back to the robot through a modified PS2 ball mouse.

The Pioneer 2 robots have standard PS2 keyboard and mouse inputs. In order to use this the 3D scanner driver has to have exclusive access to the mouse device. The robots run Red Hat Linux and mouse data is written to and read from the PS2 device '/dev/psaux'. Other programs which use the mouse on the robot have to be terminated to ensure that all mouse input is available for the laser scanner driver. The behaviour of the consumption of mouse events does seem to be Linux kernel dependent. The PCB of a standard PS2 ball mouse is removed from the mouse casing and the orthogonal x and y rotation sensors provided feedback to the computer as standard mouse movement events. These allow the angular position of the mirror to be tracked. Under the PS2 mouse specification the mouse updates at a default rate of 100Hz with a time period of the mirror rotation of T then the angular error due to the update rate is given as

$$\delta\phi = \frac{4\pi}{f_m T} \tag{4.7}$$

where f_m is the update frequency of the mouse. Note that the angular velocity of the reflected ray is twice that of the reflecting surface. Consequently the error in ϕ due to the update rate is twice that of mirror. With the default update rate for a PS2 mouse and a mirror period of 20s the angular error is 0.36°.

The other main source of error is the resolution of the angular sensor which is dependent on the number of teeth on the cog wheel b blocking the light sensor and the gearing ratio r_g between the mirror and the sensor cog. Thus the error in ϕ in radians due to discrete angular sampling is

$$\delta\phi = \frac{4\pi}{r_g b} \tag{4.8}$$

For the scanner used in the experiments $r_g = 25$ and $b = 14$ giving rise to a digitisation error of 2.1°. Clearly the angular digitisation error is dominant and so linear interpolation, (4.9), is used to help alleviate it.

The Player server (Gerkey et al., 2003) reads the SICK LMS 200 laser data and the time stamp of each scan allows association with the mouse events. When a scan is received a query to the angular sensor device driver supplying the scan time in milliseconds returns the angle the mirror was at that time. This is achieved by maintaining a history of angles and their corresponding times. The scan time is looked up in this history and the time interval t_{start} and t_{end} containing it is identified along with the corresponding angles at these times ϕ_{start} and ϕ_{end}. Standard linear interpolation based on the position of the scan time in the interval is carried out,

$$\phi(t) = \phi_{start} + \frac{t - t_{start}}{t_{end} - t_{start}}(\phi_{end} - \phi_{start}), \tag{4.9}$$

to give the linearly corrected $\phi(t)$ of the laser scan.

Another source of error is the assumption that all points within a scan are sampled at the same ϕ. This is not true as the mirror continues to move as the laser scans horizontally. The scan time for the SICK LMS 200 is dependent upon the resolution and at 1 degree resolution the scan time is 13ms. During this time the laser ray will have been rotated through 0.3°. Assuming that the time stamp on the laser data represents the time at the start of the scan a correction is applied by calculating the angular velocity for a given interval and incrementing ϕ as returns with increasing horizontal angle (θ) are considered.

In the default mode PS2 mice stream updates at 100Hz. Any packets missed by the robot's onboard computer lead to a systematic error in the angle. To ensure that updates are only received when the onboard computer is ready the mouse is set into remote mode. In this mode the mouse tracks position changes but only sends update packets upon request from the host computer.

The SICK LMS 200 has a number of range operating modes, the most suitable for these purposes is the millimetre resolution mode which has an operating range of up to 8m. The robot's own beacon is visible in the scans and so readings corresponding to this are removed, Figure 6.4.

The gear train between the mirror and the angular sensor introduces backlash errors. These are alleviated by continuously driving the mirror in one direction at a constant angular velocity rather than reciprocating it. Due to backlash, once every mirror revolution there are large variations in angular velocity which are detrimental to scanner accuracy. These occur when the centre of mass of the mirror is displaced from the axis of rotation. By weighting the mirror axis it is possible to shift these velocity variations to less important parts of the scan, in this case the floor. In the results the floor data points have been removed. Range readings when the mirror is edge on to the laser scanner are removed.

The angular sensors in a standard ball mouse only provide relative motion estimates. In order to obtain the absolute position a triggering arm connected to the remaining mouse rotation sensor falls back into position once a revolution. This sharp falling edge appears as a mouse movement event in the return PS2 update packet.

To support on demand scans which may be initiated at any time during the mirror rotation cycle the driver starts recording the laser data in spherical polar coordinates immediately and the ϕ value at which this trigger is detected is recorded. When the scan is completed all the ϕ values of the points are decremented by the ϕ offset recorded during the scan and converted to Cartesian coordinates. In this way when requesting a scan one need only wait for one revolution of the mirror to complete rather than waiting for the trigger and then a revolution.

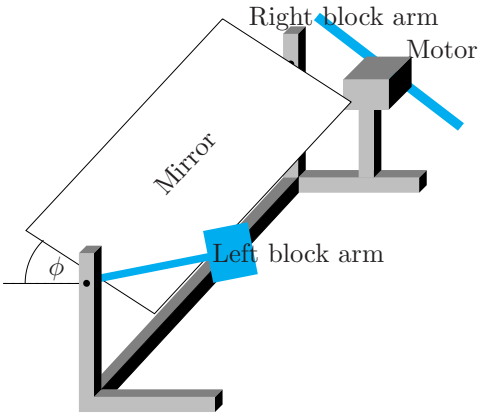

Figure 4.3: Rotating mirror mechanism diagram.

The laser scanner scans full 180° scans with the range data from side angles greater than 60° discarded. A PS2 mouse is used for feedback because it is readily available and easily interfaces with the robots which all have PS2 inputs designed for connecting a mouse and keyboard for maintenance purposes. A particular advantage of this approach is ease of deployment in that the rotating mirror mechanism can simply be placed in front of the laser scanner, PS2 mouse cable connected and no further connections are required, even the power supply is self contained in the form of a single AA battery.

The mirror is mounted in front of the scanner rather than rotating or nodding the scanner itself. This has a number of advantages namely less power consumption and simpler hardware. The disadvantages are a reduced scan angle to around 120°, a blind spot when the mirror is edge on and a requirement for more complex geometric calculations (Figure 4.4) to correct for the separation between the mirror and the scanner. In Figure 4.4 the effect of the mirror is to bend part of the xy-coordinate plane to a new elevation.

The mechanism that does the rotating may be built from anything suitable as long as the gear backlash is kept to a minimum either by reducing the number of gears in the gear train or using higher quality closer meshing gears.

4.3.2 Data Processing Software

The SICK LMS 200 laser scanner is configured to scan a 180° arc at 1° intervals. These settings are chosen to minimize the number of scan points in order to increase the complete scan frequency. The scan frequency is limited by the data rate that at the moment is 0.05Hz. This data rate is 34.8kb/s, the maximum for the serial interface, resulting in 13 horizontal scans per second. The SICK LMS 200 can support a 500kb/s data rate using a USB interface. At this data rate, full 3D 1° by 1° scans are possible at 0.5Hz (Section 4.3.3).

The mirror is displaced forward from the origin of the laser scanner. This displacement alters the range readings by a value dependent on the angle of the mirror. The following equations, which reference the values indicated in Figure 4.4, indicate the conversion between the r, θ and ϕ coordinates, measured by the laser scanner, and 3D Cartesian coordinates.

$$\mathbf{x} = \begin{pmatrix} (r\cos\theta - d)\cos\phi + d \\ r\sin\theta \\ (r\cos\theta - d)\sin\phi \end{pmatrix} \tag{4.10}$$

where the value d is the separation between the origin of the laser scanner and the axis of the rotating mirror. The range and bearing as measured by the laser scanner are r and θ. The angle of the plane (Figure 4.4) to the horizontal introduced by reflecting the laser ray from the rotating mirror in front of the scanner is indicated by ϕ.

The frequency of angular position reports can be adjusted by altering the number of teeth on the blocking wheel and by changing the gearing. A high angular position update rate leads to better accuracy in that the system can respond faster to changes in angular velocity. However, missed counts do still occur occasionally.

4.3.3 Enhanced 3D Laser Scanner

The laser scanner described in Section 4.3 can be deployed to standard mobile robots without any hardware modifications. Although, very convenient some relatively minor

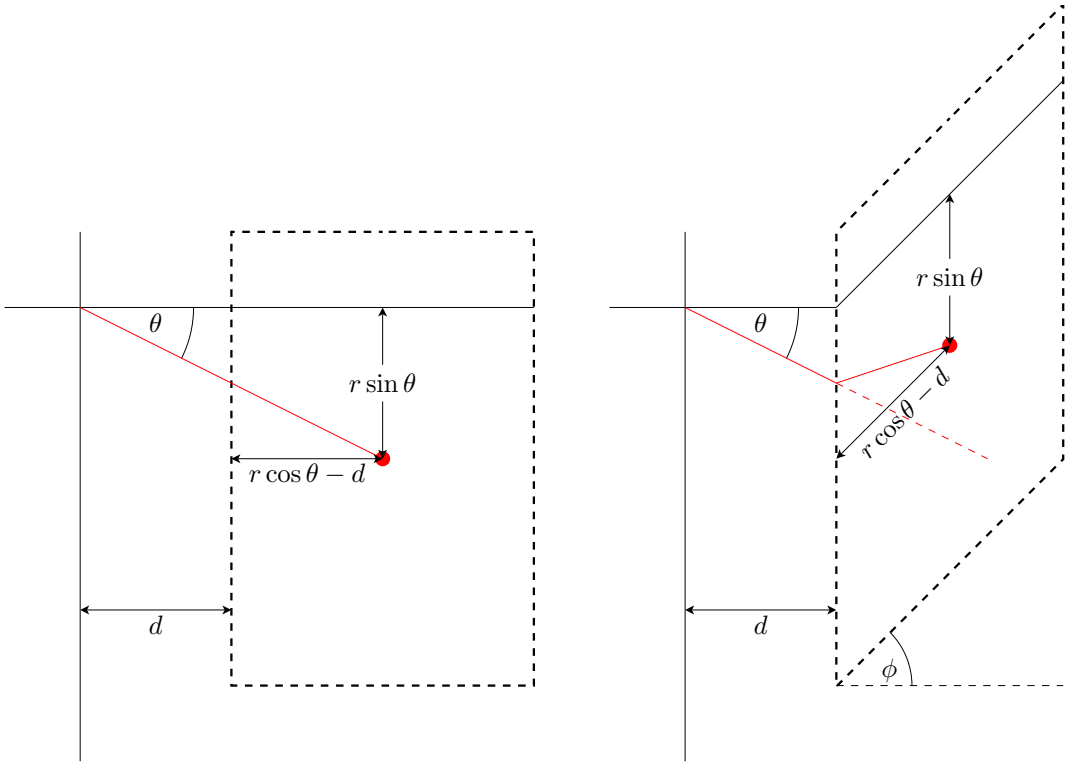

Figure 4.4: Calculating 3D coordinates from range data. The plane is elevated by an angle ϕ to the page. The line at θ to the x-axis shows the laser as it is deflected into the plane by the mirror. The range reading returned by the scanner is r and angle θ.

modifications to the hardware of the standard mobile robotic platform bring about substantial improvements. The enhanced 3D laser scanner is based on a SICK LMS 200 (Ryde and Hu, 2005b, 2006a,b) which consists of the LMS 200 facing upwards Figure 4.2 into a rotating mirror driven by a stepper motor controlled by a signal from a quartz crystal oscillator. This approach simplifies the hardware and software implementation whilst producing a 3D laser scanner suitable for real-time operation on medium sized (≈1m) mobile robots. The high update rate (75Hz) of the LMS 200 means the 3D scanner delivers scans at 1Hz with a horizontal resolution of 1° and vertical resolution of 5°. The vertical angular resolution can be controlled by the mirror rotation speed with lower speeds delivering better resolution at the expense of increased scan time.

The processing of data produced by the scanner is similar to that described in Section 4.3.2 except the resulting point cloud is rotated 90° about the y-axis and the feedback mechanism is substantially simpler due to the higher accuracy of the quartz

oscillator driven stepper motor.

Accurate determination of mirror angle is paramount to the accuracy of the 3D laser scanner. Prior approaches adopted closed loop control with a DC motor to drive the mirror through a suitable gearing ratio. There are a number of problems with this approach. The mirror has to be perfectly balanced when mounted on the spindles otherwise the rotation speed will vary throughout a complete revolution. This can be addressed by adding very frequent feedback of the mirror position and by ensuring the mirror is balanced. In practice the former closed-loop control adds significant implementation complexity and the later is difficult to achieve perfectly in practice and is likely to be a source of vulnerability in the system.

Finally the most difficult problem to address is that DC motors typically operate at high angular velocities and so they need to be geared down to the slow rotation speed required for the mirror. Thus a number of gears between the motor and the mirror introduces significant amounts of backlash. Backlash can be reduced by using higher quality gears that mesh better and by introducing rotation resistance, but in practice these efforts failed to eliminate backlash over a complete revolution or introduced rotation velocity variations of their own.

A better approach which solves these problems is to drive the mirror using a stepper motor directly coupled to the mirror spindle. Stepper motors can be operated over a range of speeds and are especially suited to low speed operation. Directly coupling the motor to the mirror eliminates backlash by removing the gear train that introduces it. As long as the stepper motor's maximum load torque is not exceeded its rotation speed is solely determined by the frequency of pulses it receives from the driver circuit. Variations in load torque do not affect its rotation speed and so closed loop control is no longer necessary and substantial hardware and software implementation complexity is removed.

The stepper motor does introduce potential problems however these are either easily surmounted or not significant in practice. The stepper motor introduces a requirement for slightly more complicated electronics versus the DC motor to drive the stepper

motor. This complexity is alleviated by having dedicated driver chips that provide the requisite waveforms for input to the stepper motor. The only input that the driver chip in this implementation requires is a clock signal which is simple to provide electronically. The accuracy and consistence of this clock signal determines the rate of rotation and in practice a clock-signal of sufficient quality can be provided by a standard 555 oscillator. It is worth noting that significant frequency drift of the 555 oscillator of the order of $0.05\%s^{-1}$ and sensitivity to temperature variations renders this approach too inaccurate for this research.

A better quality clock signal such as that provided by a quartz crystal oscillator produces superior scan results mainly because the mirror rotation frequency may be measured over many revolutions and the angular velocity calculated from the average.

It might be anticipated that the angular velocity inter-step profile would be too variable however in practice the angular momentum of the mirror coupled with non-rigidity in the physical connection between the stepper motor and the mirror means that this profile is visibly smoothed and has little impact on the accuracy of the data. More power is required in comparison to the DC motor however the mirror load is low so this is still insignificant when compared to power consumption of the 2D laser scanner. Finally an advantage of the stepper motor is that the speed of the motor may be more accurately controlled with a computer which has the potential to be exploited in future work.

The stepper motor has a 1.8° angular resolution that is half-stepped to give 400 steps per revolution. Although the stepper motor, when driven by a quartz crystal oscillator signal, has an exceedingly consistent angular velocity the mirror still needs to be approximately balanced. Significant off-centre mass distribution will cause flexing to occur between the motor connection and the mirror. This flexing causes significant distortion in the scan data which may be reduced by using stiffer materials and balancing the mirror to reduce the variation in load torque. Generally, the system is much less sensitive to variations in load torque than the DC motor driven system. A block diagram of the electronic components in Figure 4.5 illustrates how a 12MHz

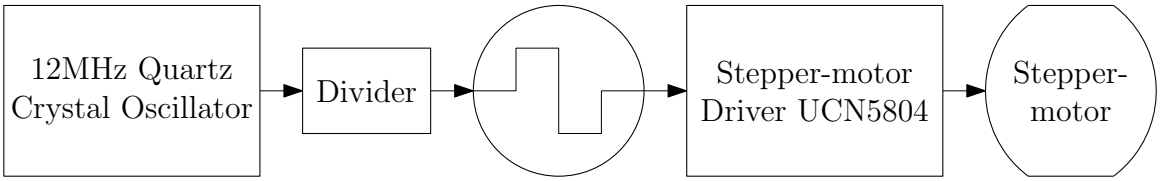

Figure 4.5: Block diagram illustrating the electronic components of the 3D laser scanner.

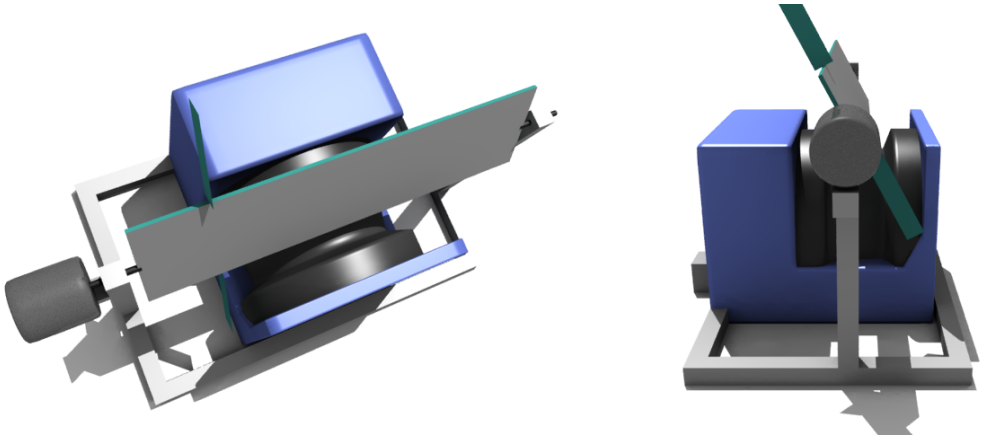

Figure 4.6: Top and side views of the enhanced 3D laser scanner.

oscillator signal is stepped down to produce a highly stable and accurate 15ms clock signal. The clock signal drives the stepper-motor driver microchip which outputs the correct waveform for the stepper-motor. This configuration coupled with an adequately balanced rotating mirror mechanism results in a very consistent rotation speed over an extended period of time. The field of view is improved by placing the laser scanner facing upwards. In this way the blind spot when the mirror is edge on is rotated to point upwards rather than in front of the robot. Most mobile robots need to see in front rather than above. The side blind lobes are unaltered and the observable volume is given by the following constraints.

$$0.5 < r < 8, 25 < \theta < 155, 75 < \phi < 105, \tag{4.11}$$

Data from a single scan of a small room is plotted in Figure 4.11. The room is a cuboid and with flat ceiling, walls and floor. The scanner records ranges from almost all directions around the robot. The quality of the data returned can be visually verified by inspection of the straight edges of the room in Figure 4.11.

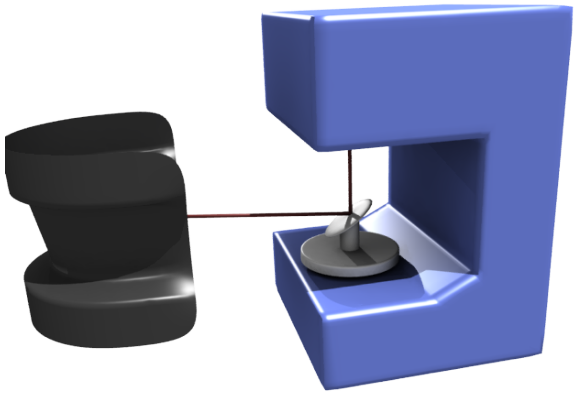

Figure 4.7: View of the SICK LMS 200 showing internal rotating mirror.

The SICK LMS 200 is capable of operating at higher data rates than can be delivered by a standard serial link. The internal mirror, Figure 4.7, rotates at 75Hz and this is the natural data frequency if the RS232 serial communications bottle-neck is removed. The SICK LMS supports a high-speed serial connection RS422 which may be connected to a standard USB port with a USB-RS422 converter. This allows 180° of range data at 1° resolution to be delivered to the host computer at 75Hz. The horizontal resolution is fixed at 1° however the vertical resolution may be adjusted by varying the speed of the external mirror rotation. Matching the vertical resolution to the horizontal resolution gives scan times of $360/75 = 4.8$ seconds. The resulting scans are high detail and can be used for map building, however for real-time obstacle avoidance a faster mirror speed is more suitable and scan frequencies of 1Hz are feasible.

Figure 4.6 indicates the positioning of the blocking arms. By blocking the -90° range side scan of the 2D scanner every half mirror rotation these arms serve two main purposes. The first is to establish the angular velocity of the mirror, although this is very consistent with a crystal oscillator, actually measuring the angular velocity makes the system more flexible and aids fault detection. The second is to establish the rotation of the scan about the y-axis so that the resulting 3D scan is correctly orientated. This does not have to be too accurate as this is affected by the inclination of the surface that the robot is traversing. In this manner the system does have an element of closed-loop feedback once per cycle but most of the fine control is an open-

loop system thus significantly reducing demands on the host hardware. If the reflecting surface used is a standard single-sided mirror then only one blocking arm should be attached otherwise there will be a 180° ambiguity in ϕ. With a double-sided mirror and two blocking arms two 3D scans are returned for every mirror revolution.

4.4 3D Sensor Model

A sensor model is required to make predictions about what data would be returned by the scanner at various poses within a map. The sensor model employed in this research is now described.

The 3D scanner returns a list of coordinates corresponding to all the range returns detected. The number of elements in this list varies because laser measurements may not be returned at all for some angles.

An approach similar to Rocha et al. (2005) where the voxels traversed by a vector is considered. This is often referred to as the ray trace sensor model and is particularly suited to laser range data even though Rocha et al. (2005) employ it for stereo vision range data, due to the collimated nature of the laser and because the beam width, Figure 4.9, is less than the map resolution.

There are a number of deficiencies, one of which is described in Section 4.4.1, with this simplified algorithm but its persistence in mobile robots is due to its simplicity of implementation and low computational requirement. The premise is that the occupancy of the terminating voxel is increased and the occupancies of the voxels on the line between the origin of the laser ray and the terminating voxel are decreased. This process is summarised in Algorithm 2.

In this way, given a 3D map, pose and sensor model the expected occupancy of voxels surrounding the pose can be calculated.

The accuracy of the sensor model is dependent upon the errors in the data returned by the 3D laser scanner. The errors of the 3D laser range finder may be modelled as Gaussian in θ, ϕ and r. Complete differentiation of (4.10) produces the following

Algorithm 2 Ray trace sensor model.

for $\mathbf{x} \in \mathbf{X}$ **do**
  Assign current $\mathbf{x}$ to voxel
  if voxel is present in occupancy list **then**
    increment terminating voxel's occupied count
  else
    add voxel to occupancy list
  end if
  determine number of steps along vector as $\frac{\mathbf{r}}{\epsilon}$
  for all steps **do**
    decrement occupied count of voxel at position $\frac{s}{n\mathbf{x}}$
    if occupied count of current voxel is ≤ 0 **then**
      remove voxel from occupancy list
    end if
  end for
end for

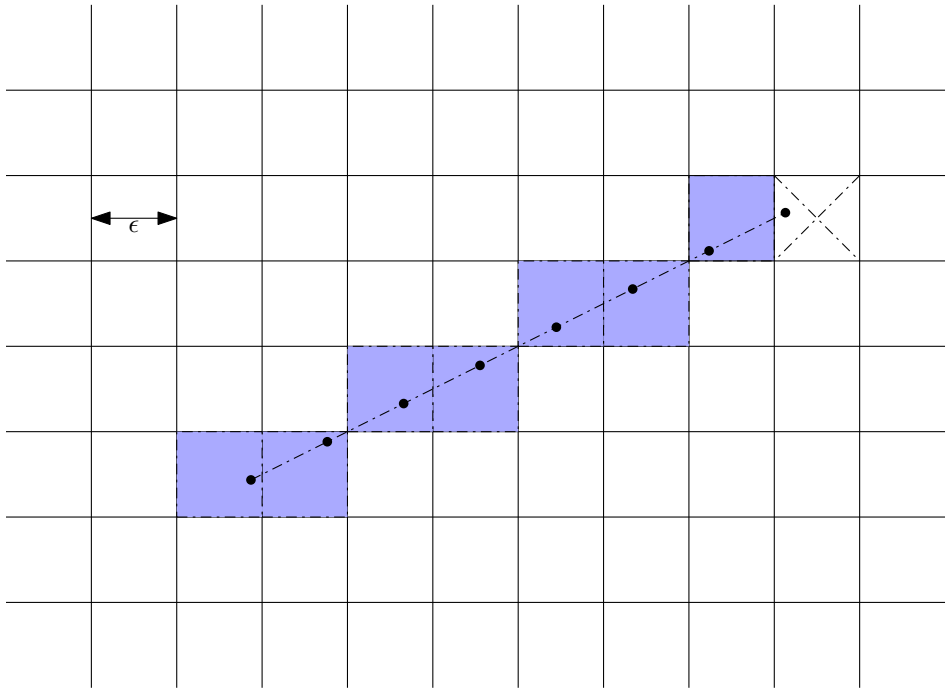

Figure 4.8: Diagram illustrating the ray trace sensor model described in Algorithm 2.

expressions for the errors in x, y and z.

$$\Delta x = \cos\phi\cos\theta\Delta r - r\cos\phi\sin\theta\Delta\theta - \sin\phi(r\cos\theta - d)\Delta\phi + (1-\cos\phi)\Delta d \quad (4.12)$$

$$\Delta y = \Delta r\sin\theta + r\cos\theta\Delta\theta \quad (4.13)$$

$$\Delta z = \cos\theta\sin\phi\Delta r - r\sin\theta\sin\phi\Delta\theta - \sin\phi\Delta d + r\cos\theta\cos\phi\Delta\phi - d\cos\phi\Delta\phi \quad (4.14)$$

Converting from the spherical polar coordinate to the Cartesian coordinate system produces errors that are linearly dependent on the distance from the origin. This is superior to the quadratic dependence (4.6) found in triangulation techniques. This spread in uncertainty is useful for inclusion in the sensor model. The unmodified 2D scanner is capable of reading ranges up to 80m. Such large distances are infrequently observed in indoor environments however and as the error scales linearly with range the accuracy of these long range readings is poor. To ease the complexity of the sensor model the readings are truncated to 8m and the upper bound of the error, namely that at 8m, is assumed for all remaining range readings. This error is then intrinsically incorporated into the sensor model by limiting the map resolution to this value.

4.4.1 Pixel Mixing

Pixel mixing is an insidious problem that arises from the width of the laser beam. The laser scanner operates at a wavelength of 880nm in the infrared part of the electromagnetic spectrum and is consequently invisible to the human eye. Figure 4.9 contains images of the laser scanner taken with an infrared sensitive camera. In order to emphasise the laser beam, two images of the scanner, off and operating, are shown for comparison. The diameter of the infrared beam is around 0.01m as seen in Figure 4.9 where the beam is incident upon the wooden block. Fig 4.9 indicates the height at which the centre of the rotating mirror needs to be placed. The beam width places a limit on the resolution of the map and plays a role in the selection of the ray trace sensor model. Due to the 0.01m beam diameter it is clearly not worth mapping to a better resolution. Indeed, experimental data supports this with range errors of the

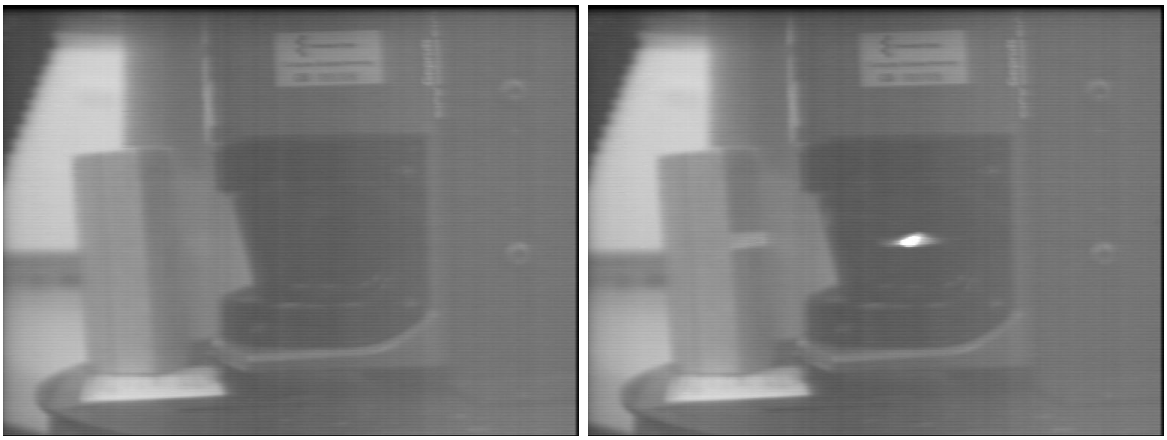

Figure 4.9: Infrared image of the SICK LMS 200 laser scanner in operation showing the width and position of the laser beam.

order of 0.02m under standard conditions, even though there is a mm range resolution setting.

At vantage points with large range discontinuities the beam falls partly on the foreground and background objects, Figure 4.10. An average of the range between the two is returned. There exist a number of attempts to improve the data by filters which recognise readings that are likely to involve pixel mixing and correct or remove the offending data. Occupancy grids should be robust to the problem of pixel mixing because the errors produced are usually random with respect to position as seen in Figure 4.10. The independence of the error upon observer position means as the map is updated from various poses truly occupied voxels are repeatedly observed. Voxels registered as a consequence of pixel mixing errors will not be repeatedly observed from other poses and will therefore have a lower occupied count. After many poses the map may be filtered to remove the low occupancy voxels.

In practice however these errors may persist for some time before they are removed. It is more efficient to remove these errors in the preliminary stages of data acquisition. In 3D occupancy grids pixel mixing tends to produce isolated occupied cells where there are discontinuities in the range readings with respect to θ and ϕ. Successful filtering to remove pixel mixing relies on the detection of these range discontinuities and corresponding isolated voxels.

For the 3D scanner pixel mixing problems are more acute for vertical range discon-

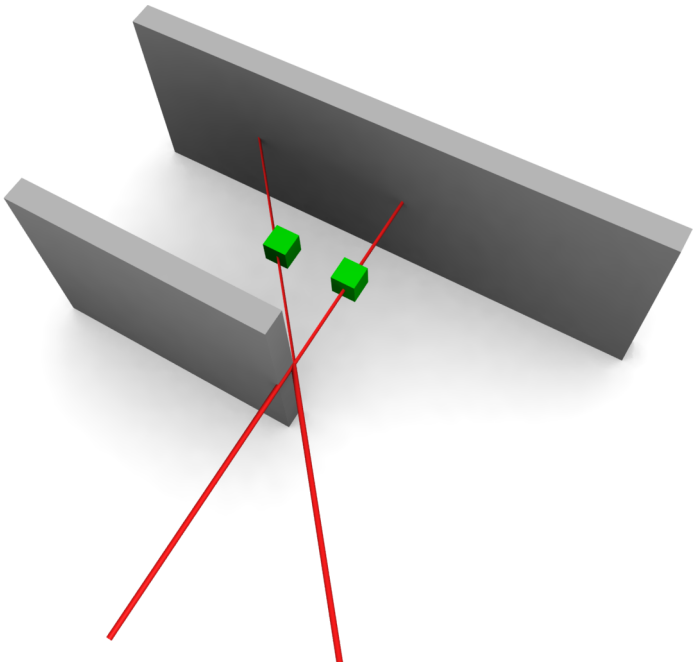

Figure 4.10: Pixel mixing explanation and illustration of observer position dependence.

tinuities. The laser scanner is designed to accommodate and correct such horizontal errors because that is the direction the laser beam is rotated by the 2D laser scanner. The vertical motion arising from a external mechanism was not anticipated by the manufactures and consequently no error correction is applied vertically by the internal electronics.

4.5 Experimental Results and Analysis

The errors in data produced by the 3D scanner are measured in a number of manners. An estimate of the errors is generated by scanning a room with ceiling, walls and floor. Systematic errors are clearly distinguished as curvature in what should be planar surfaces. Once the points are coplanar then the random error is assessed in the following manner. A subset of the data points associated with the ceiling which is 3.2m above the origin of the 3D laser scanner is extracted from the scan. The standard deviation of the z coordinates for this subset gives a good estimate of the error in z and this value. Clearly, the error may vary with z and a more thorough method is explained later. This extract indicates that the error in z is approximately 0.02m. Repeating this

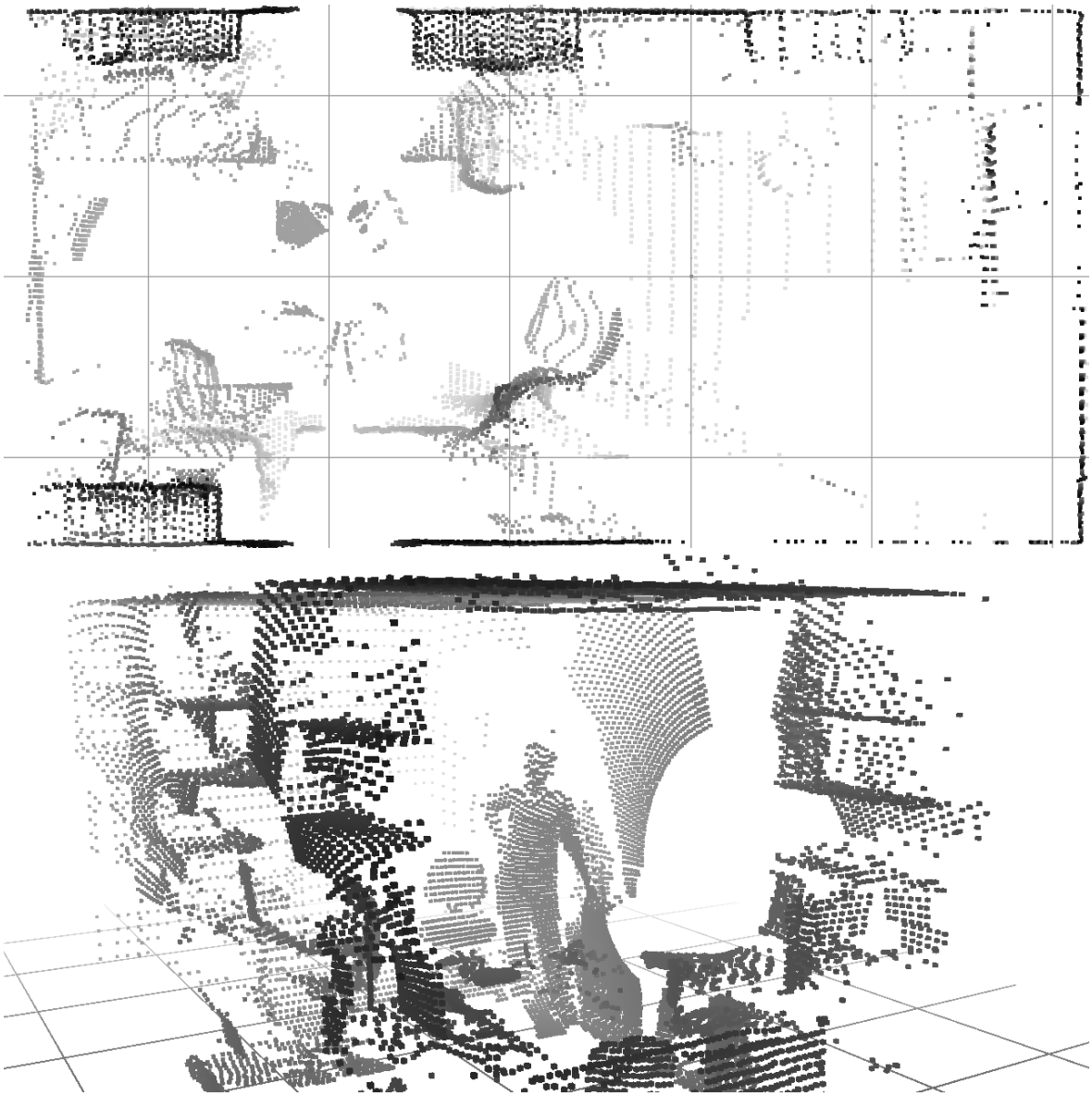

Figure 4.11: Top and side views of a single scan produced by the enhanced laser scanner for a room of dimensions 3 by 6 by 2.75m at 0.02m resolution. The points are coloured by distance from the point of view and the grid cells are 1m.

analysis for the smooth walls of the room indicate a similar error in x and y. Repeated measurements at different positions reveal that the error is relatively independent of distance over the range 0-8m implying that a simplified sensor model with constant error is appropriate. In this simplified sensor model the errors are 0.02m regardless of distance. This also places a lower bound on the map resolution.

To establish the size of the systematic error approximately ten dimensions of the experimental environment are compared to the generated map and the mean error is $0.07m$ on dimensions of $6m$. The systematic error results would seem to indicate that there is no point using map resolutions better than 0.07m and experimentally find the best localisation and mapping performance with resolutions of 0.08m, however it is worth remembering that local regions of the scans are accurate to 0.02m. For the enhanced laser scanner driven by a stepper motor the systematic error drops to $0.05m$. Example data returned from the enhanced 3D scanner is rendered in Figure 4.11.

A more thorough estimate of the random error is available by taking repeated scans at the same pose within static surroundings. Ideally the scans should be identical because nothing in the environment has changed between the scan times. The differences between the two scans are due to random error as any systematic distortions will be equally present in both scans. For each point in the first scan the distance to the closest point in the second scan is recorded. The distribution of these corresponding closest point distances is graphed in Figure 4.12.

Although the mean closest point separation in Figure 4.12 is 0.011m the median is 0.006m and the mode is 0.004m. The random error is remarkably small with very few separations venturing above 0.02m. Thus it may be concluded that the majority of the error comes from systematic errors or distortions within the laser scan. In this case flexing between the stepper motor axle and the mirror. For mobile robots the data accuracy is significantly better than that acquired by stereoscopic vision and comparable to those approaches that rotate or nod the entire 2D scanner. For medium sized robots in standard human environments this 0.02m accuracy is adequate and paves the way for equally accurate mapping and localisation.

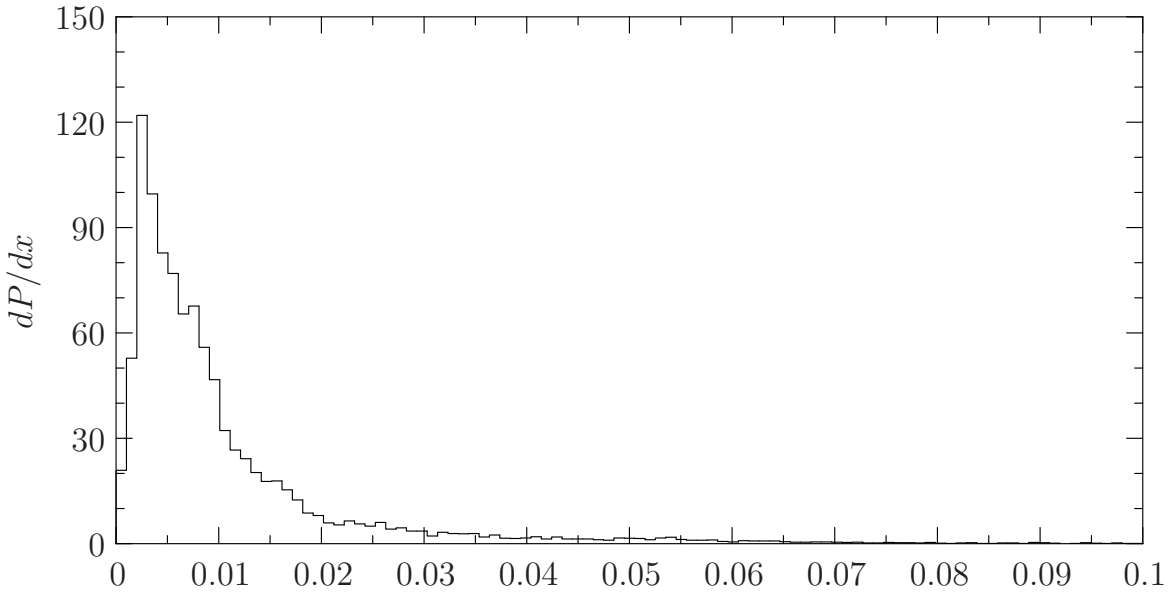

Figure 4.12: Distribution of the distances(m) to the corresponding closest points in two sequential scans.

4.5.1 Visualisation of 3D Map Data

Results are visualised using a number of methods and software programs. To visualise the data a 3D scene is generated consisting of cubes or spheres representing occupancy and the results rendered using a ray tracer, Figure 4.13. This allows the production of high quality images and videos with shadowing and perspective, visual cues that allow the brain to extract the 3D information more effectively from the 2D display media. Display of these images in black and white is greatly enhanced by using a depth map. In this representation the grey scale value of an occupied voxel is determined by its distance from the observer. This allows an object in front of a wall to be easily discerned by the viewer as it is a different colour, Figure 5.6. Note the addition of a 1m grid on the floor plane, Figure 5.6, aids the depth perception. An indication of the accuracy of the laser scanner is displayed in Figure 4.13 where the render of a 3D scan is placed with a photograph taken from the same view point. Each sphere in Figure 4.13 represents a laser range return and is coloured by its distance from the view point, with those closer being lighter. Example depth maps from different vantage points are contained in Figure 5.6. Real-time visualisation of the results as point clouds (50,000 points) is perfectly feasible on low end graphics cards.

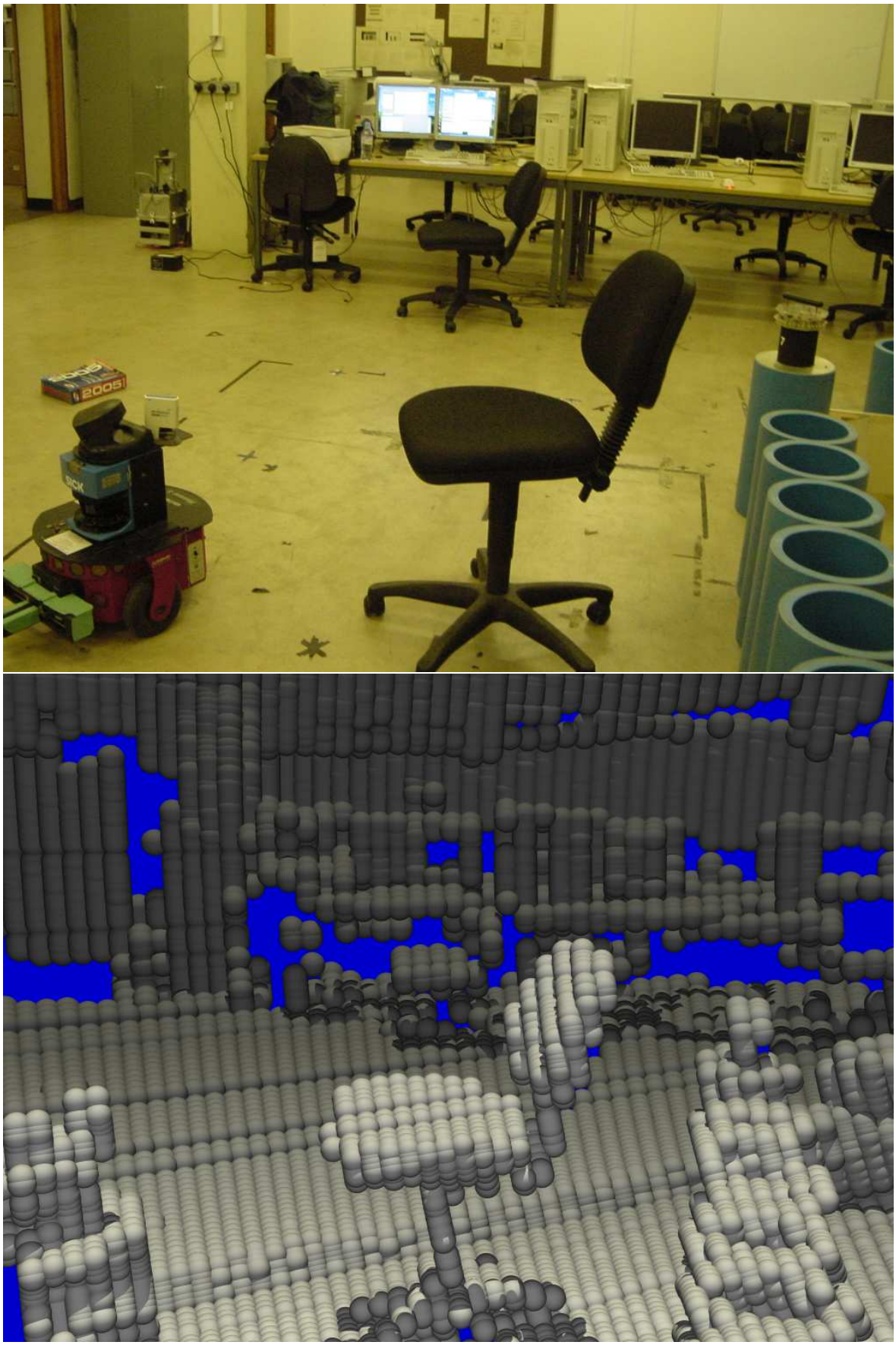

Figure 4.13: Photograph and 3D scan of mapped room from the same point of view.

4.6 Summary

This chapter describes the enhancements made to a standard SICK LMS 200 laser scanner to produce a full 3D sensor with low power consumption (20W), minimal hardware requirements and that may be easily deployed on standard Pioneer 2 and other robotic platforms with little or no modification. This is achieved in the first instance by mounting a rotating mirror above the 2D scanner to divert the 2D scans to gather range data from a volume rather than a single scan plane. Knowledge of the mirror's angular velocity and position allows this range data to be converted into 3D point clouds that contain information regarding the obstacles surrounding the robot.

Having shown that this approach is viable the 3D scanner is improved in a number of manners. First by using a USB-RS422 converter to increase the data rate from the 2D laser scanner to 75Hz. Orientating the 2D scanner to point up into the mirror delivers a wider field of view with the blind spots placed above and to the sides of the robot, rather than in front of the robot. The mirror is driven in an open control loop manner with a stepper motor clocked by a quartz crystal oscillator clock signal. The addition of the external rotating mirror has a negligible impact on the accuracy of the 2D laser scanner over the 0 to 8m range resulting in an error of 0.02m. In this way almost omni-directional 3D scans consisting of up to 5000 data points are delivered to the robot's computer at a configurable 1Hz.

The open-loop control nature of the 3D scanner means that no input from the host computer is required for the scanner to run. A useful extension would be to allow the host computer to control the angular velocity of the mirror. When the robot needs to acquire accurate data for mapping the mirror would spin slowly giving a higher density of data. However when moving, increasing the rotation speed of the mirror raises the update rate and allows obstacle detection, albeit with lower vertical resolution.

Overall the enhanced laser scanner driven by a stepper motor yields superior performance with a simpler implementation. Placing the laser scanner on its back facing upwards, into the mirror above it, enables the robot to perceive its surroundings with greater coverage.

Equipped with a sensor capable of returning 3D information of the robot's surroundings, Chapter 5 and Chapter 6 delve into the representation and processing of this data in a localisation and mapping context. Repeated scanning from a fixed position may be used to determine the short term dynamic profile of an environment. If there are people moving these will be detected as differences between successive scans and removed.

Chapter 5

3D Mapping Algorithm

This chapter introduces a multi-resolution sampling based algorithm for aligning generated range data stored in occupancy lists. Occupancy lists are voxel based data structures that efficiently represent 3D scan and map information. The process described in this chapter can align a sequence of scans to produce maps and localise within a prior global map. A 200 square metre office environment is mapped in 3D at 0.02m resolution, resulting in a 200,000 voxel occupancy list. Global localisation within this entire map with no prior pose estimate is done in 5 seconds on a 2GHz processor.

5.1 Introduction

Chapter 4 introduced the 3D laser scanner which delivers high density 3D data point clouds. The results show that these individual scans are not only of sufficient quality for local obstacle avoidance and path planning but, as will be shown in this chapter, are also suitable for 3D map building. This chapter is focused on extending the work of Chapter 4 and merging multiple 3D scans taken at different observation points. Also investigated is the best way to express the information in the combined scans in a succinct and tractable form. Throughout this chapter only single robot data is considered unlike Chapter 6 which considers multiple 3D sensing robots operating concertedly.

The chapter starts with Section 5.2 which inspects the research that others have undertaken to digitally represent environments for mobile robots. Section 5.3 contains the methods for storing and merging 3D data. This includes the representation employed in this research termed occupancy lists and an in depth description of the multi-resolution pseudo particle filter solution for aligning multiple occupancy lists within real-time constraints. Also presented is a theoretical analysis of the significance of overlap value returned by the scan matching algorithm. Section 5.4 includes results of single robot mapping experiments and for scan matching at various resolutions to test the theoretical work.

5.2 Related Work

Robust localisation is clearly necessary before any semblance of mapping can be accomplished. Once the localisation is accurate, the next challenge is to build maps. With mapping additional questions arise such as, what is the best internal representation or map format? How may new sensor data be incorporated into the map?

Powerful ideas from information theory exhibit how the value of estimators can be extracted from unknown probability distributions without complete knowledge of these distributions. Moreover, these sample based approaches converge to the true values surprisingly rapidly. Particle filters are such an approach and describe distributions by samples or particles as in Fox and Burgard (1999); Fox (2003); Fox et al. (1999); Lenser and Veloso (2000); Austin et al. (2000); Bergman et al. (2002). Particle filters track the robot pose by representing it as a set of poses with each pose in the set assigned a weighting. This weighted sample of poses tend to the probability distribution function of the pose with increasing number of particles. The particle filter evolves the pose distribution over time by updating the poses of the individual particles according to a prediction model. Information from observations is then used to correct the weights of the new particle set.

The computational overhead required by particle filters is orders of magnitude less than that required by grid based endeavours as they concentrate calculations on sig-

nificant regions of the state space.

In the robotics navigation field much emphasis has been placed upon the idea that localising and mapping have to be performed together. Advocates of simultaneous localisation and mapping (SLAM) or concurrent localisation and mapping (CML) use sophisticated probabilistic analysis to derive maps from uncertain location information.

Efforts to separate the localisation and mapping process (Wang et al., 2007) go some way to alleviate the computational burden of conventional SLAM. This becomes especially important when large numbers of landmarks are considered.

The assumption that they have to be done together makes developing mapping algorithms hard. It seems inadvisable to have a robot that is uncertain of its position build a map. The decoupling of localisation and mapping, if possible, should be sought. For once localisation can be achieved without using a map then mapping becomes a less complex and more reliable process.

Difficulties localising using natural landmarks have resulted in a resurgence in the use of artificial landmarks for localisation (Hu and Gu, 2000). Success using natural landmarks such as corners has been achieved by Wang et al. (2002). Although the approaches are varied there is a consensus for the necessity of fast, accurate and robust navigation and mapping approaches that would work in most environments. The focus tends to be for indoor environments because they cannot receive the GPS signal that outside robots can avail themselves. Research into SLAM for external robots is still necessary because GPS does not produce the accuracy or the ubiquity required for many outdoor robots. Indeed the GPS signal may be completely blocked in urban environments and those with heavy foliage.

5.2.1 Environment Representations

The question of how best to represent the surroundings of a robot in a manner which is easily stored, manipulated and processed has received much attention. A number of approaches have been proposed. Two of the main paradigms are geometric and topological representations. In this work the focus is geometric mapping and so now

work relating to this paradigm is considered.

Upon first inspection of the laser range data returned by the robot one is tempted to apply geometric models to represent the environment as in Burschka and Hager (2000) who develop a computationally efficient algorithm for determining the rotation and translation between consequent laser scans and other approaches including Wetzler et al. (1994); González-Baños and Latombe (2002). Map modelling with geometric primitives tends to be difficult to implement and rather brittle especially in environments containing features inadequately described by the set of primitives. It has become quickly apparent that probabilistic approaches garner the best experimental results. Moreover these solutions seem eminently scalable to problems involving data from different sensors and robots.

Extensions to straight line representations recruit further geometric primitives such as circles and conic sections (Vandorpe et al., 1996; Lee, 1996) to help describe a wider range of environments. The resulting maps require little storage, describe the environment well and are fast to process for path planning. However, they appear not to be as popular as occupancy grids or sampling based approaches. They are difficult to update, express probabilities poorly and are more complicated to implement.

Grid based approaches to 2D mapping involve the use of occupancy maps, first envisaged by Elfes (1987); Moravec (1988). They are powerful and capable of handling a large position uncertainty; this performance requires the storage and modification of an excessively large state space containing all possible positions and orientations of the robot throughout the environment. It seems that this full representation for the location probability distribution is not always necessary (Fox, 1998). Their ease of implementation has encouraged many researchers (Burgard et al., 1996; Fox et al., 1998a; Konolige and Gutmann, 1999; Olson, 2000) to adopt them.

Pfaff et al. (2007) introduces a modified $2\frac{1}{2}$-dimensional representation for exterior environments that is a good compromise between computational requirements and environment expression. Unfortunately this is not suitable for indoor environments which can be much more convoluted with many horizontal overlapping surfaces. These

cannot be adequately expressed with an elevation map.

Quantitative measures of map quality and localisation performance are notably absent from much of the literature as observed by Lee (1996). Some of the best efforts (Elfes, 1991; Lim and Cho, 1992) assess the map quality using an average entropy measure. Mobile robot navigation has evolved in a number of manners over the past couple of decades. Initial work with single robot SLAM has moved onto investigating multiple robots operating in a cooperative manner. Improvements in the sensors deployed have encouraged some researchers to embark upon 3D sensing using cameras or occasionally (Thrun and Montemerlo, 2006) 3D laser range finders.

Another method for determining the quality of the map is the information content as measured by the entropy of the map, (6.1) as described in Lee (1996); Rocha et al. (2005). The main disadvantage of this metric is that it depends upon the dimensions of the probability array. Increase the size of the probability array and there is an increase in unmapped volume which with its correspondingly high entropy thus increases the map entropy measure.

Geometric mapping algorithms attempt to survey the environment producing detailed measurements. If the environment consists of many planes (most indoor environments) then the straight lines appearing in 2D range scans encourage the use of polygonal maps (González-Baños and Latombe, 2002). Polygonal maps are good for path planning algorithms due to their inherent simplicity and extensive analysis in computer graphics.

Unfortunately not all environments contain straight lines; cluttered environments such as disaster scenes may not have enough straight lines for effective mapping and navigation. These are more suited to grid based approaches which can handle almost all environments as they make no assumptions about the prevalent geometry.

Techniques involving probability grids have enjoyed much success over the years taking advantage of the relentless growth in computer processing power. Some of the pitfalls of these grid techniques, also referred to as probabilistic grid maps, are detailed by Lee (1996). The most serious of which, are concerns over the probability meaning

assigned to each cell and the assumption that the occupancy probability of the cell is independent of other cells, especially its neighbours. A problem with cell probability updates arises from the situation when the robot is stationary and continues to update the map. Cells that are detected as occupied or free continue to accumulate probability weight without bound until the robot moves or the map updates cease. This makes the accuracy of the map particularly vulnerable to the systematic errors inherent in sonar sensors such as specular reflection. Although this effect is mitigated by using laser scanners the cell occupancy independence assumption is still detrimental. One apparently valid concern is that the accuracy in the pose estimate is fundamentally limited by the resolution of the grid; this is dispelled by Olson (1999) who achieves pose accuracy higher than the resolution of the occupancy grid by interpolation of a fitted parameterised surface in the region of the maximal pose likelihood.

Environmental representations using topology collapse the states representing robot position into a set of discrete, easily identifiable locations or paths within the arena of operations. The obvious advantage of this is the vast reduction in computational resources required to store and manipulate the position probability distribution. The process of path planning is also greatly simplified as the resulting networked nodes are fast to search. It should be noted that this solution requires easily identifiable features in the environment and the algorithms developed are situation specific. Implementation examples include Kuipers and Beeson (2002); Choset and Nagatani (2001) and a Voronoï based approach (Rives et al., 2003).

5.3 Method

The aim of this research is to produce three dimensional representations of indoor environments. There needs to be a way of knowing whether this has been achieved and to what extent. This area of map quality metrics tends to be ignored or considered superficially with most authors content to produce side by side comparisons of photographs of the environment and the resulting maps. These are good as a quick visual aid to the casual observer but a more quantitative, rigorous and repeatable method is

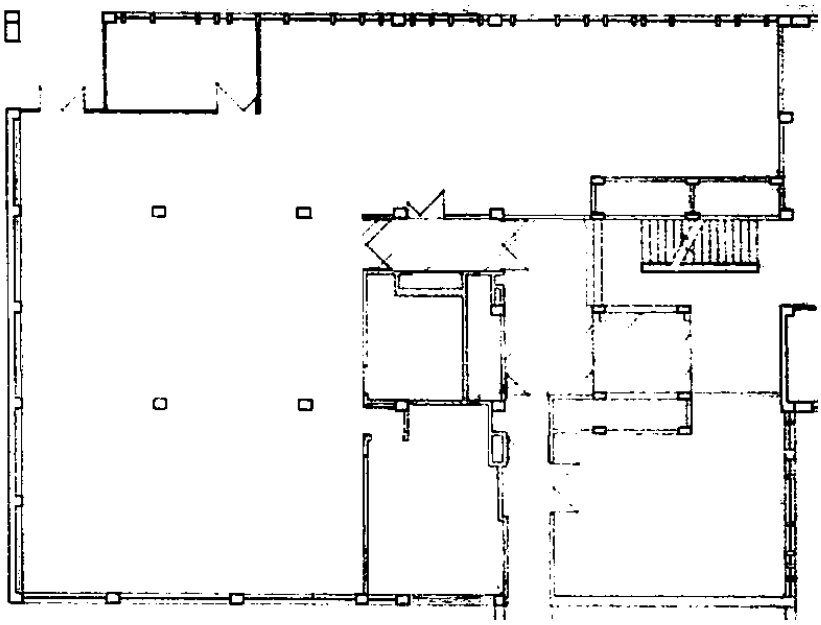

Figure 5.1: Schematic of the Brooker laboratory and surrounding area

required.

Figure 5.1 contains a schematic for the main area where some of the mapping experiments are undertaken. This can be used as an alternative check for the ground truth.

In results from experiments there are two categories of errors, systematic and random. The random error in these maps may be assessed by repeated mapping of a static environment. As long as the maps are complete then the multiple maps of the same unchanging environment should be similar and any differences would be due to errors in the mapping process. A comparison of the resulting maps will reveal the size and distribution of the random errors in the results. The systematic errors are harder to determine but mainly involve the acquisition of the true dimensions of the environment.

In this work the experimental results are tested in two manners. The first is by comparing the results against maps built by manually aligning the 3D scans. For the second various dimensions of these maps are verified against the corresponding dimensions in the environment.

5.3.1 Occupancy Lists

Ideally the map is a 3D probability density function and as a scalar field is a function of position $P(x, y, z)$ describing the probability density of occupancy at a particular position in space. Practically however, the map is stored as a discrete representation at sequential times with the kth map being

$$M_{abc} = \int_{z}^{z+\epsilon} \int_{y}^{y+\epsilon} \int_{x}^{x+\epsilon} P(x, y, z)\, dx\, dy\, dz \tag{5.1}$$

where

$$a = \left\lfloor \frac{x}{\epsilon} \right\rfloor, b = \left\lfloor \frac{y}{\epsilon} \right\rfloor, c = \left\lfloor \frac{z}{\epsilon} \right\rfloor. \tag{5.2}$$

However, the storage and processing requirements for such a representation become unwieldy even for moderately sized environments. Conventional 2D occupancy grids do not extend well into 3D in terms of their computational handling. For instance a standard room of 10m by 10m by 4m at a resolution of 0.05m requiring only a 40 thousand cell representation in 2D needs 3.2 million cells in 3D. This order of magnitudes jump has ramifications for both memory requirements and performance. Consequently, concessions are made to compress the map without an intolerable reduction in performance.

A number of observations of the nature of the 3D data typically gained from 3D mapping allows some important reductions in the size of the representation without significant loss of information, namely a lossy map compression mechanism.

The difference between the number of occupied and unoccupied cells in 3D for navigable environments is large with many more free cells than occupied. This relative rarity of occupied cells makes their information content higher. In others words, observations of occupied cells are more significant than those of free cells and consequently should be stored and manipulated in preference. The large majority of the unoccupied cells means they have low significance, especially, when it comes to matching sensor scans with prior maps. Their low significance and abundance makes them ideal candidates for removal without the loss of too much information. Unoccupied cells have been

historically maintained in occupancy grids because the grids were 2D and the sensors, often sonar, were very unreliable so every bit of information had to be maintained. Laser scanners are much more reliable and are almost guilty of producing too much data so the following lossy compression approach is feasible.

Instead of holding the cell probabilities in a 3D array all the cells that have been observed to be occupied more times than free are maintained as a list. Each entry in the list contains the voxel position and a count of the number of times that voxel has been observed to be occupied. If the occupancy of a particular voxel falls below zero, meaning that it has been observed as empty more times than occupied, it is removed from the list. This list is termed an occupancy list (Ryde and Hu, 2007) due to its list structure and similarity to conventional occupancy grids.

This additive map update is fast and is consistent with (3.28) if the map is regarded as storing a quantity proportional to the log odds occupancy. Not only is there considerable saving in storage and manipulation but there does not have to be strict bounds on the environment as is the case with a 3D array. Experimentally it has been found that the standard room considered above can be represented in this manner by 20 thousand cells at a resolution of 0.1m. The number of cells depends on the clutter of the room and grows with the number of observable surfaces but is most sensitive to the cell size.

To get an idea of the benefits delivered by this lossy compression consider a single laser scan with radius 10m and stored at a resolution of 0.02m which is the smallest sensible resolution as the laser scanner itself has an error of ± 0.02m. Assuming a typical room height of 3m the number of voxels required for this occupancy grid would be approximately 100 million. A typical scan stored to a 0.02m resolution occupancy list results in 5000 voxels or so. Thus storing only occupied cells results in a compression factor of 20,000. Secondly, the extra information stored in the occupancy grid may be easily generated from that stored in the occupancy lists for all scans by ray tracing. This is necessary for path planning algorithms that need to know the free space volume just above the floor. The information can be extracted from the map as long as the

scan poses are available. The computer iterates through nearby poses and according to the sensor model traces a ray from the pose to the nearest occupied voxel in each direction. The intervening voxels are recorded as free in a conventional 2D occupancy grid, upon which path planning may be performed.

In this instance the representation chosen is a hash table with the voxel's integer position as the index key. The value associated with each position index key is an integer count of the number of times the cell has been observed to be occupied, which as shown in Section 3.3.3 is indicative of the probability of occupancy under a number of assumptions.

The hash structure allows the probability at a given position to be looked up in constant time, on average regardless of list size, which is the most common operation on the map. Each point is quantised using (5.1) and mapped back to real space by

$$\mathbf{x} = (a\epsilon, b\epsilon, c\epsilon). \tag{5.3}$$

There is a loss of precision in this process however repeated alternate applications of (5.1) and its approximate inverse (5.3) leave the voxel coordinates unchanged. The integers a, b and c are hashed to produce the look up allowing the occupancy list to be stored in compact list representation yet still benefit from fast query times.

In this manner, the environment is represented as a dense feature set rather than an occupancy grid. Each feature is a voxel that is likely to produce a laser return. This representation scales much better with environment size. Data points are recorded for voxels containing surfaces in the environment and as these surfaces are 2D the number of voxels required to express them scales as the square of the environmental linear size. Occupancy lists are superior to conventional feature sets because no feature extraction is required. This bestows flexibility on occupancy lists enabling them to represent any environment without losing information as a consequence of featureless data.

5.3.2 Scan Matching Probability

New range scans need to be incorporated into the global map. One of the ways of finding the rotation and translation transformations necessary to merge scans, and consequently the change in pose between them, is scan matching. This is also employed for localisation to determine the pose of the robot in the environment by matching the current range data to positions in a previously generated map.

The ideal matching algorithm returns not only an estimate of the pose, but also a probability distribution function describing the pose differences between successive scans. This estimate of the pose is stored alongside the range data and the current map is the composition of these pose-scan pairs. Storing the map like this allows back propagating corrections in the light of new observations. Unfortunately this violates the real-time requirement because as more of these scans are taken, more data has to be processed. Ways of overcoming this include allowing scan data to decay over time or simply doing as much processing as possible during the available time.

As pointed out by Konolige and Gutmann (1999) when exploring areas already covered by the map care has to be taken when matching current poses with older ones and their associated range data. Whilst aligning a scan with the map it is important to determine the significance of a particular overlap value in order to assess the reliability of that localisation. Determination of the confidence in the alignment pose is required for robust map building; only those scans with high confidence value should be added to the map. The shape of the probability distribution is useful in establishing this significance. For instance if the distribution is multi-modal and the peaks are similar in height then there is doubt as to which modal peak is correct. Likewise the sharpness of the peaks indicates the accuracy of the corresponding modal poses. These measures are all relative; they say nothing about the absolute value of the overlap. This is addressed by the following analysis.

In order to determine significance of a given overlap the probability of getting an overlap at least that high by pure chance alone has to be considered. Let the volume covered by the map be M_v, the overlap o, the number of occupied voxels in the map

and scan is M and n respectively. The map and scan voxel sizes are both ϵ. The total number of possible voxels in the map volume is $V = M_v/\epsilon^3$. Thus the proportion of the map that is occupied is

$$\frac{M}{V} = \frac{M\epsilon^3}{M_v}. \tag{5.4}$$

Assuming a random distribution of occupied voxels throughout the map and that voxel occupancy is independent, (5.4) expresses the probability of a single voxel picked at random being occupied in the map. Given a scan of n voxels what is the expected value of the overlap assuming that the scan voxel occupancy probabilities are independent? This can be re-expressed in standard probability terminology as picking n voxels at random from the map without replacement. This means that the overlap counts may be modelled as a hypergeometric distribution. The mean of which is given by $\bar{o} = nM/V$, and the variance

$$\sigma^2 = \frac{n(M/V)(1 - M/V)(V - n)}{V - 1}. \tag{5.5}$$

It is important to note that for simplicity the hypergeometric distribution tends towards a normal distribution as $n/V \to 0$. The significance of the maximum overlap value may be determined by calculating its probability of occurrence using the difference of o in standard deviations, (5.5), from the mean. Figure 5.8 presents overlap distributions at a range of resolutions for a single scan matching.

5.3.3 Multi-Resolution Particle Filter

Global localisation and scan matching are achieved by an approach that is similar to a multi-resolution particle filter. In this work a particle is considered as a pose with an associated probability weight w. In the implementation these weights are not normalised, are integers and are considered proportional to the likelihood of the corresponding pose.

In Algorithm 3 the resampling step is akin to sampling importance resampling with a local uniform distribution. Typically in particle filters the initial distribution is randomly generated from a uniform distribution. In this implementation the initial pose

distribution is exhaustively generated for all possible initial poses. This is made computationally possible by the downsampling step and the overlap calculation algorithm which ensures the smoothing of the probability distribution function at low resolutions.

Algorithm 3 Multi-resolution sampling overview algorithm.

```
OccupiedList map, OccupiedList scan, Integer f
generate uniform pose distribution sampleSet
for f is 2^i where i is 6 to 0 decrement 1 do
    downsample map and local scan by f
    calculate overlaps for all poses in sampleSet
    remove from sampleSet all poses with certainty < 80% of the mode
    re-sample poses around remaining poses in sampleSet uniformly with width ε/2
spatially and Δθ/2
end for
```

The implementation of the particle set must satisfy two criteria. It must only store unique poses at the current spatial and angular resolutions and it must be able to iterate through the particle poses in angular order. The latter is not a requirement for success of the process but is an optimisation. When calculating the overlaps one of the slowest operations is the rotation step. By ordering the samples by orientation the rotation on the local scan may be done only once and then translated to poses at different positions with identical orientation. This is implemented by using a treeset structure which maintains the uniqueness and correct order of elements and guarantees insertion times of $O(\log n)$.

A typical particle distribution is visualised in Figure 5.2 where the local scan on the left is globally aligned with the map on the right. Both the local scan and the map are full 3D representations however they are shown as 2D projections in Figure 5.2 for clarity. The initial particle distribution is uniform with the spatial separation of 1.24m. The particles are resampled around the highest probability regions progressively sampling the pose space at better resolutions. The final alignment is processed with the local scan and map resolution of 0.02m. This resolution is chosen because it is comparable to the random error inherent in the 3D laser scanner.

The solution is indicated by the high cluster of poses shown in Figure 5.3. The algorithm is similar for global localisation as well as tracking with the only difference

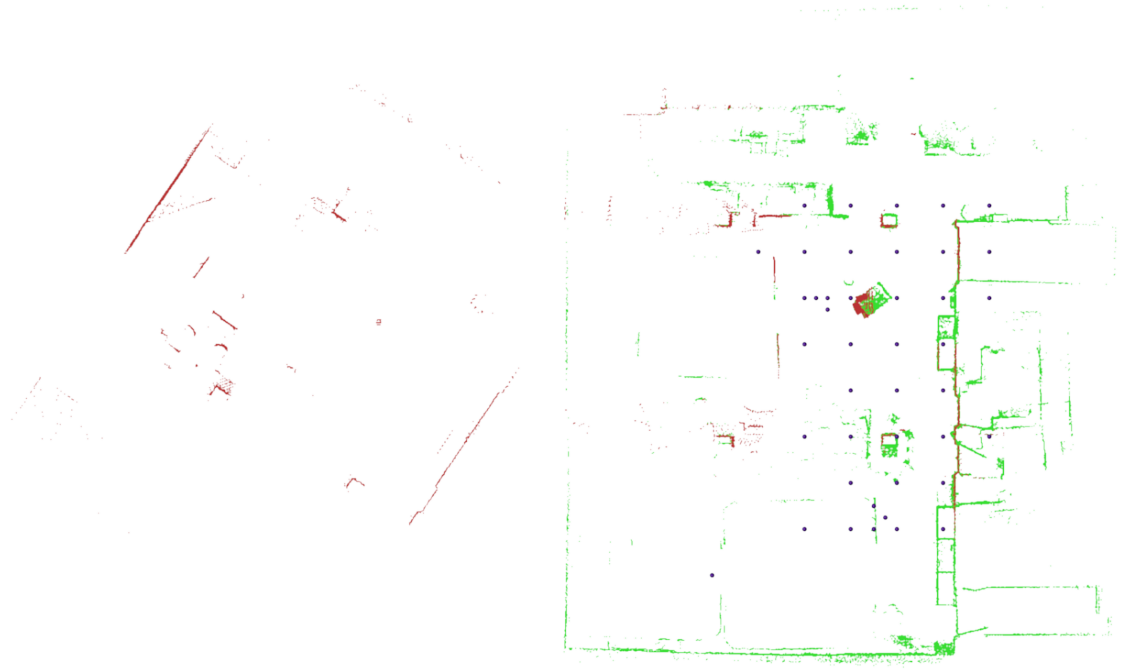

Figure 5.2: Overhead view of scan (left) and map with tested poses with $w > 0.8w_m$.

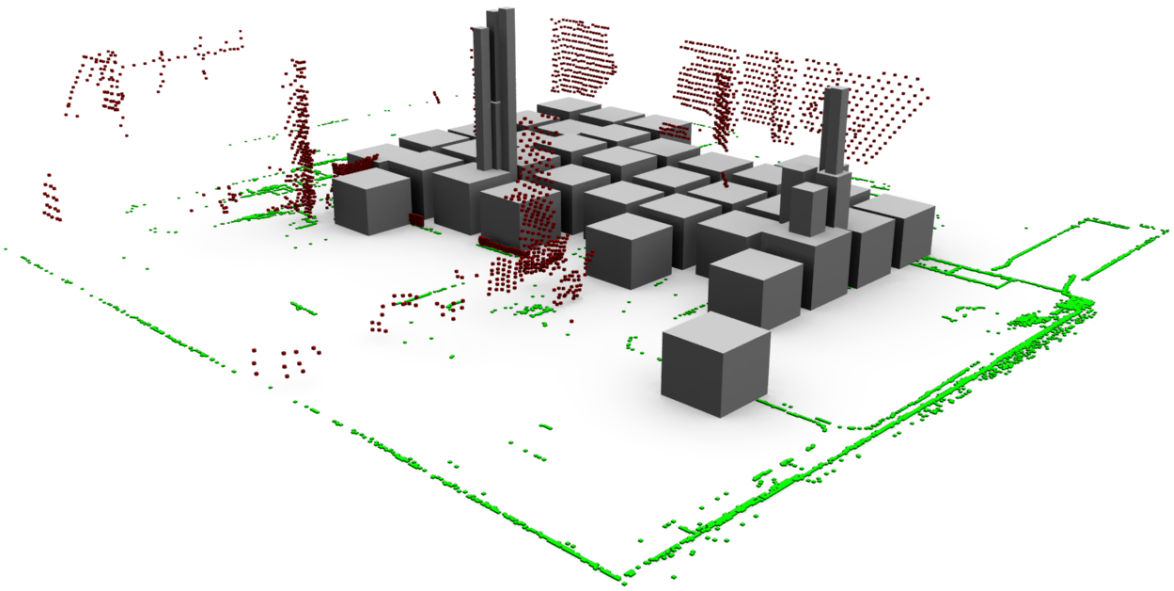

Figure 5.3: Distribution of best orientation poses as a function of 2D position with height indicating probability of pose.

being the first step of generating the initial particle distribution. For global localisation the particle distribution is over the entire possible pose space. The two main ways of generating the initial particle distribution are uniform and random. These two methods have their advantages and disadvantages and in other work the random initialisation is usually preferred. However for this algorithm it is important to use a uniform initial particle distribution with the spatial and angular resolution of the poses carefully chosen. In this way it is possible to be sure of complete coverage of the pose space. Normally this would be prohibitively costly however downsampling makes this possible. Other advantages of a uniform initial particle distribution are that it is deterministic and quick to generate.

The spatial Δx and angular $\Delta\theta$ resolutions need to be matched and are dependent on the resolution ϵ of the occupancy lists and the scan range, R. The spatial separation of the particles is dictated by the resolution of occupancy list being processed. As $\Delta x = \epsilon$ the angular separation is

$$\Delta\theta = \epsilon/R, \tag{5.6}$$

where R is the maximum range of points from the origin of the local map which need to be considered and ϵ is the voxel size. Rotations by angles exceeding (5.6) ensure that successive rotations produce new local occupancy lists at the current resolution.

Because the particles are stored in sets then the random initial particle distribution is still guaranteed to produce unique particles, therefore no computational power is wasted calculating the weighting for duplicate samples.

An important aspect of this process is the downsampling of the occupancy lists. The 3D laser scanner returns 3D scans as point clouds and these are converted into an occupancy list of resolution $\epsilon = 0.02$m. The resolution of 0.02m is chosen because this is the limit of the accuracy of the 3D laser scanner.

The downsampling process is described in Algorithm 4 and illustrated in Figure 5.4. In Figure 5.4 the point cloud and grid are represented in 2D for diagrammatic simplicity. In Algorithm 4 downsampling is done to many resolutions however only two resolution grids are shown in Figure 5.4 and Figure 5.5 for enhanced clarity. In Figure 5.4 the

crosses represent laser return points and two resolution grids are superimposed. The number of points within each cell of the corresponding grid is labelled, with numbers in larger font corresponding to the grid with the larger cells. A visualisation of the 3D downsampling of a typical scan is depicted in Figure 5.5.

Downsampling can take some time especially for large occupancy lists like the map, which normally contain substantially more voxels than the local scan. In some instances, for example repeated localisation within a map, downsampling need only be performed once on the map and then multiple localisations using that map do not have to recalculate it. This is especially useful if there are a number of scans to be merged into the map, in which case it is better to test them all and add the one with the highest overlap first.

Algorithm 4 Occupancy list downsample algorithm.

```
occupiedList L, integer factor f
new resolution ϵ = resolution L*f
if occupiedList resolution ϵ exists in cache then
    return corresponding downSampled occupiedList
end if
new occupiedList l with resolution f*resolution of L
for all voxels in L do
    add new voxel(x/f, y/f, z/f) to l
end for
store factor with l in cache
```

Once the occupancy lists have been downsampled and the particles selected the probabilities or weights for each particle are estimated. This process is finely balanced between giving an accurate weighting and computational speed as it is carried out for each particle. The process is described in Algorithm 5 and may be succinctly expressed as

$$o = |L \cap M|. \tag{5.7}$$

The overlap count is o, the number of elements or cardinality of set A is $|A|$, the map is M and the local transformed scan is L. Thus o is not only a count of the overlapping voxels but also an indicator of the probability of the associated pose particle. A statistical analysis of the significance levels for o in various situations is presented in

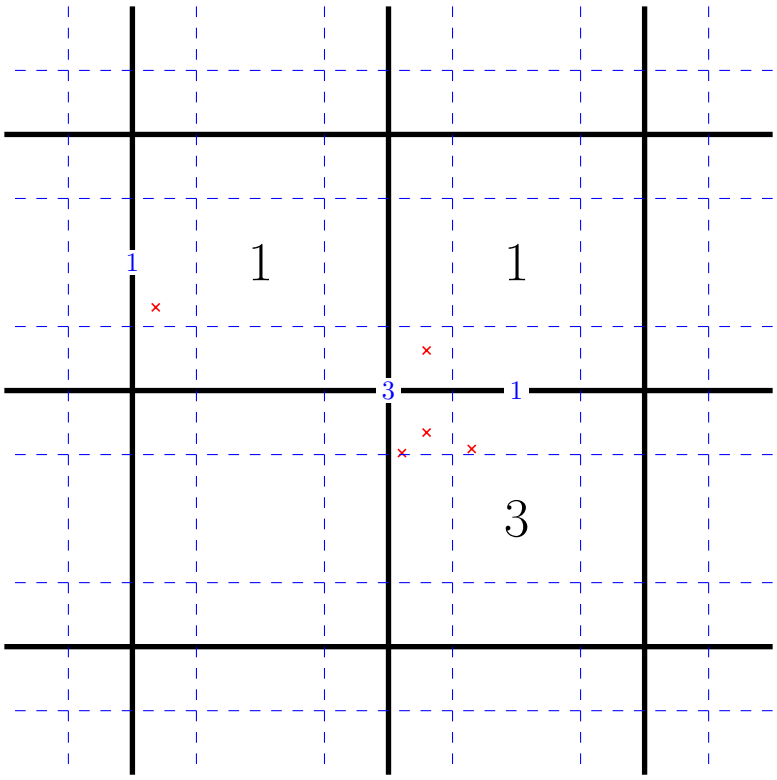

Figure 5.4: The quantisation process for converting point clouds to the occupied voxel lattice. In the diagram there are two grids superimposed upon one another with resolutions differing by a factor of two.

Section 5.3.2.

It is assumed that $|L| < |M|$ and so iteration is over the elements of L. The particles are stored in a set sorted primarily by their angle. Transformations are performed by rotation matrix multiplication for the rotations and component wise addition for the translations. Rotation operations on the local occupancy list are slower than translation operations. Rotations in 2D Cartesian coordinates require four multiplications and two additions as opposed to translations which require two additions. This processing becomes even more arduous when the alignment algorithm is extended to allow rotations due to uneven floor surface. Sorting the list by rotation groups similar rotations together and accompanied by checking to see if the rotation of the next particle is the same guarantees rotations are performed once only.

Once the particle weights have been updated, particles with a weight less than 80% of the maximum weight are removed. New particles are then generated around the remaining high-weight particles as described in Algorithm 6. This new particle set

Algorithm 5 Process for updating the weights of the sample set.

```
sampleSet poseSample, occupiedList M, occupiedList local
for each sample in poseSamples do
    o = 0
    if sample pose angle is different from previous then
        calculate new rotated local L
    end if
    translate L by (p_x, p_y)
    for each voxel in L do
        if voxel is present in M then
            increment o
        end if
    end for
    set sample weight to o
    if o > o_max then
        o_max = o
    end if
end for
```

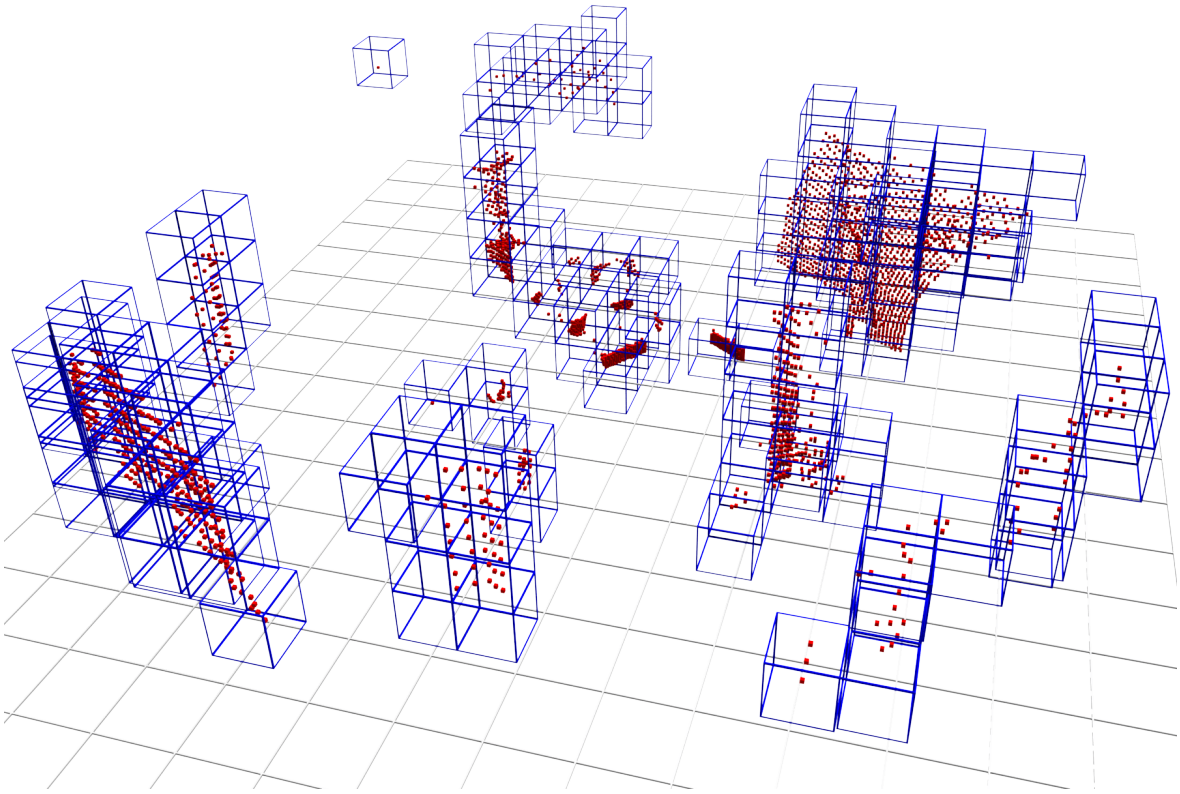

Figure 5.5: Illustration of the downsampling of a typical scan.

Algorithm 6 Re-sampling at improved spatial and angular resolution around high weight samples. This algorithm will produce repeated samples for adjacent initial samples, however samples are placed in a set to ensure uniqueness.

```
sampleSet initialSamples
integers a, b, c
create new sampleSet P
for each sample in initialSamples do
    for −1 ≤ a ≤ 1 do
        for −1 ≤ b ≤ 1 do
            for −1 ≤ c ≤ 1 do
                new sample (x + aε/2, y + bε/2, θ + cΔθ)
                add sample to P
            end for
        end for
    end for
end for
```

is then submitted for the next round of tests at better resolution. In this manner an order of 10,000 particles per second can be assessed on a standard 2GHz computer. The precise rate depends on the resolution and structure of the map and local occupancy lists.

5.4 Experimental Results and Analysis

Experiments are performed in various indoor environments, an office, a corridor and a lab. For all experiments the robots are controlled by teleoperation and are halted before each scan. The lab experiments, Figure 5.6, are testing the standard 3D scanner described in Section 4.3 rather than the enhanced scanner in Section 4.3.3, the results for which are presented in Figure 5.11 and Figure 5.10. For Figure 5.6, the map building is done by aligning horizontal cross-sections of the full 3D data. Consequently the resolution is lower at 0.08m than that in Figure 5.11 and the alignment is not as precise. Aligning the horizontal cross-sections, although faster, does not generalise to environments that do not have flat horizontal floors.

Figure 5.9 illustrates a cross-section considered for scan alignment and localisation. Spurious voxels are removed by taking a threshold of the occupancy map. Voxels that are a product of pixel mixing, the predominant error, tend to be isolated and are not

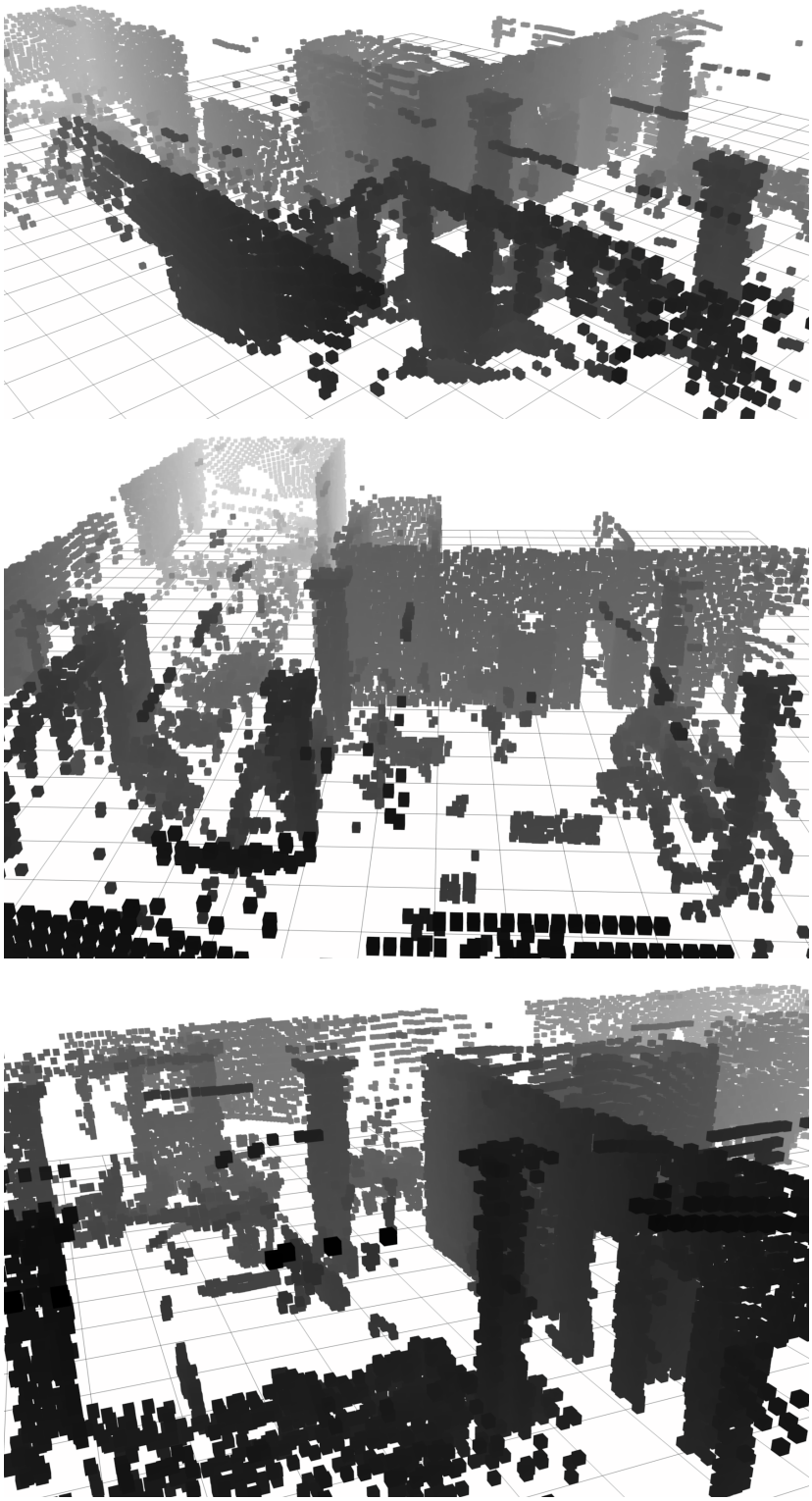

Figure 5.6: Range images of the experimental arena with a 1m grid on the floor.

observed repeatedly from different positions and hence have low occupancy counts.

The actual distributions of overlaps for a number of resolutions are shown in Figure 5.8 in which the mean overlap stays relatively constant regardless of resolution however the standard deviation grows as the resolution becomes finer. What is also clear is that the overlap count may be adequately represented by a normal distribution down to a resolution of 0.64m. The region of interest is the far right region of this graph where the significance of the maximum overlap may be established. The maximum count for the 0.16m resolution scan alignment shown in Figure 5.8 is 379 which is very significant and so the result may be regarded as reliable.

It must be recalled that one of the original assumptions is the independence of voxel occupancy. This assumption becomes more valid as ϵ, the voxel size, increases. The impact of this assumption is to increase the significance level for the overlap by increasing the spread of the distribution. Clustered voxels are more likely to produce higher overlaps. According to the distribution assuming independent voxels the maximum count of 379 for 0.16m has a vanishingly small probability and therefore it is safe to conclude that this is an actual match rather than one arising by pure chance alone.

Access to full 3D spatial information allows various subsets of the data to be selected. In Figure 5.6 the ceiling data is removed to clarify the data visualisation. Indoor environments may be extremely cluttered as is the case with the test environment used in these experiments. However, as hypothesized, this clutter is located at lower heights which is clearly demonstrated in Figure 5.7. In this figure the first map consists of the data minus the ceiling and the floor shown from above. Data points at lower heights are a lighter shade of grey. These lower points are clutter in the room, such as desks and chairs. The architectural structure of the room is emphasized by taking the points higher than 2.8m resulting in much less noise in the map and a closer agreement with the schematic in Figure 5.1. This map is immune to spurious results from the presence of humans and their actions and consequently will remain correct for longer.

The results of mapping a typical indoor office environment are displayed in Figure 5.11 and Figure 5.10. The data is gathered by the enhanced laser scanner developed

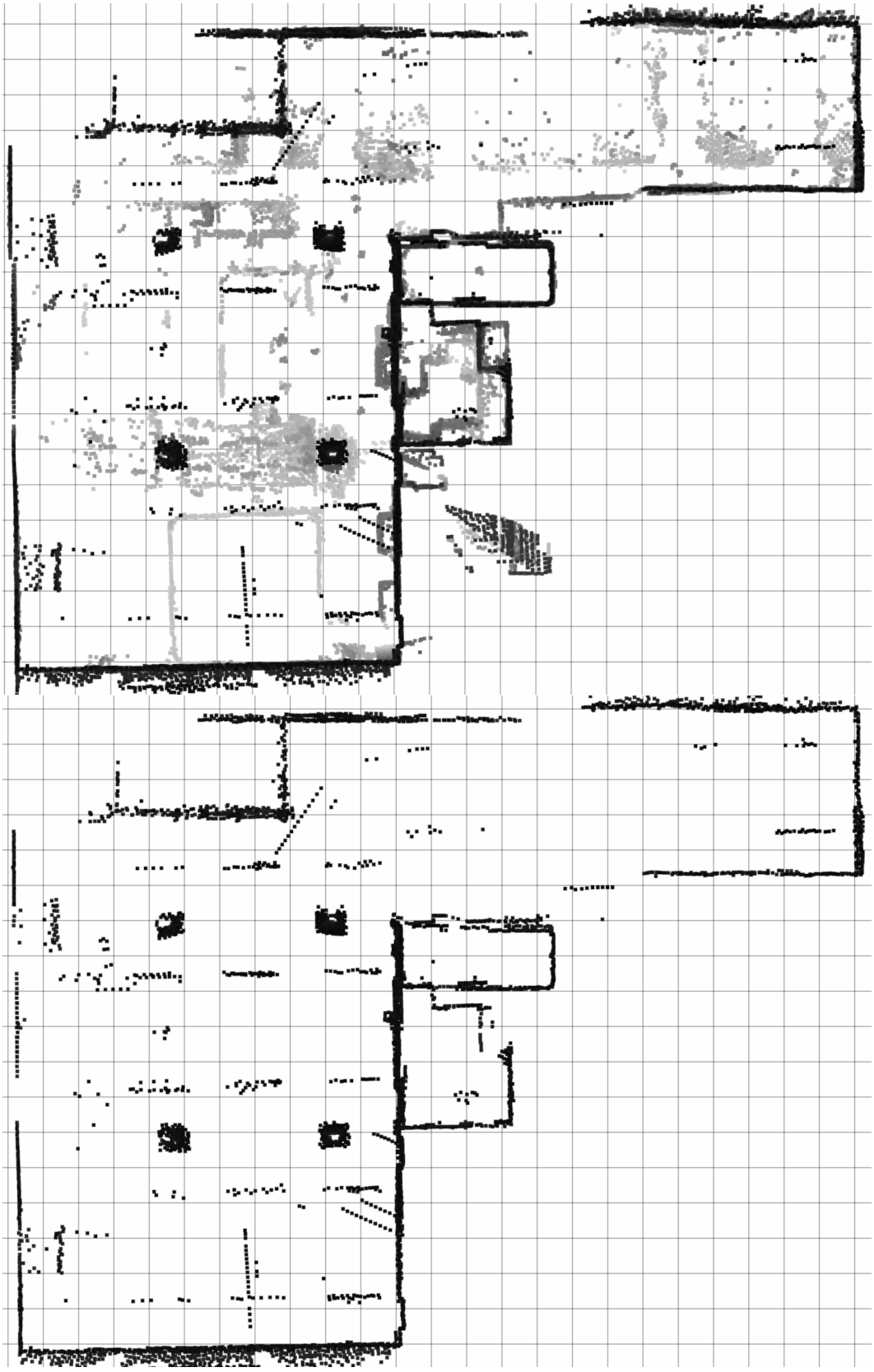

Figure 5.7: Orthographic view of arena from above with 1m grid cells. The first map shows data from all heights (higher cells are darker). The second shows only points above 2.8m.

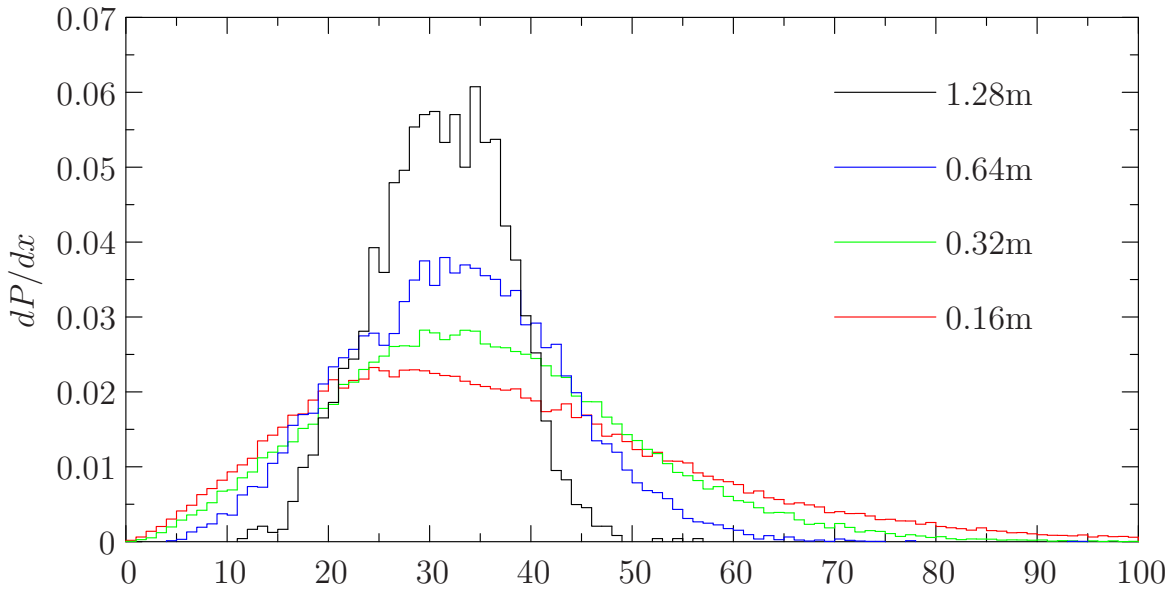

Figure 5.8: Probability distribution of overlap counts for various resolutions of the map and local scan.

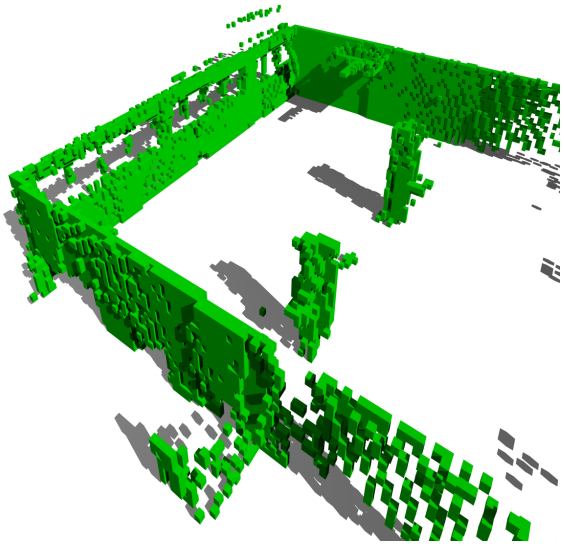

Figure 5.9: Global sub ceiling map with resolution of 0.1m. The map is a thresholded occupancy probability grid of the space 1-2.9m above the floor. The width of the room is approximately 12m. Two pillars and the windows at the back are evident.

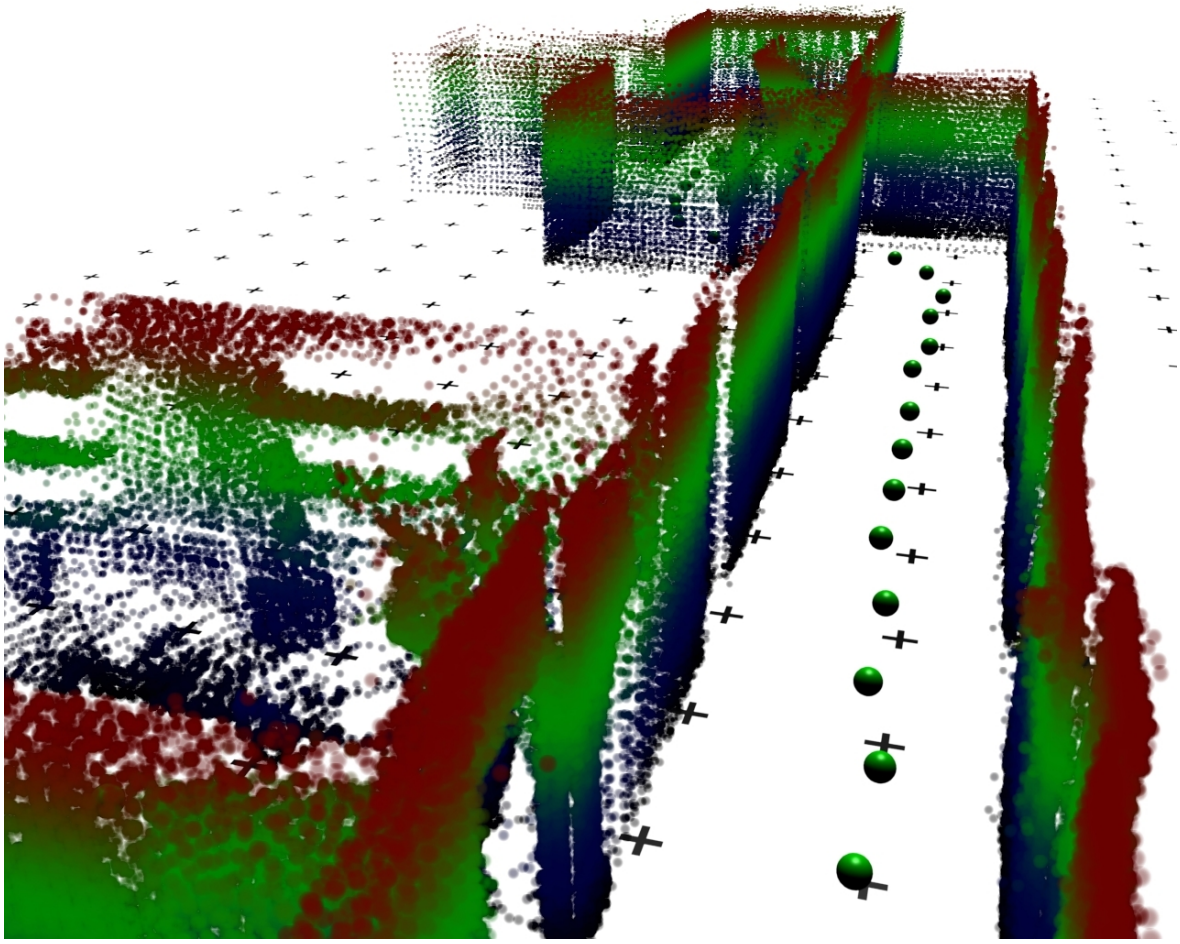

Figure 5.10: Corridor view demonstrating the mapping ability even in a relatively featureless corridor. The green spheres are the observation positions and the map voxels are coloured by height.

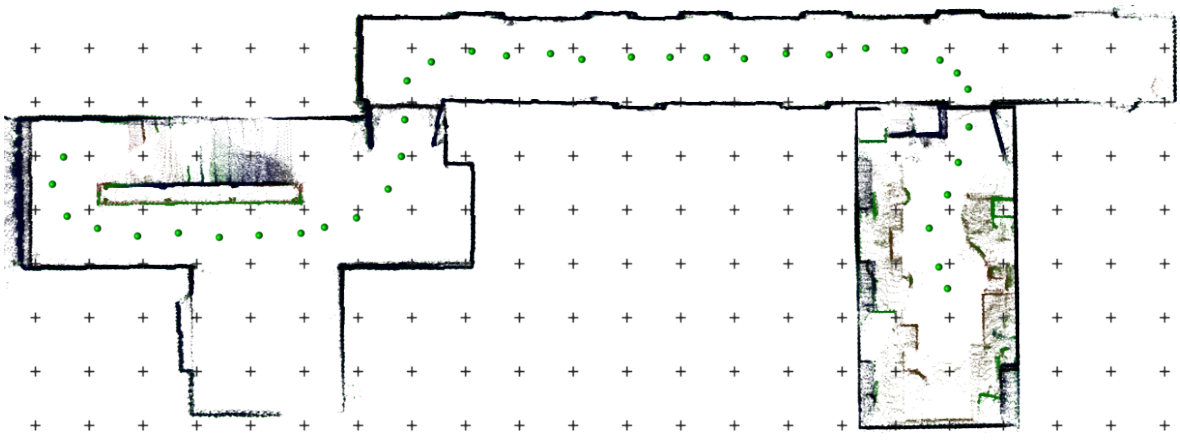

Figure 5.11: Overhead view of a 3D map containing a stair well, corridor and office at a resolution of 0.02m. The floor and ceiling points have been removed for clarity. The positions of the robot are shown as green spheres. The points are coloured by height and the grid points are at 1m intervals. No odometry is required to build this map.

in Section 4.3.3. The map is created by aligning 37 scans without odometry. Once the map shown in Figure 5.11 is created global localisation of different scans is performed within this map. The robot is assumed to be completely lost and therefore the entire area (approximately 200m^2) and all orientations is searched. Locating the 7000 voxel scan within the 200,000 voxel map takes 5 seconds on a 2GHz processor. The high data quality along with the almost omnidirectional field of view leads to good localisation and mapping performance. The occupancy list map in Figure 5.11 has a resolution of 0.02m and the position error is 0.03m. The error on the orientation may be estimated by noting that the misalignment between the corridor and the office is less than 1 degree. The alignment is of comparable quality to the manual alignment technique making detection of further improvements in accuracy difficult.

5.5 Summary

There are two main contributions covered in this chapter. First is the novel manner for representing 3D probabilistic map data which are referred to as occupancy lists. Secondly, a method for quickly aligning two occupancy lists is described which is similar to a multi-resolution particle filter. These two contributions have allowed a single mobile robot to move through an environment and generate an internal 3D representation of the environment accurate to at least 0.04m.

This chapter has concentrated on single robot mapping with a 3D sensor. As seen from the review of related work there has been a large effort with 2D maps largely due to the ubiquity of 2D sensors. Work on 3D maps is comparatively rarer. The results in this chapter however show that it is possible and clearly worthwhile to extend mobile robot sensors to incorporate 3D information.

One of the main problems of 3D sensors is the large amounts of data that they generate. This chapter elucidates an approach for mobile robots to cope with the data influx. The scan and map data is stored in occupancy lists that are fused by a multi-resolution sampling algorithm to produce 3D maps of the environment. The map allows global localisation over a 200m^2 area with no prior estimate to $\pm$0.02m

and sub-degree error in 5 seconds on a 2GHz processor. The process explained in this chapter is able to both localise and generate maps without odometry data.

The algorithm for aligning 3D occupancy lists is presented along with a theoretical treatment that gives an estimate of the significance of particular overlap values. This analysis is vital for determining if scans should be added to the map and to detect loop closures. The conclusion of this theoretical analysis is tested experimentally for a range of occupancy list resolutions. It is found that the agreement with experiment improves as the occupancy list voxel size increases. The reason, it is postulated, is due to the increase in validity of the independent voxel assumption with increasing voxel size.

Also addressed is the validity of storing occupancy as an integer number of votes. This is done by assuming voxel independence starting from the Bayesian update and applying the Markov assumption. The resulting occupancy probability update equation is re-expressed in terms of log odds probabilities and from this the validity of an additive voxel update is apparent. Under these conditions the occupancy list integer values must be considered as proportional to the log odds of the occupancy probability.

Experiments testing this approach to 3D mapping are conducted in a range of indoor environments some of which are extremely cluttered and others featureless. The idea of extracting cross-sections from the 3D map data is presented with an illustration of the reduction in environmental complexity above a floor based robot. It is shown that maps of indoor areas above 2.8m are closer to the architectural structure of the environment and it is proposed that they are more suitable for localisation over a longer term as they do not include most dynamic objects.

This chapter completes the foundations for Chapter 6 which discusses multi-robot 3D sensing and mapping.

Chapter 6

3D Cooperative Mapping

Many potential tasks for mobile robots are more amenable when tackled by teams of robots. Mapping, although possible with a single robot may be significantly improved in this way. This chapter combines the work in previous chapters to develop and test a system for multiple mobile robots equipped with 3D laser scanners that enables the robots to detect one another and map cooperatively. The performance of multiple mapping robots is tested for four robots and compared to that of a single mobile robot. It is found that when the four robots are acting in concert their localisation errors never exceed 0.3m however when acting alone 10% of their localisation errors exceed 1m.

6.1 Introduction

The mutual localisation approach discussed in Chapter 3 coupled with the 3D laser range finder prototype (Chapter 4) pushes this research into the new area of three-dimensional cooperative localisation and mapping. Combining mutual localisation with the data from multiple 3D laser scanners enables full 3D mapping of indoor environments. This would prove vital for a number of applications such as nuclear decommissioning, search and rescue scenarios, surveying as built structures and supplying prior maps for other mobile robots.

When considering multi-robot localisation and mapping a number of questions nat-

urally arise. Many of these are prompted to justify the increased cost and complexity of multi-robot systems. What intrinsic improvements to the mapping process are possible using teams of robots? Namely, improvements arising from multiple cooperating robots rather than many single robots mapping separately. Before mapping even starts robots can auto-calibrate each other by observing commands and the resulting trajectories. Through this process the accuracy of the odometry can be improved for each robot given the surface upon which it is currently operating. The accuracy of odometry is highly dependent upon the coefficient of friction between the robot wheels and the floor. If the wheels slip excessively then the accuracy of the odometry is severely degraded. For instance, one robot can drive in a test square and through comparing internal odometry readings with the path actually observed, both the odometry and command functions may be corrected. Sensor readings may be checked with less accurate sensors calibrated against the more expensive and more accurate laser scanners. In this way robots continue to benefit from a collaborative approach even if no further interaction or information exchange occurs.

How does the constitution of the team affect the mapping performance? From the simplest and cheapest team consisting of one robot with a laser scanner and one with sonar scanners to more complicated configurations. How is the mapping speed in entropic bits per second affected per unit cost?

What environmental representation should be adopted? Tied in with this question is how the sensor readings for a particular pose in the map may be predicted, a necessary process for estimating the likelihood of a pose.

How should the belief in the robot pose within that environmental representation be expressed? Are single robot probabilistic pose representations adequate for multiple robots? Do the pose distributions have to be maintained separately or might they be combined?

A sample of the literature discussing multi-robot mapping is explored in Section 6.2. This section describes the benefits that researchers have experienced with multi-robot mapping and also explores the methods others have used for localising robots relative

to one another. Section 6.3 develops the underlying theories necessary to address multi-robot mapping with sensors returning 3D data. Within this theoretical framework is the extension of mutual localisation to 3D. An overview of the information that needs to be exchanged between robots is presented, specifically the different pose constraints between robots under different combinations of observations. The implementation details and various algorithms executed for 3D cooperative mapping are included in Section 6.4. The experimental results are advanced in Section 6.5. This section includes the final experiment of this dissertation in which a team of four robots equipped with 3D laser scanners maps an indoor environment and the cooperative mapping result compared to that produced by the robots mapping independently.

6.2 Related Work

Very few research groups are performing cooperative 3D laser mapping. The closest are groups using cooperative vision and laser based mapping in outdoor environments Madhavan et al. (2004) and vision only Little et al. (1998); Rocha et al. (2005). Localisation and mapping using heterogeneous robot teams with sonar sensors is examined extensively by Grabowski and Khosla (2001); Khosla et al. (2003). Using more than one robot enables easier identification of previously mapped locations, vastly simplifying the loop-closing problem (Konolige and Gutmann, 1999).

Multi-robot cooperation brings enhancements as well as disadvantages. Combining data from multiple robots or other sources, often referred to as data fusion, needs careful treatment and benefits from the results of sensor network research. Although sensor networks tend to deal with many static sensing devices some parallels are important and may usefully be adopted.

Table 6.1 summarises the findings of Rekleitis et al. (2002) that describe the positional accuracy of robot team members for a range of team and cooperative location configurations. The rows of Table 6.1 correspond to different co-location schemes describing what relative positional information is extracted from the sensors. Range is when the only information available is the separation distances between robots, azimuth

Number of Robots	3	5	10
Range (ρ)	38.80	21.63	8.13
Azimuth (θ)	27.06	32.20	33.72
Position (ρ, θ)	34.25	21.79	7.50
Full Pose (ρ, θ, ϕ)	28.73	16.71	6.05

Table 6.1: The mean error in position estimation (cm) after 40m of travel over 20 trials (Rekleitis et al., 2002).

is the angular separation, position is both and full pose incorporates the orientation of the target as well. This table highlights a number of intriguing phenomena. The most accurate three robot localisation occurs when only angle information is available, which seems rather counter-intuitive as one would expect better accuracy with more information. Another surprising result is the deterioration of azimuth position accuracy with increasing robots. The azimuth row aside, the remaining results follow a consistent pattern with more robots and more information about relative poses producing better accuracy. It is encouraging to note that the most marked improvement in accuracy happens for smaller teams. This means that it is not necessary to employ substantial numbers of robots to achieve significant decreases in location error.

The ability to exchange maps and sensor data between mobile robots gives rise to substantial improvements in mapping efficiency. Rekleitis et al. (2000); Simmons et al. (2000) illustrate the benefits of multi-robot mapping in terms of increased mapping speed and improved reliability. Fox and Burgard (1999, 2000) succinctly highlight the benefits of sharing sensors amongst multiple robots. Most of the multiple robot mapping literature is focused on scaling an algorithm essentially designed for single robot operation (Burgard et al., 2000). These methods often fail to take advantage of the intrinsic benefits arising through the deployment of robot teams.

Using more than one robot enables easier identification of previously mapped locations, simplifying the loop-closing problem. The loop-closing problem is addressed by Konolige and Gutmann (1999) with a typical loop and its closure demonstrated in Figure 6.1. Complex cyclic environments may also be mapped through modification of the environment itself. If simple markers are used then the possibility of external influences tampering with these markers can never entirely be dismissed. If the markers are

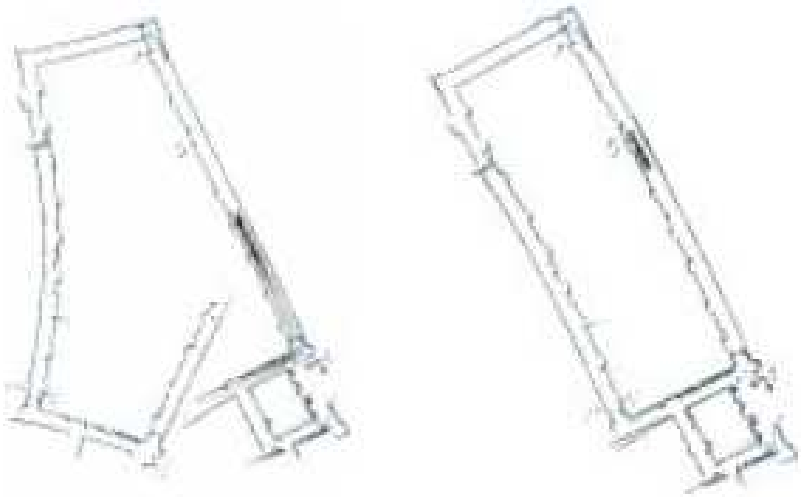

Figure 6.1: Closing of an open loop in a large (80m by 25m) cyclic environment from Konolige and Gutmann (1999)

instead other mobile robots that can detect changes in their pose then this undetected tampering is eliminated.

6.2.1 Homogeneous and Heterogeneous Teams

Many implicit assumptions based on homogeneity of the robot team fail under the more general condition of cooperation between different robots. One example would be the height of the laser range finder from the ground. It is generally true that there is a correspondence between range data taken at different heights however this needs to at least be considered.

Large heterogeneous teams (Howard et al., 2004a) allow the sensor payload to be distributed amongst many robots. As long as each robot can localise, move and communicate then sensors which are power hungry, heavy and expensive, like laser range finders, can be distributed sparingly amongst team members. Yet the benefits of these more cumbersome sensor payloads are available to all team members. This specialisation is possible once the constraints placed by assumptions of homogeneity are removed.

A number of ideas may be extracted from experiments with large (greater than 100) heterogeneous mobile robot teams (Konolige et al., 2002) performing indoor mapping and surveillance. Questions such as what is each robot's role, how these roles cooperate

to form teams and how these teams are coordinated overall become important as the size of the robot team grows. One team configuration is a master-slave set-up involving one complex, sophisticated and possibly expensive robot controlling and coordinating minion robots who may be simple and only satisfy two criteria. These slave robots must be able to stay stationary and detect if they are moved; it must be possible to locate the slaves relative to the master robot. These robots then allow the use of methods and techniques developed for active beacon localisation to be applied in an unmodified environment. These mobile robot beacons are relatively expendable and so are suited for application in dangerous environments, such as nuclear decommissioning. Contemporary research indicates (Konolige et al., 2002) that any mapping methodology must be eminently scalable. This means that separate maps generated by different teams of robots must be fused accurately. To test this the team of robots must be able to fuse two previously generated maps. This fusion process could be, scan matching the maps if there is sufficient overlap, followed by explore and rendezvous steps to verify satisfactory map cohesion.

Figure 6.2 illustrates the robots tested by Howard et al. (2003) to aid co-location of robots carrying retroreflecting cylinders. These cylinders allow not only the relative position to be measured but also the orientation which has a 180° ambiguity. The experiments in Section 2.5.1 employed a plastic cylinder and detection is successful up to 5m; using retroflective cylinders yields a substantial increase in the detection range.

6.2.2 Odometry

Odometry is often used as an initial estimate for most localisation algorithms, consequently the more accurate the initial guess, the quicker these algorithms will converge on their solutions. Therefore, there have been many attempts by researchers to improve odometry through cooperating robots. Calibration is often done to remove systematic errors in robot odometry and has a number of benefits that are covered by Borenstein and Feng (1996). Attempts to classify and benchmark (Borenstein and Feng, 1995) systematic errors have met with varied success. Borenstein and Feng (1995) lists the

Figure 6.2: A team of Pioneer 2DXs used by Howard et al. (2003) equipped with SICK LMS200 scanning laser range finder and pair of retroreflective cylinders

following odometric errors.

- Systematic Errors

  Average of both wheel diameters differs from nominal diameter

  Misalignment of wheels

  Uncertainty about the effective wheelbase

  Limited encoder resolution

  Limited encoder sampling rate

- Random Errors

  Travel over uneven floors

  Travel over unexpected objects on the floor

  Wheel-slippage

It is interesting to note that all of these are ameliorated with the multiple mobile robot approach. The usual deterioration of odometry accuracy with increasing speed of operation may be readily quantified. For a recent in depth analysis of odometry errors

the reader is referred to Kelly (2004). Borenstein and Wehe (1997) have the interesting idea of correcting odometry errors by having two drive platforms loosely coupled so that any momentary wheel-slippage experienced by one is picked up and accounted for by the other. The results achieved are rather impressive with an order of magnitude improvement in the odometry realised. This comes back to the following question. How can multiple robots working in concert improve the map building process? A way of achieving this is by improving the odometry of all robots in the team which may be done via automatic calibration. In this way odometry may be calibrated for particular floor surfaces and current operating conditions.

6.3 Theoretical Framework

Figure 6.3 presents an overview of the cooperative 3D map building system. Any robot can establish itself as the observer and then remains stationary whilst a local map is generated. As the mapper robots map around the observer various combinations of sightings are possible as explained in Section 6.3.3. The mathematical constraints these sightings place on the poses of the mapper robots are labelled in Figure 6.3.

The 3D scans are taken by each robot and the resulting point cloud stored locally. These point clouds are combined into the map with the aid of the inter-robot sightings. This combination may be done on any of the robots, but is done on a central computer for these experiments because processing the 3D data is computationally intensive and the robots have limited computing power. The main computer stores a probabilistic digitised representation of the environment, receives 3D sensor scans and issues commands. Communication between the robots and the control computer is through wireless LAN.

6.3.1 Robot Information Interchange

In any multi-agent system it is important to consider both how and what information is exchanged between the agents. The latter is important because any physical commu-

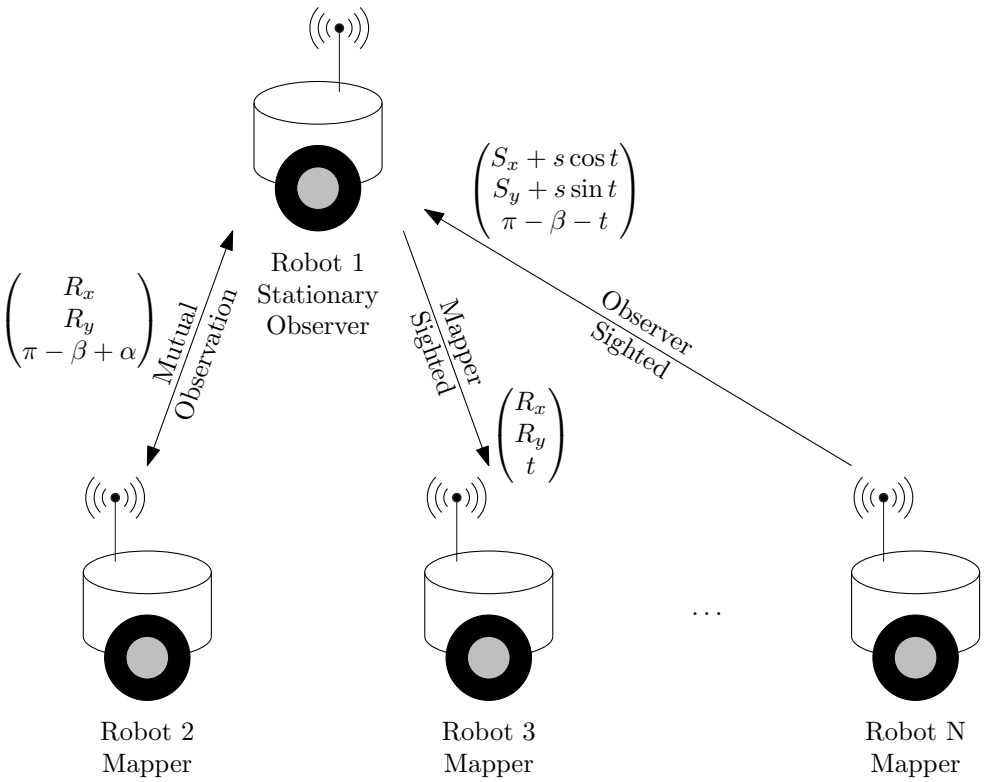

Figure 6.3: Overview of the map building system. The pose constraints are labelled for each type of observation in the temporary reference frame of the stationary observer. The arrows indicate the direction of sighting.

nication link has a bandwidth and so excessive data exchange requirements will limit real-time operation. The choice of communication mechanism will determine how much information can be transmitted between agents which will affect the information selected for transmission. This is especially important for 3D mapping where the amount of data involved in map sharing is potentially very large. These data transfer requirements are substantially reduced by representing 3D maps as occupancy lists, which are explained in Section 5.3.1.

In mobile robotics there are a number of communication methods depending on the amount of internal processing and electrical power that is available. The nature of the operational environment needs also to be considered.

- Does the link have to be encrypted for security?
- Does it have to be local?
- Is there are wireless communication infrastructure already present?

- Does it have to be peer-to-peer?

As is often the case in mobile robotics there is no one ideal solution for all deployment situations and compromises must be made. For medium sized robots (0.1m-1m) standard wireless networks are applicable especially if a wireless network is already present. If there is no wireless network present then peer-to-peer wireless networking is possible but adds significant operational complexities and constraints. In this situation it is another reason for the robot to have a good awareness of its location so that it can return to a location known to have communication available. In this way semi-autonomous operation is possible. The robot is issued with a mission such as explore and map and then after a certain interval will return to its base for further orders or maintenance.

In this work wireless networking is used for data transmission. Having briefly addressed communication hardware requirements and criteria the question of exactly what information should be transmitted is now addressed.

There is substantial flexibility of the system regarding which processes run where. For instance the mapping algorithm may be run in real-time on a single robot or the data offloaded to a central machine where the mapping algorithm takes advantage of a faster CPU. In the experiments in this research all raw data is recorded and then analysed later. Care is however taken to try and match the processing capabilities with the incoming data rate to ensure that the algorithms are as close to real time as possible. This aim is aided by the multi-resolution sampling approach discussed in Section 5.3.3 which always has an answer available, tuning it over time to increase the accuracy and reliability of the result. Such algorithms are known as anytime algorithms and are suited for operation under real-time restrictions.

The transmission of information, especially sensory data, throughout an ad hoc network of mobile agents is sometimes referred to as sensor networks and achieving this in a peer-to-peer manner is termed decentralised sensor networks. Extensive work on information interchange of multiple 3D mapping robots Rocha (2005) focuses on minimising the mutual information of maps before transmission thus alleviating the

burden on the transmission channel. Information entropy as a measure of the utility of the robot's sensory data aids this effort. In Rocha (2005) localisation and mapping is considered as an application domain example for efficient multi-robot cooperation. Indeed an external absolute localisation mechanism is assumed, simplifying map building.

6.3.2 Multi-robot Mapping

As explained in Fox and Burgard (2000) maintaining a probability distribution of poses across all robots is unfeasible because the dimensionality of this distribution would be three times the number of robots in 2D. Instead a simplifying assumption brings this into the realms of tractability. The pose probability distribution for each robot is maintained separately and thus the poses of robots are assumed to be independent which is valid for low numbers of small robots. As long as the volume of robots in the environment is much smaller than the environment itself then the bulk of sensor data will be from the environment rather than from nearby robots. The impact of this assumption is also reduced by having a reliable robot detection process and once a robot senses another removes the corresponding volume of data from the laser scan.

Mapping inaccuracies are divided into systematic and random errors. Random mapping errors may be assessed by repeatability; map the environment multiple times and test the map similarity. This will not find systematic map errors however these may often be calibrated out by comparison with known maps. The use of multiple heterogeneous robots may enable the automatic detection and removal of these systematic errors which would otherwise only be detected through comparison with the true map. Acquiring this ground truth is often a time consuming and arduous process. The errors of the 3D sensor are determined in Chapter 4. The ground truth for the 3D map is acquired by careful manual alignment of the 3D scans which produces a reference map accurate to 0.04m.

Mapping performance also needs to reflect the mapping speed, which is proposed as the entropic measure of information extracted from the environment per unit time

(bits/s).

$$H(X) \equiv -\sum_{x} P(x) \log_2 [P(x)] \tag{6.1}$$

The Shannon entropy of a variable X is expressed by (6.1) and gives a good measure of the data content of a signal measured in bits; $P(x)$ is the probability that X is in state x. Using this, rather than the raw number of bits, helps differentiate that data acquired from simple environments is not as informative as data from complex environments.

Whenever possible raw data should be stored for any mapping situation. With this raw data available post mapping off-line algorithms will always produce superior maps. There are a number of reasons for this. Off-line algorithms can be run on more powerful computers without mobile robot constraints such as power and physical size. When considering a particular scan knowledge of future scans is helpful, for instance in loop closing which although may be approximated in real-time is inherently amenable to a batch data algorithm.

The aim of real-time mapping is to ensure that the robot knows its location and for path planning. As long as the map is of sufficient quality to satisfy these criteria than high map fidelity may be acquired off-line later. In real-time mapping if the robot is unsure of its pose the option remains to move in such a way as to increase its knowledge and certainty of its position. This form of active localisation is not available when post processing the raw data.

A scan stored as an occupancy list with the floor and ceiling points removed for a medium sized room at a resolution of 0.02m is around 5000 voxels and so each voxel has four integer components, three for position and one for occupancy. Typically at four bytes for each component a typical scan stored in a 0.02m resolution occupancy list is about 80kB. Assuming that a robot can move to a new position and scan every 5 seconds then raw data recorded for one hour, stored in this manner, does not normally exceed 100MB per robot. This value will depend on the shape and size of rooms as well as the particulars of the mapping mission, however as an estimate of the storage required it suggests the feasibility of logging the 3D scan data as occupancy lists on Pioneer 2 robots.

It is possible to irreversibly merge multiple nearly sequential occupancy grids with high match certainties however there is a loss of data with this approach and there still exists the risk that an incorrect merge takes place especially in highly symmetric environments. Secondly when loop closures are detected it is often necessary to back propagate the changes in the intervening pose estimates along the loop. These poses changes cannot be made if the scans have been irreversibly merged. Removing the floor and ceiling voxels achieves a large data reduction. These are superfluous for localisation in typical indoor environments because the robots are often operating on a smooth horizontal floor. Points belonging to flat featureless horizontal surfaces contribute little information to the robot's knowledge of its pose in a horizontal plane.

In order for robots to operate in an ad hoc manner in an environment lacking wireless access points each robot stores a list of all occupancy list along with their time stamp and which robot recorded them. Whenever robots are within wireless networking range of each other they synchronise their occupancy list data stores and if it is the first time they have met they align their maps using mutual localisation. Associated with each occupancy list is a particle set of poses that locate that occupancy list within the global frame.

6.3.3 Mutual Localisation with 3D Scanners

Laser range finders are not the only way for robots to locate other robots. Others include a heterogeneous team consisting of a leader equipped with a scanning infrared detector and simple robots carrying infrared beacons. The disadvantage of this method is that only angle information is available, it requires many beacon robots and has problems resolving position in certain robot distributions. The advantage is that this information is easily obtained by the sensors with no computational overhead unlike cylindrical retroflective beacons that require processing to find in the range data. The multi-resolution sample based mechanism described in Chapter 5 easily accommodates this extra information. So if for instance infrared beacons are deployed on the robots this angle information constrains the set of initial poses to be distributed around a

line rather than an area. The effectiveness of mutual localisation with 3D scanners and retroreflective beacons described in this chapter meant there is no need to deploy alternative sensors for robot detection.

The 3D laser scanner deployed on the four robots is described in Section 4.3 and has a blind spot directly in front of it with an angular width of 20° thus higher mounted beacons will be visible from further afield. As shown in Chapter 4 in practical situations it is better to have the laser scanner mounted facing upwards so that this blind spot is located in a less significant volume of the scan. For the experiments in this chapter it is necessary to deploy scanners on four robots and so the lesser hardware intrusive mechanism is employed. Most clutter in rooms is quite low, for example chairs and tables, so even when the line of sight between the robots is interrupted, the beacons, being some way from the ground, are still visible to laser scanners on other robots.

Mutual localisation is accomplished by first projecting the observed beacon's 3D position onto the xy-plane which is parallel to the floor of the room. Once projected onto the horizontal plane the 2D mutual localisation algorithm from Chapter 3 is applied. This method assumes that the beacons are always the same height above the floor, reasonable in the case for many indoor environments with flat horizontal floors. Given the relative pose of the mapping robot with respect to the observing robot multiple 3D laser scans are then combined to produce a local map.

In the situation where the horizontal separation of the beacons and the laser scanner is small the separation is calculated from the robot observations by the cosine rule for triangles, otherwise a more complicated approach is necessary as described in Chapter 3. For 3D mutual localisation the beacons are positioned directly above the origin of the laser scanner thus minimising the horizontal separation and simplifying the mutual localisation process.

Much effort has been expended in the area of scan matching and global localisation (Kwok et al., 2004; Jensfelt and Kristensen, 2001), mostly to improve performance in generating the position probability density function through a variety of sample techniques. This localisation or scan matching has to be performed on a fairly regular

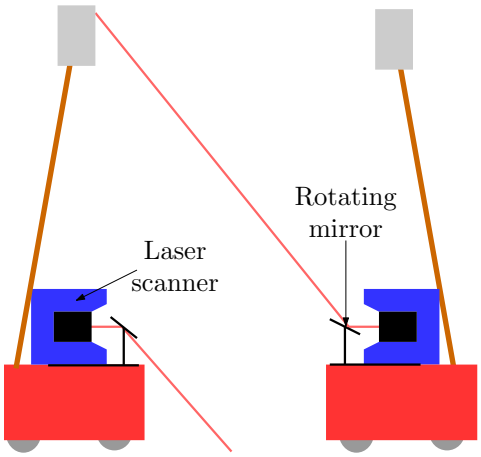

Figure 6.4: Beacon positioning for 3D mutual localisation. The retroreflective cylinders are supported directly above the origin of the laser scanner.

basis (every meter or so) because the accuracy of odometric data deteriorates so rapidly. Frequent pose updates mean that the previous pose is a reliable initial guess value which is required by most algorithms. Because inter-robot sightings are often available, relative pose determination is much faster. Consequently, reliable local maps can be generated quickly and robots have a better chance of unambiguously localising in the global map. Scan matching depends on a good overlap between subsequent scans which comes from good sensor coverage and good visibility of features, both of which are delivered by the higher regions of a 3D scan. It is noted that in experiments when robots move through doorways or other such thresholds in the environment the overlap between scans is sometimes too small for reliable scan matching. It is in these instances that cooperative localisation is especially helpful.

Cooperating robots simplify and improve the map building process (Ryde and Hu, 2005b, 2006a) as elaborated in Chapters 2 and 3. For the cooperation of robots to be most successful they must have the ability to observe one another in a reliable fashion. This is accomplished by mounting retroreflective cylinders above the laser scanner as shown in Figure 6.4 the details of which are expounded in Chapter 2.

These inter-robot observations may be divided into two categories. Normal observations and simultaneous mutual observation events where the robots observe one another simultaneously are explained in detail in Section 3. The latter are more pow-

erful in terms of the reliability and accuracy of the relative positional information that they deliver. One of the problems of the simultaneous mutual observations is that they are rarer. Therefore, it is beneficial to incorporate the information delivered by both types of observation events into the map building process. A standard observation, although not uniquely identifying a single pose, vastly reduces the number of poses that need to be considered to a locus of points in pose space.

Consider two robots, the stationary robot, S and the mapping and moving robot R. There are three observational situations to consider in descending order of occurrence frequency in mapping missions. The first is if S observes R at a position of (R_x, R_y) then the locus of poses is given by

$$\begin{pmatrix} R_x \\ R_y \\ t \end{pmatrix} \tag{6.2}$$

where t is a parametric variable used to describe the locus that varies from 0 to 2π. This in effect says that the position of R is fixed and its orientation is unknown. Secondly, if R observes S at (s, β) then the locus of poses of R in the coordinate frame of S gives rise to a locus of poses, the positions of which are along the circumference of a circle and with each pose there is associated a single orientation and this is expressed in parametric form as

$$\begin{pmatrix} S_x + s\cos t \\ S_y + s\sin t \\ \pi - \beta - t \end{pmatrix} \tag{6.3}$$

again with the parameter t. Finally, the rarest and most informative observation is when both S and R observe each other and the relative pose is determined with high certainty (Ryde and Hu, 2005a).

$$\begin{pmatrix} R_x \\ R_y \\ \pi - \beta + \alpha \end{pmatrix} \tag{6.4}$$

The pose constraints are summarised in Figure 6.3. In the situations where observations give rise to a range of poses, then localisation is performed by map matching along the locus of poses by simply testing values of the parameter t. This not only dramatically reduces the search space for possible poses, but also decreases the likelihood of high correspondence false map matches. This approach gives localisation in four degrees of freedom x, y, z and θ the horizontal orientation. The two degrees of freedom missing for a fully unconstrained 3D pose may be acquired from either the assumption that the robot is on a flat floor or that the robot is capable of determining the direction of acceleration due to gravity.

6.4 Cooperative Map Building

The accomplishment of reliable and accurate cooperative localisation is easily extended to allow the construction of maps. The pose change gives the rotation and translation between successive laser range scans. This information allows all the scans to be converted into one reference frame resulting in the production of a global map of the environment.

One of the main complaints levelled at cooperative map building is the need for multiple robots to be in communication and line of sight of one another, this increases the difficulty of path planning and control. Thus any cooperative approach must be robust to single robot operation and be able to fluidly combine the sensor data as it arrives from possibly disparate robots. Two elements which enable this are the environment representation and the map update process.

6.4.1 Concurrent Processing

An exhaustive search algorithm of localisation is unfeasible because it does not run fast enough to provide continuous real-time pose updates. However for the experiments in this chapter the pose updates are supplied by cooperative localisation. The global localisation is run in parallel with the local map build process as depicted in Figure

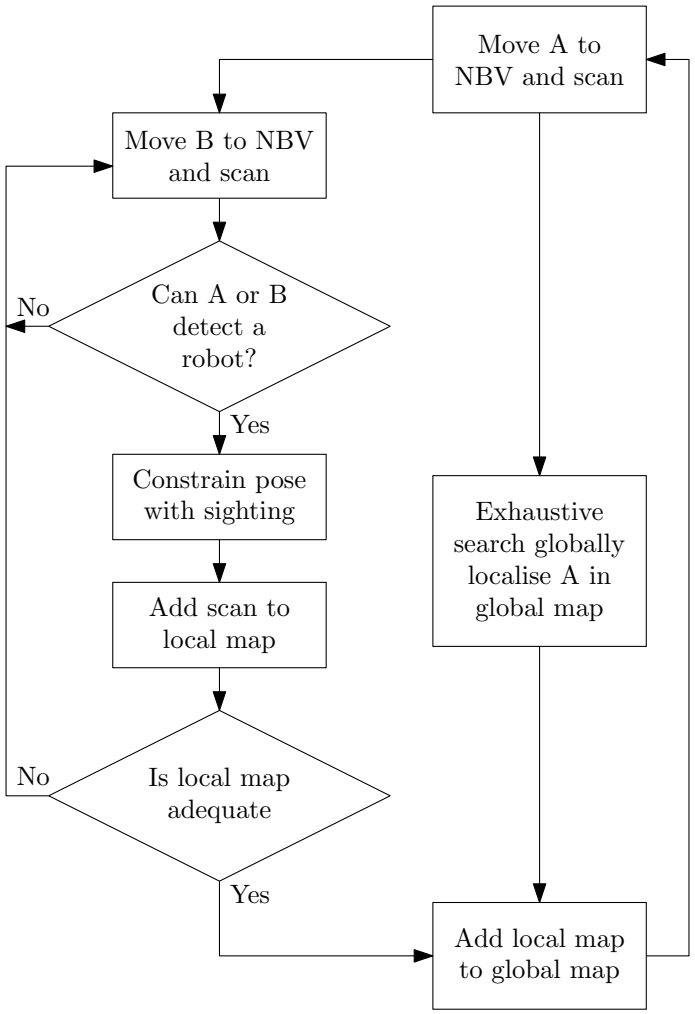

Figure 6.5: The 3D mapping process for two robots A and B. The constraints placed on relative poses due to sightings are in Figure 6.3.

6.5. In Figure 6.5 the first robot, A, moves to the next best view (NBV). The NBV is the next position which is most likely to yield useful observations for updating the map and in these experiments is estimated manually. Robot A is acting as the stationary robot providing a reference frame for B whilst B maps the local area around A. When the local map is sufficiently good it can be added to the global map. Whilst the local map around A's new pose is built the pose of A within the global map is calculated. This pose along with (3.22) enables the merger of the local and global maps. The process is repeated until the global map is of sufficient quality.

The laser scanner can detect returns for distances approaching 80m subject to the reflectivity of the target. Angular error in the laser scanner means that the error increases linearly with range. This information could be incorporated into the sensor

model and map update process however long range returns would require the increment of many voxels by a small amount over a large volume. This could be accommodated by a multi-resolution occupancy list however it adds significantly to the complexity of implementation. This loss of information is insignificant when regarding the large amounts of data already returned by the 3D scanner. A balance is struck in the sensor model which assumes constant errors over a relatively small range of 8m equal to the largest error in that range which occurs at 8m.

Part of the pose search procedure in Algorithm 7 is the calculation of the difference metric. This calculation is embedded in several loops. Thus the dependence of the localisation on this aspect means that performance of the difference or overlap metric is paramount. For this reason the overlap is simply calculated by counting the number of coincident occupied voxels between the two occupancy grids. When comparing the two probability grids the cell probabilities fall broadly into three main categories, occupied, empty and uncertain. The uncertain probabilities usually indicate areas that have yet to be observed. There is a small chance that numerous conflicting observations produce a cell probability that is equal to the grid initialisation probability however the vast majority of cases are due to unobserved cells. The accuracy of the 3D scanner combined with the effectiveness of mutual localisation means that the probabilities coalesce quickly to high certainty of occupancy or vacancy after several observations.

Digitising the space of an average indoor environment produces a majority of unoccupied voxels. This is the case because in order for an environment to be suitable for operational movement it must have a large proportion of connected free space. The robot can only observe within the free space so large volumes of occupied space may only be detected as large volumes of uncertainty bounded by occupied voxels because the sensor cannot see into the interior of these volumes. The occupied voxels essentially lie on surfaces where as the free voxels occupy volumes. This provides the source of the asymmetric balance between occupied and unoccupied voxels. Another way of looking at this is to consider that for each scan ray emitted by the laser scanner only one terminating voxel is observed but it contributes to the observation of many free

Algorithm 7 Exhaustive search global localisation.

```
initialise all elements of overlap array P to 0
for θ is 0 to 359 increment 1 do
    rotate sensor grid by θ°
    for x ∈ X do
        for y ∈ Y do
            if current grid element is unknown then
                consider next θ
            end if
            calculate overlap metric o between map M and rotated sensor grid S with
offset (o, p) as |M ∩ S|
            if o > P_op then
                P_op = o
            end if
        end for
    end for
end for
```

voxels. This asymmetry is taken into account when considering differences between probability grids. Because occupied voxels are rarer then their appearance is more significant so comparisons involving occupied voxels must be given more weight.

The exhaustive search approach described in Algorithm 7 is rather slow and unfortunately has a quartic dependence on resolution. Although at first disappointing this quartic dependence confers an advantage. By varying the resolution a compromise between speed and accuracy may be made. An initial low resolution (0.5m) localisation is done to the nearest degree in the coarse map. This gives a number of likely relatively large cells which can then be investigated further at a better resolution (0.05m). This delivers a vast increase in the speed of localisation prompting the iterative particle filter method detailed in Chapter 5. A variable resolution particle filter operating on conventional occupancy grids is described by Verma et al. (2003).

The single robot and cooperative maps are built in a similar manner. Although adequate results are possible with a real-time algorithm, it is found that the comparison between multiple and single robot mapping is easier with a batch algorithm producing more accurate maps. In this batch algorithm every scan is tentatively aligned with the first scan. The one with the highest overlap is taken and combined with the first scan to produce the current global map. All remaining scans are then aligned with

this global map and the best one selected for combination with the global map. Scans are combined with the global map by adding the occupancy of the globally aligned scan voxels to the occupancy of the corresponding voxels in the global map. This is repeated until all scans have been aligned overcoming one of the main problems with sequential real-time scan alignment, that of ensuring sufficient overlap. Sometimes there is not enough overlap between sequential scans. In the experiments scans are taken approximately every two meters. The most problematic situations occur when the robot travels from one room to another or around a corner. By comparing the map with all scans at each update step these situations are overcome. It should be noted that the sequential scan matching performs significantly better when the laser scanner is facing upwards as in Chapter 5 because the scanner returns a more omni-directional scan therefore the probability of two sequential scans having significant overlap is much higher.

6.4.2 Iterative Multi-resolution Localisation

Many methods exist for global localisation within a prior map however in this chapter an iterative multi-resolution sampling algorithm tuned for occupancy lists is implemented. The occupancy list maps are generated from the 3D laser scans at a resolution of 0.08m. The first scan is considered the frame for the global map of that robot until a mutual sighting is made. All subsequent scans are aligned with the first scan and the one with the highest overlap is taken and combined with the first scan. All remaining scans are then tested against this composite scan and the process repeated until all the scans have been added. At this resolution, an exhaustive search takes too long, so the occupancy list maps are downsampled by a factor of 16. The alignment is done at this downsampled resolution and a list of likely candidate poses are returned. The map is then downsampled by a factor of 8 and more uniformly distributed samples are taken around the likely candidate poses. The process is repeated until the alignment is done at 0.08m and the top pose from the candidates at this resolution is taken as the current map alignment. This process is summarised in Algorithm 8 and detailed

further in Section 5.3.3.

Algorithm 8 Iterative multi-resolution alignment.

```
for factor is 16, 8, 4, 2, 1 do
    Down sample both map and local by factor
    Align downsampled occupancy lists using guess poses as initial estimates
    Assign poses whose overlap exceeds 80% of the modal pose to guess poses
    set position error to resolution*factor*0.5
end for
```

To investigate whether the downsampling is valid for the pose probability distribution an exhaustive search global localisation within a map of size 20m by 20m with pose resolution of 0.1m and 1° is completed. This localisation takes of the order of minutes on a 2GHz processor; an example of the results of such an alignment are presented in Figure 6.6. In Figure 6.6 the two submaps (which are 3D but shown from overhead for clarity) are shown before and after alignment along with a surface plot showing the overlap count for the best orientation as a function of position. An exhaustive search is done in this situation for graphing the data to generate a smooth surface. This exhaustive search is not necessary for localisation and mapping but illustrates how the pose probability distribution has large regions of low probability and varies smoothly.

Each scan is converted into a submap at 0.08m resolution and the number of voxels overlapping are calculated by stepping through various positions and rotations with respect to the map. The poses that exceed an overlap of 80% of the modal pose are then returned for further inspection. The multi-resolution algorithm aligns scans of 20,000 points at 0.08m resolution to the nearest 0.25 degrees in 0.1 to 1 seconds which compares favourably to the minutes required for the equivalent exhaustive search. The resolution of the occupancy grid is an important consideration. Make the resolution too fine and storage and handling of the map are more difficult, too coarse and the efficacy of the map is reduced.

As discussed earlier one of the purposes of a good global representation scheme is to eliminate errors introduced by the sensor. In experiments the LMS performs exceedingly well however the main problem is pixel mixing. The removal of this error is a good case study of the use of occupancy maps in enhancing map quality. The

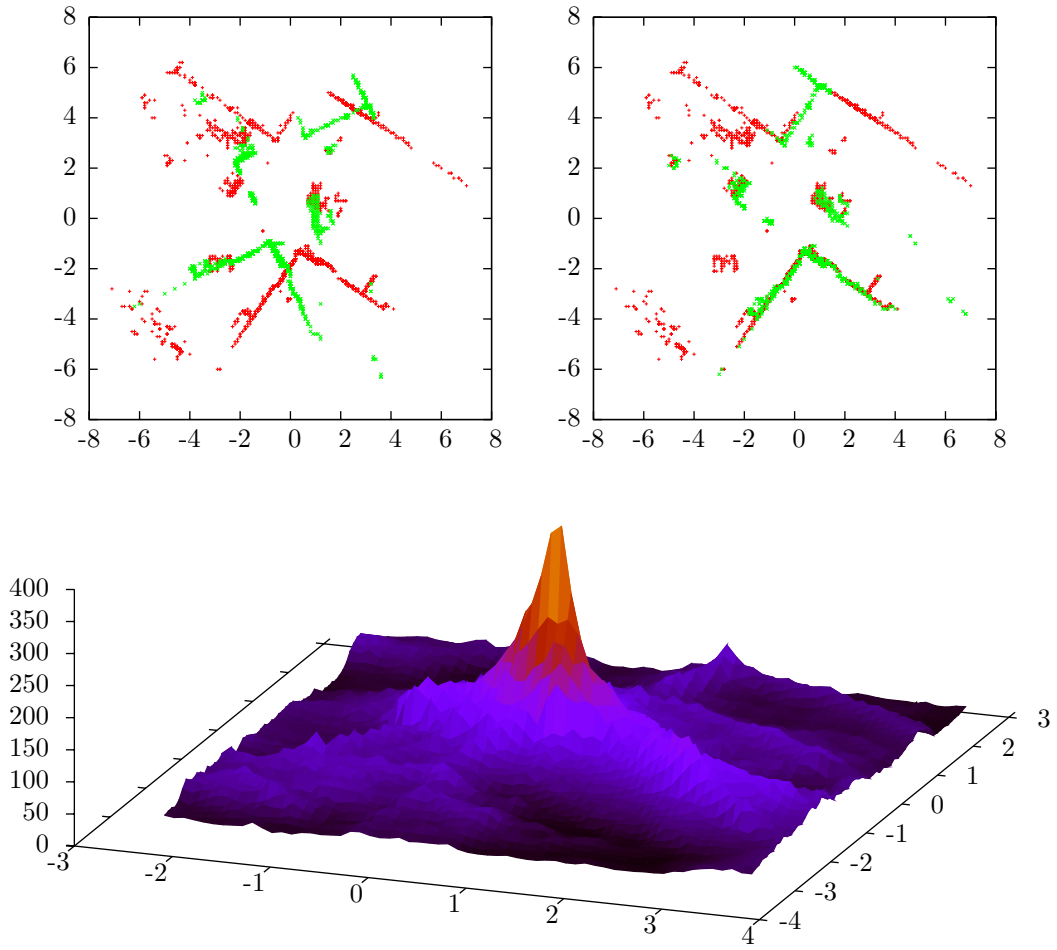

Figure 6.6: Alignment of two submaps with angular resolution of 0.5° and cell size of 0.1m. The scans are 3D however the cross-sections are shown containing voxels from 0-2.8m. The surface plot shows the voxel overlap count for the best orientation as a function of position.

Bayesian update process used to fuse new laser scans to the global map relies on various assumptions. One of the most important is the assumption that errors are random. Errors due to pixel mixing are only random if observations are taken from different positions. Many readings from positions giving rise to spurious pixel mixing errors could swamp the proper readings obtained from other positions. Thus it is necessary to take into account the position of each scan before it is added to the map checking that no other scan has been taken from a nearby position.

The computer graphics community has done copious work in producing compact representations of 3D environments that also allow fast operations such as nearest neighbour and orthogonal range searches. Octrees, binary space partition trees and kd-trees are all data structures capable of representing maps and further compression of the 3D data sets is possible by extracting features that are common in indoor environments, however this results in a loss of generality. These representations however do not have an impact on the theoretical basis of the map as they are lossless optimisations which increase the speed of operations. Moving from occupancy grids to occupancy lists, although lossy, has a far greater impact on the operation speed and negligible effect on the accuracy as explained in Section 5.3.1.

A horizontal cross-section of the map, the volume from 2.3 to 2.9m in height, is chosen for localisation. This vertical range is chosen for localisation however the global map has all the map data below 2.9m. In this way different levels of the map may be used for different purposes. The data above 2.3m is used for localisation whilst that below or near the floor is essential for path planning and obstacle avoidance. Selection of data from this height range removes most dynamic elements that exist in such environments such as other robots, people, doors and other moveable objects.

Several global localisation algorithms given a prior map exist however in this research two main approaches are implemented. The first exhaustive search has some advantages, the entire pose probability density function may be recorded and inspected, it is very robust and suffers only from its slow execution speed. It is useful for testing the validity of other approaches as it is guaranteed to find the best match. The impact

of this on real-time operation may be vastly reduced especially if a position estimate is available which is almost always the case.

The second approach is the multi-resolution particle filter which is described in Chapter 5. It is found that the execution speed of an unoptimised exhaustive search algorithm meant that global localisation could take place on maps of size 20m by 20m with pose resolution of 0.1m and 1° in the order of minutes on a 2GHz processor. For the implementation of the exhaustive search algorithm the 3D sensor scanner is converted into a 2D occupancy grid by extracting the cross-section near the ceiling. A correlation metric is then calculated iteratively for the sensor occupancy grid by stepping through various positions and rotations with respect to the map. This process is summarised by the pseudo code in Algorithm 7.

Typically, pose resolutions of 0.1m and 1° are used. The pose with the best correlation metric is then chosen with an example unnormalised position probability distribution plotted in Figure 6.6.

Figure 6.6 shows the voxel overlap count as a function of 2D position for an example scan alignment scenario. The reliability and accuracy of the pose may be ascertained by inspection of this overlap function. The true position is easily discerned and the dominance and shape of the peak denotes the reliability and error respectively. The number of poses with overlap counts exceeding a proportion (for example 80%) of the modal pose is a good indicator of the reliability of the modal pose with fewer poses indicating a single sharp peak and therefore a higher certainty that the modal pose is correct.

Usually mobile robots operate on the floor, consequently the 2D assumption may be enforced for global localisation. In most typical indoor environments moveable objects tend to lie on the floor. Floor-level 2D maps (as most in previous 2D localisation mapping research have been) are unreliable and quickly become out of date when furniture is moved and doors are opened or closed. To avoid these problems 2D map data is extracted as a horizontal cross-section just below the ceiling. This volume of a room changes infrequently and is more easily observed, especially in cluttered

environments.

6.5 Experimental Results and Analysis

As the thrust for this research is that cooperative mapping is more reliable and accurate than single robot mapping so experiments to test this are performed. The experiments consist of mapping an indoor environment with four robots. The number of robots chosen is limited by the experimental resources available. The odometry, scan data and inter-robot sightings are recorded. Maps are then generated using two algorithms. For the first, the map is generated for each robot individually and the presence of other robots and their data is ignored. This produces the four separate maps in Figure 6.7. The second algorithm considers the robots as cooperating sharing their data and generates the map in Figure 6.9. The true trajectories (as established by manual alignment of their 3D scans) of the robots during the experiment are plotted in Figure 6.10. All the maps are 3D with the floor and ceiling voxels removed to aid both performance and visualisation, however they are displayed as 2D plan views to improve the accuracy of comparison with voxel colour determined by its height from the floor.

6.5.1 Cooperative Scan Alignment

The ability for robots to discern others improves maps not just through cooperative and mutual localisation, but also allows the independence assumption of the robot scans to be upheld when updating the map. Before the map is updated with a scan in which another robot has been detected a cylindrical region of space (radius 0.2m and height 2m) around the robot sighting is cleared of data points. These points are likely to be due to another robot and consequently are not included in the map.

Figure 6.7 displays the maps resulting from each of the four robots individually without any cooperation or sharing of data. From Figure 6.7 it is clear how false alignments severely and almost irrevocably degrade the map quality when compared with the schematic of the arena, Figure 5.1, and the map built involving data from all

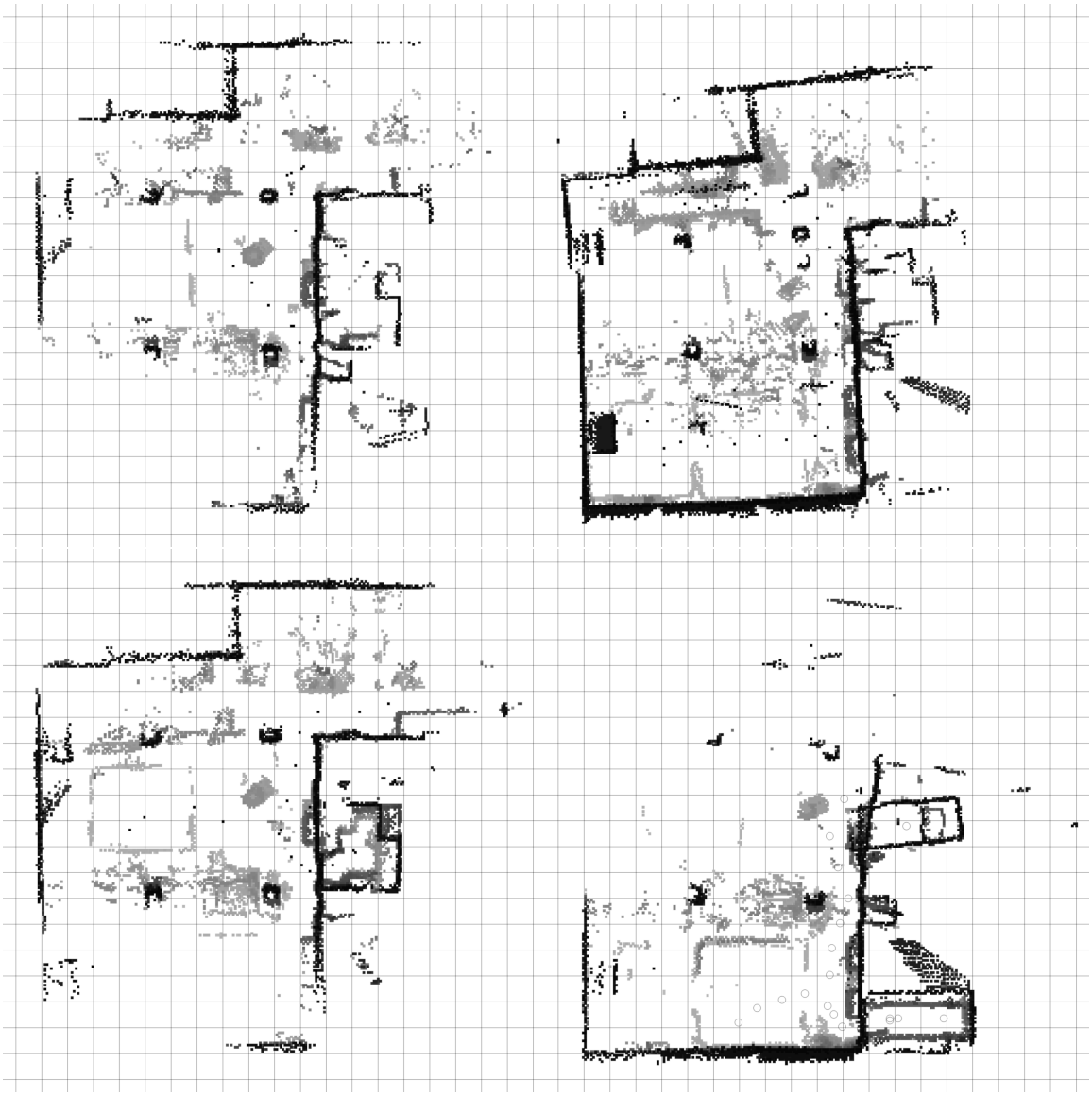

Figure 6.7: Single robot maps built using the iterative multi-resolution scan alignment described in Algorithm 8. The maps have been built in a batch manner by taking each scan and testing alignment overlap with every remaining scan, the best match is then combined into the map. Note the occasional large misalignment errors due to poor overlap of some scans with the map.

four robots, Figure 6.9. The individual paths of the robots are plotted in Figure 6.10 which shows regions where the robots are close and hence benefiting from cooperative mapping and points where a robot is isolated and effectively operating alone.

It is not so much accuracy that is important for mapping but rather reliability or certainty. It is preferable to align to a lower accuracy but with higher reliability and such considerations are necessary when deciding upon the alignment and map resolution. Loop-closing is another case in mapping where certainty is paramount, for an incorrect loop closure decision damages the large scale map structure, potentially irrevocably.

Single robots may maintain an approximate track of their position using odometry, however when multiple robots are involved there is usually no estimate for the relative positions between the robots. Before their maps may be combined the robots have to be sure of their relative positions and such certainty can be delivered by mutual localisation. Once the relative pose between the map frames of the robots has been initially established relative position estimates are available from odometry.

The robot odometry is artificially degraded, Figure 6.8, and this degraded odometry is supplied to both the single robot and multi-robot mapping algorithms. The original odometry is too accurate and both algorithms perform equally well. However, in other environments with different floor surfaces odometry accuracy is much worse. By increasing the odometry error it can be seen that the cooperative algorithm can tolerate odometry with 0.5m position error and 10° orientation error and the difference in performance between single and cooperative mapping is more apparent.

Mutual localisation produces composite scans with wide observation baselines where as sequential scan matching often aligns scans with small separations. Observing the environment from vantage points with larger separations results in better coverage. Through mutual localisation the large scale structure of the map can be established hence providing a framework within which further scans may be aligned to enhance map detail.

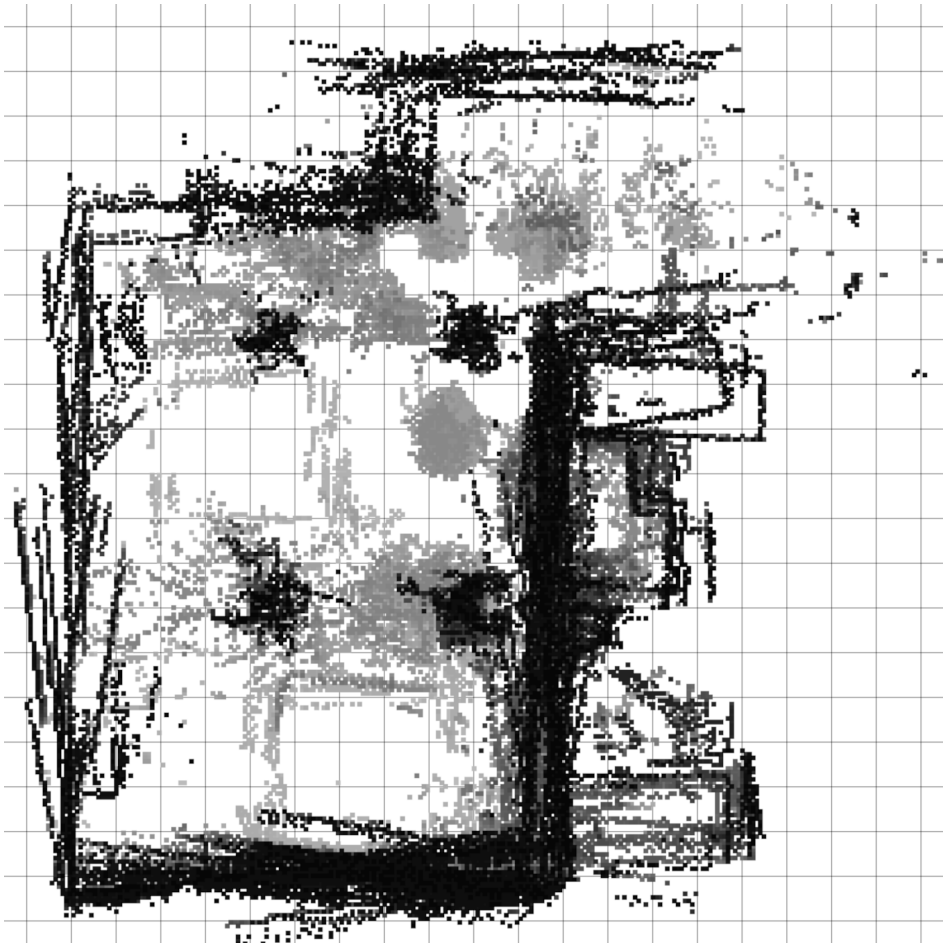

Figure 6.8: The first map built by combining scans from all robots with artificially degraded odometry having 0.5m position error and a heading error of 10°. Higher cells are darker.

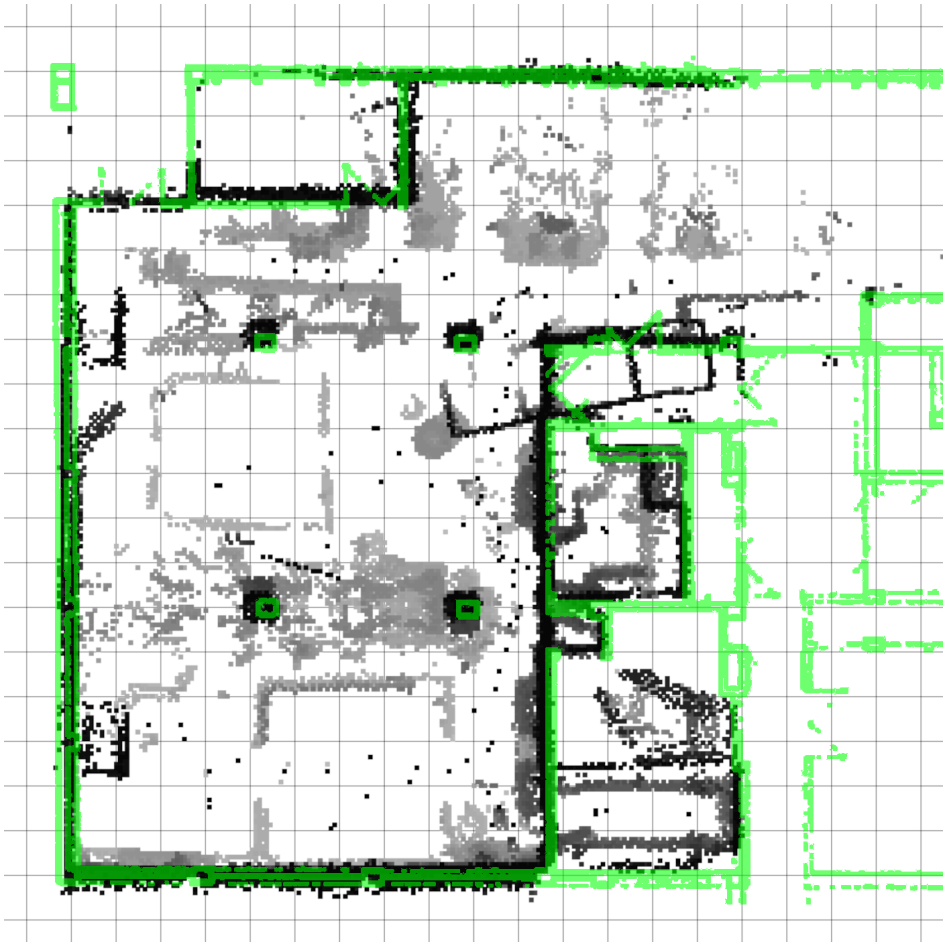

Figure 6.9: All scans for the trajectories in Figure 6.10, aligned using the iterative multi-resolution scan matching algorithm. Angular resolution of matching is carried out at 0.25°. The occupancy map has a cell resolution of 0.08m and includes 20 scans from each of the four robots. Higher voxels are darker. The ground truth schematic is superimposed in green.

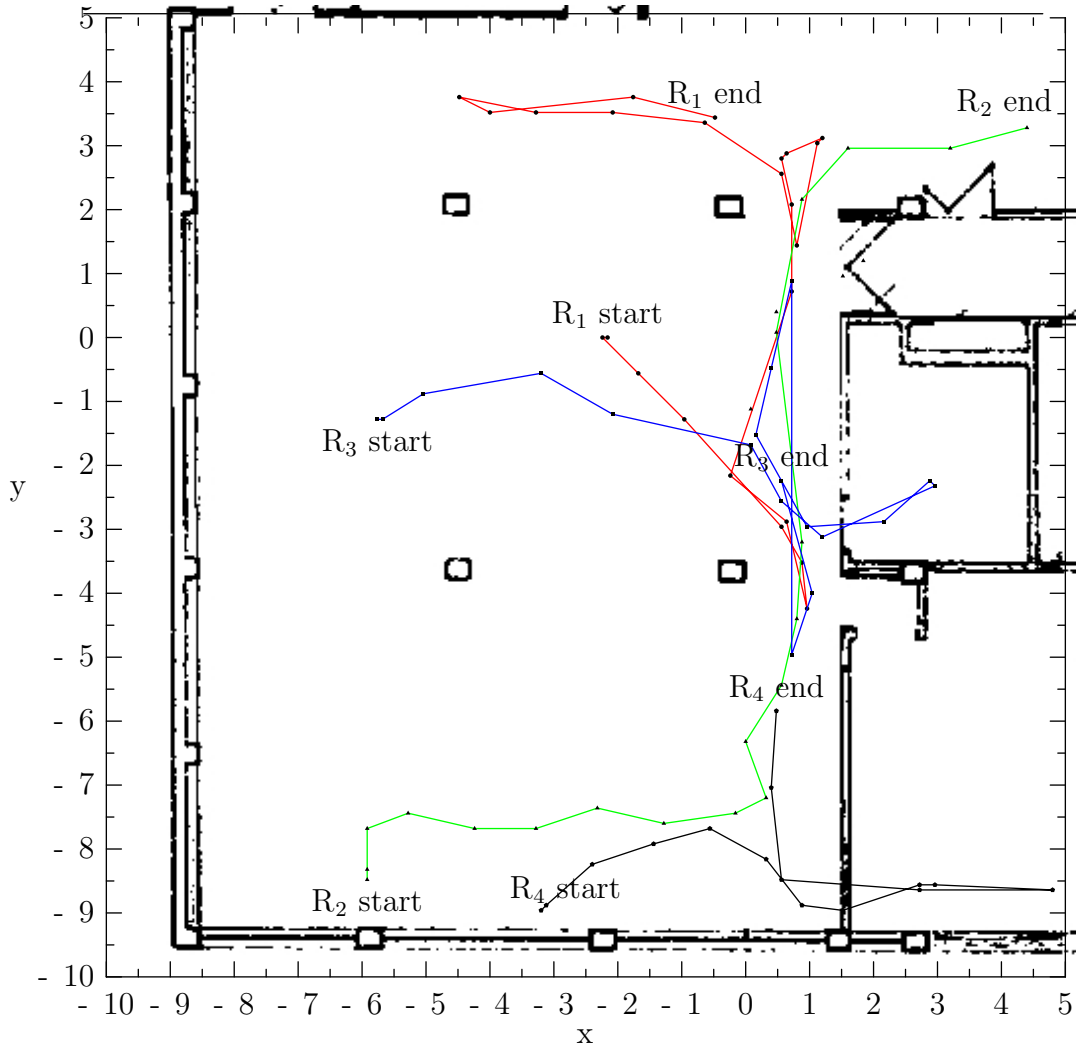

Figure 6.10: Paths of the four robots, $R_{1...4}$, and the laboratory schematic during cooperative mapping experiments for the maps in Figure 6.7 and Figure 6.9.

6.5.2 Map Quality Assessment

The accuracy of map building is assessed by comparing the robot positions with their true positions. The true positions for the robots are ascertained by manually aligning the scans and are plotted in Figure 6.10 along with a schematic of the laboratory. The error in individual scans has already been assessed and incorporating additional information (for instance the rectilinear nature of the walls) means these poses are within 0.07m of the true poses. Each robot performed single robot mapping in isolation and the results compared to the positions derived from cooperative mapping. The average position error over 80 scans for single robot mapping is 0.35m whilst that for cooperative mapping is 0.1m. This increase in accuracy is in itself a significant improvement, however what is also evident from Figure 6.11 is that cooperative mapping is more reliable. Figure 6.11 details the difference in position error distributions between maps built using single robot mapping and those produced cooperatively. Often the accuracy of the single robot mapping versus that for cooperative mapping is similar, however the single robot mapping is much less reliable, with large misalignments occurring which manifest themselves as a tail in the single robot distribution in Figure 6.11. These misalignments are clear when viewing the map from above as in Figure 6.7. This shows that one of the main areas of concern for mapping, reliability, is significantly improved by multiple robots sighting each other and sharing information in the map building process.

6.6 Summary

The ability of a mobile robot to develop an awareness of their surroundings and locate themselves within their operating environment is one of the fundamental problems currently thwarting autonomous mobile operation. A solution for an individual robot with a 3D laser scanner is presented in Chapter 5. This chapter explores the intricacies and effectiveness of extending single robot 3D mapping to multi-robot mapping when four are equipped with the 3D laser range finders described in Chapter 4 and retroreflective

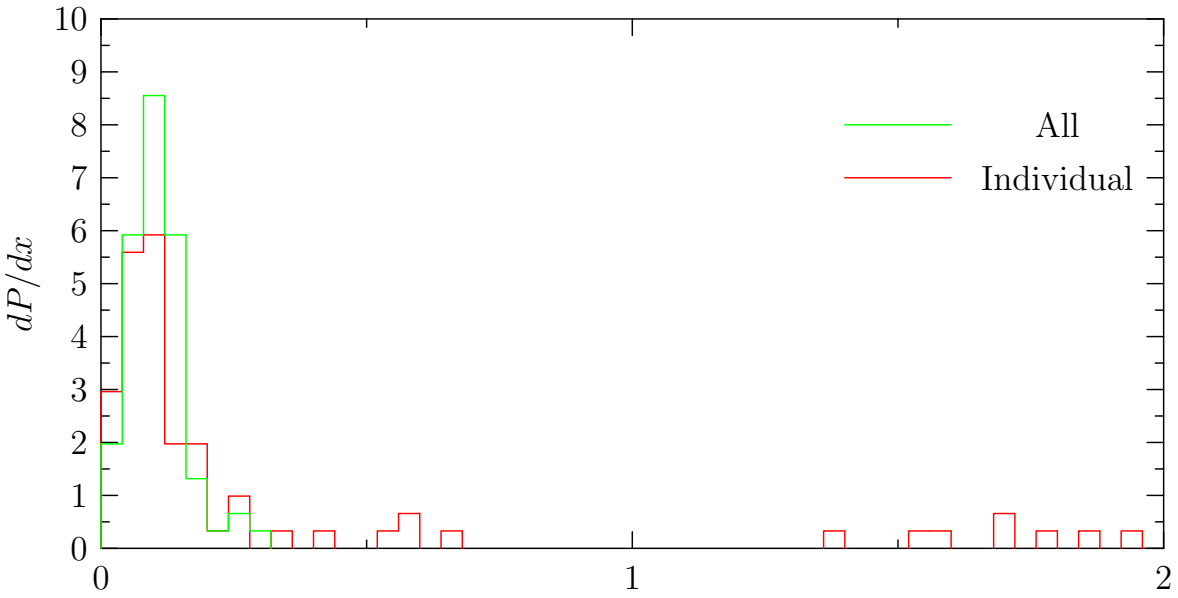

Figure 6.11: Distribution of robot position errors for single and multiple robots.

beacons. Related literature in the areas of how robots work as teams specifically how they may identify and locate one another is explored. The theoretical framework for cooperative mapping with mutual localisation in 3D is presented.

The method and algorithms for localisation and mapping are described and applied to a team of four robots mapping a large (20 by 20m) room. The quality of the maps are assessed in a number of ways. The schematics of the room are presented for comparison and the scans are aligned by hand to produce the true poses. The accuracy of the poses compared to the true ones along with the accuracy assessment of the 3D scanner undertaken in Chapter 4 gives a good indication of map quality. Finally one of the most important results is illustrated in Figure 6.11 which plots the distribution of position errors of robots mapping alone versus mapping as part of a team. For Figure 6.11 it is clear that although the modal position error is comparable for single and multi-robot mapping far fewer gross poses errors occur with multi-robot mapping where as with single robot mapping there are a number of positional errors exceeding 1.5m. Thus it may be concluded that although the accuracy of multi-robot 3D mapping is comparable to a robot mapping alone the reliability is far superior. This reliability is important when constructing maps for one large error can easily corrupt the map beyond usefulness especially if later alignments compound the problem.

The methods proposed in this chapter can be improved in a number of ways by future research. Currently all path planning and movement commands are given by a human. Implementing local obstacle avoidance and path planning would increase the mobile robots autonomy. This would also bring about the need for real time operation.

Chapter 7

Conclusions and Future Work

This chapter includes a description of the principle original contributions of this dissertation by chapter. A summary of the research undertaken and the reasons for it are presented as well as a look towards future work that awaits investigation.

7.1 Summary

For decades mobile robots have been shackled by the 2D assumption, which has been necessary due to the absence of suitable 3D sensors. Even today 3D sensing technology is problematic and expensive. In this research a rotating mirror mechanism has been added to a standard SICK LMS 200 laser scanner thus creating a 3D laser range scanner. These 3D scans coupled with mutual localisation have produced full 3D environmental representations inexpensively. The fidelity of the representations is validated in various quantitative ways, by performing global localisation within them, comparing dimensions with their real-world equivalents, comparing maps to those made by manually aligning scans and comparison of repeated scans. Although these manual alignments cannot be done in real time, all data from the mapping experiments is logged and manual alignment is undertaken afterwards to establish the ground truth.

The ultimate aim of this research is to enable mobile robots to sense and map their environment in full 3D, alone and in conjunction with other robots. This has largely been achieved however before this was possible a number of research contributions had

to be established which are detailed in Section 7.2.

The problem of extending traditional occupancy grids into 3D mapping is tackled with a representation dubbed occupancy lists, which maintain all the occupied cells in a hash map structure. An iterative multi-resolution sampling alignment algorithm has been developed to correctly combine these scans to produce accurate environment representations. The fidelity of these representations has been validated by comparing the localisation results produced during map building with those derived from aligning the scans manually.

The resulting maps are 3D occupancy lists and although they are interesting for humans to inspect their primary purpose is to provide information to robots for executing various indoor tasks. Thus the accuracy of global localisation performed with them is a good indicator of their fidelity and suitability for further use by mobile robots. Global localisation in these maps delivers poses accurate to 0.03m and 1°.

The 3D scanner has been deployed on four mobile robots so that the 3D scans coupled with mutual localisation produce full 3D environmental representations at low cost. An important result central to the thesis of this research is a quantitative comparison of single robot versus multi-robot 3D mapping. What is observed is that although the local accuracy of the map is similar the reliability is improved. This reliability is especially important for map building because the occasional large misalignment affects the usefulness of the map much more significantly than more frequent smaller errors in localisation. This result is succinctly summarised in Figure 6.11 which shows that using four robots improves the reliability of map building over single robots by a factor of 3.5.

7.2 Contributions

The aim of this research is to tackle the perception end of the sense-decide-act robotic chain. The focus is on two premises. The first, that multiple robots can map more accurately and reliably. The second, that sensors that deliver true 3D data from the environs of a robot improve the operation of mobile robots especially in indoor

environments. These have been successfully verified experimentally. An example of the latter is shown in Figure 5.7 which displays how maps consisting of only data above 2.8m are more similar to the schematic ground truth. These maps will be less susceptible to becoming more inaccurate over time in dynamic cluttered environments. The following contributions are made in each chapter.

Chapter 2 develops a novel algorithm for the fast detection of circular features from range data. The algorithm is based on least squares fitting however is highly optimised for the nature of sensor range data. The performance of the algorithm is compared to that of a conventional Hough transform based approach. The least-squares based approach is not only more accurate but is also more reliable. The computational performance is better in terms of both processing time and memory usage.

The main contribution of Chapter 3 is an extension of cooperative localisation to a process referred to has mutual localisation. Mutual localisation is when robots localise relative to one anther by observing other robots and being simultaneously observed themselves. An extensive error analysis of mutual localisation is conducted and compared to conventional cooperative localisation.

Chapter 4 explains the novel 3D laser scanner which enabled the collection of 3D data. Whilst 3D range sensors are not entirely new other researchers either use some form of stereoscopic vision or rotate the entire 2D laser scanner. The approach described in this chapter is different in that the 2D scanner is stationary and a rotating mirror is placed above it. This physical configuration has distinct advantages in terms of simplicity of hardware required, lower power consumption and expense.

Chapter 5 proposes two main contributions. The first is a compact representation suitable for probabilistic storage of the large amounts of 3D data produced during the mapping process. These so called occupancy lists are similar to occupancy grids however they only maintain lists of high probability occupancy voxels and so are eminently scalable to 3D. Secondly, discussed in Chapter 5 is a mechanism for both the fast alignment of sequential occupancy lists and for fast localisation within a global map occupancy list. This mechanism is similar to a multi-resolution particle filter.

The contributions conclude with Chapter 6 which builds upon much of the work described in the preceding chapters. The chapter concludes with a comparison of the mapping performance between a single robot mapping in 3D and multiple robots working in concert and establishes that although multi-robot mapping has similar errors to that done by a single robot the reliability is markedly improved.

7.2.1 Publications

A number of peer review publications have arisen from this research and are listed.

- J. Ryde and H. Hu, Mobile Robot 3D Perception and Mapping with Multi-Resolution Occupancy Lists, IEEE International Conference on Mechatronics and Automation (ICMA 2007), Harbin, China, 5-8 August 2007.

- J. Ryde and H. Hu, Cooperative mutual 3D laser mapping and localization, IEEE International Conference on Robotics and Biomimetics, Kunming, China, 17-20 December 2006. *Awarded Best Conference Paper.*

- J. Ryde and H. Hu, Mutual localisation and 3D mapping by cooperative mobile robots, The 9th International Conference on Intelligent Autonomous Systems (IAS-9), The University of Tokyo, Tokyo, Japan, 7-9 March 2006.

- J. Ryde and H. Hu, Laser based simultaneous mutual localisation for multiple mobile robots, Proceedings of IEEE International Conference on Mechatronics and Automation, Niagara Falls, Canada, 29-31 July 2005. *Finalist for Best Student Paper.*

- J. Ryde and H. Hu, Fast circular landmark detection for cooperative localisation and mapping, IEEE International Conference on Robotics and Automation, pages 2756-2761, Barcelona, Spain, 18-22 April 2005.

- H. Hu, J. Ryde and J. Shen, Chapter 4 Landmarks and Triangulation, in the Book of Autonomous Mobile Robots, Marcel Dekker Inc., pages 149-186, April 2006.

7.3 Future Work

The future work will investigate further the deployment of the enhanced 3D laser scanner developed in Chapter 4 in both indoor and outdoor environments. Success with mapping using the sub-ceiling cross-section is noted and it would be beneficial to test the robustness of this to dynamic environments where it is expected to excel.

The scanner is capable of recording reflectance data. Incorporating this data into the occupancy list maps would make certain aspects of the environment more distinguishable, improving reliability and aiding in the pending loop closing work.

Recent research into the human brain is beginning to show how difficult it is to separate perception and action activity within the brain. Indeed the best approach may be to work from either end of the perception-actuation coupling and eventually the gap will be small enough to close. This dissertation focuses on the perception end of the mobile robotics problem and it is important to realise the distinction between perception and sensing. Whilst sensing is the recording of data from the environment surrounding a robot, perception involves the production of information and combining it with what is already known. In a sense perception is the combination of sensing and the internal representation. In this dissertation both sensing and representation are tackled in order to improve the perception of mobile robots. It is believed that with improved perception it will be easier to improve the autonomy of mobile robots so that they may undertake useful tasks outside the laboratory setting.

It is shown throughout this research that robots with 3D sensors can localise and map much more effectively and that occupancy lists are an adequate representation for these purposes. What still needs to be tested is the performance of conventional path planning algorithms with occupancy lists. It is hypothesised that occupancy grids suitable for path planning can be generated from the occupancy lists and observation poses. This would be done by ray tracing, taking into account the sensor model including the field of view from each observation pose, to the nearest voxel along each ray and generating a free space map for a given volume, for instance that just above the floor to the height of the robot. This would produce a 2D occupancy grid of the

navigable free space available for robot motion.

Due to the fast execution speed of the multi-resolution particle filter for 3D localisation it would be possible to pre-process occupancy list maps for regions of direct isometry. This would help avoid false loop closures and would allow for active localisation algorithms. In terms of algorithm performance the multi-resolution particle filter although adequately fast on a standard CPU is amenable to parallel computation and so could easily be implemented on a parallel architecture such as a FPGA or as is becoming increasingly popular, due to their cost and prevalence, on 3D graphics cards. Real time performance is not directly tested in this research, however the multi-resolution particle filter is capable of executing each update in under a second on a standard desktop computer and the 3D scanner can be configured to produce a 3D scan each second. Therefore, although the method is close to real-time performance further investigation into the real time performance and ways of improving it would be most beneficial.

An important study would be a comparison of the multi-resolution particle filter with competing mechanisms for aligning multiple 3D scans, for instance iterative closest point (ICP) based algorithms. The particle filter approach is effective for aligning when there is large error on the initial estimate and ICP performs well for fine alignments when the true result is close to the initial estimate. Therefore it can be expected that a hybrid approach would deliver superior performance. In this hybrid approach the approximate alignment is achieved by the multi-resolution particle filter and then this alignment is fine tuned by ICP.

As discussed in Section 6.3.2 map entropy is an important quantity,

$$H(M) = -\sum_{x \in S} p(x) \ln p(x). \tag{7.1}$$

If the probability of occupancy of unlisted voxels is p_0 then the entropy of the occupancy list is

$$H(M) = -\left(\frac{V}{\epsilon^3} - |M|\right) p_0 \ln p_0 - \sum_{i,j,k} p(M_{ijk}) \ln p(M_{ijk}) \tag{7.2}$$

where V is the volume of the corresponding occupancy grid, ϵ is the voxel size and M_{ijk} is a particular occupancy list map voxel occupancy.

The occupancy list stores the log odds of the probability therefore the probability of occupancy of a listed map voxel is

$$p(M_{ijk}) = \frac{1}{e^{-M_{ijk}} + 1}. \tag{7.3}$$

This results in the final expression for the occupancy list entropy as

$$H(M) = \left(|M| - \frac{V}{\epsilon^3} \right) p_0 \ln p_0 + \sum_{i,j,k} \frac{\ln \left(e^{-M_{ijk}} + 1 \right)}{e^{-M_{ijk}} + 1}. \tag{7.4}$$

It would be worth investigating how this entropy and other statistics of the occupancy list global map varied over time throughout various mapping missions.

The physical implementation can be improved so that ad hoc networking is used rather than relying on an existing wireless infrastructure and would thus enable true autonomous operation in any indoor environment.

The immediate future of mobile robotics is difficult to foretell. At some point mobile robotic technology will be ubiquitous, however the question remains, when will this be? It is tempting to draw an analogy between the mobile robots of today with the first personal computers of the 1980s and therefore infer the corresponding impact on human society during the following decades.

Like many fledgling technologies it requires a coincidence of both abilities and needs. Finding these coincident points where the abilities of mobile robots are sufficient to perform tasks that people desire in a manner better than existing solutions is paramount. Otherwise mobile robots will be constrained to the laboratory with limited interest from and impact on non-specialists.

Mobile robotics, as a research field, suffers from a large disparity between the expectations of those outside the field to the abilities of the mobile robots currently possible. We must close this disparity not by lowering expectations but by surpassing them.

Appendix A

3D Vision

This appendix briefly introduces a novel and inexpensive mechanism for acquiring 3D data suitable for indoor mobile robots. The scanner is currently capable of ranging up to 4m with an accuracy of 0.1m and costs less than $200. The data is returned from all directions around the robot from floor level to a height of 1m. This data would be eminently suitable for navigation, mapping and obstacle avoidance. The scan rate is dependent on the frame rate of the camera and the rotation speed. Sub-degree angular resolution (both horizontally and vertically) scans are possible in 4 seconds with a conventional webcam.

A.1 Introduction

The 3D sensors in this dissertation are modified from standard 2D laser scanners. They produce good results and are substantially less expensive than commercial 3D scanners. The hardware improvements cost approximately $300 and this is in addition to the price of the 2D laser scanner. They are ideally suited for research groups which already own the 2D scanners.

There is an even less expensive alternative which employs a laser line, camera and triangulation to determine range which is described in this Appendix. As has been discussed in Section 4.2.1 the main problem with triangulation ranging is the rapid deterioration of accuracy with range. Consequently, this scanner is less accurate but

various design decisions have improved this accuracy to a level that is acceptable (0.1m error up to 4m range) for indoor mobile robots. The cost for the system consisting of laser line projector, camera and rotating mechanism varies depending on the quality of components however is realised for under $200 in these experiments.

A.2 Scanner Design

For ranging via triangulation the accuracy falls rapidly with range and thus it is important to maintain as large a baseline separation as possible. Environments designed for humans to operate within tend to have more space vertically. Thus vertical separation is preferred because a separation of up to 2m is possible for most human environments whilst horizontally separations of only 0.5m are possible. Taller robots interact better with humans and have improved observational capability. As there is usually more clutter on the floor sensors close to the floor tend to be blocked more often than those higher up.

In camera and laser line projection systems the laser plane is often perpendicular to the separation between the camera and laser. The sensitivity of the range error on the separation means that it is important to maintain the greatest accuracy in the angular difference between the camera and laser. Scanning a horizontal laser plane up and down is less accurate than maintaining a fixed angle between the camera and laser. Secondly, if there is relative motion between the camera and laser then during the exposure time of the camera the laser line image will move across the sensor. This results in blurring the perceived line which reduces its effective brightness and the added apparent width is detrimental to the accuracy. Thus it is preferable for the angular velocity of both the camera and the laser to be identical. Finally, as is demonstrated in Chapter 5 omni-directional sensors are ideal for robots. They allow them to map more effectively by improving the overlap between successive scans and will be useful for planning and improving situational awareness.

The constraints of vertical separation, fixed relative angular difference between camera and laser and a preference for an omnidirectional sensor are the reasoning for

the scanner in Figure A.1. The scanner consists of a camera and laser line projector both mounted on a rotating platform and separated vertically. An encoder disk is placed just below the camera such that its circumference is visible at the bottom of the camera frame. The laser line is projected at an angle rather than horizontally so that a volume is swept as both the camera and laser line projector rotate about the vertical axis.

The two main differences for this range imaging device compared to conventional devices are the projection of the laser plane at an angle (rather than perpendicular to the camera-laser separation vector) and the feedback of angular position via the camera image itself.

The image received by the camera is shown in Figure A.1. At the bottom of the image is an encoder wheel divided into 32 sectors with binary labels. By reading the binary code and determining the offset the angular position of both the camera and laser is determined by the robot to sub degree accuracy. Indeed for a camera the angular accuracy approaches the angular resolution of the camera if the offset is pixel accurate. In experiments pixel accuracy is rarely achieved because the camera is focused further away thus blurring the encoder disk which is much closer to the lens (0.1m). The question of which point to focus the lens is important. As the range accuracy is much poorer at larger distances it is vital that the camera is focused at these distances (4m) in order to extract the angle of the laser line to the nearest pixel. If the laser line appears blurred at close distances this is not so problematic because the reduction in accuracy due to blurring is more than compensated for by the increase in range accuracy due to the improved geometric configuration. Thus for a camera with horizontal resolution of 640 pixels and field of view of 35° the angular resolution is potentially as good as 0.05°.

Including the horizontal angular feedback in the image itself rather than via more conventional manners confers a number of advantages. Firstly the inclusion of feedback means that the motion does not have to be precisely controlled and therefore there is substantial flexibility in the choice of scanning actuation mechanisms. The main

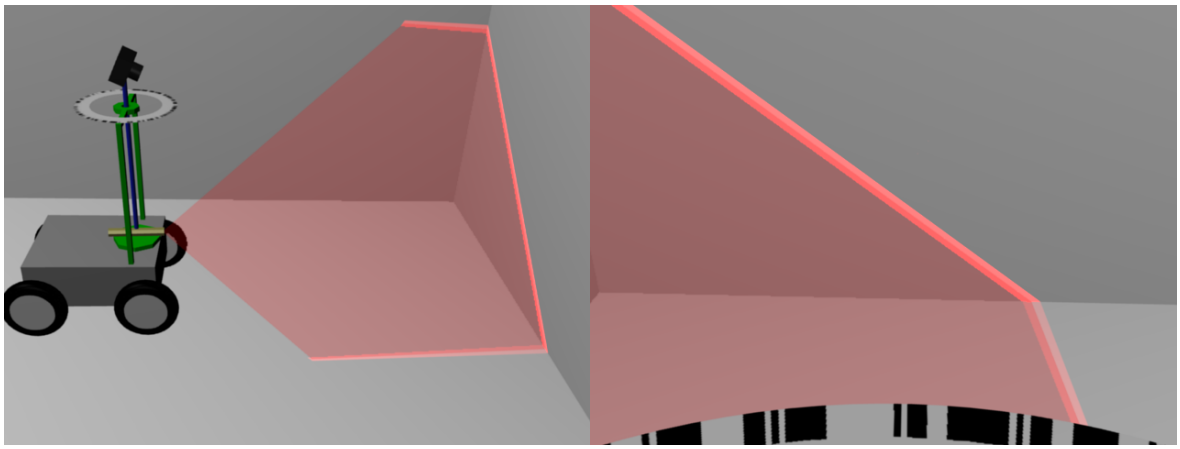

Figure A.1: Diagram of the 3D vision system operating on a robot and image from the point of view of the camera.

practical constraint is that the motion should be smooth so as to reduce errors in the relative angle due to the twisting of the rotating mechanism joining the camera and laser. The advantages over more conventional rotation sensors are the very high angular resolution, guaranteed data synchrony and reduced cost and complexity in feedback of the rotation sensor data to the robot. The data synchrony refers to the problem of ensuring which angular data matches a particular camera frame and although simple theoretically can be difficult to implement in practice. This problem becomes more difficult to overcome at higher scan speeds. With all the data necessary included in one camera frame, the range and angular reading are guaranteed to have been acquired simultaneously.

By projecting the laser line at an angle, different columns in the image correspond to different angular separations between the camera and laser point. This is much more accurate than actuating the laser to different angles.

The 3D position of each point in the image is determined by intersecting the camera ray with the laser plane. The camera ray starts at the camera position and the angle calculated from the pixel row and column. The laser plane may be established in a number of ways but for this work it is defined by the coordinates of three points. These may be obtained by projecting the line onto a vertical wall and measuring the position and height of two illuminated points on the wall. The coordinates of these two

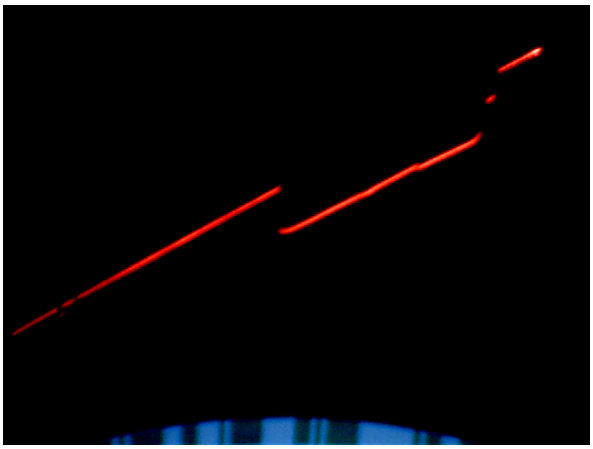

Figure A.2: Single webcam frame during scanning.

points coupled with the position and height of the laser aperture itself provide enough information to specify the plane. Given the plane and ray the intersection point is calculated by established 3D geometric methods.

A.3 Preliminary Results

A single camera image from from the operating scanner is displayed in Figure A.2. This experiment was done in complete darkness to enhance the reliability of laser line extraction, however scans with subdued indoor lighting have been successful. The bottom of Figure A.2 shows the stationary encoder wheel, the position of which is extracted using edge detection which works satisfactorily despite the blurring. The camera is focused at the limit of the range accuracy at about 4m. Also note from Figure A.2 that despite the rotation motion of the scanner the laser line is perfectly focused with very little motion blur. This is due to the smoothness of the surface being scanned and that there is no relative rotation between the laser and the camera.

A single scan is displayed from two points of view in Figure A.3. From Figure A.3 it is possible to see the diagonal line stripes each of which correspond to one camera frame. The overhead view in Figure A.3 is presented to give an indication of the scanner accuracy. The walls are vertical and should appear as thin lines in the overhead view. The width of these lines represents the random error and with the grid size of 1m it can be seen that this error is less than 0.1m.

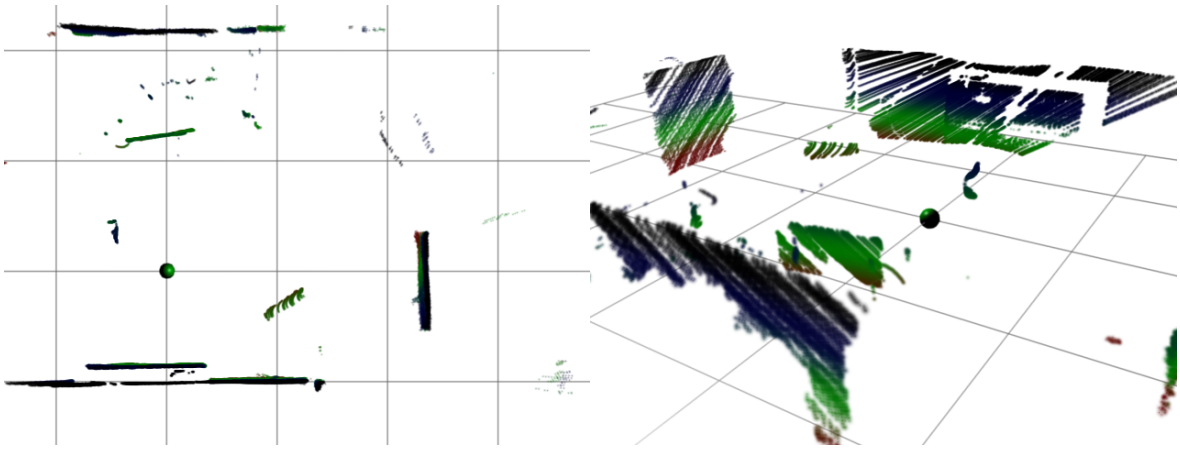

Figure A.3: Overhead and side views of a single scan from the camera scanner. The robot position is marked by a green sphere and the scan points are coloured by height. The grid lines are 1m apart.

A.4 Conclusions and Future Work

Although the accuracy degrades quite rapidly for larger ranges (> 4m) the laser scanner described in this Appendix offers some significant advantages. First and foremost it is currently two orders of magnitude less expensive than commercial time of flight scanning laser range finding systems. It also has the capacity to return the texture information associated with each range reading. This colour information is useful not only to the mobile robot but also for humans interpreting the data and matching it to what they themselves observe. It is accurate over the range most often required of indoor mobile robots to a level necessary to perform the usual tasks requested of a mobile robot. The camera sensor may also be used by standard vision algorithms where necessary such as face recognition or when the robot is to be teleoperated. Without the laser line panoramic images can be acquired with the multiple frames easily stitched together because the camera angle for each frame is known precisely.

Currently the laser line is produced via a cylindrical lens from a $<$1mW laser for eye safety. The scans are done in low light conditions to simplify the laser line extraction image processing. Whilst operation in darkness is an advantage, future work would extend operation to brighter ambient light conditions. This could potentially be achieved by increasing the brightness and quality of the laser line.

Bibliography

K. Åström. Automatic mapmaking. *IFAC Int. Workshop on Intelligent Autonomous Vehicles*, 1993. 14, 46

A. Aboshosha and A. Zell. Robust mapping and path planning for indoor robots based on sensor integration of sonar and a 2D laser range finder. *IEEE 7th Int. Conf. on Intelligent Engineering Systems*, 2003. 18

P. Ahman, J. Forsberg, and A. Wernersson. The Hough transform inside the feedback loop of a mobile robot. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, volume 1, pages 791–798, 1993. 17

R. Alur, A. Das, J. Esposito, R. Fierro, Y. Hur, G. Grudic, V. Kumar, I. Lee, J. Ostrowski, G. Pappas, J. Southall, J. Spletzer, and C. Taylor. A framework and architecture for multi-robot coordination. *The Int. Journal of Robotics Research*, 21 (10-11):977–955, 2002. 46

K. O. Arras and S. J. Vestli. Hybrid, high-precision localization for the mail distributing mobile robot system MOPS. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 1998. 44

D. Austin and P. Jensfelt. Using multiple gaussian hypotheses to represent probability distributions for mobile robot localization. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2000. 44

D. Austin, P. Jensfelt, O. Wijk, and M. Andersson. Feature based condensation for mo-

bile robot localization. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2000. 97

N. Bergman, F. G. F. Gustafsson, and U. Forssell. Particle filters for positioning, navigation and tracking. *IEEE Trans. on Signal Processing*, 2002. 45, 97

P. Besl and N. McKay. A method for registration of 3-D shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992. 7

M. Betke and L. Gurvits. Mobile robot localization using landmarks. *IEEE Trans. on Robotics and Automation*, 13(2):251–263, 1997. 14, 15

J. Borenstein and L. Feng. UMBmark: A benchmark test for measuring odometry errors in mobile robots. *Proc. of the 1995 SPIE Conf. on Mobile Robots*, 1995. 127

J. Borenstein and L. Feng. Measurement and correction of systematic odometry errors in mobile robots. *IEEE Trans. on Robotics and Automation*, 12(6):869–880, December 1996. 127

J. Borenstein and D. K. Wehe. Internal correction of odometry errors with the omnimate. In *Proc. of the Seventh Topical Meeting on Robotics and Remote Systems*, pages 323–329, 1997. 129

J. Borenstein, H. R. Everett, and L. Feng. *Navigating Mobile Robots: Systems and Techniques*. A. K. Peters, Ltd., Wellesley, MA, 1996. 15

M. Brady. Artificial intelligence and robotics. *Artificial Intelligence*, 26:79–121, 1985. 1

W. Burgard, D. Fox, D. Hennig, and T. Schmidt. Estimating the absolute position of a mobile robot using position probability grids. In *AAAI/IAAI, Vol. 2*, pages 896–901, 1996. 41, 99

W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, 114(1–2):3–55, 1999. 71

W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun. Collaborative multi-robot exploration. *IEEE Int. Conf. on Robotics and Automation*, 2000. 125

D. Burschka and G. Hager. Laser-based position tracking and map generation. In *Proc. of the IASTED Int. Conf. on Robotics and Automation*, Aug. 2000. 99

J. A. Castellanos, K. O. Arras, and R. Siegwart. Feature-based multi-hypothesis localization and tracking for mobile robots using geometric constraints. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2002. 44

N. Chernov and C. Lesort. Least squares fitting of circles and lines. *Computer Vision and Pattern Recognition*, 2003. 25

H. Choset and K. Nagatani. Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization. *IEEE Trans. on Robotics and Automation*, 17(2):125–137, 2001. 101

I. J. Cox. A review of statistical data association techiques for motion correspondence. *Int. Journal of Computer Vision*, 10(1):53–66, 1993. 44

I. J. Cox and J. J. Leonard. Modeling a dynamic environment using a bayesian multiple hypothesis approach. *Artificial Intelligence*, 66:311–344, 1994. 44

M. C. Deans. Maximally informative statistics for localization and mapping. *IEEE Int. Conf. on Robotics and Automation*, 2002. 44

S. J. G. A. Doucet and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000. 45

G. Dudek and M. Jenkin. *Computational Principles of Mobile Robotics*. Cambridge University Press, 2000. 3

H. F. Durrant-Whyte, M. W. M. Dissanayake, P. Newman, S. Clark, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Trans. on Robotics and Automation*, 17(3):229–241, 2001. 44

A. Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal of Robotics and Automation*, 1987. 99

A. Elfes. Dynamic control of robot perception using stochastic spatial models. *Information Processing in Mobile Robots*, 1991. 47, 100

J. Forsberg, U. Larsson, and A. Wernersson. Mobile robot navigation using the range-weighted Hough transform. *IEEE Robotics and Automation Magazine*, 2(1):18–26, 1995. URL `http://robotics.caltech.edu/readinggroup/forberg.pdf`. 22

D. Fox. Adapting the sample size in particle filters through KLD-sampling. *Int. Journal of Robotics Research*, 22(12):985–1003, Dec. 2003. 41, 97

D. Fox. *Markov Localization: A probabilistic framework for mobile robot localization and navigation.* PhD thesis, University of Bonn, 1998. 41, 99

D. Fox and W. Burgard. Collaborative multi-robot localization. *Kunstliche Intelligenz*, 1999. 97, 125

D. Fox and W. Burgard. A probabilistic approach to collaborative multi-robot localization. *Autonomous Robots*, 8(3):325–344, 2000. 125, 132

D. Fox, W. Burgard, A. Derr, and A. B. Cremers. Integrating global position estimation and position tracking for mobile robots: the dynamic Markov localization approach. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 1998a. 99

D. Fox, J. S. Gutmann, W. Burgard, and K. Konolige. An experimental comparison of localization methods. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 1998b. 44

D. Fox, F. Dellaert, W. Burgard, and S. Thrun. Using the condenstation algorithm for robust, vision-based mobile robot localization. In *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 1999. 97

B. Gerkey, R. Vaughan, and A. Howard. The player/stage project: Tools for multi-robot and distributed sensor systems. In *Proc. of the 11th Int. Conf. on Advanced Robotics*, pages 317–323, Coimbra, Portugal, 2003. 57, 76

H. H. González-Baños and J. Latombe. Navigation strategies for exploring indoor environments. *The Int. Journal of Robotics Research*, 21(10-11):829–848, October-November 2002. 99, 100

R. Grabowski and P. Khosla. Localization techniques for a team of small robots. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2001. 47, 124

D. Hähnel, D. Schulz, and W. Burgard. Map building with mobile robots in populated environments. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2002. 70

P. V. C. Hough. Methods and means for recognising complex patterns. *U.S. Patent 3 069 654*, 1962. 17, 21

A. Howard, M. J. Matarić, and G. S. Sukhatme. Localization for mobile robot teams: A distributed MLE approach. *Experimental Robotics VIII*, pages 146–155, 2003. x, 47, 127, 128

A. Howard, A. Parker, E. Lynne, and G. S. Sukhatme. The SDR experience: Experiments with a large-scale heterogenous mobile robot team (extended abstract). *Int. Symposium on Experimental Robotics*, June 2004a. 70, 126

A. Howard, D. F. Wolf, and G. S. Sukhatme. Towards 3D mapping in large urban environments. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 419–424, Sendai, Japan, Sept. 2004b. 70

H. Hu and D. Gu. Landmark-based navigation of industrial mobile robots. *Industrial Robot: An Int. Journal*, 27(6):458–467, 2000. 46, 98

C. Jennings, D. Murray, and J. J. Little. Cooperative robot localization with vision-based mapping. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, May 1999. 46

P. Jensfelt and H. Christensen. Pose tracking using laser scanning and minimalistic environmental models. *IEEE Trans. on Robotics and Automation*, 2001. 17

P. Jensfelt and S. Kristensen. Active global localization for a mobile robot using multiple hypothesis tracking. *IEEE Trans. on Robotics and Automation*, 17(5):748–760, 2001. 41, 44, 135

W. Jeong and K. Lee. CV-SLAM: A new ceiling vision-based SLAM technique. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2005. 71

R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 82:35–45, 1960. 43

A. Kelly. Linearized error propagation in odometry. *The Int. Journal of Robotics Research*, 2004. 17, 129

P. Khosla, R. Grabowski, and H. Choset. An enhanced occupancy map for exploration via pose separation. In *Proceedings of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2003. 47, 124

P. Khosla, R. Grabowski, and H. Choset. Development and deployment of a line of sight virtual sensor for heterogeneous teams. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2004. 46

K. Konolige. Improved occupancy grids for map building. *Autonomous Robots*, 4(1): 351–367, 1997. 54

K. Konolige and S. Gutmann. Incremental mapping of large cyclic environments. *Int. Symposium on Computer Intelligence in Robotics and Automation*, 1999. x, 45, 99, 106, 124, 125, 126

K. Konolige, C. Ortiz, R. Vincent, A. Agno, M. Eriksen, B. Limketkai, M. Lewis, L. Briesemeister, and E. Ruspini. DARPA software for distributed robotics. Technical report, Artificial Intelligence Center, SRI Int., 2002. 126, 127

B. Kuipers and P. Beeson. Bootstrap learning for place recognition. In *Proc. of the National Conf. on Artificial Intelligence*, 2002. 101

R. Kurazume and S. Nagata. Cooperative positioning with multiple robots. *IEEE Int. Conf. on Robotics and Automation*, 1994. 45

C. Kwok, D. Fox, and M. Melia. Real-time particle filters. *Proc. of the IEEE*, 92(2), 2004. 135

D. Lee. *The map-building and exploration strategies of a simple sonar-equipped mobile robot*. PhD thesis, University College London, 1996. 47, 99, 100

S. Lenser and M. Veloso. Sensor resetting localization for poorly modelled mobile robots. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2000. 97

J. J. Leonard and H. F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Trans. on Robotics and Automation*, 7(3):376–382, 1991. 41, 44

J. J. Leonard and H. J. S. Feder. A computationally efficient method for large-scale concurrent mapping and localization. In *Proc. of the 8th Int. Symposium on Robotics Research*, 1999. 44

F. Lewis, editor. *Autonomous Mobile Robots - Sensing, Control, Decision Making and Applications*. CRC Press, 2006. 3

J. H. Lim and D. W. Cho. Physically based sensor modelling for a sonar map in a specular environment. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 1992. 47, 100

K. Lingemann, H. Surmann, A. Nüchter, and J. Hertzberg. Indoor and outdoor localization for fast mobile robots. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 2185–2190, Sendai, Japan, Sept. 2004. 70

J. Little, C. Jennings, and D. Murray. Vision-based mapping with cooperative robots. In *Sensor Fusion and Decentralized Control in Robotic Systems*, volume 3523, pages 2–12, Oct. 1998. 124

F. Lu and E. Milios. Robot pose estimation in unknown environments by matching 2D range scans. *Journal of Intelligent and Robotic Systems*, 18:249–275, 1997a. 44

F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4:333–349, 1997b. 45

R. Madhavan, K. Fregene, and L. Parker. Distributed cooperative outdoor multirobot localization and mapping. *Autonomous Robots*, 17:23–39, 2004. 124

M. Montemerlo and S. Thrun. Large-scale robotic 3D mapping of urban structures. In *Proc. of the Int. Symposium on Experimental Robotics (ISER)*, 2004. 70

H. Moravec. Sensor fusion in certainty grids for mobile robots. *A.I. Magazine*, 1988. 99

A. I. Mourikis and S. I. Roumeliotis. Performance analysis of multirobot cooperative localization. *IEEE Trans. on Robotics*, 22(4):666–681, Aug. 2006. 16

L. Navarro-Serment, R. Tinos, and P. Khosla. Fault tolerant localization for teams of distributed robots. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2001. 47

L. E. Navarro-Serment and P. K. Khosla. A beacon system for the localization of distributed robotic teams. *Int. Conf. on Field and Service Robotics*, 1999. 45

J. Neira, J. A. Castellanos, J. M. M. Montiel, and J. D. Tardós. The SPmap: A probabilistic framework for simultaneous localization and map building. *IEEE Trans. on Robotics and Automation*, 15(5):948–952, 1999. 44

P. Newman, D. Cole, and K. Ho. Outdoor slam using visual appearance and laser ranging. In *Proc. of the Second Int. Conf. on Robotics and Automation*, Florida, USA, 2006. 45

A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann. Heuristic-based laser scan matching for outdoor 6D SLAM. In *Advances in Artificial Intelligence. 28th annual German Conf. on AI*, Sept. 2005. 7, 70

C. Olson. Subpixel localisation and uncertainty estimation using occupancy grids. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, May 1999. 101

C. Olson. Probabilistic self-localization for mobile robots. *IEEE Trans. on Robotics and Automation*, 16(1):55–66, 2000. 99

P. Pfaff, R. Triebel, and W. Burgard. An efficient extension to elevation maps for outdoor terrain mapping and loop closing. *The Int. Journal of Robotics Research*, 26(2):217–230, 2007. 70, 99

I. M. Rekleitis, G. Dudek, and E. E. Milios. Graph based exploration using multiple robots. *Fifth Int. Symposium on Distributed Autonomous Robotic Systems*, 2000. 125

I. M. Rekleitis, G. Dudek, and E. E. Milios. Multi-robot cooperative localization: A study of trade-offs between efficiency and accuracy. *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2002. 124, 125

P. Rives, A. C. Victorino, and J. Borrelly. Safe navigation for indoor mobile robots. *The Int. Journal of Robotics Research*, 22(12):1005–1039, Dec. 2003. 101

R. Rocha. *Building volumetric maps with cooperative mobile robots and useful informatio sharing*. PhD thesis, University of Portugul, 2005. 131, 132

R. Rocha, J. Dias, and A. Carvalho. Cooperative multi-robot systems: a study of vision-based 3-D mapping using information theory. *Robotics and Autonomous Systems*, 53(3-4):282–311, 31 Dec. 2005. 85, 100, 124

S. I. Roumeliotis and G. A. Bekey. Bayesian estimation and kalman filtering: A unified framework for mobile robot localization. *IEEE Int. Conf. on Robotics and Automation*, 2000. 44

J. Ryde and H. Hu. Fast circular landmark detection for cooperative localisation and mapping. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 2756–2761, Barcelona, Spain, April 2005a. 13, 27, 137

J. Ryde and H. Hu. Laser based simultaneous mutual localisation for multiple mobile robots. In *Proc. of IEEE Int. Conf. on Mechatronics and Automation*, pages 404–409, Niagara Falls, Canada, July 2005b. 80, 136

J. Ryde and H. Hu. Mutual localization and 3D mapping by cooperative mobile robots. In *Proc. of Int. Conf. on Intelligent Autonomous Systems*, The University of Tokyo, Tokyo, Japan, Mar. 2006a. 80, 136

J. Ryde and H. Hu. Cooperative mutual 3D laser mapping and localization. In *IEEE Int. Conf. on Robotics and Biomimetics*, Kunming, China, Dec. 2006b. 80

J. Ryde and H. Hu. Mobile robot 3D perception and mapping with multi-resolution occupancy lists. In *Proc. of IEEE Int. Conf. on Mechatronics and Automation (ICMA 2007)*, Harbin, Heilongjiang, China, Aug. 2007. 104

B. Schiele and J. L. Crowley. A comparison of position estimation techniques using occupancy grids. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, volume 2, pages 1628–1634, May 1994. 17

T. Schmitt, R. Hanek, M. Beetz, S. Buck, and B. Radig. Cooperative probabilistic state estimation for vision-based autonomous mobile robots. *IEEE Trans. on Robotics and Automation*, 18(5):670–684, Oct. 2002. 46

O. Serrano, J. M. C. nas, V. Matellán, and L. Rodero. Robot localization using wifi signal without intensity map. *WAF*, 2004. 2

R. Simmons, D. Apfelbaum, W. Burgard, D. Fox, M. Moors, S. Thrun, and H. Younes. Coordination for multi-robot exploration and mapping. In *AAAI/IAAI*, pages 852–858, 2000. 125

J. Spletzer, A. Das, R. Fierro, C. Taylor, V. Humar, and J. Ostrowski. Cooperative localization and control for multi-robot manipulation. In *Proc. of the Conf. on Intelligent Robots and Systems*, 2001. 47

S. Thrun. Learning occupancy grids with forward sensor models. *Autonomous Robots*, 15:111–127, 2003. 54

S. Thrun and M. Montemerlo. The graphSLAM algorithm with applications to large-scale mapping of urban structures. *The Int. Journal of Robotics Research*, 25:403–429, 2006. 100

J. Vandorpe, H. V. Brussel, and H. Xu. Exact dynamic map building for a mobile robot using geometrical primitives produced by a 2D range finder. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 1996. 99

V. Verma, S. Thrun, and R. Simmons. Variable resolution particle filter. In *In Proceedings of Int. Joint Conf. on Artificial Intelligence*. AAAI, August 2003. 141

P. Štěpán, M. Kulich, and L. Přeučil. Robust data fusion with occupancy grid. *IEEE Trans. on Systems, Man and Cybernetics*, 35(1):106–115, 2005. 53

L. Wang, L. Yong, and M. A. Jr. Mobile robot localisation for indoor environment. *SIMTech Technical Report*, 2002. 98

Z. Wang, S. Huang, and G. Dissanayake. D-SLAM: A decoupled solution to simultaneous localization and mapping. *The Int. Journal of Robotics Research*, 26(2):187–204, 2007. 31, 98

T. Weigel, J. S. Gutmann, and B. Nebel. Fast, accurate, and robust self-localization in polygonal environments. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 1999. 44

J. Weingarten, G. Gruener, and R. Siegwart. A state-of-the-art 3D sensor for robot navigation. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct. 2004. 70

E. W. Weisstein. Least squares fitting–polynomial. MathWorld–A Wolfram Web Resource, 2005. URL `http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html`. 29

C. Wetzler, G. Weiss, and E. V. Puttkamer. Keeping track of position and orientation of moving indoor systems by correlation of range-finder scans. *IEEE Int. Conf. on Intelligent Robots and Systems*, pages 595–601, Sept. 1994. 99

O. Wulf, K. O. Arras, H. I. Christensen, and B. Wagner. 2D mapping of cluttered indoor environments by means of 3D perception. In *IEEE/RAS Int. Conf. on Robotics and Automation (ICRA)*, New Orleans, USA, 26 Apr. 2004. 4, 71

S. I. Y. Yagi and M. Yachida. The integration of environmental map observed among multiple mobile robots with omnidirectional image sensor. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, volume 2, pages 640–647, 1996. 46

B. Yamauchi. Frontier-based exploration using multiple robots. In *Proc. of the Second Int. Conf. on Autonomous Agents*, pages 47–53, 1998. 46

C. Ye and J. Borenstein. Characterization of a 2-d laser scanner for mobile robot obstacle negotiation. *Proc. of the 2002 IEEE Int. Conf. on Robotics and Automation*, 2002. 20, 35