Stereo Panoramic Vision for Obstacle Detection

A thesis submitted for the degree of
Master of Philosophy
at The Australian National University

Leanne Matuszyk

February 2006

Statement of Originality

This masters project was conducted under the supervision of Professor Alexander Zelinsky. The work submitted in this thesis is a result of original research carried out by myself, except where duly acknowledged, while enrolled as a Masters student in the Department of Systems Engineering at the Australian National University. It has not been submitted for any other degree or award.

Leanne Matuszyk

Abstract

Statistics show that automotive accidents occur regularly as a result of blind-spots and driver inattentiveness. Such incidents can have a large financial cost associated with them, as well as injury and loss of life. There are several methods currently available to assist drivers in avoiding such collisions. The simplest method is the installation of devices to increase the driver's field of view, such as extra mirrors and wide angle lenses. However, these still rely on an alert human observer and do not completely eliminate blind-spots. Another approach is to use an automated system which utilises sensors such as sonar or radar to gather range information. The range data is processed, and the driver is warned if a collision is immiment. Unfortunately, these systems have low angular resolution and limited sensing volumes. This was the motivation for developing a new method of obstacle detection.

In this project, we have designed, built and evaluated a new type of sensor for blind spot monitoring. The stereo panoramic sensor consists of a video camera which views a convex mirrored surface. With the camera and mirror axes aligned, a full 360 degrees can be viewed perpendicular to the sensor axis. Two different mirror profiles were evaluated - the constant gain, and resolution invariant surfaces. It was found that the constant gain mirror was the most effective for this application.

It was shown that the sensor can be used to generate disparity maps from which obstacles can be segmented. This was done by applying the v-disparity algorithm, which has previously not been utilised in panoramic image processing. We found that this method was very powerful for segmenting objects, even the case of extremely noisy data. The average successful obstacle detection rate was found to be around 90%, with a false detecion rate of 8%.

Our results indicate that range can be estimated reliably using a stereo panoramic sensor, with excellent angular accuracy in the azimuth direction. In ground truth experiments it was found that the sensor was able to estimate range to within 20cm of the true value,

and a maximum angular error of 3°. Through experimentation, we determined that the physical system was approximately half as accurate in comparison to the simulations. However, it should be noted that the system is a prototype which could be developed futher. Nevertheless, this sensor still has the advantage of a much higher angular resolution and larger sensing volume than the driver assistance systems reported to date.

Acknowledgments

To begin with I would like to thank my supervisor Professor Alex Zelinsky for providing me with some amazing opportunities over the past few years. His guidance and innate ability to motivate have been invaluable.

Thanks are extended to my co-supervisors Lars Nilsson, and Magnus Rilbe from Volvo Technology Corporation. Working with them in Gothenburg was both productive, and thoroughly enjoyable. And thank you for making time for the many early morning teleconferences due to the time difference to Canberra. Thanks also to Martti Soinenen from Volvo Cars Corporation. This work has been partially sponsored by both Volvo Technology Corporation and Volvo Car Corporation. The funding is gratefully acknowledged.

To my fellow students at Systems Engineering, thanks for being such great company: Kirill Kouzoubov, Gareth Loy, Nick Apostollof, Arvin Dehghani, and the Swedish imports, Peter Lundgren, Anton Falk, Carl Bark-Holste, and Magnus Salfstrom. In particular, thanks to my cubicle buddy and travel companion Grant Grubb; it's been a pleasure working with him. Special thanks to Luke Fletcher for his code, and for answering all my questions, and to Dr Lars Petersson for helping with field experiments and so much more. Thank you also to the administrators, James Ashton and the lovely, and amazingly efficient Rosemary Shepard .

Thank you to Jason Chen for his help at the farm, and making sure I didn't machine off any fingers. Thanks also to Ben Nash from the Engineering Department, for making version 2.0 of the mounting equipment for this project and preventing me from over-engineering everything.

To my best buds Naomi C and Diana, thanks for being there to listen. And to Dennis for keeping me company as I stumbled towards the finish line. But most importantly, I would like to thank my fabulous family - Andre, Tim, Dina and Eugene - for all their love and support.

Publication

Leanne Matusyk, Alexander Zelinsky, Lars Nilsson, and Magnus Rilbe. Stereo Panoramic Vision for Monitoring Vehicle Blind-spots. *In Proceedings of the IEEE Intelligent Vehicles Symposium*, Parma, June 2004.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

UNKOWN to his parents, a 20-month-old boy had wandered out the back door of his house. As there was no fence to prevent him, the boy was able to walk through the backyard and play on the driveway. The father then began reversing his four-wheel-drive vehicle out of the garage. He stopped the car when he felt a bump, and was horrified when he found his son under the vehicle between the tyres. The child was rushed to the hospital where his injuries were treated. Fortunately the boy was able to completely recover from his injuries within three months. However the parents required extensive counselling, and they no longer own the vehicle. Although there are many cases of accidents with much graver outcomes, this particular incident is an example of a typical driveway motor vehicle injury (Holland et al. 2000).

There are several products currently available to help prevent rear collision accidents of the kind described above. It is possible to install blind spot mirrors, which can be seen by the driver through the rear window of the car. However, larger vehicles are often designed to carry payloads which can obscure the driver's view of these mirrors. Another option is to use an automated system that monitors the area behind the vehicle. There are several commercially available products that detect obstacles and warn the driver in the event of an impending collision. These products utilise sonar and radar sensors. While these obstacle detections systems are fairly effective, there are certain situations in which such sensors fail, as will be shown in Chapter 2.

Hence, there was interest in developing a more reliable means of detecting obstacles. This project was a collaborative effort between the Australian National University, Volvo Technology Corporation and Volvo Car Corporation, to develop a blind spot monitoring system. Such a system can not only improve safety around cars, but could also be

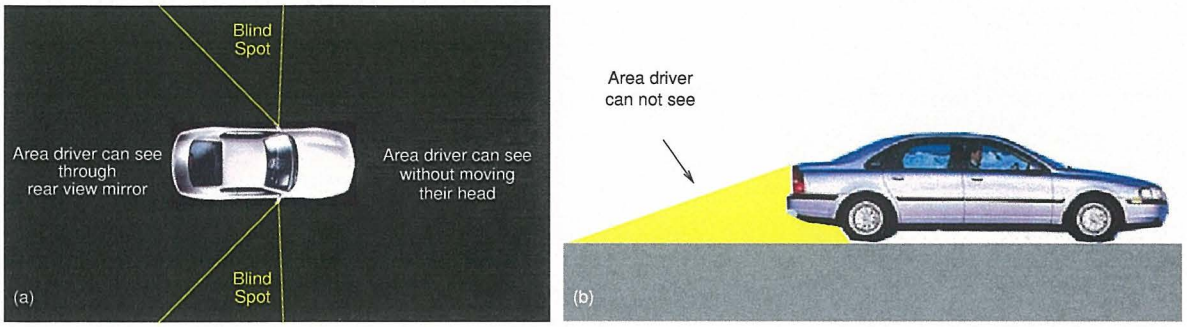


Figure 1.1: (a)The traditional definition of a blind-spot is an area that a driver cannot see when only looking in their mirrors. (b) The area behind the car that is completely obscured from the driver's view point.

implemented on other vehicles such as buses, trucks and construction vehicles.

The rest of this chapter describes the motivation behind this project, the scope of the project and concludes with an overview of the entire thesis.

1.1 Motivation

Accidents can occur at the rear of motor vehicles for a variety of reasons. The view of the driver may have been obscured, or the accident may have been caused by inattentiveness or misjudgment of distance. There are several collision warning systems currently available; however, there are situations where these devices can fail to prevent an accident.

1.1.1 Accidents due to Blind-spots and Human Error

The traditional definition of a blind-spot is any area that a driver cannot see when only looking in their mirrors (Department of Motor Vehicles, California 2001). A driver must turn their head in order to see an obstacle in blind-spot areas, which are shown in Figure 1.1(a). In the context of this project we are also concerned with areas that are completely obscured from the driver's field of view. In particular, the area directly behind the vehicle (Figure 1.1(b)).

Many accidents have occurred even though the driver had the ability to see the obstacle. The driver may not have turned his/her head to monitor a certain area, or attention was elsewhere. Many collisions occur when a driver simply misjudges the distance from the vehicle to an object in the environment. A survey of accidents in the UK over a three

year period revealed that 24% were due to inattentiveness, while 15% were caused by misjudgement (Maycock 1996).

There is often a large financial cost involved with blind-spot collisions. Automobile liability collision data was collected in Texas, USA for the first nine months of the 2001-2002 financial year (Thompson 2002). From a total of almost 900 accidents, 25% of claims were due to lane encroachment, while 17.5% were backing accidents. The average cost per claim was approximately $US 2000, in both cases.

In 2002, 58 children were backed over and killed in the US by pickup trucks or sports utility vehicles, in what has been dubbed the "bye-bye syndrome" (Koeppen 2003). In these accidents, a child runs out onto the driveway to wave goodbye to their parents. The driver is unaware of the presence of the child, and accidentally reverses over them. Over the past 10 years there have been 294 documented cases of this situation in the US, with 179 of these resulting in the death of the child. A study was completed in Australia to examine driveway motor vehicle accidents involving children (Holland et al. 2000). It was found that these injuries accounted for 12% of all children admitted to hospital with pedestrian motor vehicle injuries and deaths - a significant proportion. The report concluded that "prevention represents the only effective approach to reducing deaths from this cause".

A particular blind-spot problem occurs in trucks. Many accidents occur when the truck turns in a direction opposite to the driver side of the vehicle. In the European Union, all truck drivers are obliged to install three mirrors to increase the field of view. However, even with these aids, there is still a substantial "dead angle" on that side (Figure 1.2). Every year a large number of people, the majority being cyclists and pedestrians, are killed or seriously injured in these situations. The European Commission for Transport reports that in Belgium, an average of one cyclist is killed every month by turning trucks (European Commission 2003). The European Commission estimates that 500 fatalities per year occur in the EU due to blind-spots on trucks.

The National Occupation Health and Safety Commission has completed a study on work-related deaths in Australia (Driscoll et al. 1998). It was found that vehicles were the major cause of accidents in the 13 industry and occupational groups surveyed, with the worst area being forestry. Of all the documented work-place accidents, just under 35% were due to "being hit by moving objects", which encompasses both falling objects and pedestrian incidents involving vehicles. In the report, one of the more common situations which have led to work-related deaths is cited as "construction and mining labourers on work sites being run over by reversing vehicles from which the driver's vision was restricted

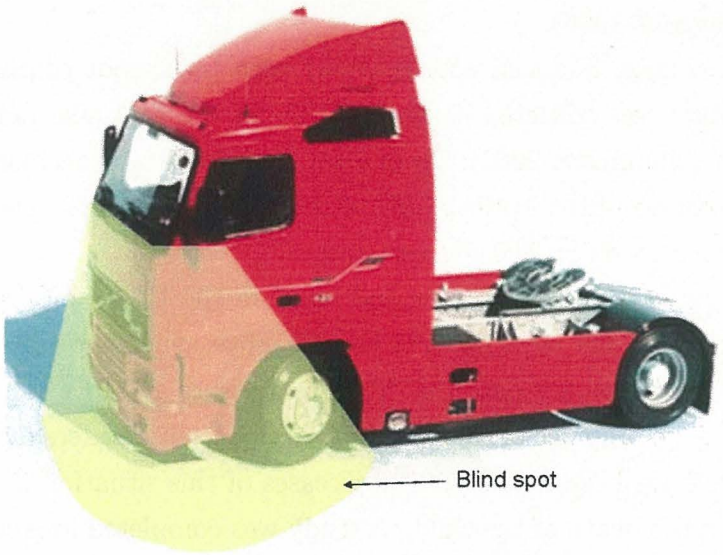


Figure 1.2: Blind spot for a right hand drive truck.

because of blind spots".

1.1.2 Rear Collision Avoidance

It is apparent from the statistics presented in the previous section that there is a need for the prevention of rear and blind-spot collisions. An effective way to avoid such accidents is to monitor the area behind the vehicle. In recent years, many products have been devised which attempt to solve this problem. There are two main methods currently available:

- **Human Observation** - Blind spot mirrors are the cheapest attempt to solve the problem. This strategy decreases the blind spot area depicted in Figure 1.1. There are also wide angle mirrors that can be mounted over the rear window of vans, which enable the driver to view the area directly behind the vehicle. An extension of this approach would be to use a conventional video camera or cameras, mounted to the rear of the vehicle. A monitor placed inside the vehicle cabin would allow the driver to monitor the external environment.

- **Automated Monitoring Systems** - Systems are available that use sonar or radar to detect obstacles. In the event that a detected object comes too close to the rear of the vehicle, the driver is provided with an audio and/or a visual warning.

While these approaches can be effective, certain situations exist where these systems can fail. These methods, as well as their advantages and disadvantages, are examined in more detail in Chapter 2. Due to the drawbacks of the various existing rear collision avoidance systems, there is a strong motivation for the development of a new technology solution.

1.2 Objectives and System Specifications

The aim of this project is to design, build and test a sensor for a new category of automated monitoring systems. The sensor will be used for detecting obstacles in close proximity to the rear of an automotive vehicle, to improve the level of safety. Although the prototype system initially is used on an experimental car, the sensors can easily be modified to suit trucks, buses and construction vehicles.

In order to develop an effective means of collision warning, it is necessary to carefully consider the general requirements of such a system. As mentioned in Section 1.1 many accidents occur as a result of areas around the car obscured from the drivers field of view. Therefore, an obstacle detection system should monitor these regions. Due to the possibility of driver inattentiveness or human error, it was decided that the work space also should be extended to areas visible to the driver.

When designing a new system, it is desirable that the performance should at least meet, or ideally exceed that of it's predecessors. In Chapter 2, the currently available obstacle detection methods have been examined, and their limitations are discussed. This information contributed to the decision to develop the specifications that would endow the prototype system with an accuracy equal to existing methods. Most current systems would fail to meet the full set of specifications. The specifications formulated for this project are as follows:

1. The zone behind the vehicle should be monitored, as close to the vehicle as possible.
2. The system is to have at least 180 degrees of sensing (Figure 1.3).
3. The working volume should have a radius of five metres, and height of three metres above the ground.
4. The system should be accurate to within 20 centimetres.
5. The system should operate in real-world environments.

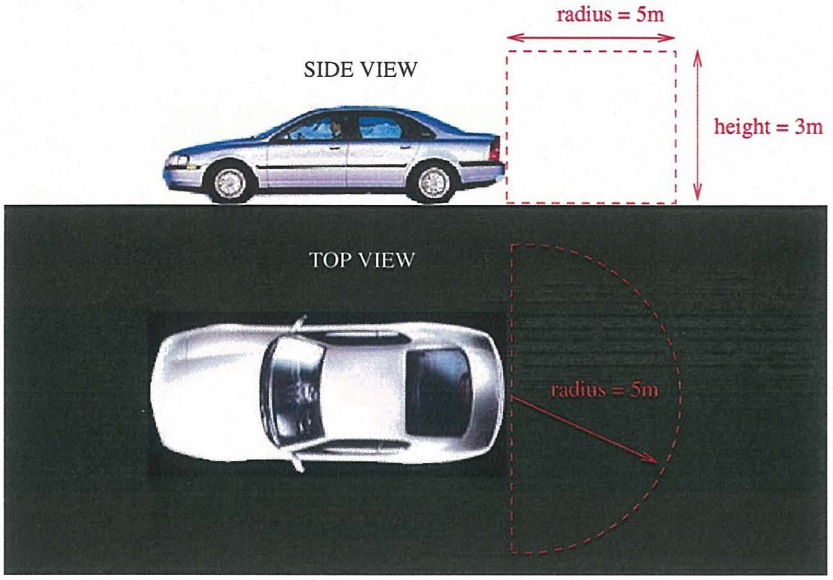


Figure 1.3: Requirements for the obstacle detection system. The sensing area should have a radius of five metres and a height of three metres.

One approach that has not yet been widely investigated for blind spot monitoring is the use of vision for an automated system. Therefore, we have decided to design, and build a stereo panoramic sensor, and have evaluated the range estimation capabilities of such a system.

1.3 Contributions and Achievements

In this research work, contributions have been made in the following areas:

- The design of a system to detect and locate obstacles using a stereo panoramic vision sensor.
- Implementation, and evaluation of the *ground plane subtraction* algorithm for obstacle segmentation from panoramic images, using ground truth data and field experiments.
- Implementation, and evaluation of the v-disparity algorithm to segment obstacles from panoramic images, using ground truth data and field experiments.

- Investigation to determine stereo panoramic sensor characteristics for optimal range estimation and obstacle segmentation. These included the use of different mirror profiles, baselines, and sensor configurations.
- The selection of sensor parameters to achieve most desirable characteristics for our application.
- The use of temporal filtering to reduce the false detection rate.
- Theoretical evaluation of the system, using Matlab, and ground truth data generated by the POV-Ray raytracing software.
- Real-world evaluation of a stereo panoramic vision system. The range accuracy was measured, and a quantitative analysis was carried out to investigate the presented obstacle detection algorithms.

1.4 Outline of Thesis

This section outlines chapter by chapter the contents of the remainder of this thesis. The body of this work is organised as follows:

- **Chapter 2 - Related Work**
  In this chapter we examine the methods for monitoring blindspots that are currently available. Their advantages are highlighted, as well as the inadequacies, particularly with respect to the system specifications given in Section 1.2. The way in which animals solve similar problems is discussed, and using this as inspiration, a vision based sensor solution is suggested. This is followed by a survey of the work that has been completed using the alternative sensor.
- **Chapter 3 - System Overview**
  A summary of the implementation of the stereo panoramic vision system is presented. Details of the hardware and software developed in this project are provided. In particular, the design of the mirror profiles and camera mounts are presented. Conceptual drawings are presented to show that it would be possible to integrate such a sensor into cars, buses and construction vehicles.
- **Chapter 4 - Obstacle Detection**
  Two different algorithms are employed to detect obstacles in the panoramic images

- ground plane subtraction and v-disparity. These techniques are explained, and the details of their implementation in this project are discussed.

- **Chapter 5 - Theoretical Analysis and Simulation**
  The stereo panoramic vision system was evaluated using a ray tracing software package, POV-Ray. The software was used to create artificial images of virtual environments, which was then used to provide ground truth data. The different sensor characteristics have been compared, including the significance of the mirror profile, baseline and orientation. The two different obstacle detection methods described in Chapter 4 were also tested. The results of these experiments are then summarised.

- **Chapter 6 - Field Experiments**
  The equipment built for this project was tested in a series of experiments. The sensor characteristics, and obstacle detection methods were evaluated. These results were compared to those obtained from the theoretical experiments. Finally, a summary of the results is provided.

- **Chapter 7 - Conclusions**
  A summary of the thesis is given highlighting the main issues and contributions. The chapter concludes with suggestions for further work in this area.

Chapter 2

Related Work

In the first chapter we discussed the definite need to improve safety around cars, buses, trucks and construction vehicles, particularly in blind-spot areas. There are large financial costs associated with these rear collisions and blind-spot accidents, not to mention the severe injuries and fatalities that can be caused. This chapter presents the currently available technologies which attempt to solve to the problem, and an alternative approach will then be introduced.

We will begin by providing a survey of the products available to help drivers avoid rear collisions and other blind-spot accidents. The characteristics of each system will be discussed, and their limitations highlighted. The second section explores another possibility that has not yet been utilised in this application, which has been biologically inspired. The rest of the chapter provides the details of the sensor to be used in this project, and discusses the work that has been conducted in this area. The chapter ends with a summary of the advantages and disadvantages of each of the driver assistance systems.

2.1 Current Rear Collision Avoidance Methods

There are several methods currently available to assist drivers in avoiding rear end collisions. These systems fall into two main categories. The first utilises the human observer. Drivers are presented with raw visual data, and this information is used make decisions on how to control the vehicle. The second category is automated monitoring systems. These use sensors to determine the position of objects around the car, and warn the driver in the event of an impending collision.

The Motor Accidents Authority of NSW, Australia (MAA) commissioned a study on

devices that could be used to prevent young pedestrians from being involved in an accident behind a vehicle (Paine & Henderson 2001). The various visual aids will be discussed, and the relevant conclusions of the MAA study will be highlighted.

2.1.1 Human Observation

The usual method of collision avoidance relies on the human observer - the driver of the vehicle. However, there are many areas around the vehicle where the driver cannot see, even when turning their head. The most common way of avoiding collisions in these areas is to extend the driver's field of view. This can be done by installing extra mirrors, wide-angle lenses or video cameras around the vehicle.

Blind Spot Mirrors

One of the cheapest approaches is the installation of blind spot mirrors. Most products are designed for installation on the sides of the vehicle, to decrease the unseen area depicted in Figure 1.1. Obviously the side mirrors are inadequate for detecting most rear end collisions, as they do not provide the driver with the necessary field of view. In the case of trucks, there is a significant blind-spot at the front side, opposite the driver (Figure 1.2).

There are also wide angle mirrors that can be mounted behind the rear window of vans to view the area directly behind it (Figures 2.1 and 2.2). The MAA study found that the images reflected by wide angle mirrors were usually highly distorted and upside down. This requires the driver to learn how to interpret the images, in order to use the information for controlling the vehicle effectively (Paine & Henderson 2001). It was reported that the maximum distance that could be viewed was about 1.5m behind the vehicle. Another drawback of such mirrors is that they are of no use if the rear window is obscured by a payload inside the vehicle. This a common occurance in vehicles such as vans.

Wide Angle Lenses

Another low cost aid that can be used to provide a larger field of view to the driver is the wide angle lens. These are designed to be adhered to the rear window of the vehicle. An example of a wide angle lens product, produced by Bushranger Auto Gear(FastLane Communications 2003) is shown in Figure 2.3. However, unless the rear window is almost vertical to the road surface, the lens becomes completely ineffective. There are other similar products available, which include a hinge, to provide the required angle for viewing the

Figure 2.1: A convex blind-spot mirror mounted to the exterior of a van.

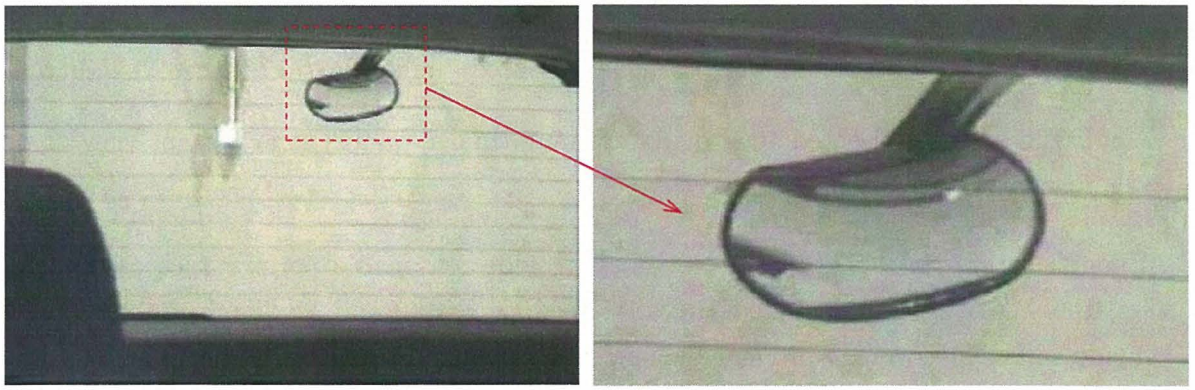


Figure 2.2: (a) The rear window of a van with an external convex mirror, similar to that shown in Figure 2.1, as viewed from the driver's seat. (b) A close up of the convex mirror, showing a bird's eye view of the area directly behind the car. These images are from a study commissioned by the Motor Accidents Authority of NSW, Australia.

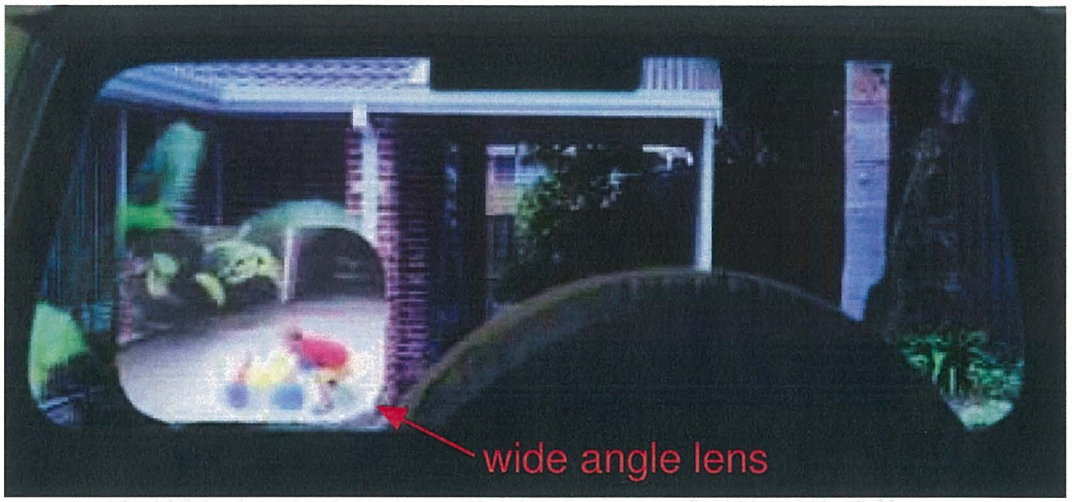


Figure 2.3: A wide angle lens fixed to the rear window of a four-wheel drive vehicle. Produced by Bushranger Auto Gear.

scene behind the car. The MAA evaluated several products, and their main observations are as follows:

- The image viewed in the lense was often too small.
- The image tended to be distorted and fuzzy.
- At certain angles of sunlight, the lens appeared "milky" , and reduced visibility.
- If the lens was located to provide optimum image quality, the normal view to the rear of the vehicle was obscured.

Although certain products in this category provide reasonable performance as reversing aids, the MAA study concluded that "wide angle lenses and auxiliary mirrors do not provide sufficient coverage or clear enough images to enable drivers to reliably see that a child is to the rear of the vehicle".

Conventional Video Cameras

A more expensive device that can be used to prevent reversing accidents is a video system. These usually consist of a video camera mounted to the rear of the vehicle, and a monitor mounted inside the cabin for the driver to view. One product that is currently on the

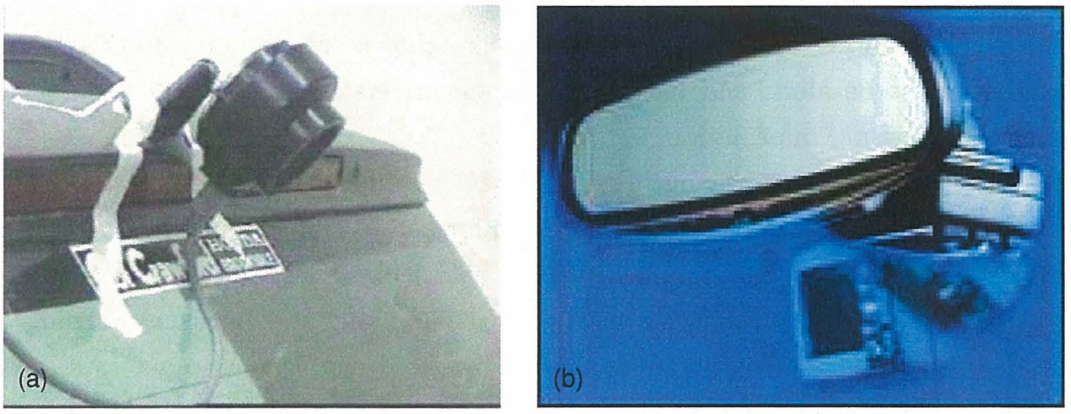


Figure 2.4: (a) The camera used by the MAA to evaluate video systems as driver visual aids. It is mounted above the rear window, and views the area behind the vehicle. (b) *Donnelly's VideoMirror* is attached to the rear vision mirror and can be flipped down when the driver is reversing the vehicle.

market is the *Donnelly VideoMirror with ReversAid,* shown in Figure 2.4. The system is sold as a self-installation kit, with a rear-mounted video camera and flip-down monitor to be attached to the rear view mirror (Trilogix Electronic Systems Inc. 2003).

Standard video cameras have a field of view between 30° and 45°. This makes it difficult to monitor all the danger zones behind the vehicle. The MAA report concluded that currently available video devices did not meet all safety requirements, but could be used in conjunction with another preventative method to provide a complete driver assistance system. In particular, the authors suggest the use of the *Guardian Alert* Microwave sensor [1] to patrol areas closest to the car, where their video system could not view. It also suggests that it could "be possible to develop a visual aid system that covers all blind spots".

2.1.2 Automated Monitoring

The driver assistance methods presented in Section 2.1.1 have the aim of increasing the driver's field of view. The extra visual information is provided directly to the human observer, to be used in decision making to control the vehicle. A drawback of such methods is that the implemented systems do not extend the field of view to include all possible danger zones. In the case of a video system, if cost was not a consideration, one could monitor a wider area by installing several video cameras all around the vehicle. However,

[1] www.guardianalert.com

the main disadvantage of these systems remains - that is, relying on human observation. The driver must be alert, and have an adequate understanding of how to interpret the supplementary visual information.

Another approach to reducing blind spot accidents is by using automated monitoring systems. These consist of a sensor to gather information from around the vehicle, a processor to perform the necessary calculations to detect obstacles, and a device to notify the driver of an impending collision, for example, with an audio and/or visual warning. These systems can be used to monitor areas unseen by the driver, and also act as a secondary observer to ensure the driver is aware of the presence of obstacles within their field of view. We will now review the different types of sensors that have been used for automated monitoring.

Sonar

Sonar range sensors emit ultrasonic signals to determine the distance to an object. A sonar pulse is emitted from the device, and the time taken for the first echo to return is measured (Figure 2.5). Using knowledge of the measured time of flight and the speed of sound in air, it is possible to infer the distance to the object reflecting the pulse, described for example in (Dudek et al. 1996). These are the most common sensors used in the automotive industry due to their relatively low cost. Car manufacturers such as Volvo, Mercedes-Benz and BMW have integrated sonar sensors into their production cars, and post-production installation kits are also readily available. In Australia, Holden has released several vehicles with "rear park assist" as a standard feature (Holden 2003). The Holden Commodore Acclaim has four ultrasonic sensors mounted in the rear bumper. When reverse gear is selected, a speaker emits an audible sound which increases in frequency as the vehicle moves closer to an obstacle behind it.

"Sonar devices are well known for the apparent unreliability of their readings" (Dudek et al. 1996). These devices rely on sound waves traveling through air, reflecting off the surrounding objects. As a result there are several potential sources of error when calculating range from sonar measurements:

- **Low angular resolution:** The angular resolution of a sonar sensor is defined by the beam cone angle, depicted in Figure 2.5. Outside this angle, the sonar pulse exists, but is too weak to return a measurable echo. When an echo is received, this could have been caused by an object anywhere on a spherical surface depicted in Figure 2.6.

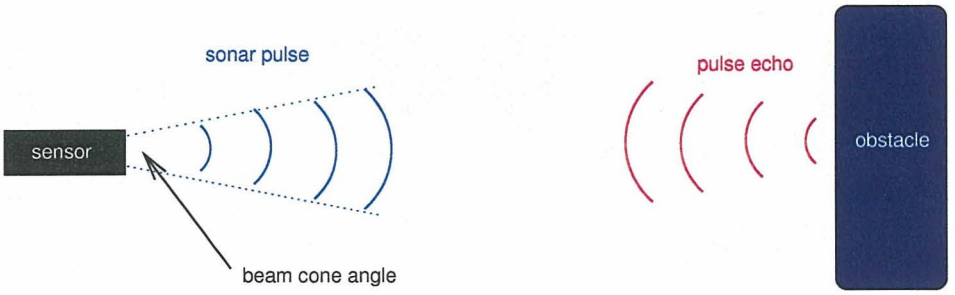


Figure 2.5: Distance to an object is calculated by measuring the time taken for a sonar pulse to return to the sensor from that object.

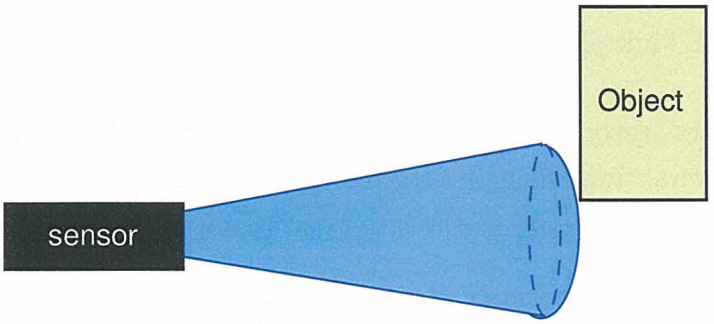


Figure 2.6: An echo returning to the sensor could be from an object anywhere on a spherical surface in the beam cone.

Therefore the angular resolution of the measurements are determined by the sensor, which can range from about 10° to 125° (Polaroid OEM Components Group 1999).

- **Cross talk:** Many vehicle reverse assistance systems utilise more than one sonar unit to help determine the angular position of an object. However this causes yet another problem which is known as cross talk. This is when a sensor receives a sonar pulse of the same frequency from another sensor, and confuses this with it's own signal. This can be solved by using sonars of different frequency or sound wave "signatures". However, this requires the use of sensors with differing characteristics or a substantial amount of signal processing of the echos (Wijk 2001).

- **Specularity:** Many surfaces act as specular surfaces at ultrasonic frequencies. This means that a sonar wave that hits the surface will follow the law of reflection - the angle of incidence is equal to the angle of reflection. It then follows that the surface needs to be almost facing the sonar sensor, otherwise the sonar beam will be reflected

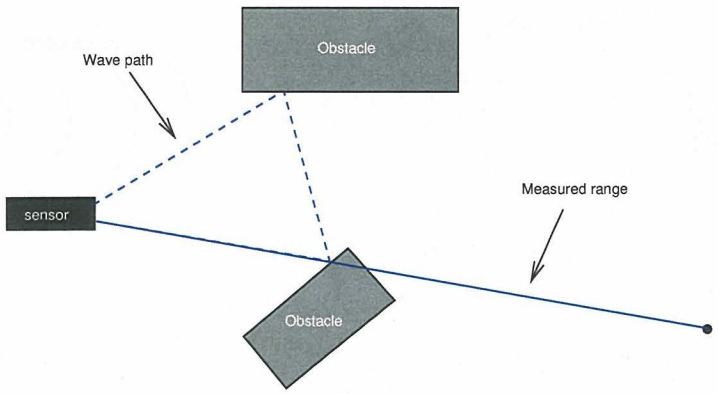


Figure 2.7: Multiple sonar reflections cause erroneous range measurements.

away from the sensor. In which case no echo will be detected, or an echo that has been reflected several times may be detected, providing an inaccurate reading (McKerrow & Zhu 1996). An example of the latter can be seen in Figure 2.7.

- **Diffraction:** Another source of error in sonar measurements is due to diffraction. This happens when the sonar beam hits the corner of an object, which causes the echo wave to spread out cylindrically, with less energy (Lacroix & Dudek 1997). As a consequence, the corner is less likely to be detected and may not be detected at all.

- **Dead zone:** There is also a minimum detection distance. When a sonar pulse is emitted by the sensor, any returning echos will be ignored because the device is still transmitting and not yet listening for the returning pulse. This distance is usually between 15 to 35 centimetres (Sensing and Control, Honeywell Inc. n.d.).

In the study for the MAA (Paine & Henderson 2001), three different brands of driver assistance packages which utilised sonar were tested and evaluated. The products tested were as follows:

- *Reverse Sensor* from Williams Geddes & Co, Belrose NSW, Australia [2].

- *Safe reverse* from Global Accents, Canada [3].

- *Smart Park* from MKP, Neutral Bay NSW, Australia.

[2] www.reversesensor.com.au
[3] www.savereverse.com

The *Smart Park* had the longest range, which was 1.5 metres. It was possible to extend this slightly by adjusting the sensitivity but this would increase the incidence of false alarms, in all the sonar products tested. In general, the authors reported that too many "nuisance alarms" would occur during operation. The alarm would not sound when reversing over smooth surfaces such as concrete, while more coarse surfaces such as bitumen would set off the alarm. The study also found that small dips in the road would cause the alarm to sound at high priority. Such false alarms could be extremely detrimental, as drivers could become desensitised and ignore the warning signs of a real accident.

Radar

The MAA study (Paine & Henderson 2001) tested the *Guardian Alert* product by S & S Distributing, USA. This device had the longest range of the products tested, with a 3 metre maximum but this was reduced to 2.2 metres on the side opposite to the mounting. As intended, the alarm only sounded when there was relative movement between the vehicle and the object, which helped to minimise false alarms. The main drawback of this system is the large vertical angular range, making it impossible to distinguish between objects of varying size and position.

In September 2004, the automotive supplier Delphi gave a demonstration in Washington D.C. of their latest active safety technologies (Bishop Consulting 2004). One of the systems that was presented was the *Dual Beam Back-Up Aid.* This system uses radar sensors to detect obstacles up to a range of 5 metres. An audio, and on some models, also a visual alert notifies the driver of objects in the rear path of the vehicle. The system emits beeps that vary in volume and frequency to indicate the distance to the closest obstacle. This system is scheduled to be released on the market in 2006. Delphi are also hoping to release a variation of this system, which includes a wide field of view camera. As well as the active safety system, the camera enables the driver to visually inspect the area immediately behind the vehicle. The *Integrated Back-Up Aid with Camera* is anticipated to be released on the market in 2008. There was no information provided indicating the accuracy of either system.

Infrared Intensity

Infrared sensors are another relatively cheap device that could be used. These sensors emit light from the infrared spectrum, and measure the distance to objects based on reflected light intensity. As a consequence, they give different results depending on the colour of an

obstacles surface. This type of sensor is only effective for determining short range distances of no more than about 1 metre (Wijk 2001). The sensor characteristics make it unsuitable for this project.

Laser

Laser range sensors are quite effective at accurately determining ranges to objects in the environment. These sensors release a thin beam of light, and in a similar manner to sonar sensors, measure the time of flight for the light beam to return. It can measure distances up to 100 metres away with centimetre precision. Laser range sensors are able to sample the environment much faster than a sonar sensor, as it relies on light rather than sound.

A driver assistance system requires a sensor to monitor a wide area behind the vehicle. A system utilising lasers would therefore require the use of several lasers, or a single laser which is scanned across the area behind the car - a laser scanner. Several passive lasers would not be able to detect obstacles between laser beams, so a laser scanner would be most suitable for this application. However, the laser sensor does have certain drawbacks. A laser scanner:

- is quite expensive.
- can only detect objects within a single plane or possibly multiple planes (depending on the sensor), and will miss any object in between.
- has difficulty detecting transparent objects such as glass windows.
- is affected by dirt or water on the lenses, which can lead to false readings.
- requires moving parts, which decreases the robustness of the sensor and increases cost.

An example of such a system is the IBEO, LD Automotive laser scanner (IBEO Automobile Sensor 2004). It has a scanning frequency from 10Hz to 40Hz, and can determine the range to objects between 0.3m and 256m with an accuracy of 5cm. It has a maximum field of view of 270° in the horizontal direction, and an angular resolution of 1°. However, a basic system experimental currently costs 27 600 Euro. It will be some time before such a sensing device will be cost effective.

Infrared Time-of-Flight

This is another time-of-flight sensor that uses an infrared light source to provide a burst of light. The device measures the time taken for the light to return to the sensor to determine the depth to surrounding objects. One of the leading technology providers is a company called Canesta (Canesta, Inc 2006). They produce a CMOS-based single-chip sensor, with a 64 by 64 pixel array (Gokturk et al. 2004). Depth and intensity images are generated up to a maximum rate of 30 Hz. Most sensors of this type have problems dealing with strong ambient light, which can cause the depth images to saturate. The Canesta chip tries to overcome this problem using their patented Common Mode Reset (CMR) function. Unfortunately, even with this improvment, the infrared time-of-flight sensor still has several characteristics that make it undesirable for our application. The main disadvantages are as follows:

- When used outdoors, even with Canesta's patented CMR technology, the maximum range of the sensor is 1.3 metres.
- The range is also limited by the strength of the infrared source.
- The maximum field of view is 80 degrees.
- The sensor is currently only available to the general public as part of a development kit, which costs US$7 500

2.1.3 Summary

This section has reviewed the various rear collision avoidance methods currently available on the market. The most important characteristics of each approach has been summarised in Table 2.1.

2.2 Vision

In the previous section we presented the various methods for rear collision avoidance. Each approach was able to perform effectively under certain conditions, however in each case there was a definite need for improvement. One avenue that has not yet been pursued for this application is the use of vision in an automated monitoring system.

System Characteristics	Human Observer			Automated				
	Blind Spot Mirror	Wide Angle Lens	Video System	Sonar	Radar	IR Intensity	Laser Scanner (IBEO)	IR Time-of-Flight
Provide driver warning	×	×	×	√	√	√	√	√
Adequate field of view	×	×	×	×	√	×	×	×
Angular resolution (degrees)	n/a	n/a	n/a	10-125	180	180	0.5 along a single or multiple planes	1.25
Range (m)	1.5	Depends on human observer	Depends on mounting	1.5	3	1	100	1.3
Cost ($AU)	~20	~20	~780	60-649	~700	~50	~44 000	~10 000
Found to be effective by MAA	×	×	√	×	√	n/a	n/a	n/a
Relies on human observer	√	√	√	×	×	×	×	×

Table 2.1: Summary of the currently available obstacle detection methods.

Figure 2.8: Examples of biological vision systems.

2.2.1 Biological Inspiration

Vision is often the primary sensor used in the animal kingdom (Figure 2.8[4]). Biological vision systems have taken millions of years to evolve and now provide animals with efficient ways in which to perceive the world around them. Each vision system is perfectly adapted to enable an animal to carry out tasks that ensure their survival including hunting, evading predators and foraging for food. Therefore, researchers have used these highly evolved sensors as inspiration for the design and implementation of artificial vision. This area of research is known as computer vision, which is the endevour to enable computers to extract useful information from visual data. There are several forms of artificial vision sensors available, and it was important to determine which would be most suitable for this particular application.

One of the main characteristics exploited in computer vision to infer range information is the use of stereoscopic sensors. Stereoscopy is when depth is determined by viewing an object from two different positions in space. This requires at least binocular vision, with some overlap between the field of view of each sensor. Since the sensor developed in this project is required to determine the range to objects in the environment, it was decided that the use of a binocular vision system would be most suited to this application. The methods for extracting depth information from visual inputs will be discussed further in Section 4.1.

[4]Images from www.califex.com, www.ebiomedia.com and www.acriticaldecision.org .

2.2.2 Improvements on Nature

Humans have quite a large stereoscopic field of view. The overlap between the field of view of the left and right eyes lies between 120° and 180°, with a combined field of view between 160° and 208° (Chandler 1997). There are many other animals that have larger total fields of view, and a particularly good example of this is in the case of insect vision. Many insects see the world through compound eyes, which are composed of thousands of tiny hexagonal facets. These are packed together in a hemispherical shape and protrude from insect's head to provide visual field which can exceed 180° for each eye, creating a total field of view in excess of 270° (Mazokhin-Porshnyakov 1969). However, the angular width of the field of stereoscopic vision is much less, and rarely reaches 90°.

Ideally a driver warning system should have a wide field of view, with a large portion of this contributing to stereoscopic vision. If possible the field of view should exceed 180°, which would enable the system to monitor the entire area behind the vehicle, and in certain situations also the side of the vehicle (Figure 3.13). There are several methods for panoramic imaging which could be used for this purpose. This is discussed in detail in the following section.

2.3 Panoramic Imaging

There are three main approaches to obtaining images with a wider field of view. One could use multiple images from several cameras, or a single rotating camera and combine this visual information. However, it may be preferable to use a single static camera with a wider field of view. This could be achieved through the use of special lenses or convex mirrors.

2.3.1 Multiple Images

It is possible to take several images, and combine them into a single panoramic image. This process of stitching the images together is known as "mosaicing" (Szeliski 1994). A mosaic can be created using either a multiple camera system, or a rotating camera.

Rotating Camera

This method uses a single camera which is rotated around a vertical axis. The edges of the different images are matched and then stitched together to form a panoramic image. The

main advantage of this method is that it is possible to obtain high resolution panoramic images depending on the angular resolution of rotation. On the other hand, there are several significant disadvantages of this method:

1. Long image acquisition time as many images are required. As a result, this approach is unsuitable for real-time applications.

2. This system requires moving parts, which decreases it's robustness and increases cost.

3. Although it is possible to view 360° in the azimuth direction, the elevational field of view is limited to that of a conventional camera.

It is also possible to perceive depth information using this technique (Peleg & Ben-Ezra 1999, Peer & Solina 2002). This can be achieved by using two camera, one above the other, rotating around the same axis. Alternatively, range data can be determined by placing a single camera a certain distance away from the axis of rotation. If an object can be viewed in two images acquired at different points in the rotation, the range to that object can be determined using stereopsis.

Multiple Cameras

A simple approach is to use several conventional cameras (Yagi 1999). Depending on the field of view of the cameras used, it would be necessary to use at least four cameras to cover 360° in the azimuth direction. If we would like to increase the elevational field of view, this would increase yet again. The more cameras required, the less attractive this option become, due to the increased cost. Another drawback of the multiple camera approach is the difficulty in making such a system compact.

2.3.2 Wide Angle Lenses

We have learned from the previous section that systems with rotating or multiple cameras are less attractive for driver assistance systems. It would be more appropriate to use a sensor that had no moving parts, and captured the necessary field of view with a single camera. This could be achieved through the use of an ultra wide angle lens. In particular fish-eye lenses and panoramic annular lenses have been used to increase the area viewed by conventional cameras.

A fish-eye lens can provide a field of view in excess of 180°, effectively a complete hemisphere with a single camera. However, these lenses are bulky and expensive (Chahl

& Srinivasan 1997). Furthermore, fish-eye lenses suffer from angular distortion, which can be difficult to remove. Images taken with these cameras typically have good resolution in the centre, but this deteriorates in the periphery. These lenses have been used to control the position of a mobile robot using given targets (Cao et al. 1985), and following lines (Elkins & Hall 1994).

Another type of refractive optic that can be used to widen the area viewed by a camera is a Panoramic Annular Lens (PAL). This has been used in mobile robotics for the localisation of moving obstacles (Zhu et al. 2000). The PAL is fairly complicated, consisting of a single glass section with two reflective and two refractive planes. It has a 360° field of view in the azimuth, but only about 40° in elevation. Unfortunately it is difficult to increase the viewing region of this sensor in the elevation direction.

2.3.3 Convex Mirrors

Convex mirrors are one approach to panoramic imaging that has been utilised widely in the field of robotic navigation (Yagi et al. 1994, Matsumoto et al. 1999). The sensor consists of a video camera situated below a convex mirrored surface. The optical axis of the camera is aligned with that of the mirror, which enables the camera to view a full 360° of the surrounding environment. The minimum and maximum angles of elevation captured are dependant upon the profile of the mirror surface.

This method of panoramic imaging has several distinct advantages:

1. It is a passive sensor and as a result, the power requirements are minimal.
2. The lack of moving components means that the sensor could be made cheaply, in a robust manner, and require little maintenance.
3. The optical distortion caused by the mirror can be removed easily.

An example of an image from a panoramic sensor is shown in Figure 2.9 (a). Such raw images are difficult for humans to understand, and require the modification of conventional image processing techniques. However, it is possible to unwarp them to create a more intuitive panorama, as seen in Figure 2.9 (b). The distortion can be removed using several methods:

1. Transformation from a Cartesian to polar coordinate system.
2. Projection of the warped image onto a shape in 3D space, such as a cylinder.

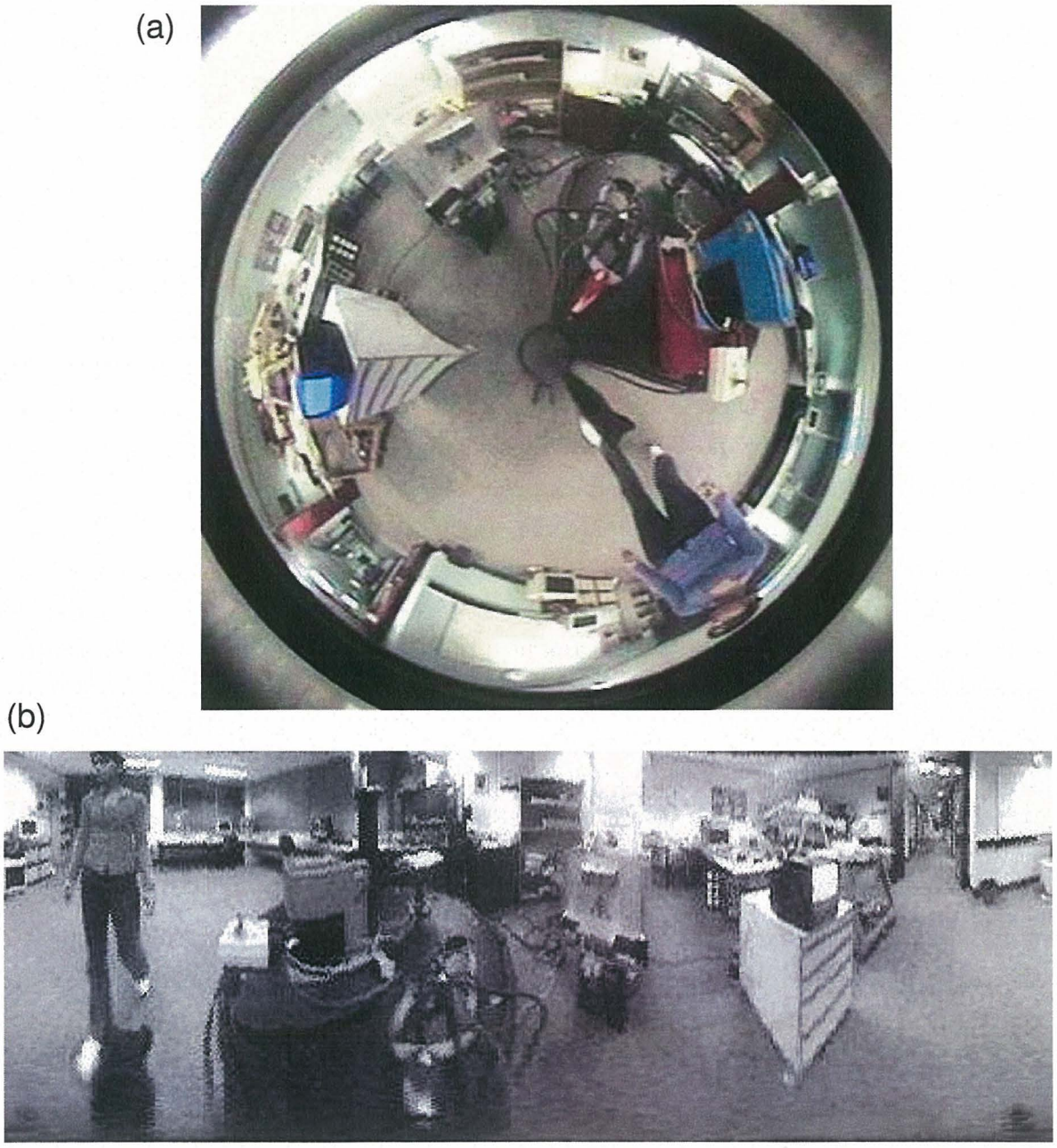


Figure 2.9: (a) An example of an image taken from a panoramic sensor that uses a convex mirror. (b) The corresponding unwarped image.

3. Radial correction around the image centre to create a "bird's eye view".

The effect of the first method is to unwrap the raw image about the origin (Chahl & Srinivasan 1997). Assuming that the centre of the image is the origin (0, 0) and all variables are scaled to range between 0 and 1, the equations for unwarping are as follows:

$$\begin{aligned} x &= y_u \cos(2\,\pi\, x_u) \\ y &= y_u \sin(2\,\pi\, x_u) \end{aligned} \tag{2.1}$$

where (x_u , y_u) is a point in the unwarped image, and (x, y) is a point in the raw image.

The second method is similar to the first, except that the image is also unwarped along the radial lines in the image. This is done by projecting each pixel in the warped image back onto a theoretical cylinder in three dimensional space, using a method similar to that described in (Gluckman et al. 1998). Each pixel value p' on the panoramic cylinder is back projected to the mirror, and finally onto the original image to determine p, to obtain the panorama, as shown in Figure 2.10. This is done using Equation 2.1 to unwarp the image.

The third method of unwarping provides a *bird's eye view* of the scene reflected by the convex mirror. This has been used in the field of mobile robotics to provide a means of navigation (Gaspar et al. 2000). In this case, unwarped images are generated by radial correction around the image centre, which results in a scaled orthographic projection of the ground plane. It is also possible to produce these ground plane unwarped images through the use of a specially designed convex mirror profile (Hicks & Bajcsy 1999).

In our project, we have chosen to use the second method for unwarping panoramic images. This approach not only unwarps about the mirror axis, it removes the warping introduced by the curved mirror profile. Thus, the unwarped image is independent of the shape of the mirrored surface used by a panoramic sensor and enables the use of conventional image processing techniques.

2.3.4 Summary

Panoramic sensors with rotating cameras would require maintenance, due to the moving components, and the time to capture one image makes this method unsuitable for real-time applications. Wide angle lenses do not require moving parts, however they are bulky, expensive, and suffer from wide angle distortion that cannot be easily removed. Of the three panoramic imaging methods discussed in Section 2.3, the use of convex mirrors was

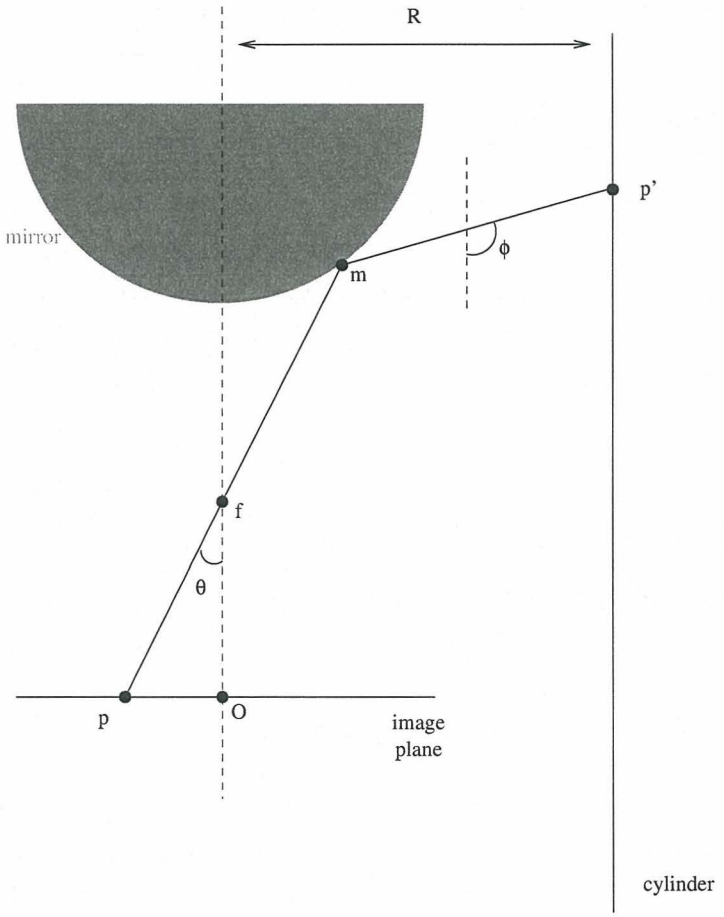


Figure 2.10: Mapping an image point p to a point in the unwarped image p'. The image point is back projected onto a cylinder in three dimensional space, with a radius R.

the most appropriate method for monitoring vehicle blind-spots. It is a passive sensor, with small power requirements, and is robust and lower in cost in comparison to a sensor with moving components. The images captured using mirrors can also be unwarped to remove optical distortion that is typically seen in systems using lenses. The sensor is quite versatile, as it can be placed in a number of positions around various types of vehicles in a horizontal or a vertical configuration. This will be discussed in more detail in Chapter 3.

2.4 Mirror Profiles

When using a convex mirror for panoramic imaging, the camera is pointed towards the mirror with the camera axis aligned with that of the mirror. This provides the camera with a 360° view of the surrounding environment in real time, as shown in Figure 2.9. However, the field of view in the elevation direction is dependant upon the profile of the mirror surface. We will now discuss some of the various mirror profiles that have been developed in the past.

2.4.1 Conical

One of the simplest shapes of convex mirror is the cone. Given the field of view of the camera, the angle of the apex of the cone can be determined using Snell's law of reflection:

$$\phi = \pi - 2\beta + \theta \tag{2.2}$$

Where ϕ is the angle of elevation, β is the half-angle of the apex of the cone and θ is the angle of the reflected ray entering the camera, as shown in Figure 2.11.s Although 360° can be imaged in the azimuth direction, the field of view in the radial direction is limited to half the field of view of the camera. This can be increased by making the convex mirror profile curved.

2.4.2 Hyperbolic and Parabolic

There are two types of curved convex mirror that have a *single viewpoint constraint* - hyperbolic (Baker & Nayar 1999) and parabolic (Gluckman et al. 1998) surfaces. The single viewpoint constraint requires that the rays of light that are reflected from the mirror into the camera intersect at a single focal point, as shown in Figures 2.12(a) and (b). As a result, the effective pinhole camera model (see Section 4.3) may be used, and therefore

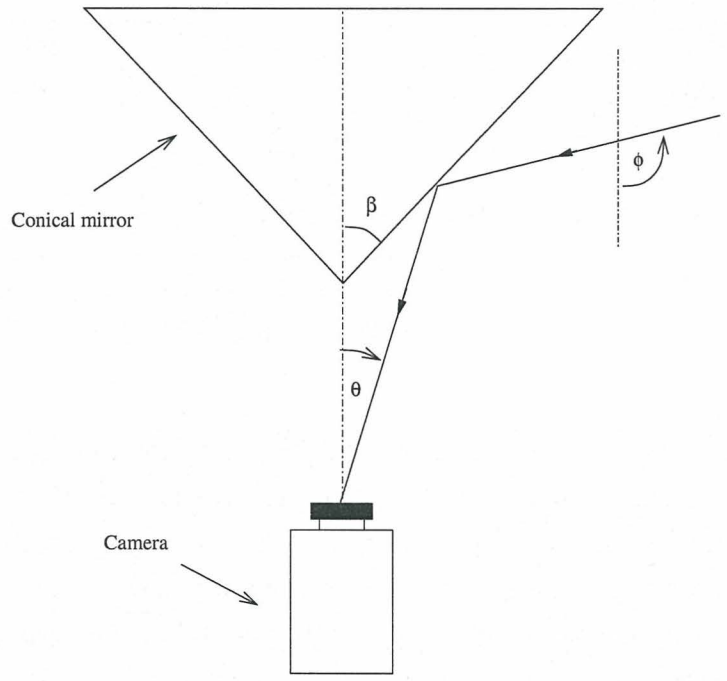


Figure 2.11: A conical convex mirror to produce panoramic images (Chahl and Srinivasan 1997).

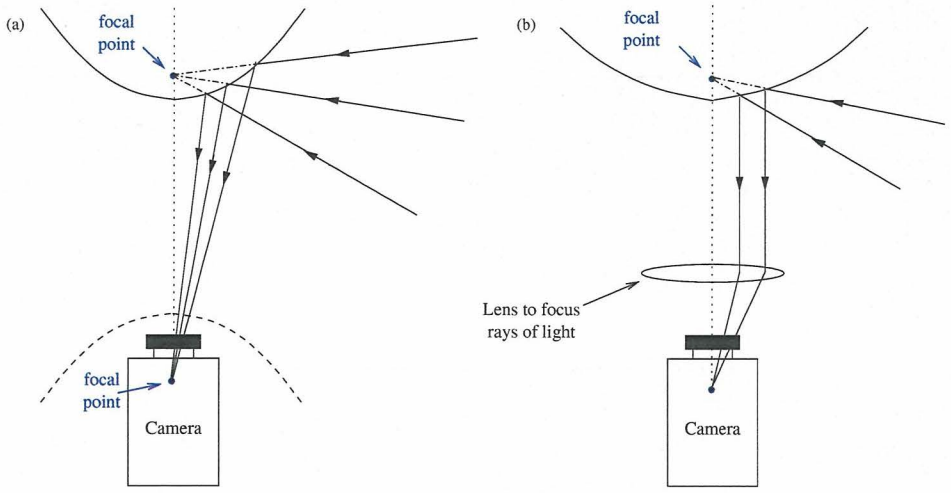


Figure 2.12: Panoramic mirrors which conform to the single view point constraint. That is, all rays of light reflected from the mirror to the camera would all intersect at a single focal point. (a) A hyperbolic profile - the camera must be placed at the second focus of the hyperbole. (b) A parabolic profile - a lens is required to focus the rays of light into the focal point.

perspective images can be generated from the warped panoramic images. This property is useful for applications such as virtual tours of an environment. The sensor can be used to capture an image of a scene, then sections of the image can be unwarped to provide the user with a perspective image of an area of interest. However this is not necessary for pure range sensing applications, and requires further calibration and equipment. In the case of the hyperbolic mirror, the camera focus must coincide with the focus of the second hyperbola (which is not used in reality), otherwise the single viewpoint assumptions is violated. The parabolic mirror causes the reflected rays of light to become parallel, which requires the use of a special lens to focus the light rays, or an orthographic camera with built in tele-centric optics.

2.4.3 Constant Gain

An important parameter for describing the profile of a mirror is known as the *mirror gain.* This is the relationship between the changes in the incidence angles of light rays and the changes in the direction of the reflected rays. If the relationship is linear, it is then known as constant mirror gain, and can be defined as follows:

$$\alpha = \frac{\delta\phi}{\delta\theta} \tag{2.3}$$

where α is constant, ϕ is the elevation angle of a ray of light reflected off the mirror, and θ is the angle viewed by the camera, as shown in Figure 2.13.

Basically, this parameter specifies the angular magnification of the surface in the vertical direction. The conical mirror described in the Section 2.4.1 has a unity gain, and is therefore also a planar mirror. By increasing the gain of the mirror, we can increase the field of view of the panoramic sensor.

For a full derivation of the constant gain surface profile, see (Chahl & Srinivasan 1997, pp 8277-8278) and (Conroy 2000, pp 64-67). To design a convex mirror, the first step is to decide the maximum and minimum light ray angles of elevation ($\overline{\phi}$ and $\underline{\phi}$ respectively) and to determine the maximum and minimum angles viewed by the camera ($\overline{\theta}$ and $\underline{\theta}$). The mirror gain can then be found as follows:

$$\alpha = \frac{\overline{\phi} - \underline{\phi}}{\overline{\theta} - \underline{\theta}} \tag{2.4}$$

Consider a mirror profile (r,θ) in polar coordinates, where r is the radial distance to the camera, and θ is the angle from the optical axis of the camera to the reflection point on the mirror surface as shown in Figure 2.13. The general equation for the constant gain mirror is:

$$r = r_0 \left[\frac{\sin\underline{\gamma}}{\sin\left[\underline{\gamma} + \frac{(\underline{\theta}-\theta)(1+\alpha)}{2}\right]} \right]^{\frac{(1+\alpha)}{2}} \tag{2.5}$$

where $\underline{\gamma} = \frac{\pi-\underline{\theta}-\underline{\phi}}{2}$, and r_0 is the distance from the camera focal point to the mirror assembly when $\theta = \underline{\theta}$.

2.4.4 Resolution Invariant

The vast majority of the work completed in the field of panoramic imaging has utilised CCD camera technology. These CCD cameras have been used to capture images reflected from a convex mirror, which is essentially a polar image of the surrounding environment. In practice, this means that a Cartesian array of pixels has been used to capture the polar image, and as a result the pixel density per angle increases with the radius of the polar image. That is, a ray of light reflected from the edge of the mirror will be captured with

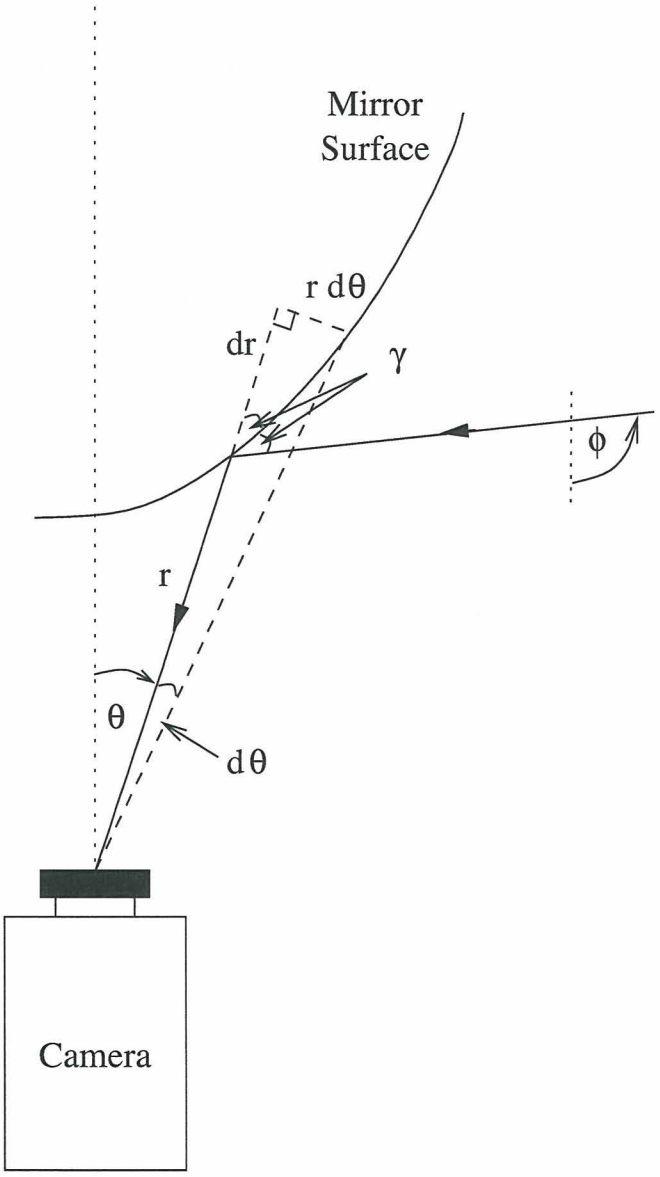


Figure 2.13: Relationships used to derive a family of constant gain mirror profiles (Conroy 2000).

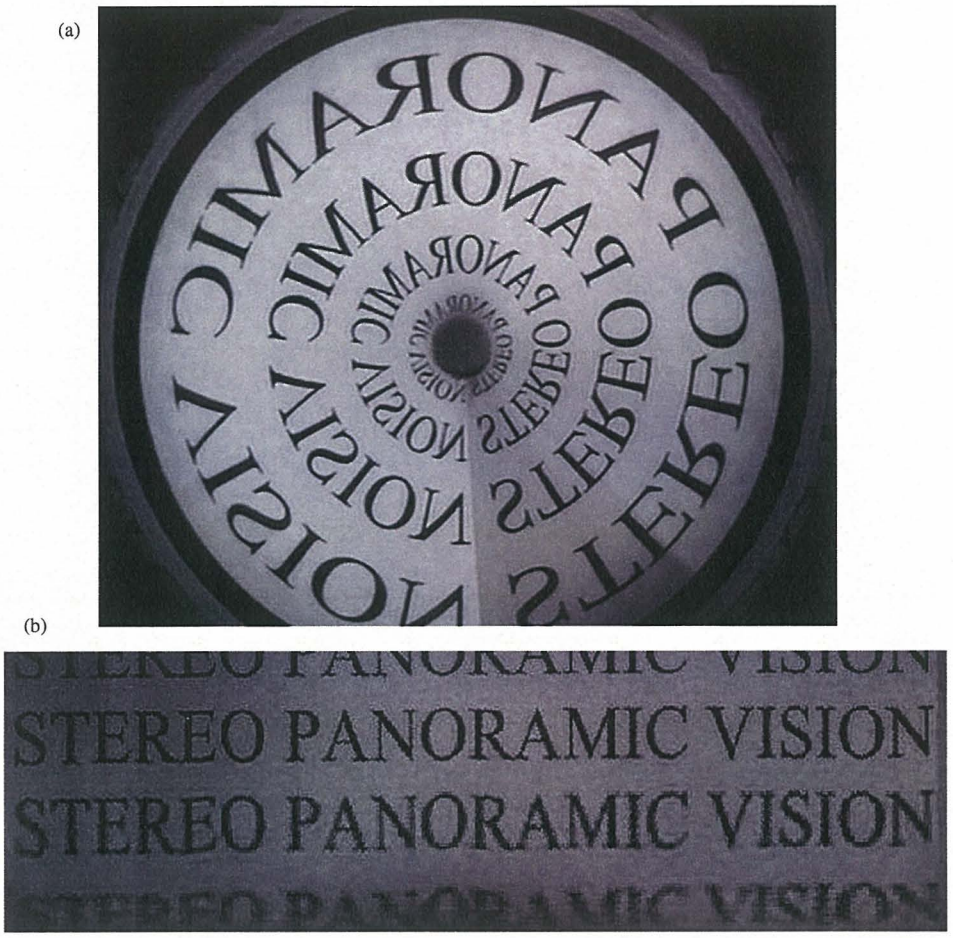


Figure 2.14: (a) The raw image captured by the cameras. (b) The unwarped image, showing the variation in image quality.

a higher density of pixels than a ray of light reflected towards the centre of the mirror, as shown in Figure 2.14.

One way to compensate for this problem would be to use a specially designed camera with a polar array of pixels (Panerai et al. 1995), as shown in Figure 2.15(a). The pixel density is greatest in the centre of the CCD array, and decreases towards the edge. The main advantage of these cameras is that the panoramic images can be unwarped in hardware by scanning the pixels radially. However, these cameras are relatively expensive, and since the unwarping of images in software is not demanding, the costs cannot be justified. Furthermore, the alignment between the camera and the mirror would require a high level of accuracy, and further complicate the calibration process.

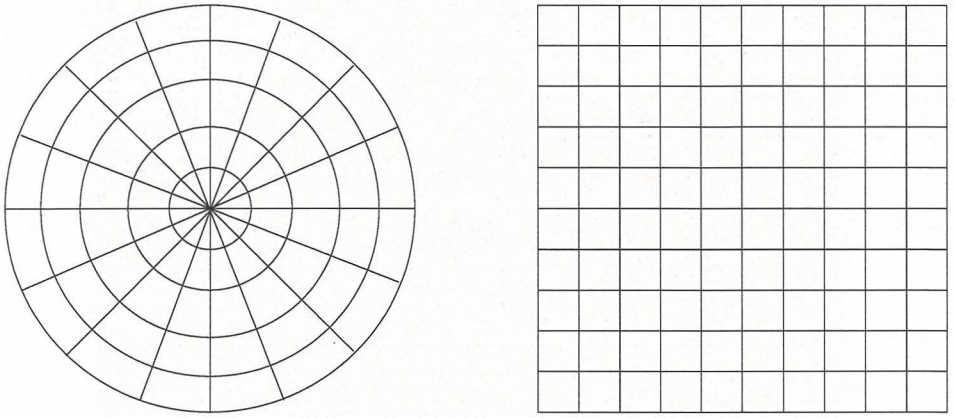


Figure 2.15: (a) A polar array of pixels and (b) a conventional rectangular CCD array.

Another approach is to use a mirror profile which has been designed to compensate for the change in pixel density - the *resolution invariant* surface, developed by (Conroy & Moore 1999). The resolution invariant mirror profile can be determined by solving the following equation numerically:

$$\frac{d\,r}{d\theta} = r\ \cot\left[-\frac{1}{2}\int\left(1+\alpha(\theta)\right)\,d\theta\right] \tag{2.6}$$

where r is the distance from the camera focal point to the mirror surface, θ is the angle viewed by the camera, and $\alpha(\theta)$ is the mirror gain given θ, as shown in Figure 2.13. The variable mirror gain $\alpha(\theta)$ is given by:

$$\alpha(\theta) = B_\alpha\left[\tan\theta + \tan^3\theta\right] \tag{2.7}$$

where

$$B_\alpha = \frac{2\left(\overline{\phi}-\underline{\phi}\right)}{\tan^2\overline{\theta}-\tan^2\underline{\theta}} \tag{2.8}$$

2.4.5 Summary

Conical mirrors are the simplest convex mirrors that have been utilised in panoramic sensing. Although the field of view is limited to half that of the camera, in the elevation direction. Hyperbolic and parabolic surfaces have an increased field of view, and also have

an effective single camera viewpoint. However, this property is not necessary for range sensing, and requires further calibration and specialised optics or cameras. We decided that two of the presented mirror profiles would be tested in this project, as they both have interesting properties that could be useful for range sensing. The constant gain mirror profile was chosen because of the constant angular resolution, while the resolution invariant profile produces a panoramic image with invariant pixel density.

2.5 Panoramic Vision Range Sensors

Biologically inspired binocular vision systems are used extensively in computer vision, as mentioned previously. However, such systems typically utilise cameras with small field of views to minimise wide angle lens distortion. It is also possible to estimate range using stereo panoramic systems. There has been research results published in this area, however, the majority of the results presented have been theoretical. The published work to date is surveyed as part of this research project.

In (Bunschoten & Kröse 2001), the authors described a method for estimating range from a pair of omnidirectional images, captured from a single panoramic sensor on a moving platform. They determined the epipolar geometry for a panoramic sensor, using unwarped images that had been projected on to cylinders in three-dimensional space. The system was tested only on theoretical, computer generated images captured with hyperboloidal mirror profiles. The virtual stereo sensor had a horizontal baseline of 50cm, and was placed at the centre of a box (6m x 6m x 6m) with textured sides and bottom. The panoramic images used were 720 x 200 pixels, and were unwarped from raw images of 640 x 480 pixels. The distance to the walls were estimated with an average error of approximately 0.2m.

A virtual stereo sensor was placed in a synthetic environment, consisting of four textured walls, at a distance of three metres from the sensor. Their results showed that range to the walls could be reliably determined.

(Ollis et al. 1999) examined the several different possible camera/mirror configurations for a stereo panoramic sensor, as shown in Figure 2.16. It was reported that the range precision was determined by the lower resolution image portion, and so the detail provided by the higher resolution image portion would be wasted. Therefore, camera configuration (1) was found to have the highest range accuracy, as any object in space would appear with approximately the same resolution in both panoramic images. However, only results from theoretical analysis were presented. The experiments used constant gain mirrors ($\alpha = 3$),

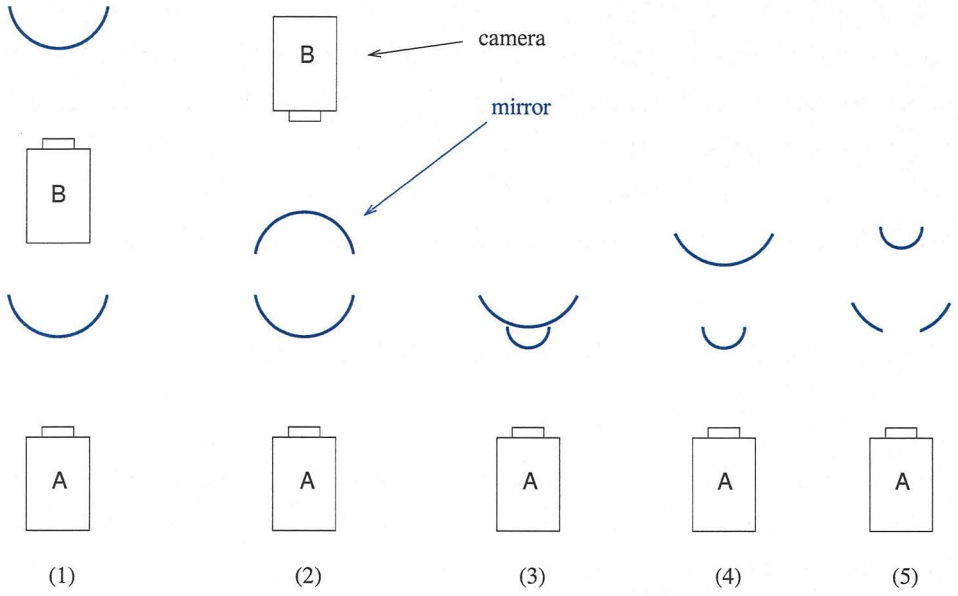


Figure 2.16: The five sensor configurations examined by Ollis *et al.*

with a field of view of 90 degrees in the vertical direction, and a baseline of 30cm. At a range of 5m the measurement error was reported to be 12cm.

In Figure 2.16, configuration (3), (4), and (5) have the advantage that they only require the use of a single camera, which decreases the cost of the sensor. Configuration (3) was investigated in detail by (Conroy 2000), again using virtual images. The main drawback of this configuration was the small baseline (only a few centimetres), which limited the range finding capabilities to a maximum of about 1.5m. Conroy presented some preliminary theoretical range estimation results, using resolution invariant mirror profiles. At a range of 1.5m, the maximum measurement error was reported to be 45cm.

The baseline could be increased to improve range estimation by using configurations (4) or (5). However, the range accuracy is still severely decreased due to the image resolution, since two panoramas are viewed by a single video camera.

(Gandhi & Trivedi 2004) have presented an obstacle detection system that uses a single panoramic camera. The system compensates for ego-motion, then detects objects that have independent motion. The sensor was mounted 60cm above the roof of the test vehicle, to provide a 360° view of the surroundings. The road plane motion parameters were estimated using apriori knowledge of camera calibration and vehicle speed. These parameters were then used to compensate for motion in consecutive frames, and any remaining motion is

deemed to be an object. The accuracy and reliability of the system was not reported. The disadvantages of this system are the impractical sensor mounting, as well as the requirement of vehicle odometry.

(Ng et al. 1998) and (Gluckman et al. 1998) presented experimental results gained from two different panoramic sensor systems. The first system consisted of four panoramic sensors placed at equal height around a room. They were able to generate perspective images of the surrounding environment, and preliminary range estimation, showing only relative depth was presented. The second system consisted of two panoramic sensors, with the camera and mirror axes aligned. Again, only preliminary range estimation results were discussed. In both cases, no evaluation of the range finding capabilities were carried out, and there was no further processing of the data.

Stereo panoramic sensors have also been used in mobile robot navigation (Koyasu et al. 2003). The authors presented a system similar to that described by (Gluckman et al. 1998) with the camera and mirror axes aligned. The range information computed from the sensor was used for path planning and avoidance of both static and moving objects. An experiment was carried out in which the robot successfully avoided a person walking across the planned path. Diagrams and photographs were used to describe the experimental results. However, no scale was provided for the diagrams and the range accuracy of the sensor was not discussed.

In (Sogo & Ishiguro 2000), the authors describe an extension of the system presented in (Ng et al. 1998). They used the same multi sensor system placed at equal height around a room, rather than aligning along the camera axes. Obstacles were detected using a background subtraction method, where the image intensities of all pixles are subtracted in two sequential frames. Any non-background objects would remain. The azimuth angle of the obstacles detected from each sensor were combined to determine the position of the obstacles in space. The system is able to determine the position of a target to within 5cm for targets within 3m from the sensors. However, this system has several disadvantages:

- Added cost of the extra sensors.
- Detection of false objects, as shown in Figure 2.17, due to the fact that obstacle position is determined by the azimuth angle from three or more sensors.
- Obstacle detection method used is unsuitable for a sensor located on a moving platform, as it relies on a static background.

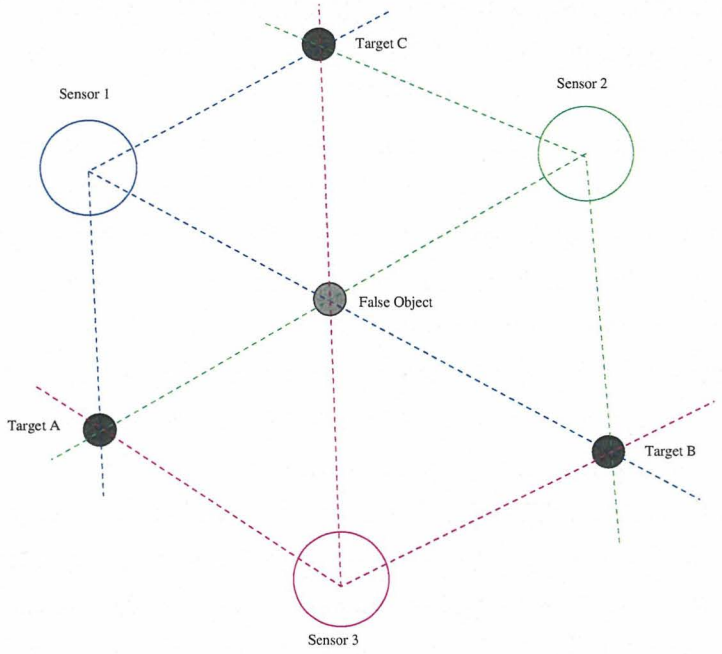


Figure 2.17: False matchings for an N-ocular panoramic vision system, with the sensors aligned along equal heights rather than camera axes.

Examining the research results published by (Ollis et al. 1999), we decided to use camera configuration (1), as shown in Figure 2.16. This configuration was found to have the highest range accuracy, as the panoramic images would have approximately the same resolution at each pixel. To date, there has been a distinct lack of results describing the performance of panoramic sensors in range estimation, in real-world experiments. The majority of the results have been either purely theoretical, or only preliminary range estimation studies. We decided to investigate a stereo panoramic sensor, to evaluate the range estimation capabilities and process the captured images in order to segment obstacles from a moving platform.

2.6 Chapter Summary

In this chapter, we have presented the various products currently available to help drivers avoid rear collisions and other blind-spot accidents. Visual aids increase the driver's field of view, so they can monitor problem areas. The main disadvantage of these devices is that they rely on the human observer to be alert, and understand how to interpret

the supplementary visual information provided. The second approach is the use of an automated monitoring system. These utilise range sensors, such as sonar and radar to gather range data, and the system warns the driver in the event of an impending collision. However, these systems have a low angular resolution and/or a limited sensing volume.

We then suggested the use of computer vision. To utilise a vision system in this application, it would be necessary to extend the field of view of the sensor. Multiple cameras, rotating cameras, and wide angle lenses were all considered. We decided that the use of a convex mirror would be the most suitable approach to panoramic imaging. As a passive sensor, power requirements are small, and the lack of moving components increases robustness. Unlike wide angle lens distortion, it is comparatively easy to remove the optical distortion caused by the curved mirror profile.

A survey review of stereo panoramic vision was undertaken. The results presented in the literature have been either theoretical, or preliminary range estimation studies, with only raw disparity maps. Only one system (Sogo & Ishiguro 2000) was tested to determine the accuracy of the range finding capabilities. However, this system relies on three or more panoramic sensors, increasing the cost of the system. Also, the obstacle detection method cannot be used for a sensor on a moving platform, as it requires a static background.

In this research project we decided to use a panoramic vision system which utilised convex mirrors, containing only two sensors to reduce economic costs. The camera axes are aligned, with each sensor oriented in the same direction, as suggested by (Ollis et al. 1999), to provide the highest combined resolution for range finding. The sensors are quite versatile can be placed in a number of positions and configurations. The following chapter will discuss these possibilities and provide an overview of the prototype vision system that was built for this project.

Chapter 3

System Overview

OUR stereo panoramic vision system consists of three main parts, as shown in Figure 3.1. The stereo panoramic sensor is mounted on a vehicle, in a position to view an area of the environment that is unseen by the driver. The sensor captures monochrome images of the blind-spot, which are fed to the host computer. These images are processed using an on-board PC to determine where obstacles are situated in the work space. The results can be sent to another system which would provide a warning to the driver. However, the issue of providing a driver warning system is outside the scope of this research project.

In this chapter, details of the hardware components used in the obstacle detection system are given. This is followed by an overview of the software structure used to process the stereo images. The driver assistance system designed and built in this project is a prototype, and therefore needs to be refined before would be suitable for inclusion in a commercial vehicle. The final section of this chapter presents some conceptual designs, to show that it is possible to integrate a stereo panoramic system into several types of vehicle, in a practical and useful manner.

3.1 Hardware

This section provides the details of the major hardware components required for the prototype stereo panoramic vision system. The system contains two panoramic sensors, a mounting device to attach the sensors to the car, and an on-board PC to process the images and display the experimental results.

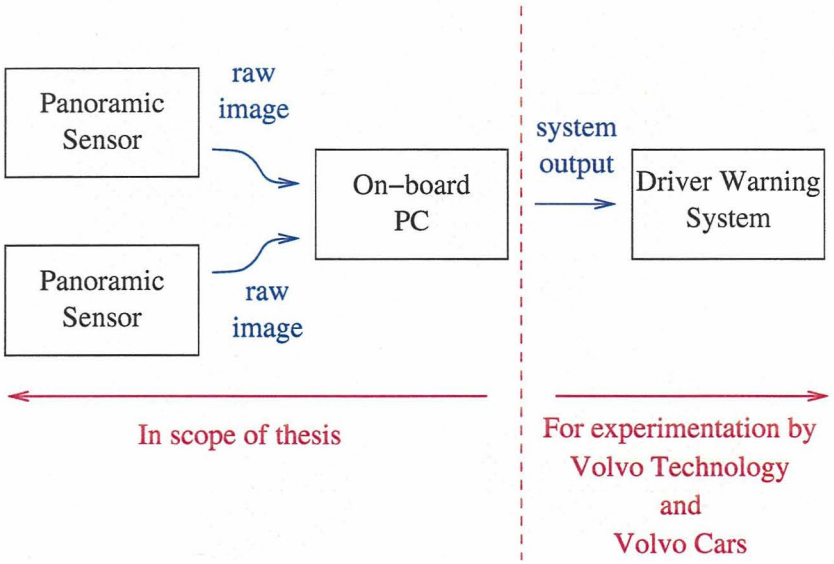


Figure 3.1: Overview of the stereo panoramic vision system.

3.1.1 Panoramic Sensor

The panoramic sensors designed for this project consist of the following:

- Curved mirror surface
- Video camera
- Glass cylinder
- Aluminium collars
- Needle

We have already discussed the purpose of the curved mirror surface situated over a video camera to increase the field of view in Section 2.3.3. Two off the shelf Sony FCB-EX470L analog video cameras were used to capture images of the curved surface. One of the sensor requirements was that the mirror must be aligned with the axis of the camera. This is done by using a glass cylinder with aluminium collars at either end to attach to the mirror and the camera. The glass cylinder also keeps the mirror at a constant distance from the camera. Although the needle does slightly decrease the panoramic sensor field of view, it is useful for eliminating the internal reflections caused by the glass cylinder (Sogo &

Figure 3.2: (a) An assembled panoramic mirror. (b) An exploded view of the panoramic mirror assembly.

Ishiguro 2000). The technical drawings for each of these components can be viewed on the attached CD.

As mentioned in Chapter 2, a system containing two panoramic sensors was built. The camera and mirrror axes are aligned to decrease computational costs, which is explained further in Section 4.1. Each sensor will be oriented in the same direction to ensure the highest possible combined resolution at each image point, to improve the range accuracy, as suggested by (Ollis et al. 1999).

3.1.2 Mounting Equipment

One of the aims of this research project was to ensure that the driver assistance system developed should operate in a real world environment (Section 1.2). Therefore, the prototype system was tested by mounting it to a test vehicle, to evaluate the performance under conditions relevant to the intended application. The mounting equipment had the following design requirements:

- Variable base line between the sensors
- Camera and mirror axes aligned
- Horizontal and vertical configurations of the sensors

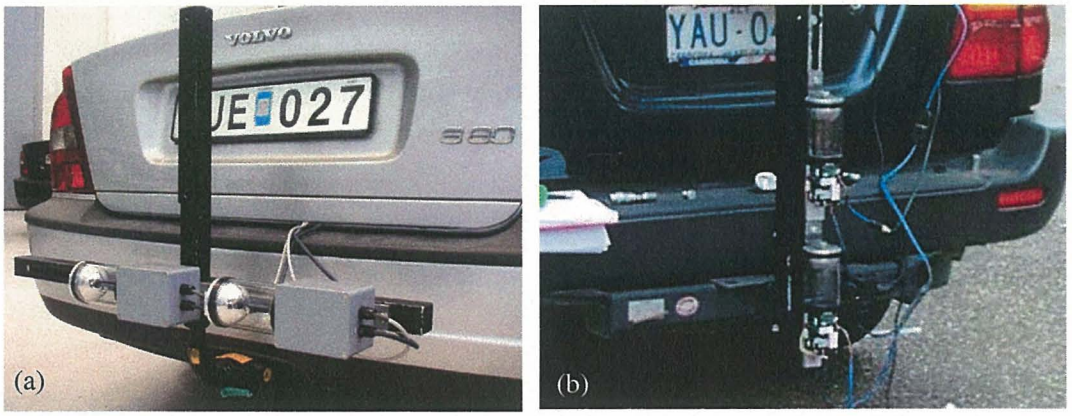


Figure 3.3: The stereo panoramic sensor, attached to the test vehicle in (a) a horizontal configuration, (b) a vertical configuration.

- Mount to tow bar, to enable the equipment to be removed easily, yet secured firmly.

The prototype equipement can be viewed in Figures 3.3(a) and (b). The research for this thesis was carried out in both Canberra, at the Australian National University and Gothenburg, at Volvo Technology Corporation. As a result, it was necessary to build two camera mounts, to attach the sensor to two different test vehicles.[1]

3.1.3 Frame Grabbers and Host Computer

Two *Imagenation* PXC-200AL frame grabbers were installed to capture images from the video cameras (i.e. one for each camera). These devices were used to capture the analogue signal recieved from the cameras and convert it to a digital image to be processed by the on-board PC. The experimental computer used in this thesis was a 2.4 GHz Intel Pentium IV.

3.1.4 Driver Warning System

In a more developed version of the stereo panoramic vision system, given the information from the processed images, the driver would be warned if a collision is likely to occur (Figure 3.1). This would be in the form of a simple audio or visual signal. It is also concievable that a more sophisticated visual display could be employed to provide the

[1]The technical drawings can be found on the attached CD in the *.tif* files in the directories `/Mirror Design/ProEngineer/` and `/Mirror Design/ProEngineer/camera_mount/`.

driver with more specific information. The driver would not only know how close the objects were, but also where they were, and they could use this information to decide how to control the vehicle to avoid the object.

However, the design an implentation of such a system was outside the scope of this project. Instead, for experimental purposes, the output of the obstacle detection and range estimation was displayed on the monitor of the on-board PC.

3.2 Software

The software for this project has a simple cascade design, as shown in Figure 3.4. The raw panoramic images are captured, then unwarped using a pre-computed lookup table, to decrease the loop execution time. The unwarped images are used to detect obstacles, and finally the range the each obstacle is estimated. This will be discussed in detail in Chapter 4.

A Graphical User Interface (GUI) was also implemented to display the results, and enable the user to modify the algorithm parameters at run time. For further documentation, see the *User Guide* on the attached CD.

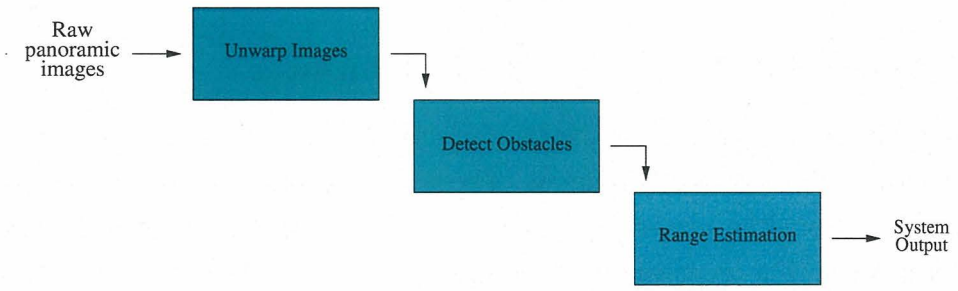


Figure 3.4: The software structure.

3.3 System Calibration

Before the stereo panoramic sensor can be used to accurately determine the range to obstacles, it is necessary to calibrate the various hardware and software components of the system. Firstly, the video cameras must be calibrated to calculate the intrinsic parameters

used in range estimation. Finally, the correct unwarping parameters are required to produce accurate disparity maps.

3.3.1 Camera Calibration

It is necessary to calibrate the video cameras to determine the internal parameters, such as the focal length, which will be used in the range calculations. The cameras must be calibrated for each zoom setting to be used. Calibration was done using the MatLAB Camera Calibration Toolbox[2] to calculate the values of the intrinsic parameters of the two video cameras. The toolbox requires 10 to 20 images of a calibration checker board with squares of a known dimension, in various orientations. Examples of calibration images are shown in Figure 3.5.

The MatLAB Camera Calibration software uses a linear technique to estimate the internal parameters, in a similar manner to that described by (Zhang 1999). However, a different intrinsic model is employed, with the model presented by (Heikkila & Silven 1997) being used instead. A Maximum Likelihood estimation is then used for final refinement of the internal parameters.

The cameras used for the stereo panoramic vision system were calibrated using the Matlab Camera Calibration Toolbox, as described above. The calibration results were sufficiently accurate, with the focal length and principle point determined to within 0.22 pixels in the x direction, and 0.21 pixels in the y direction.[3]

3.3.2 Unwarping Parameters

System calibration is required to generate panoramic images that produce accurate disparity maps. The centre of the mirror is not always located at the centre of the raw image. There may also be a slight misalignment in the rotational position of the two panoramic sensors about the mirror/camera axis. Both of these problems would cause a misalignment of the epipolar lines, decreasing the likelihood of an accurate match between corresponding features in the stereo images. Therefore, a calibration tool was integrated into the software developed for this system. Sample images from the graphical user interface (GUI) are shown in Figure 3.6. This enables the user to determine the necessary parameters,

[2]The Matlab calibration tool box is freely available at: http://www.vision.caltech.edu/bouguetj/calib_doc/

[3]The detailed calibration results can be found on the attached CD in the folder `/Matlab code/Camera Calibration/`. The files containing the calibration output are `leftCameraCalib.rtf` and `rightCameraCalib.rtf`.

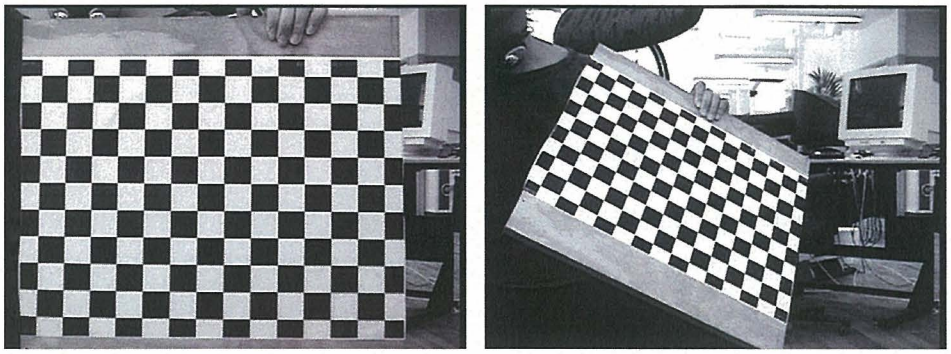


Figure 3.5: Examples of camera calibration images.

which are then used to compensate for such misalignments when generating the panoramic images.

As discussed in the following section, certain areas of the warped image may not be useful for the purpose of obstacle detection. In our application, the vehicle the sensor is mounted on obscures a portion of the field of view. The calibration tool is also used to select an area of interest in the image, and only this segment is unwarped into the panorama. The area between the green and red lines in Figure 3.6 show the angles where unwarping begins and ends. This enables the user to remove the region of the image containing the vehicle the sensor is mounted to. The upper and lower edges of the unwarped image are defined by the dark and the light blue circles.

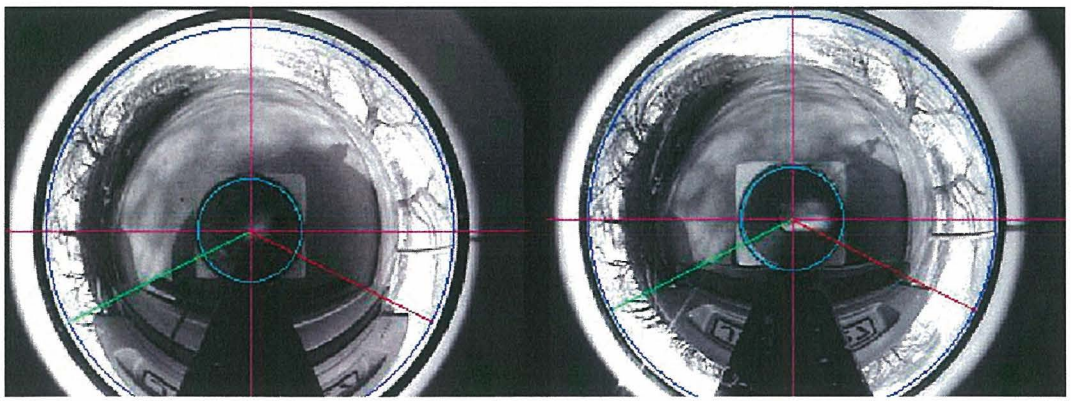


Figure 3.6: Sample images from the graphical user interface used for calibrating the stereo panoramic vision system.

An aid that can be used in conjunction with the calibration GUI is shown in Figure 3.7.

This equipment was designed in our laboratory for the calibration of the panoramic cameras. When the mirror is placed in the centre of the barrell, the stripes in the cylinder will be unwarped to horizontal lines in the panoramic images. The unwarping parameters were changed using the GUI until it was apparent that the system parameters were set correctly. That is, the centre of the warped image was changed until the black and silver sections of the calibration cylinder were transformed into straight, parallel lines in the unwarped image, as shown in Figure 3.8. An example of the use of incorrect unwarping parameters is shown in Figure 3.9.

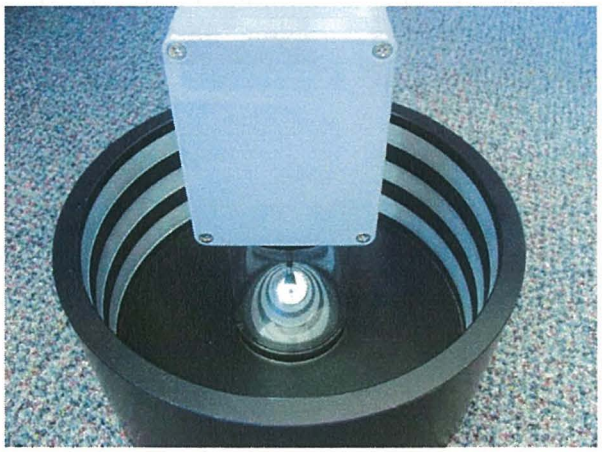
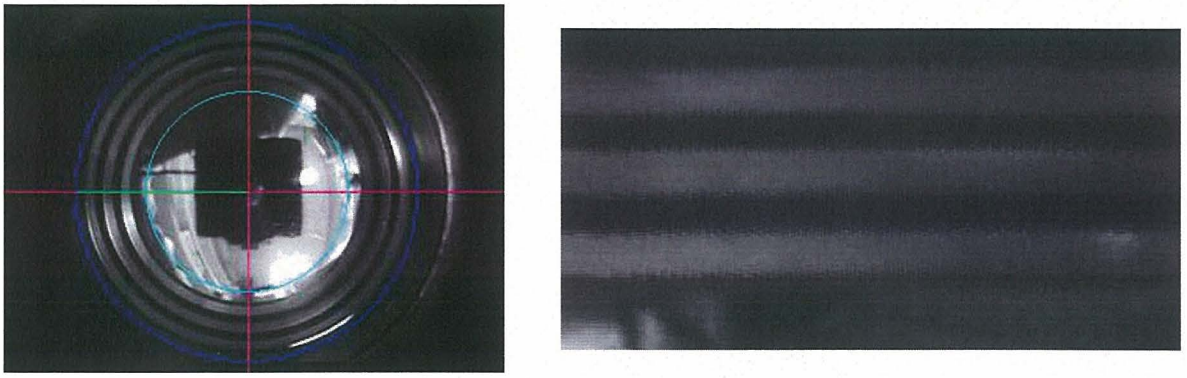


Figure 3.7: Precisely machined equipment for calibrating the panoramic cameras. When the mirror is placed in a hole in the centre, the stripes machined in the cylinder should unwarp to horizontal lines.

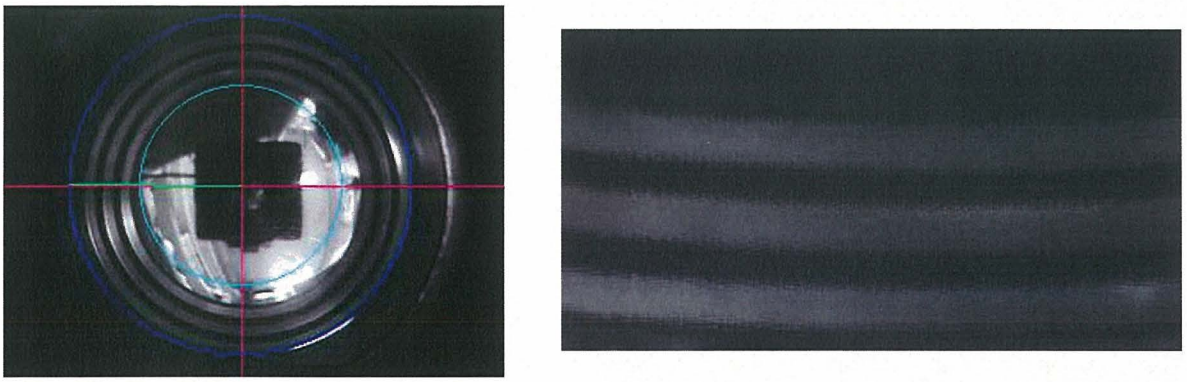
3.4 Conceptual Drawings

As mentioned in Section 1.2, the aim of this project is to develop a vision system to distinguish obstacles from the surrounding environment, with the goal of improving vehicle safety. However, as discussed earlier, it is only a prototype system that requires further development. If our system were to be used on production vehicles, it must be possible to integrate the hardware into the existing designs, in a position where the system can be most effective. This section presents several ideas on how the equipment could be mounted on different types of vehicles.



(a) (b)

Figure 3.8: (a) Warped image showing unwarping parameters. (b) Unwarped image. The stripes on cylinder appear as straight lines. Therefore the unwarping parameters have been set correctly.



(a) (b)

Figure 3.9: (a) Warped image showing unwarping parameters. (b) Unwarped image. The stripes on cylinder appear as curved lines. Therefore unwarping parameters need to be modified until the stripes are no longer curved.

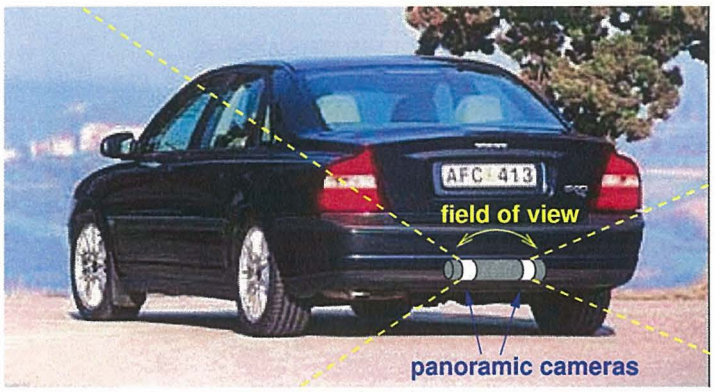


Figure 3.10: A conceptual drawing showing how the final panoramic sensor could be mounted to a production car.

Figure 3.11: The sensor could monitor underneath the rear of the vehicle, as well as the area behind.

3.4.1 Cars

The stereo panoramic sensor can be kept inside a tube, with clear sections around the mirrors to view the vehicle's surroundings. The most convenient way to mount this would be horizontally, with the tube being placed inside the car's bumper bar. See Figure 3.10.

The motor vehicle will be obscuring part of the panoramic sensor's field of view. However, by placing the sensor lower down on the car it is possible to increase the possible work space. Figure 3.11 shows that the sensor could monitor the area behind, and underneath the rear of the car.

Figure 3.12: The stereo panoramic sensor mounted at the rear of a bus.

3.4.2 Commercial Vehicles

Due to the size of a commercial vehicle, such as a bus, truck or construction vechicle, compared to that of a car, it is not necessary to constrain the cameras to a horizontal position. The tube containing the panoramic cameras could be mounted as shown in Figure 3.12. In this position, the system can view the area behind the bus without obscuring the driver's view.

It would also be possible to mount the tube on the rear corner of the bus, as seen in Figure 3.13. The system could view the rear and the side of the vehicle. This would enable the system to monitor not only behind the bus, but also the driver's blind spot. In the case of a public transport bus, the system could be used to alert the driver if a passenger has not completely disembarked. Alternatively, as shown in Figure 3.14, two panoramic sensors could be mounted on opposite corners to cover all blind areas around the vehicle.

Many accidents also occur in the area in front of trucks opposite to where the driver sits. This area is completely obscured from the driver, and the placement of a stereo panoramic sensor here would prevent damage to the vehicle, as well as injury or death. There would also be obvious benefits if the system were placed on construction vehicles, as there are many blind spots on the various types of machinery. The larger the vehicle, the larger the unmonitored areas, and the more damage can be caused due to the size of the moving parts.

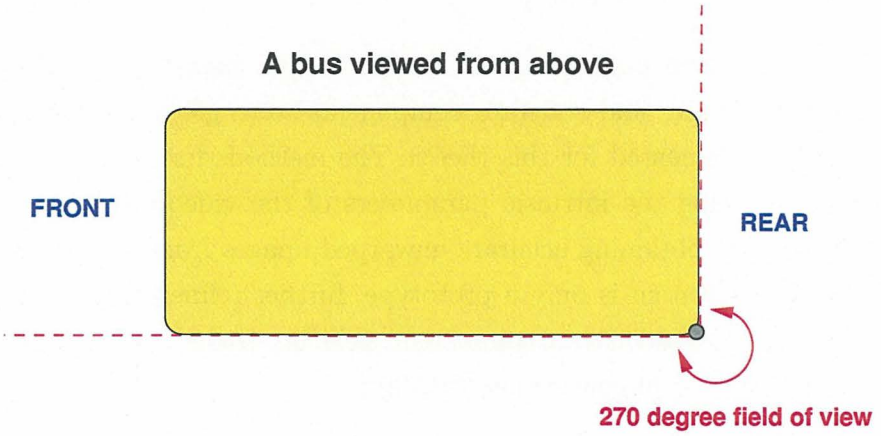


Figure 3.13: The panoramic sensor could also be placed in the rear corner of the bus. This enables the system to view the rear and the side of the vehicle.

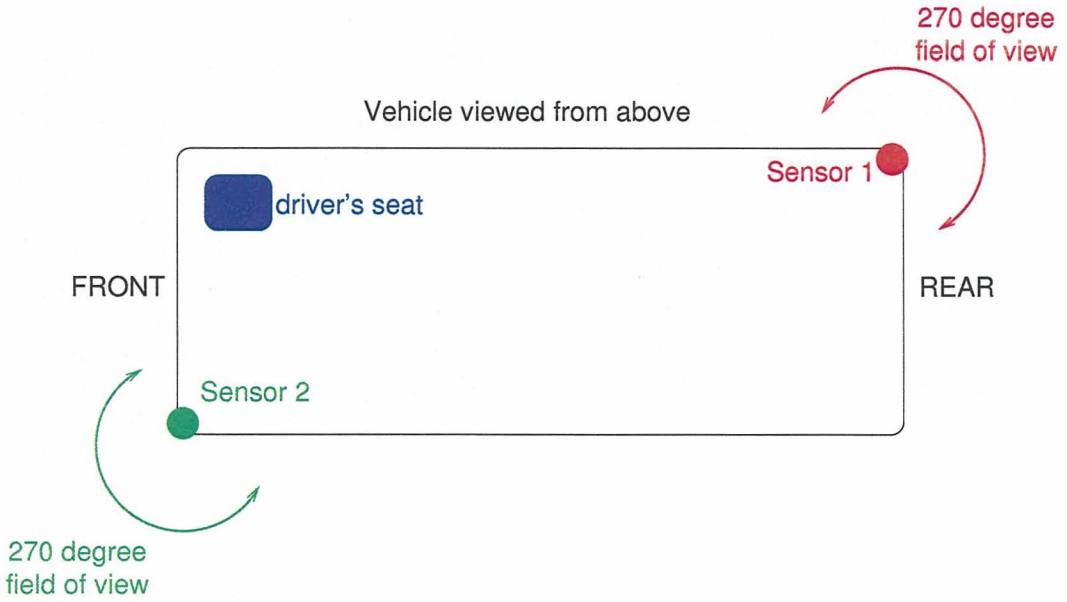


Figure 3.14: If two panoramic sensors were placed on opposite corners of the vehicle, all blind spots would be covered.

3.5 Chapter Summary

An overview of the stereo panoramic vision system was presented in this chapter. A summary of the hardware, and software components were given for the prototype system designed and implemented for this thesis. The methods for camera calibration were presented for determining the intrinsic parameters of the video cameras, as well as the parameters required for obtaining accurate unwarped images from the panoramic sensors.

Since the designed system is only a prototype, further refinements would be required before it could be integrated into a commercial vehicle. Some possible implementations were presented in the form of conceptual drawings.

Chapter 4

Range Estimation and Obstacle Detection

In Chapter 3 we presented an overview of the stereo panoramic system to be used in the automated driver assistance system to monitor vehicle blind-spots. In order to make use of this sensor, the system must be able to process the image data to extract the appropriate information. Firstly, the system needs to segment objects from the environment around the vehicle, then determine the position of each object in space. If this reveals that the vehicle is likely to collide with an obstacle, then an adequate action must be taken. The entire process consists of the following steps:

1. Estimation of range to the surrounding enviroment.
2. Detection of obstacles.
3. Finding the position of the objects in three-dimensional space.
4. Using temporal consistency to reduce the noise in obstacle detection.

In this chapter, we describe the methods for estimating range and obstacle detection. We begin with a description of the image processing necessary for finding the correspondences between the same feature in stereo images - that is, the generation of disparity maps. This is followed by an explanation of two methods for segmenting obstacles from the surrounding environment. Then the equations are presented that are required to calculate the position of a point in three-dimensional space, when using a stereo panoramic sensor. Finally, we discuss the use of a temporal consistency constraint to increase the reliability of obstacle detection, followed by a summary of the chapter in Section 4.6.

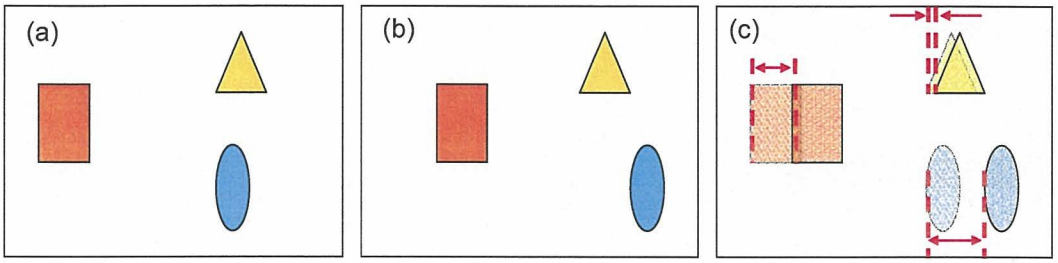


Figure 4.1: An illustration of disparity for conventional horizontally aligned stereo cameras. (a) Left image. (b) Right image. (c) Left and right images superimposed on each other. The disparity is the change in position of the same object in the images. The larger the disparity, the closer the object to the cameras and vice versa.

4.1 Disparity Maps

In Chapter 2, we discussed the influence of biological binocular systems on computer vision research. Given two or more artificial vision systems viewing the same object, it is possible to estimate the range to that object using a property known as *disparity.* When an object is imaged from a different viewpoint, it will appear in a slightly different location in each image. This discrepancy in location is the disparity, and provides a means for estimating range to the object. The larger the disparity, the closer the object is to the cameras, and vice versa. See Figure 4.1. Given the disparity, and the internal and external camera parameters, the three-dimensional position of a point can be estimated as described in Section 4.4.

The *epipolar constraint* limits the search required to locate corresponding features for determining disparity. That is, a feature in one image must lie along a unique line (an *epipolar line*) in the other stereo image (Trucco & Verri 1998). In the case of conventional stereo cameras which are usually aligned horizontallly, these epipolar lines or *epipoles* are horizontal parallel lines. This is shown in Figure 4.1 by the objects appearing at the same height in the left and right images. In the case of the panoramic stereo system used in this project, the camera and mirror axes are aligned. therefore the epipoles appear as radial lines in the warped image, as in Figure 4.2. When the images are unwarped, the epipoles are mapped to vertical parallel lines.

Therefore, to compute disparity it is necessary to find a point in one image, then search

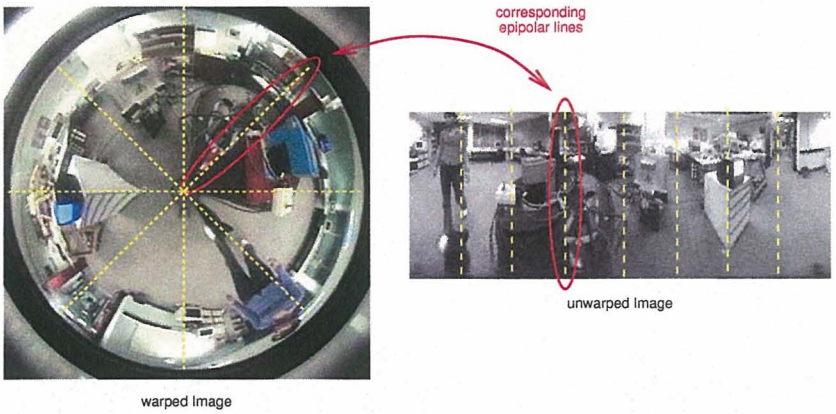


Figure 4.2: Epipolar lines are mapped from radial lines in the warped image to parallel lines in the unwarped image.

for the corresponding point in the other stereo image. A single pixel does not contain enough information to uniquely identify a feature. Most matching algorithms used for this purpose are area based and extract a window of pixels, then search for a similar window along the epipole in the corresponding image, as in Figure 4.3. There are many methods for determining the correlation between two windows. Two similarity measures that have been considered for this project are the Sum of Absolute Differences (SAD), and Normalised Cross Correlation (NCC). The simplest measure is SAD, which yields a result of zero for identical image regions:

$$SAD = \sum_{i=1}^{M} \sum_{j=1}^{N} |I1_{ij} - I2_{ij}| \tag{4.1}$$

where $I1$ and $I2$ are M by N pixel windows. However, since SAD is merely a pixel subtraction of the windows, there is no compensation for changes in lighting between the images. As the name suggests, NCC normalises the windows, and therefore helps to remove inaccurate similarity measures due to differences in average intensity. The NCC measure is +1 for identical image regions, and -1 for regions that are exactly opposite:

$$NCC = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} (I1_{ij} \cdot I2_{ij})}{\sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} (I1_{ij})^2 \cdot \sum_{i=1}^{M} \sum_{j=1}^{N} (I2_{ij})^2}} \tag{4.2}$$

As can be seen above, the NCC measure is significantly more complicated than the SAD measure. It requires several multiplications and divisions, and is therefore more com-

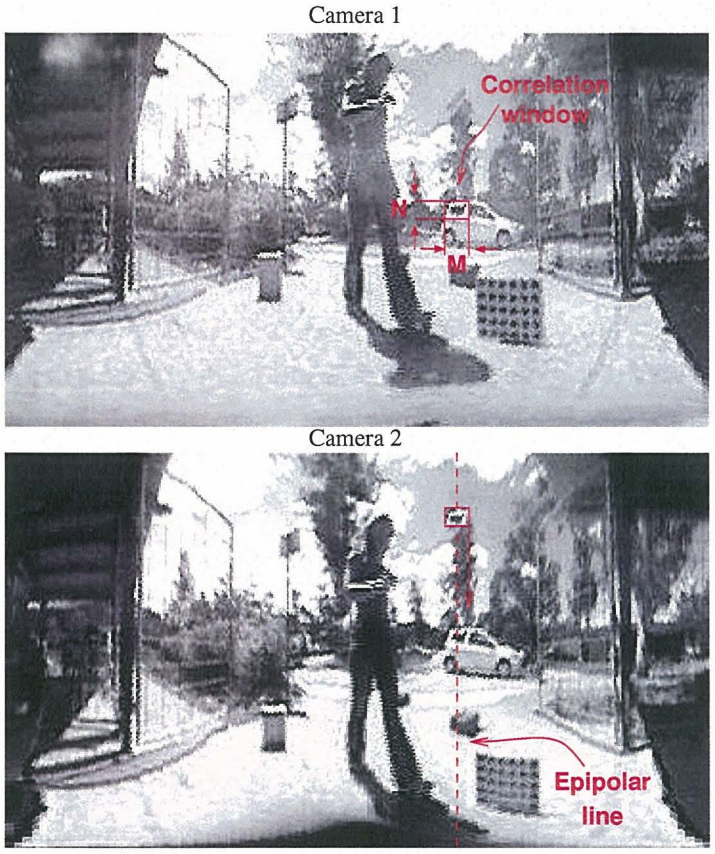


Figure 4.3: Computing disparity using an area based method.

putationally expensive. (Banks & Corke 2001) completed a qualitative and quantitative comparison of various similarity methods, including SAD and NCC. They reported similar matching performances for these two algorithms when there were no differences in lighting conditions between the stereo images. However, in outdoor environments, a stereo sensor will often provide images of different brightness due to the varying view points of the cameras. In these situations, the authors reported that the NCC algorithm was more successful in computing the pixel disparities. In the most difficult example, the SAD algorithm was only able to match 27% of the pixels correctly, while the NCC algorithm correctly matched 72% of the pixels. Therefore, it was decided that the increased computational burden of the NCC algorithm would be justified by the improved robustness of the disparity maps.

The window correlation is done for every possible pixel in the image, and this creates a *disparity map.* This map gives an indication of the depth at each pixel, except for a strip around the edge of the image where there are not enough pixels to extract a complete window. In this project we use an Intel MMX/SSE(R) multimedia instruction set implementation of the NCC disparity map algorithm presented by (Fletcher et al. 2001). This was based upon an implementation by (Kagami et al. 2000) for real-time disparity map generation, which uses a recursive correlation technique (Faugeras et al. 1993), with increased memory efficiency, consistency checking and utilises MMX/SSE(R).

There is also a trade off between searching for larger possible disparities, and computational time. For smaller disparities, the computational burden is less. However, if the disparity search is increased, the system will be able to detect objects at closer ranges. In this project a 64 pixel search was carried out around each pixel (i.e. ± 32 pixels). A search of this size adequately determined the location of an object in the desired work space, while ensuring that the execution time for the algorithm was reasonable. This characteristic will be discussed further in Chapter 5.

4.2 Obstacle Segmentation

A driver assistance system for monitoring blind-spots should only issue a warning if the vehicle is likely to hit an obstacle. In order to make this decision, the system must have information on the position of obstacles around the vehicle, and their dimensions. In the case of a vision system, this can be done by extracting such useful information from the disparity maps generated from the stereo images. Two approaches to obstacle segmentation have been investigated for this project - *ground plane subtraction*, and *v-disparity.*

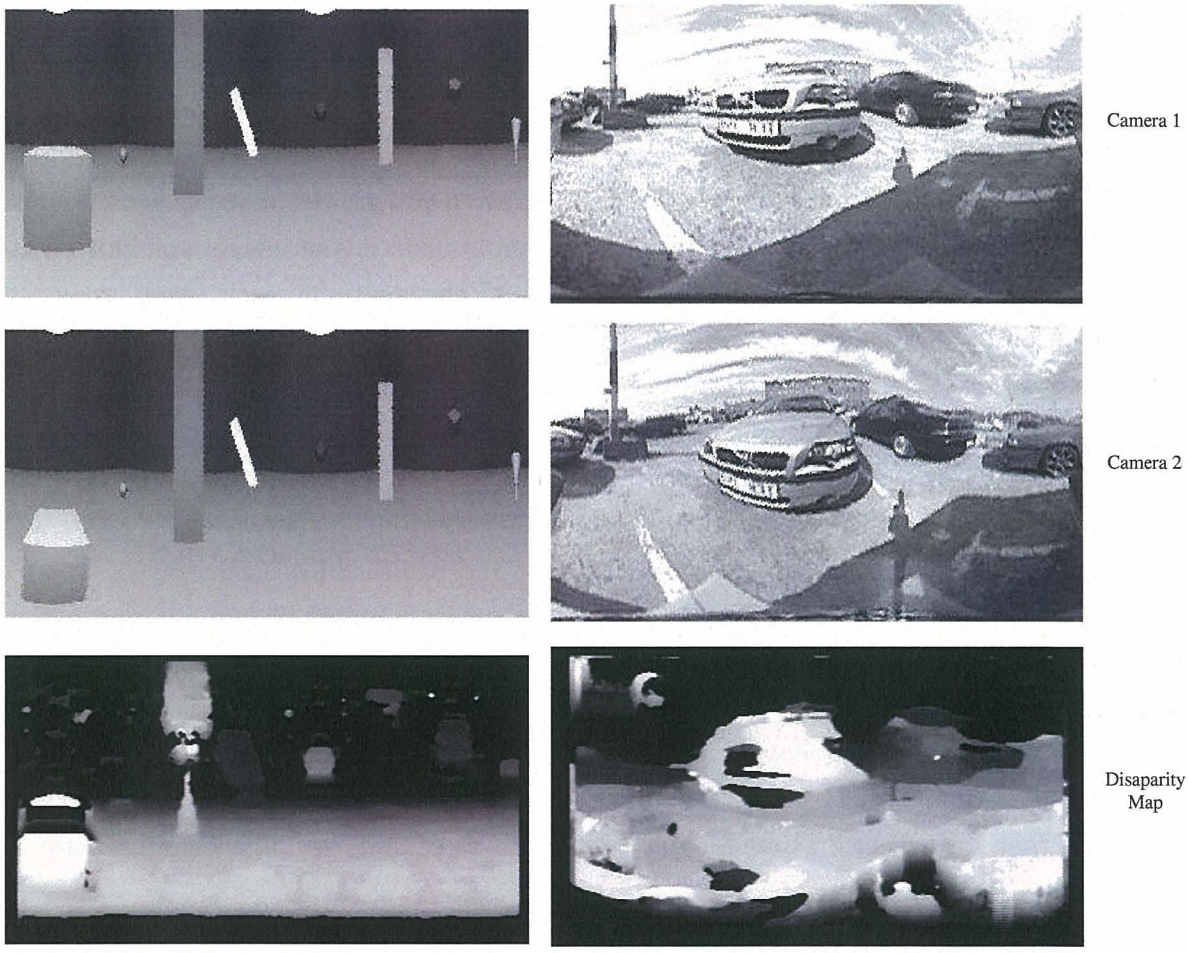


Figure 4.4: Examples of stereo image pairs with their corresponding disparity maps. (a) Ray traced images. (b) Real world images.

Algorithm 1 Computation of the theoretical disparity image of the ground plane

Assume that the position and orientation of the ground plane relative to the cameras is known. That is, in Figures 4.5 and 4.6, h is known. For each pixel in Camera 2:

1. Compute the image coordinates of the point P on the assumed ground plane, as viewed by the pixel current in Camera 2.
2. Determine the image coordinates of the pixel in Camera 1 that corresponds to point P on the assumed ground plane.
3. The disparity is computed as the difference between the pixel positions of point P as viewed in each camera.

4.2.1 Ground Plane Subtraction

The ground plane subtraction approach to obstacle detection is similar to the work presented by (Badal et al. 1994). This method builds a model of the surrounding environment without any objects present, and any deviation from the model indicates a potential obstacle. This method assumes that the ground can be modelled as a flat plane, and that the position of the ground plane relative to the cameras is known. A theoretical disparity map is calculated using this assumption, and this is subtracted from the disparity map generated from the stereo images captured from the panoramic sensors. The result is a disparity map containing only non-ground pixels.

The theoretical disparity maps were calculated off-line for both vertical and horizontal configurations of the panoramic sensor as shown in Figures 4.5 and 4.6 respectively. Assuming that the position of the ground plane is known relative to the cameras, the theoretical disparity maps can be computed using Algorithm 1.

A pixel in the image captured by Camera 1 is mapped back to point P on the ground plane. The pixel in Camera 2 that also maps to point P is then determined, and the difference between these two pixels is the disparity. The entire theoretical disparity map is calculated by repeating this procedure for all pixels in the image captured by Camera 1. Example disparity maps calculated for the raw panoramic images can be seen in Figures 4.7(a) and (b).

Once the theoretical disparity maps have been calculated, they can be unwarped, and subtracted from the real-world disparity maps generated from the stereo panoramic images. The results of ground plane subtraction can be seen in Figure 4.8.

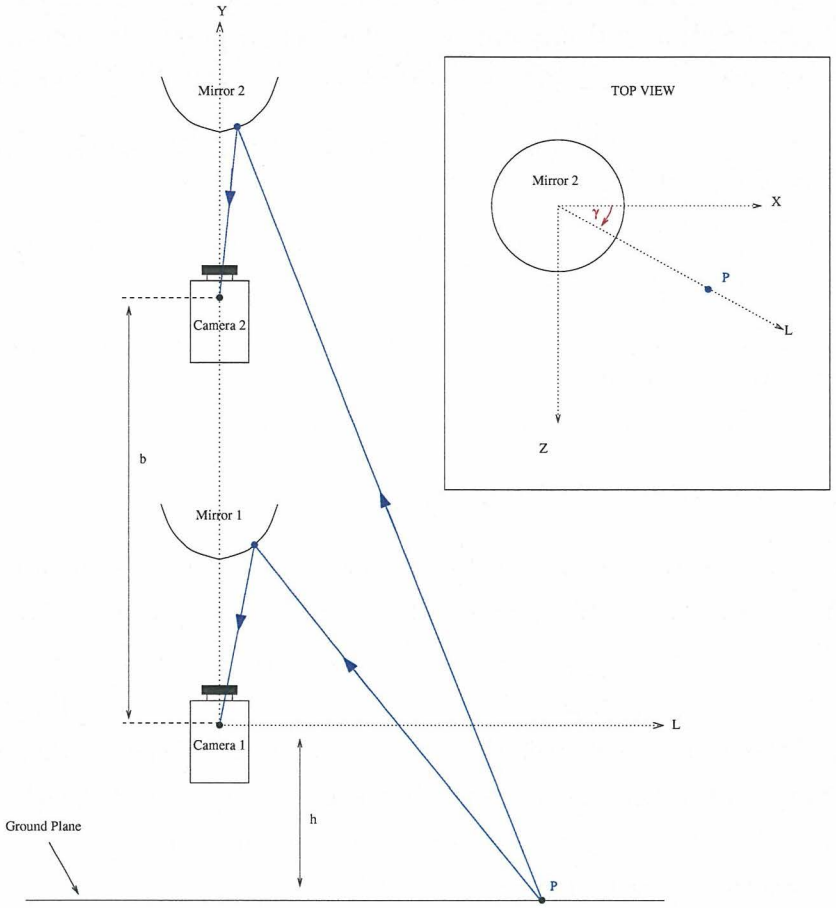


Figure 4.5: Ground plane subtraction parameters for a vertical sensor configuration.

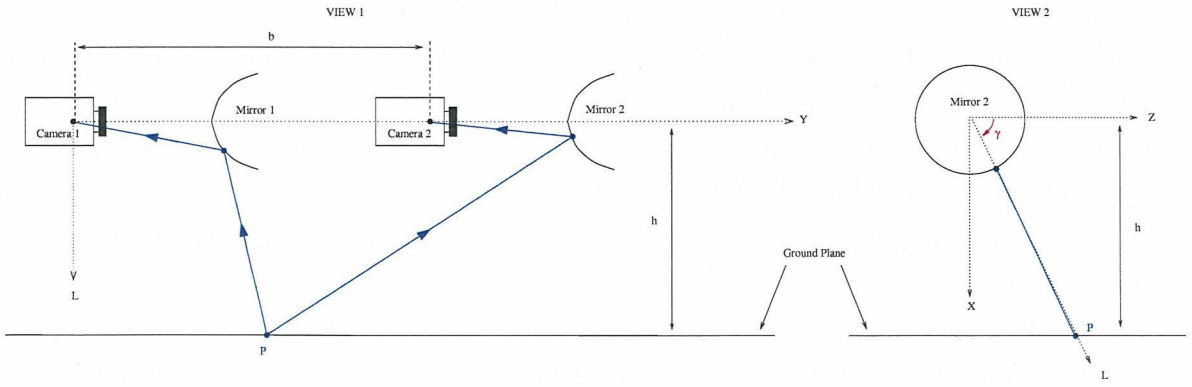


Figure 4.6: Ground plane subtraction parameters for a horizontal sensor configuration.

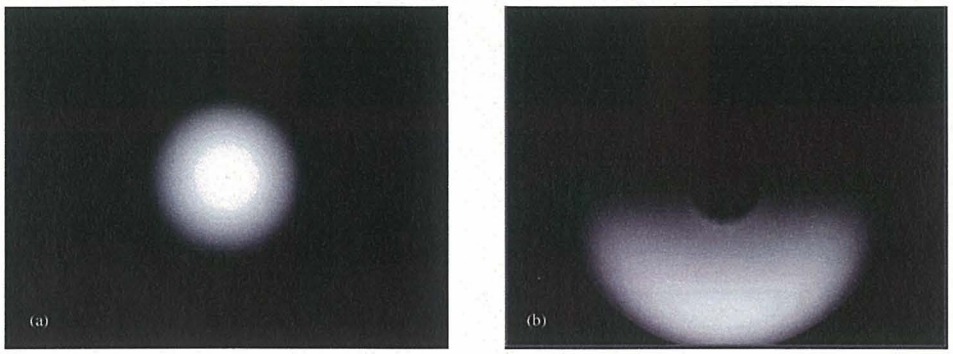


Figure 4.7: Theoretical disparity maps containing only the ground plane for (a) vertical and (b) horizontal stereo panoramic sensors. These were computed using Algorithm 1.

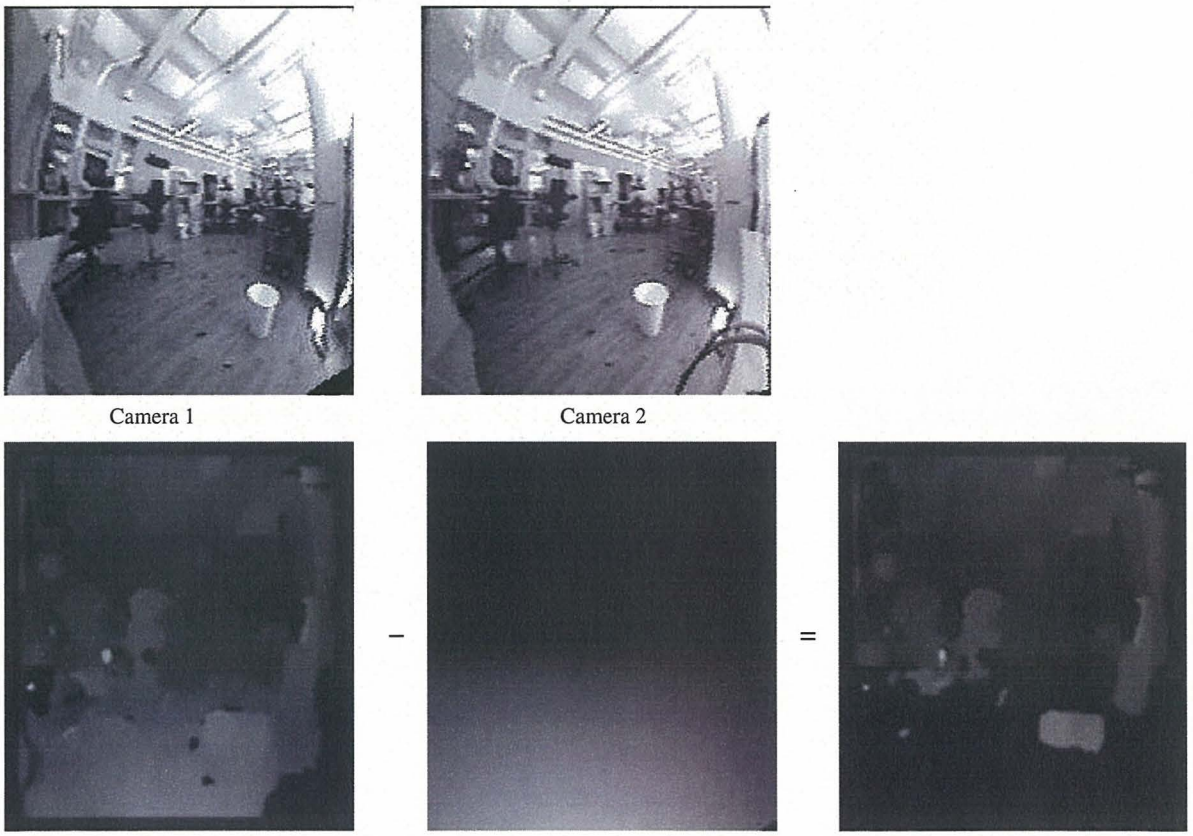


Figure 4.8: The results from the ground plane subtraction algorithm.

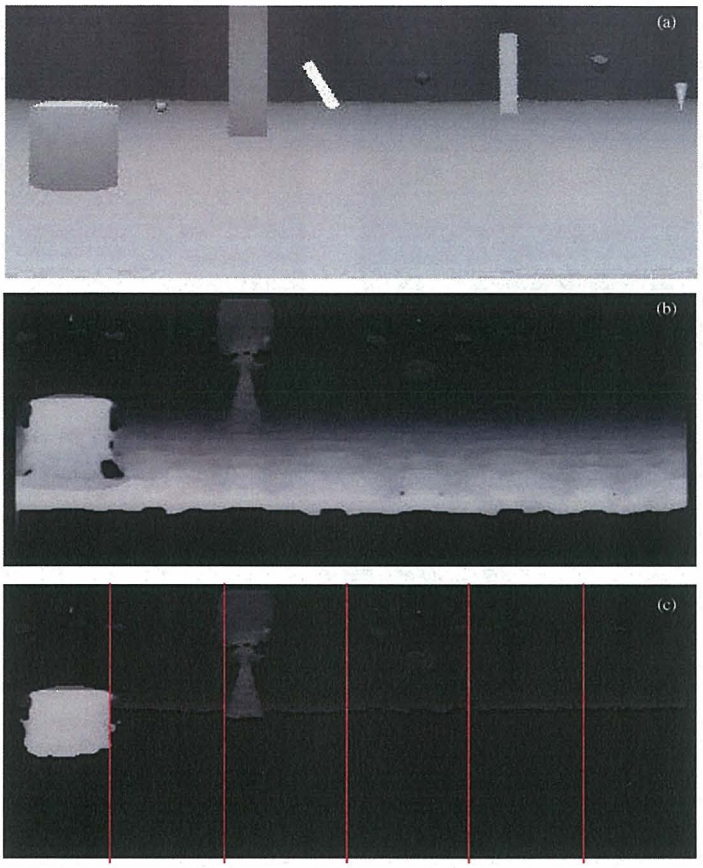


Figure 4.9: (a) The unwarped image from one of the panoramic sensors. (b) The disparity map. (c) The disparity map with the ground plane subtracted. To create a range scan of the environment, the disparity map is separated into sections. The range to the closest object in each section of the environment is reported, as shown in Figure 4.10.

Obstacle Segmentation

Once the ground plane has been subtracted, this leaves the obstacles and background in the resulting disparity map. One method of determining the position of obstacles is to complete a range scan of the environment. This is done by first subtracting the ground plane from the disparity map, as shown in Figure 4.9. The disparity map is then split into sections, effectively dividing up the environment into separate regions. The distance to the closest object in each section is calculated, and displayed in a top view of the area around the sensor, as shown in Figure 4.10.

The disadvantage of this method is that any noise in the disparity map would cause

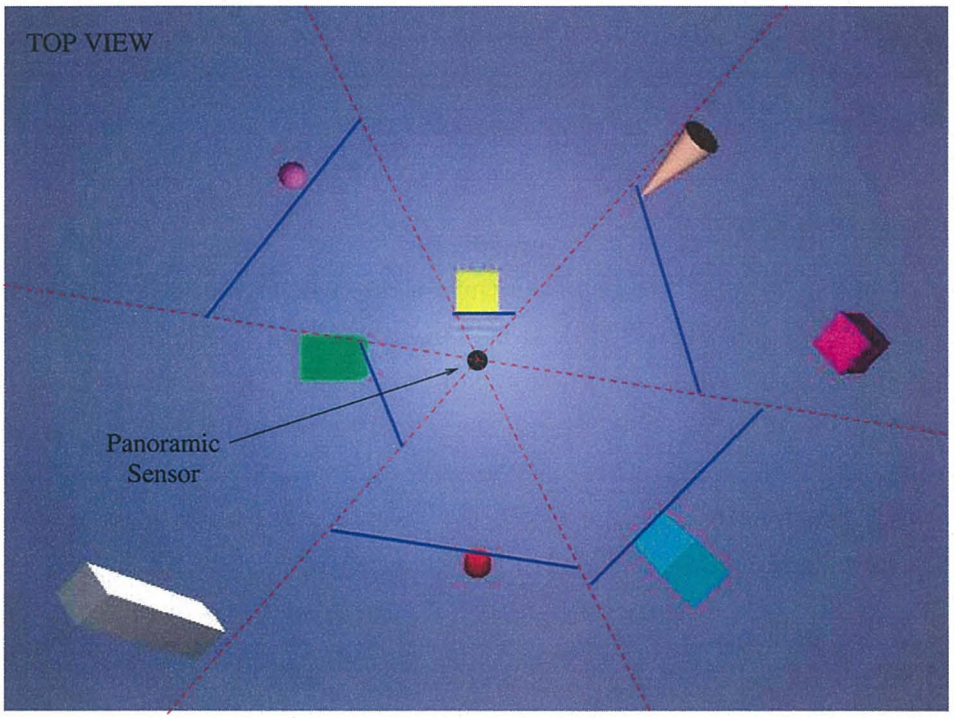


Figure 4.10: An example of a range scan of the environment (seen from above), after ground plane subraction has been applied. The red lines show the sections from Figure 4.9. The blue lines show the depth that would be reported for each section of the surrounding environment.

a false reading for an entire region. This could be over come by instead employing a connected component analysis (Ballard & Brown 1982). In this manner, individual objects could be segmented from the disparity map by detecting a region of connected pixels at approximately the same depth.

However, the connected component analysis would still not overcome the lack of robustness that is introduced by the ground plane assumption. That is, this method requires *a priori* knowledge of the orientation of the ground plane with respect to the sensor. In our application, the stereo panoramic sensor is mounted on a moving vehicle, and therefore there will be cases where such an assumption will fail. This issue is discussed further in Section 6.2.1.

4.2.2 V-Disparity

The second method of obstacle detection investigated utilises an algorithm known as v-disparity (Labayrade & Aubert 2003). It is more robust than the ground plane subtraction method, as it requires no *a priori* knowledge of the exact orientation of the ground plane with respect to the cameras. The ground plane can be reliably segmented from each pair of stereo images, which makes this method well suited to our application. In particular, this algorithm can reliably segment obstacles, even in the case of extremely noisy input data. The v-disparity algorithm requires the following steps:

1. Generation of a disparity map.
2. Mapping the pixels of the disparity map to u- and v-disparity images, as in Figure 4.11.
3. Segmentation of the obstacles from the u- and v-disparity images.

The v-disparity image is created by accumulating all pixels from the disparity map according to their v ordinate from the (u,v) coordinate system, and their intensity (which is proportional to disparity). See Figure 4.11(d). Similarly the u-disparity image is generated by placing the disparity map pixels into bins according to their u ordinate and intensity, shown in Figure 4.11(e).

The ground plane appears as a downward sloping line in the v-disparity image, while the background is represented by the left-most vertical line (Figure 4.12). The back ground can also be seen in the u-disparity image as the highest horizontal line. These lines can be detected by the use of a Hough transform, as suggested by (Labayrade & Aubert 2003).

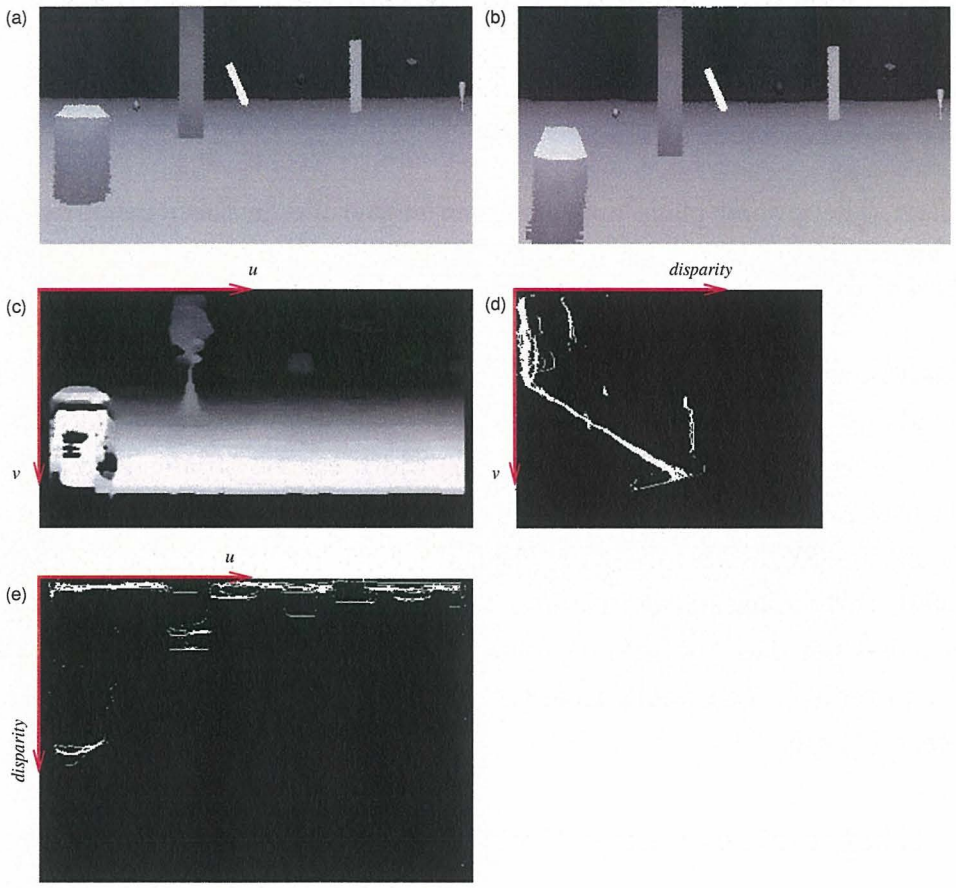


Figure 4.11: (a) Unwarped image from camera 1. (b) Unwarped image from camera 2. (c) The disparity map produced from (a) and (b). (d) The v-disparity image. (e) The u-disparity image.

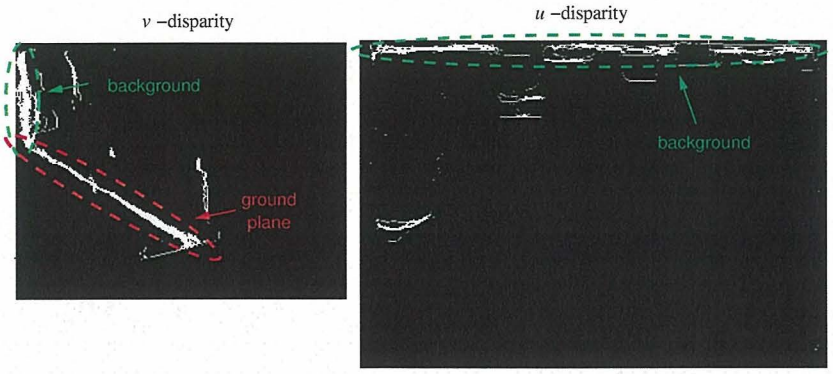


Figure 4.12: The ground plane appears as an angled line in the v-disparity image. The background is the left-most vertical line in the v-disparity image, and it corresponds to the highest horizontal line in the u-disparity image.

Obstacles are any near vertical lines above the ground plane, and to the right of the background in the v-disparity image. The height of an obstacle is determined by the length of the line in the v-disparity image, while the width can be obtained from the length of the horizontal line in the u-disparity image. See Figure 4.13.

It should be noted that since the v-disparity algorithm was first published (Labayrade et al. 2002), this representation has been tested on conventional stereo camera systems. However it has not been utilised or evaluated for a stereo panoramic vision system until this research project. The results from the application of this algorithm will be presented in Chapters 5 and 6.

4.2.3 Obstacle Detection Summary

This section has introduced two methods for segmenting obstacles using a disparity map generated from a pair of stereo images. Ground plane subtraction and the v-disparity representation have both been evaluated on the stereo panoramic system, and the results of this will be presented in Chapters 5 and 6. The following section will discuss the process of determining the three-dimensional position of a point of interest (i.e. an obstacle) and the disparity of that point.

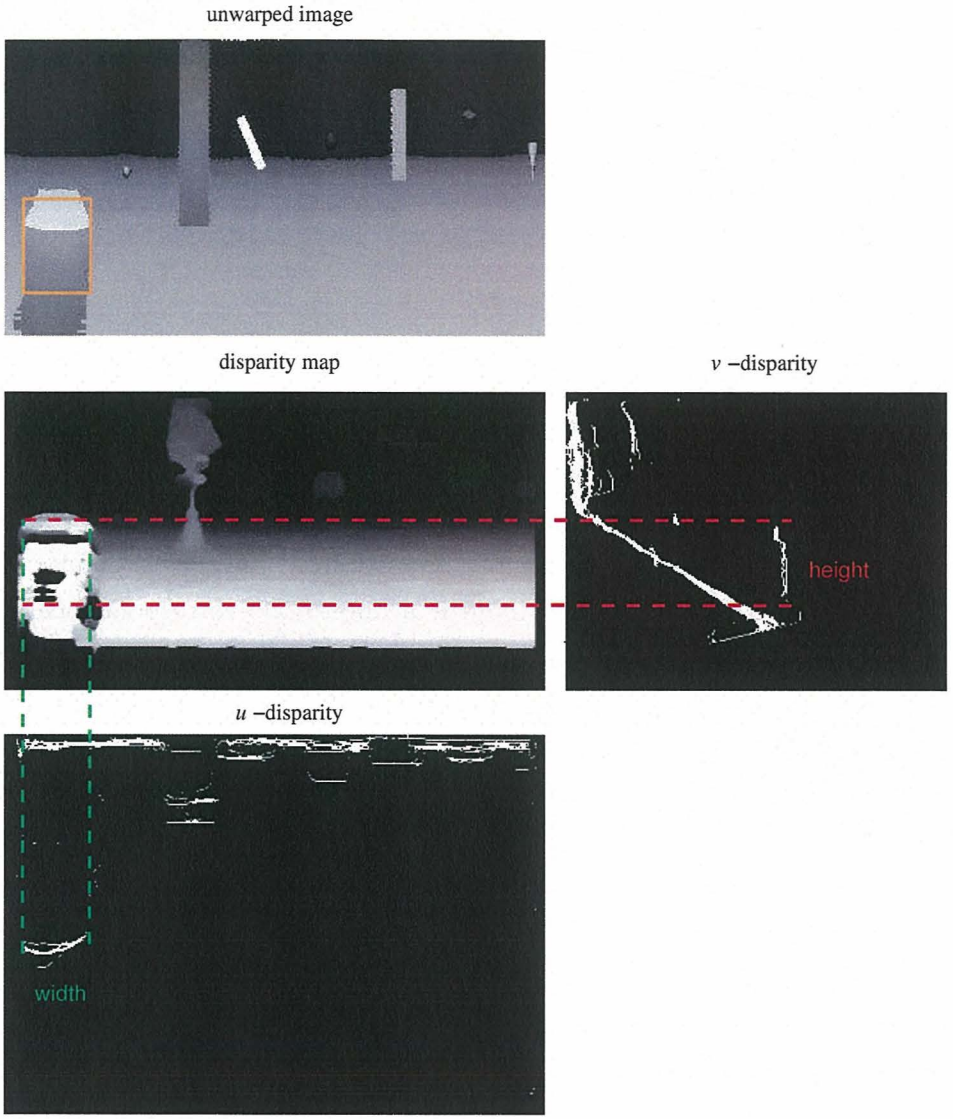


Figure 4.13: The v-disparity image provides the height of an object, while the u-disparity image reveals the width of the object.

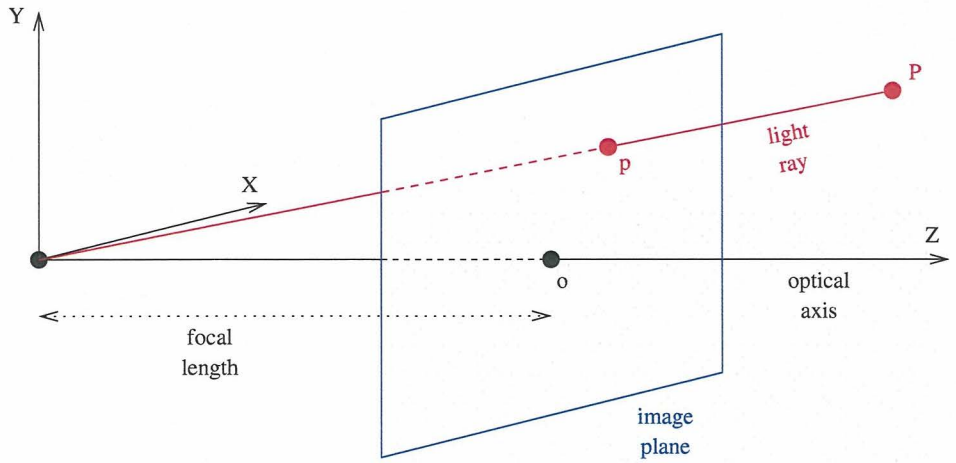


Figure 4.14: The pinhole camera model.

4.3 Pinhole Camera Model

In the pinhole camera model, only one ray of light from any given point can enter the camera (Trucco & Verri 1998). Thus there is a one-to-one correspondence between the points in an image $\mathbf{p}$, the light rays, and the visible points $\mathbf{P}$, as shown in Figure 4.14. Let $\mathbf{p}=\begin{bmatrix} x & y \end{bmatrix}^T$ and $\mathbf{P}=\begin{bmatrix} X & Y & Z \end{bmatrix}^T$. The fundamental equations of the pinhole camera model are as follows:

$$\begin{aligned} x &= f\,\frac{X}{Z} \\ y &= f\,\frac{Y}{Z} \end{aligned} \tag{4.3}$$

Where f is the focal length.

Using the assumption of the pinhole camera model, and knowledge of the mirror profiles and assuming the mirror-camera axes are aligned, it is possible to estimate range to points in the surrounding environment.

4.4 Range Calculations

The mirror profiles chosen for this project do not conform to the single viewpoint constraint (this was defined in Section 2.4.2). This means that the rays of light that reflect off the mirror into the camera, do not converge to a single (virtual) focal point inside the mirror, as shown in Figure 2.12. As a consequence, if we were to apply the range finding methods

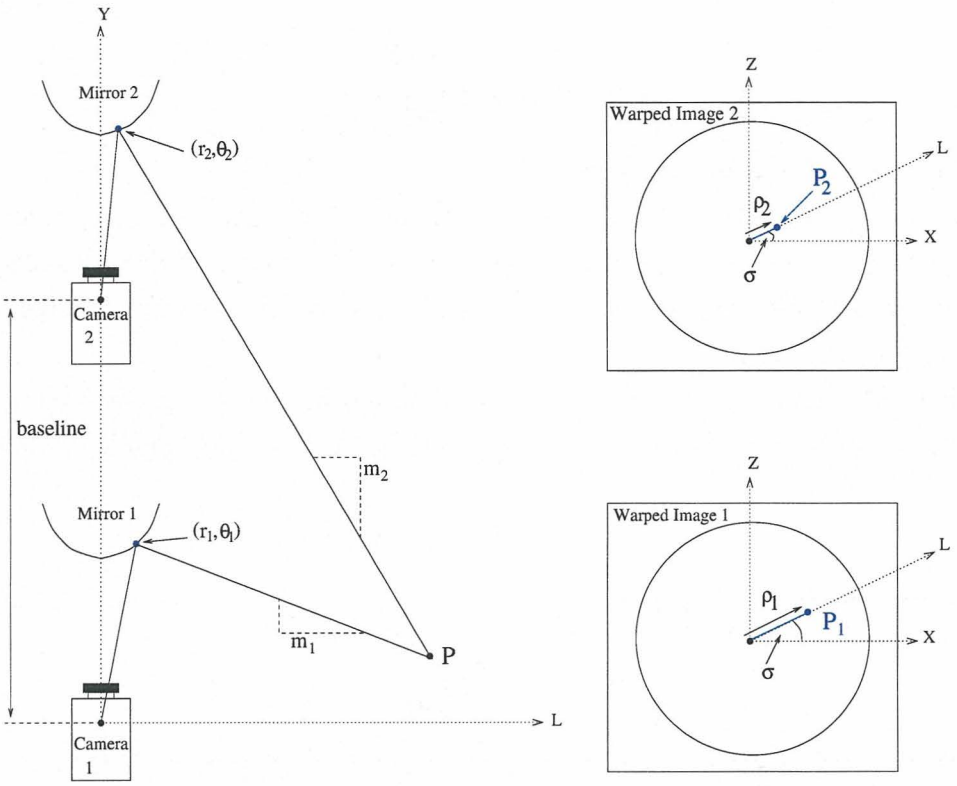


Figure 4.15: Depiction of range estimation parameters. A point P in space will be projected onto P_1 and P_2 in the warped panoramic images.

used in conventional stereo, the results not be accurate. However, depth can be computed by using triangulation methods, similar to those presented by (Conroy 2000).

That is, by finding the intersection of the rays of light from the mirror surfaces to the point P in three-dimensional space, as in Figure 4.15. Since the stereo configuration will not change during run time, it is useful to pre-compute a lookup table of all possible image correspondences off line to increase the speed of execution (see Appendix A for comments).

Given the correspondence between two image points, as calculated for the disparity map, the position of the point can be determined in three-dimensional space. In the warped images, these points are P_1 and P_2, and they are a distance of ρ_1 and ρ_2 from the image centre as shown in Figure 4.15. The direction of the radial line line which is situated along the L-axis, is represented by the angle σ. The distance ρ_1 can be used to calculate the radial angle of the reflected ray of light into Camera 1 as follows:

$$\theta_1 = \arctan\left(\frac{\rho_1}{f_1}\right) \tag{4.4}$$

where f_1 is the focal length of Camera 1.

Using this information, we can then determine the distance from the camera focii to the the reflection points on the mirror, r_1 and r_2. In the case of a constant gain mirror these can be calculated using Equation 2.5, and similarly for a resolution invariant surface by solving Equation 2.6.

To find the intersection point P, we need to determine the line gradients m_1 and m_2 of the incident ray of light. For a constant gain profile, this can be calculated as follows:

$$m_1 = -\cot\left(\alpha\left(\underline{\theta} - \theta_1\right) + \underline{\phi}\right) \tag{4.5}$$

For a resolution invariant mirror:

$$m_1 = -\cot\left(\frac{B_\alpha}{2}\tan^2\theta_1 + \phi_{\theta=0}\right) \tag{4.6}$$

Given that the gradients m_1 and m_2, and the reflection points (r_1,θ_1) and (r_2,θ_2) are known, the linear equations for the incident beams of light can be defined. The constant for the equation for Mirror 1 is then

$$C_1 = r_1\cos\theta_1 - m_1\, r_1 \sin\theta_1 \tag{4.7}$$

As mentioned above, both light rays pass though the point P, so this point can be found by solving the two equations simultaneously. In this manner, we can determine the distance along the camera axis, y_P and the radial distance from the camera axis l_P.

$$\begin{bmatrix} y_p \\ l_p \end{bmatrix} = \begin{bmatrix} 1 & -m_1 \\ 1 & -m_2 \end{bmatrix}^{-1} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{4.8}$$

Point P can be described in Cartesian coordinates as (x_P, y_P, z_P) where the focus of Camera 1 is the system origin. From Equation 4.8, y_P is known, and the remaining coordinates can be determined using the image angle σ :

$$\begin{aligned} x_P &= l_P\cos\sigma \\ z_P &= l_P\sin\sigma \end{aligned} \tag{4.9}$$

The point P has now been located in three-dimensional space, in relation to the stereo

panoramic system, where l_P is the *range* to the object, while (x_P, y_P, z_P) is its position.

4.5 Temporal Consistency

In general, the v-disparity approach to obstacle detection is reliable even in the case of extremely noisy input data. However, the system occasionally detects obstacles where there are none, if the depth maps are inaccurate. One method of increasing the robustness of the algorithm is to introduce temporal consistency. An obstacle is not considered to be detected until it meets the following criteria:

- The obstacle appears in at least two consecutive frames.
- It has no significant changes in dimensions.
- The obstacle must appear within a specified search window in the second frame to be considered as the same object.

In Chapter 6, we will present results using temporal filtering which show a significant decrease in the rate of false detections.

4.6 Chapter Summary

In this chapter we have presented the methods for extracting useful information from the data provided by the stereo panoramic vision system. Given a pair of images from the sensor, depth can be estimated through the generation of a disparity map. These can been calculated using area-based correlation methods such as Sum of Absolute Differences (SAD), and Normalised Cross Correlation (NCC). (Banks & Corke 2001) published results showing that the NCC algorithm was more robust to differences in lighting between the stereo image pair. Since we will be using our sensor in an outdoor environment, it is likely that there will be brightness variations between the stereo images. Therefore, we decided to use the NCC matching algorithm over the SAD method to improve the robustness of the disparity maps.

Once the disparity map has been calculated, obstacles can be segmented from the environment. In this project, two obstacle detection algorithms have been evaluated to find the most suitable for our application - *ground plane subtraction* and *v-disparity*. The first algorithm relies on *a priori* knowledge of the ground plane with respect to the sensor.

Whereas the *v-disparity* representation enables the detection of the ground plane in each frame, and therefore a more accurate segmentation of obstacles. This algorithm seems robust to significantly more noise. A thorough analysis and comparison of these two obstacle detection methods is presented in Chapter 5

Finally, the calculation for range and bearing estimation were presented. Given the position of objects around the vehicle in three-dimensional space and their dimensions, a decision can be made about when to warn a driver of an impending decision. We have presented a theoretical framework for the implementation of a driver assistance system for monitoring blind-spots. The following two chapters present theoretical and experimental results from the evaluation of such a system.

Chapter 5

Theoretical Analysis and Simulation

In this research project, we proposed a new approach to the problem of obstacle detection for monitoring vehicle blind-spots. The prototype system consists of two panoramic video cameras, which include curved mirror surfaces to provide a wide field of view. In Chapter 4, the algorithm used for range finding was presented, as well as two methods for detecting obstacles known as *ground plane subtraction* and *v-disparity.* However, the performance of these algorithms when applied to stereo panoramic vision has not previously been tested.

This chapter discusses the results from the theoretical analysis and simulation experiments that were carried out to evaluate the prototype system. A ground truth analysis was performed by creating artificial panoramic images using a ray-tracing software package (the Persistence of Vision Ray tracer, POV-Ray). Objects were placed in a virtual environment, and the software was used to capture panoramic images from a known location. A theoretical analysis was completed to determine the resolution of the stereo panoramic sensor for range estimation, and the advantages and disadvanges of the various sensor characteristics are presented. The range estimation performance was tested using the ray-traced image, first without the use of the obstacle detection algorithms to determine the theoretical capability of the system. Then finally the obstacle segmentation was introduced, and evaluated. These experiments were carried out using different mirror profiles, baselines, and sensor configurations, and the results of the comparison of sensor characteristics have been summarised. This has guided us in the selection of suitable parameters and algorithms to achieve the goals of the project.

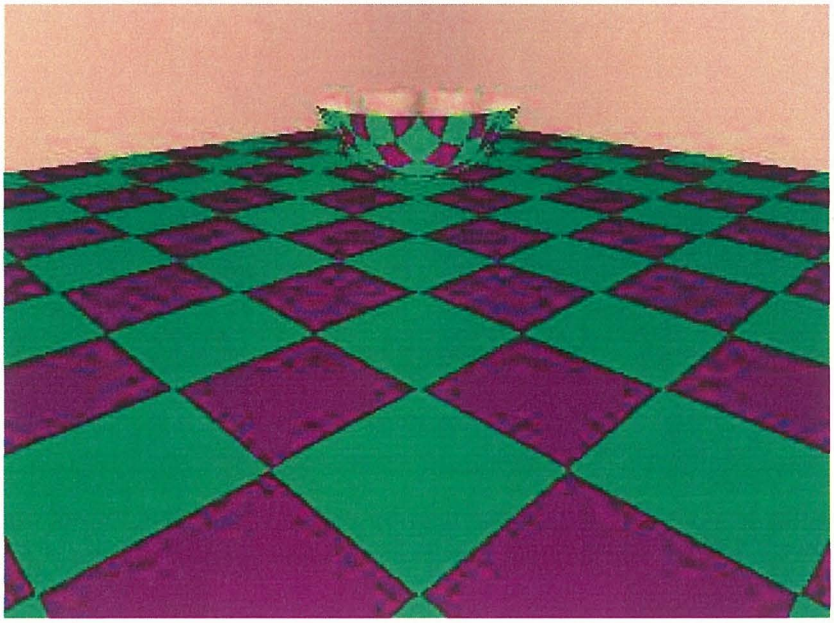


Figure 5.1: An example of a ray traced image, showing a room with a panoramic mirror in the centre.

5.1 Ray Traced Images

The Persistence of Vision Ray tracer (POV-Ray) is a software tool for producing high-quality computer graphics. It is possible to download the package over the internet, and use it free of charge, [1] subject to conditions in the authors' license. Ray-tracing is a method of calculating an image of a scene by simulating the way rays of light travel. In the real world, light rays are emitted from a light source and bounce off objects until finally reaching the camera. The POV-Ray program works backwards, by tracing the rays of light from the camera back to the objects in the scene to determine the colour of each image pixel. An example of a ray traced image is shown in Figure 5.1.

The POV-Ray software was used to create artificial panoramic images, to enable a ground truth analysis to be carried out. Various objects were placed in a virtual environment at known locations, including a panoramic mirror, and virtual panoramic images of the scene were captured. A top view of the scene used in the ground truth experiments is shown in Figure 5.2. The panoramic mirrors were placed in the centre of the environment, and images of the surroundings reflected off the mirrors were ray traced. Two

[1] www.povray.org/download/

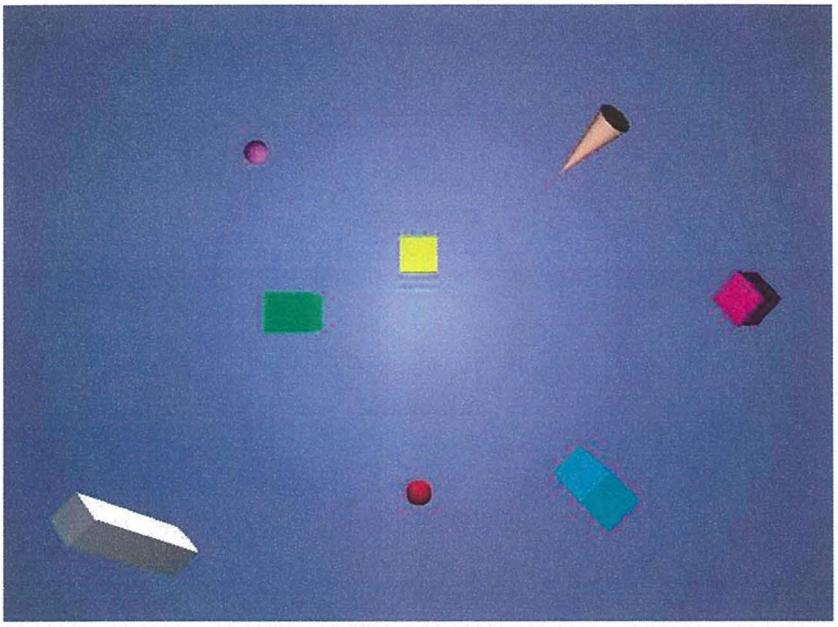


Figure 5.2: Top view of the virtual environment used in the ground truth analysis.

different mirror profiles were used (constant gain and resolution invariant), and two sensor configurations were used (horizontal and vertical.)

5.1.1 Constant Gain Mirrors

The images generated by the POV-Ray software are shown in Figures 5.3 (a) and (b), with the stereo sensor was placed in the environment in a horizontal configuration. In this configuration, the stereo sensor also captures several areas that do not need to be monitored. These include the ceiling of the virtual room, and areas that would be obscured by the vehicle the device will be mounted to. An example of a useful section of the image has been unwarped. These images are displayed in Figures 5.3 (c) and (d). The resulting depth map is shown in Figure 5.3(e).

Figures 5.4(a) and (b) are ray traced images of a stereo sensor in a vertical configuration. The corresponding unwarped images are displayed in Figures 5.4 (c) and (d), while the disparity map is shown in (e). In this configuration, over 50% more of the image is utilised for monitoring areas of interest. As a result, the panoramic images produced are of a higher resolution.

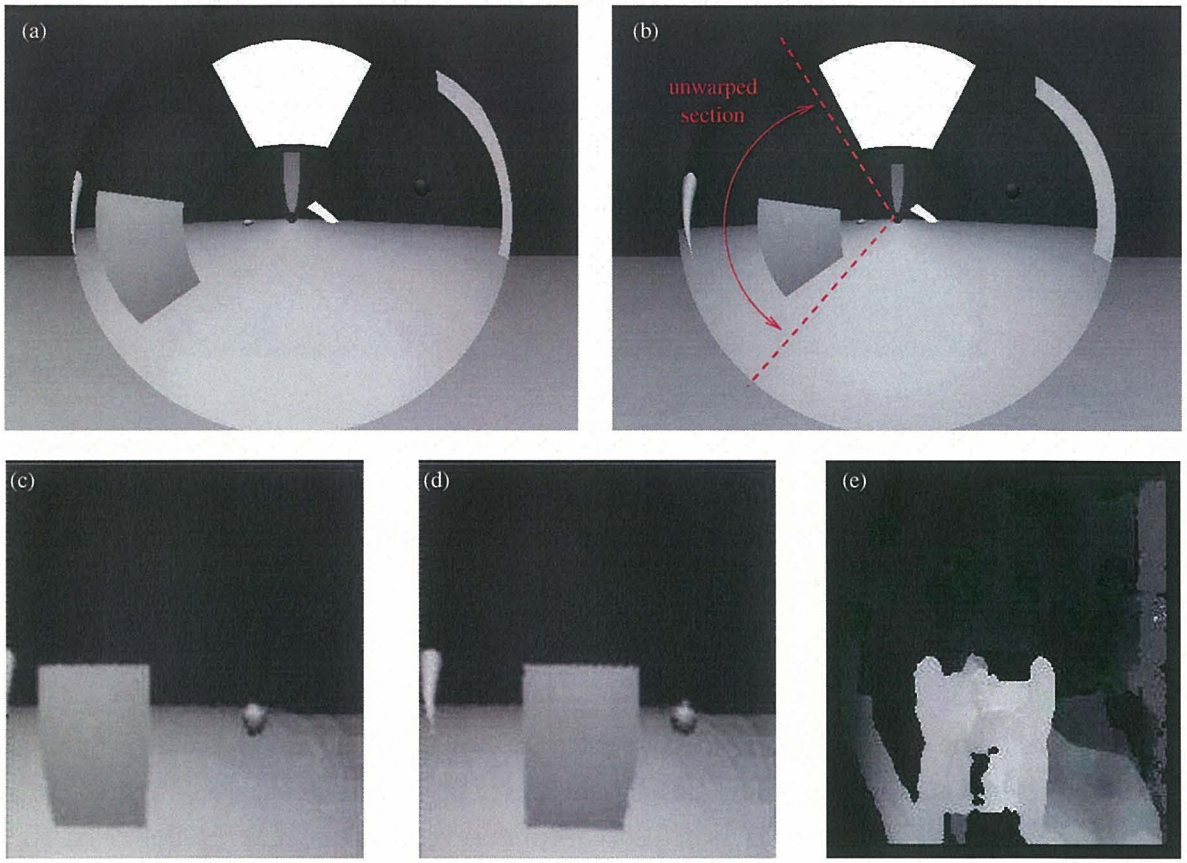


Figure 5.3: Images captured by a stereo sensor using constant gain mirrors, placed in a horizontal configuration. (a) The raw image from camera 1. (b) The raw image from camera 2, with the unwarped section depicted in red. (c) The unwarped image generated from (a). (d) The unwarped image generated from (b). (e) The disparity map calculated from (c) and (d).

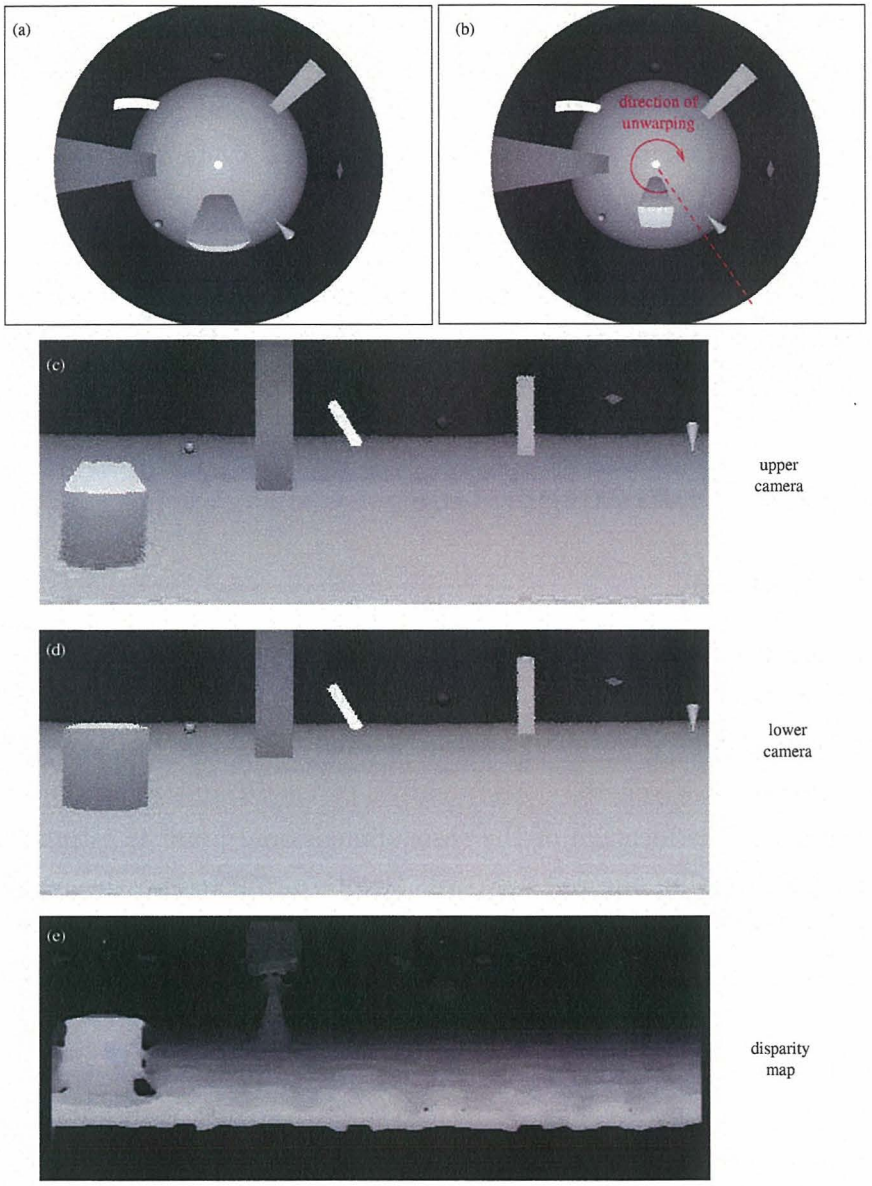


Figure 5.4: Images captured by a stereo sensor using constant gain mirrors, placed in a vertical configuration. (a) The raw image from the upper camera. (b) The raw image from the lower camera. (c) The unwarped image generated from (a). (d) The unwarped image generated from (b). (e) The disparity map calculated from (c) and (d).

5.1.2 Resolution Invariant Mirrors

Figures 5.5 and 5.6 show the same virtual environment displayed in Figure 5.2. In this case, the images have been captured using resolution invariant mirrors in a horizontal and vertical arrangement, respectively. In these figures, (a) and (b) are the raw ray-traced images captured to simulate a stereo panoramic vision system. The unwarped images calculated from (a) and (b) are displayed in (c) and (d). The unwarped images were then used to calcluate a disparity map, as shown in (e).

In this section we presented the ground truth data generated by the POV-Ray software. The following sections discuss the various sensor characteristics, and present the methods and results of the ground truth experiments carried out using the ray traced images.

5.2 Sensor Characteristics

5.2.1 Resolution

An image is captured by a video camera using a uniform Cartesian array of pixels. The discretisation of the image space results in discontinuities in the ranges estimated by a stereo vision system. The position of a point can only be estimated to a certain degree of accuracy. For example, given the corresponding pixels P_1 and P_2 in the warped images, shown in Figure 5.7, the location of the three-dimensional point is estimated to be at P. However, any point within the red area will also be projected to the pixels P_1 and P_2 . The discrepancy between the point P and the actual location of an object in space is the discretisation error. This error is characterised by the depth resolution of the stereo vision system.

The depth resolution for a stereo panoramic vision system is more complicated than for conventional stereo imaging. Use of the curved mirror surfaces results in a non-uniform spatial resolution, and this varies according to the mirror profile. The resolution was calculated in two-dimensions using the equations from Section 4.4. The graphs were generated by computing the depth for every possible pair of image pixel correspondences - that is, for every combination of P_1 and P_2 - up to a maximum disparity of 64. The resulting plots of the two-dimensional depth resolutions for systems using constant gain and resolution invariant mirrors are shown in Figures 5.8 and 5.9 respectively. Only one half of the graph is shown, since the depth resolution is symmetrical about the camera-mirror axis.

As discussed in Section 4.1, the maximum disparity search needs to be limited in order

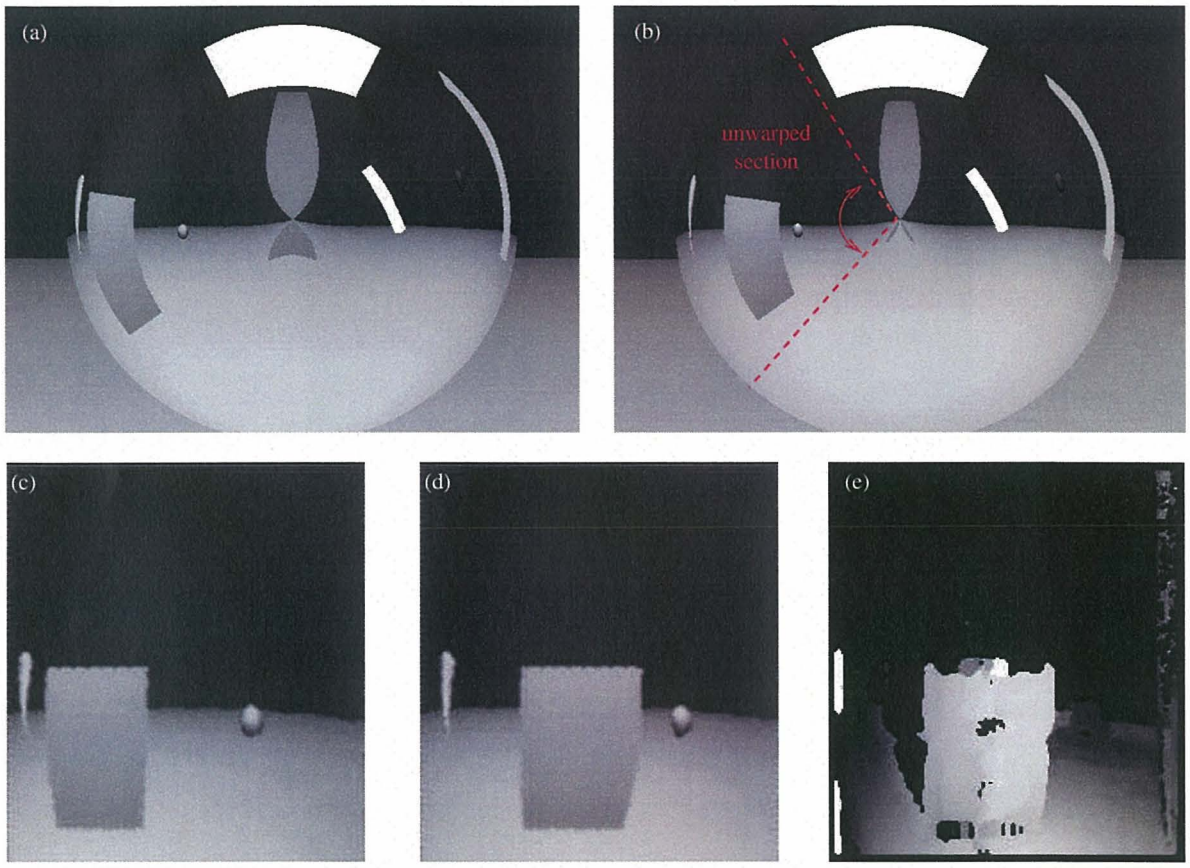


Figure 5.5: Images captured by a stereo sensor using resolution invariant mirrors, placed in a horizontal configuration. (a)The raw image from camera 1. (b) The raw image from camera 2, with the unwarped section depicted in red. (c) The unwarped image generated from (a). (d) The unwarped image generated from (b). (e) The disparity map calculated from (c) and (d).

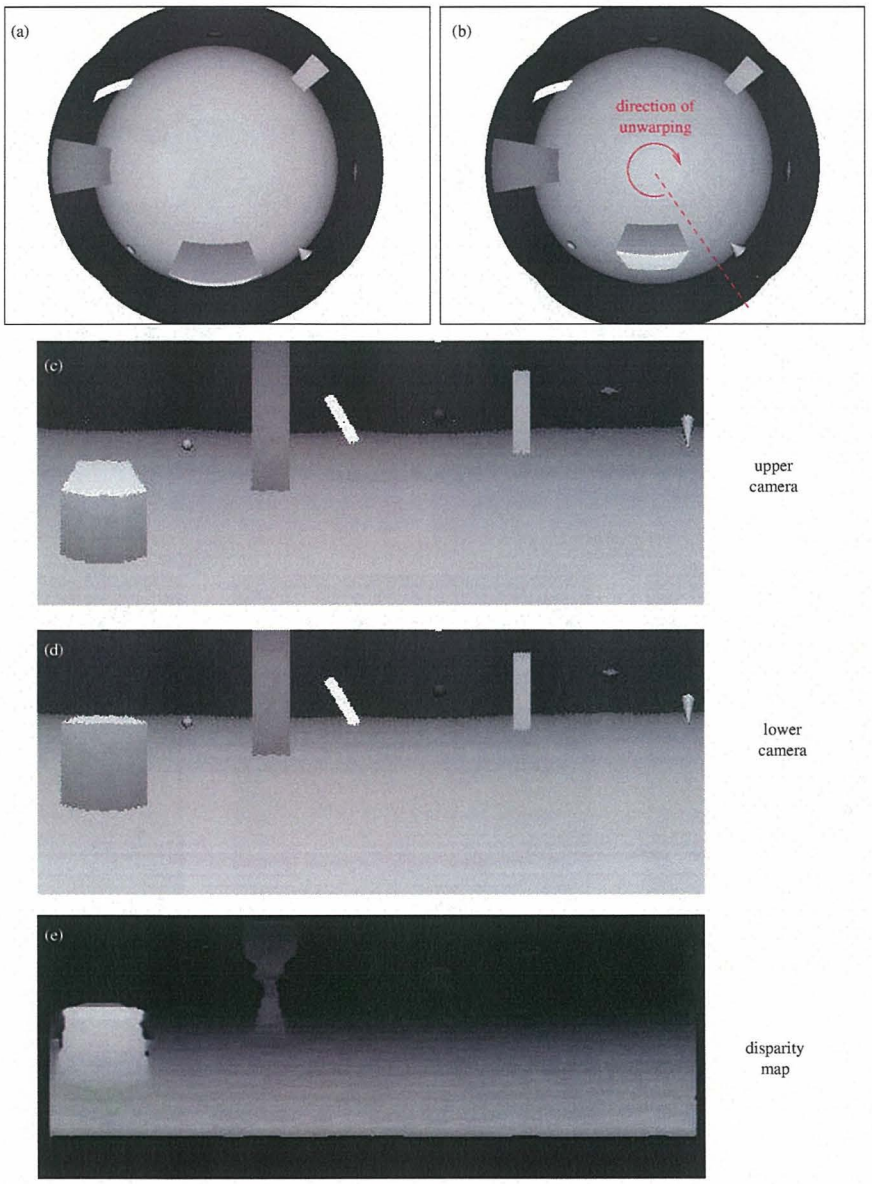


Figure 5.6: Images captured by a stereo sensor using resolution invariant mirrors, in a vertical configuration. (a) The raw image from the upper camera. (b) The raw image from the lower camera. (c) The unwarped image generated from (a). (d) The unwarped image generated from (b). (e) The disparity map calculated from (c) and (d).

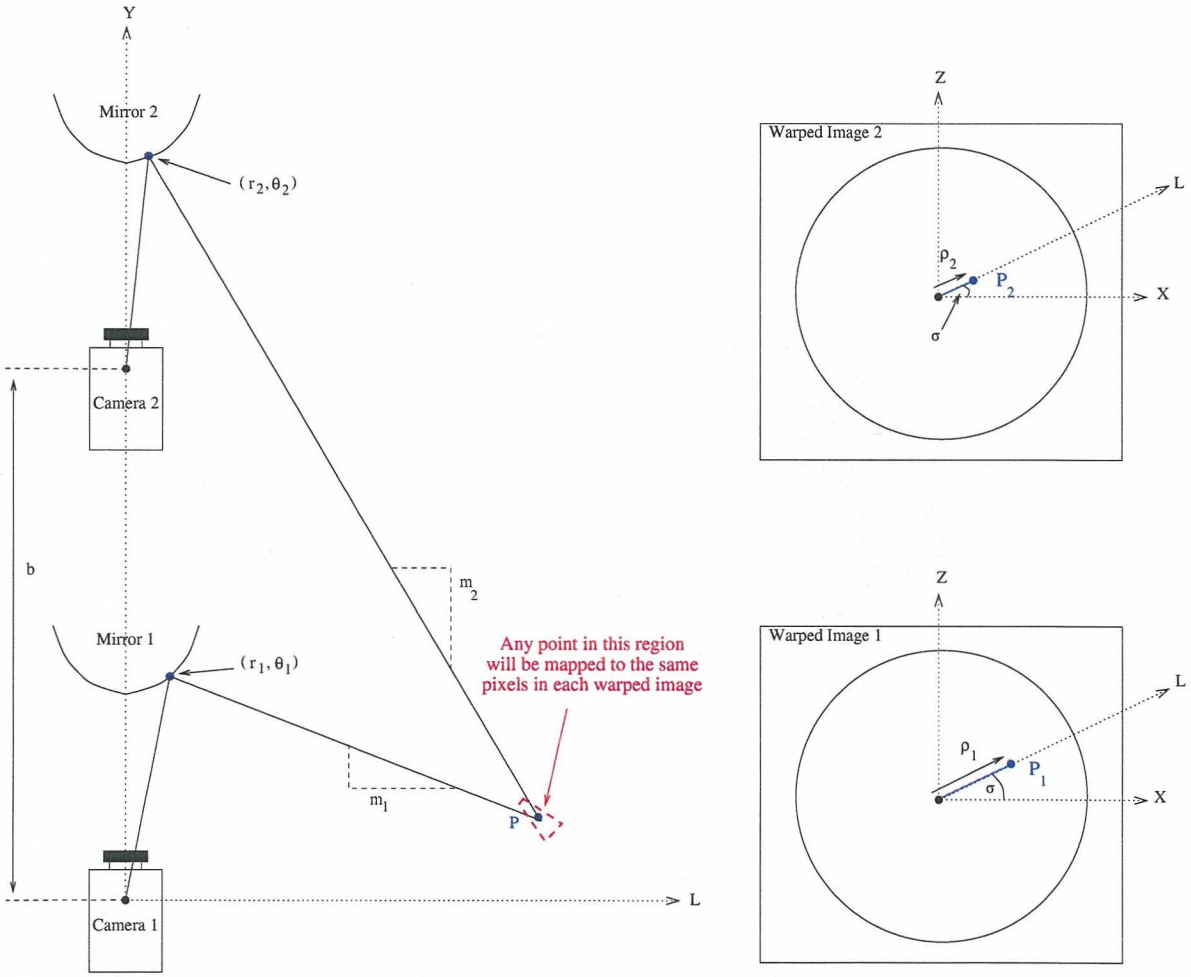


Figure 5.7: Any point within the red dotted line will be mapped to P_1 and P_2 in the warped images. As a result there will be a certain degree of error in range estimation, due to the effective discretisation of three-dimensional space.

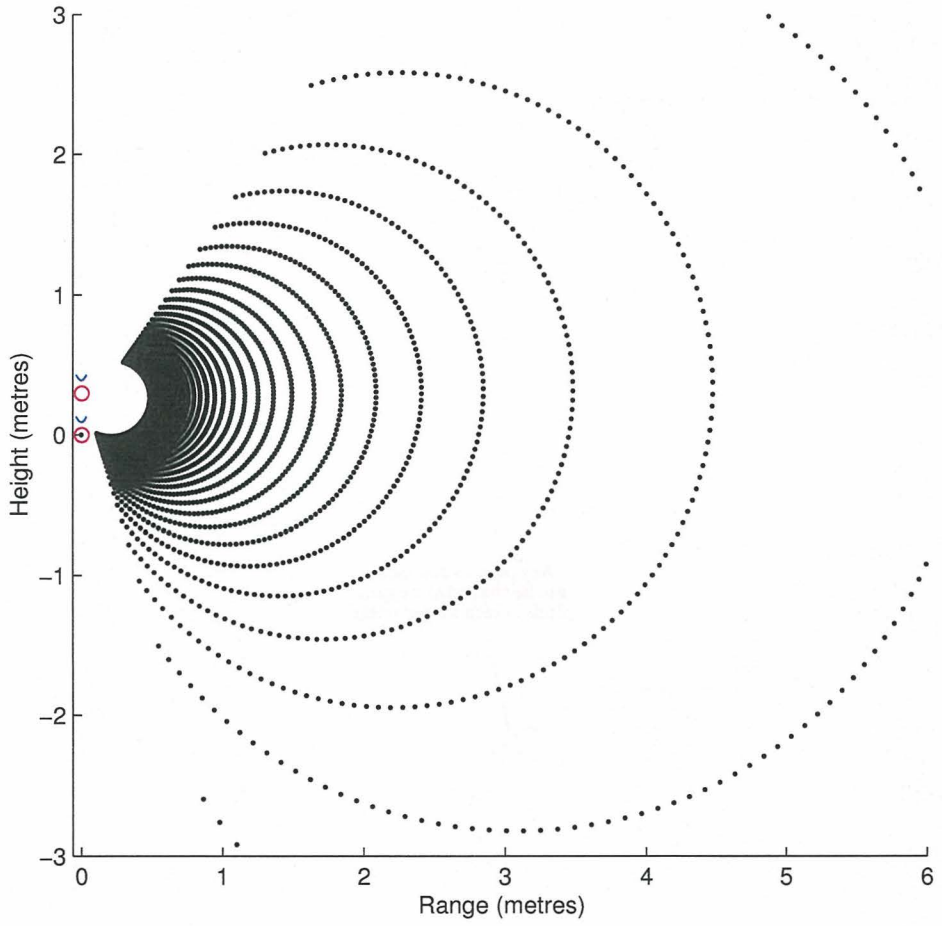


Figure 5.8: Depth resolution of a stereo panoramic sensor, using constant gain mirrors with a baseline of 30cm.

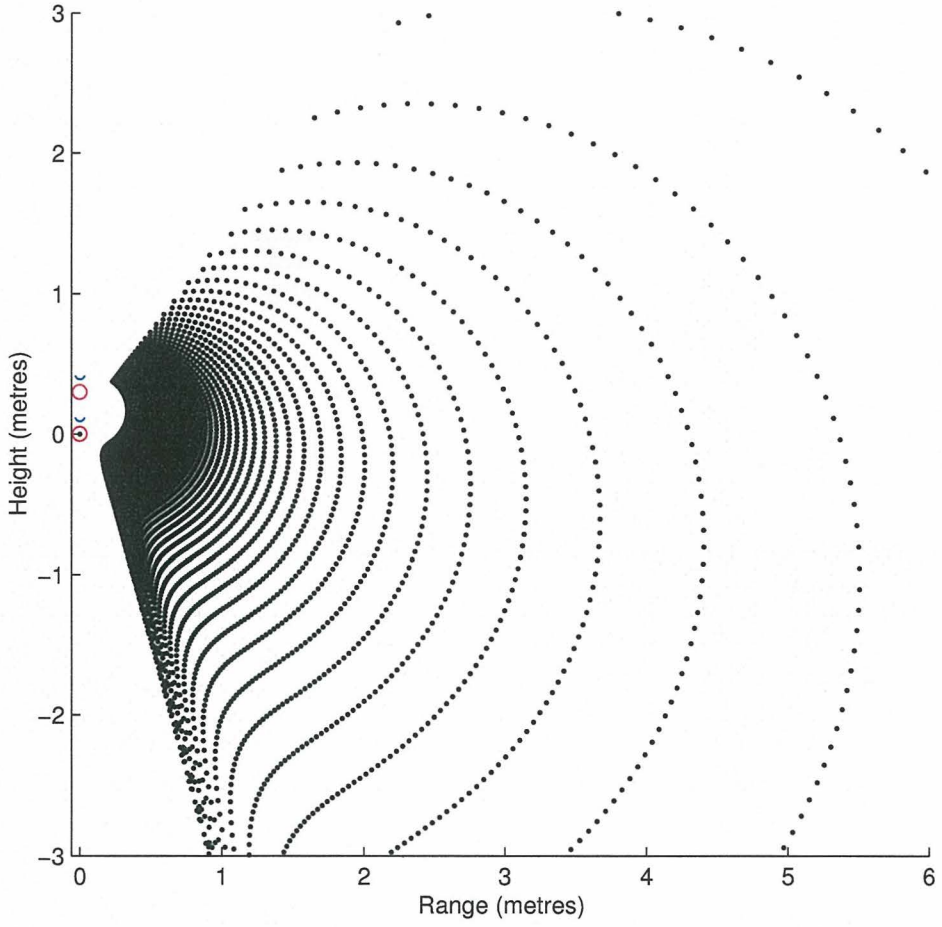


Figure 5.9: Depth resolution of a stereo panoramic sensor, using resolution invariant mirrors with a baseline of 30cm.

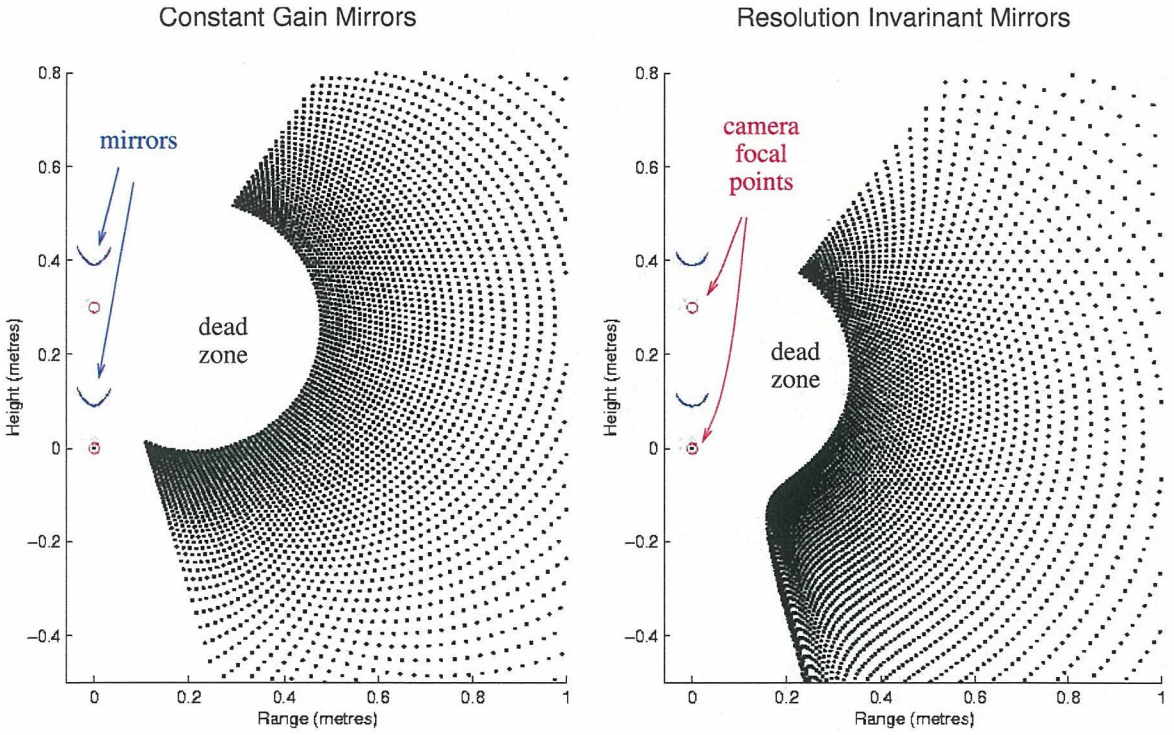


Figure 5.10: An enlargement of Figures 5.8 and 5.9, showing the *dead zone* where range can not be estimated.

to ensure that the disparity map can be produced at a rate that is appropriate for real-time applications. However, limiting the pixel search creates a *dead zone* area, in which the range to obstacles can not be estimated. This is due to the fact that any object appearing in this region will have a disparity greater than this value, and will therefore not be detected. A maximum disparity search of 64 pixels was chosen because it provided a good trade off between the size of the dead zone, and computational burden. The graphs in Figure 5.10 are an enlargment of the area around the stereo sensor from Figures 5.8 and 5.9, showing the extent of the dead zone for the two mirror profiles tested. It should be noted that such dead zones are also present in the other sensors discussed in Chapter 2. This is due to the time taken to switch from transmitting a ranging signal, to recieving the returning signal.

The depth resolution was calculated for stereo panoramic systems with various baselines, and mirror sizes. The following parameters were used in the simulation experiments:

- Baselines of 20cm, 30cm, 40cm and 50cm were used
- Constant gain mirror with a gain of 7.3 and a resolution invariant mirror, with field of views from 30 degrees to 140 degrees
- Mirror sizes with outer diameters of 10cm, 13cm,17cm, and 20cm

It was found that as the camera baseline increases, the size of the dead zone increases. However, at the same time the sensing range increases and the depth resolution becomes denser for further distances - that is, the baseline effects the scale of the graphs in Figures 5.8, 5.9, and 5.10. Therefore, there is a trade off between having a higher depth resolution for longer ranges, and an increasing dead zone. The depth resolution can also be increased by enlarging the mirror size. In this case, the trade off is between depth resolution and having a compact sensor.

As mentioned previously, a stereo vision system effectively discretises three-dimensional space when estimating the position of a point from two corresponding pixels. As a result, an object in space will be mapped to the closest point in Figures 5.8 and 5.9. Figures 5.12 or 5.11 show examples of the possible errors in estimating position using constant gain and resolution invariant mirror systems, respectively. The graphs on the left in each figure show the maximum possible discretisation error in range estimation along a ray of light with an elevation angle of 90° to the lower mirror axis. The discretisation error in the height of an obstacle along this same ray of light is displayed on the right hand side. The figures

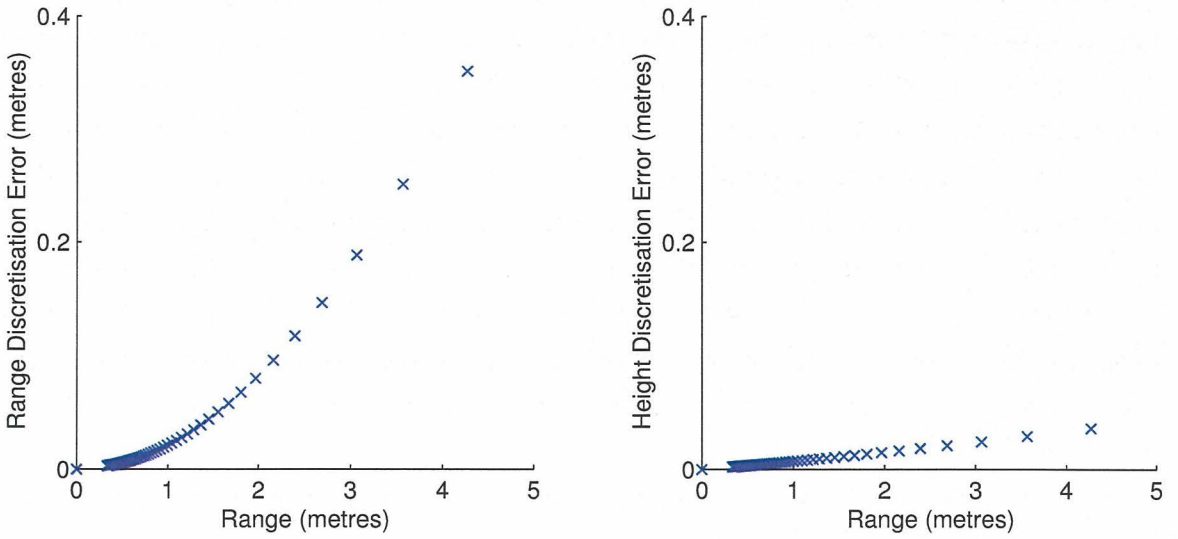


Figure 5.11: Maximum discretisation error of a resolution invariant stereo panoramic sensor along a ray of light perpendicular to the camera-mirror axis.

show that the range discretization error increases at a faster rate than the error in height estimation. The range error for the resolution invariant mirror was found to be higher than that for the constant gain mirror. For example, at a range of 3m the discretisation error was 0.2m for the resolution invariant profile. While the error for the constant gain mirror did not reach a value of 0.2m till a range of 3.75m.

The graphs presented in this section have shown the depth resolution of the stereo panoramic sensor in two-dimensions. However, there are further changes in resolution when the third-dimension is included. At the edge of a warped image, where the mirror radius is larger, there is a higher number of pixels available to capture the rays of light reflected off the mirror, as shown in Figure 5.13(a). On the other hand, in the centre of the raw image there are much fewer pixels to capture the rays of light from the mirror. This means there are more pixels per azimuth degree as the mirror radius increases. The number of pixels per degree in the azimuth direction can be calculated as follows:

$$p(\rho) = \frac{\pi \kappa w^2 [2\rho - 1]}{360} \tag{5.1}$$

where $p(\rho)$ is the number of pixels per degree in the azimuth directions, at a radius of ρ pixels. κ is the number of pixels per unit area, and w is width of a single pixel. This was plotted against the elevation angle for the graphs in Figure 5.14. For a constant gain

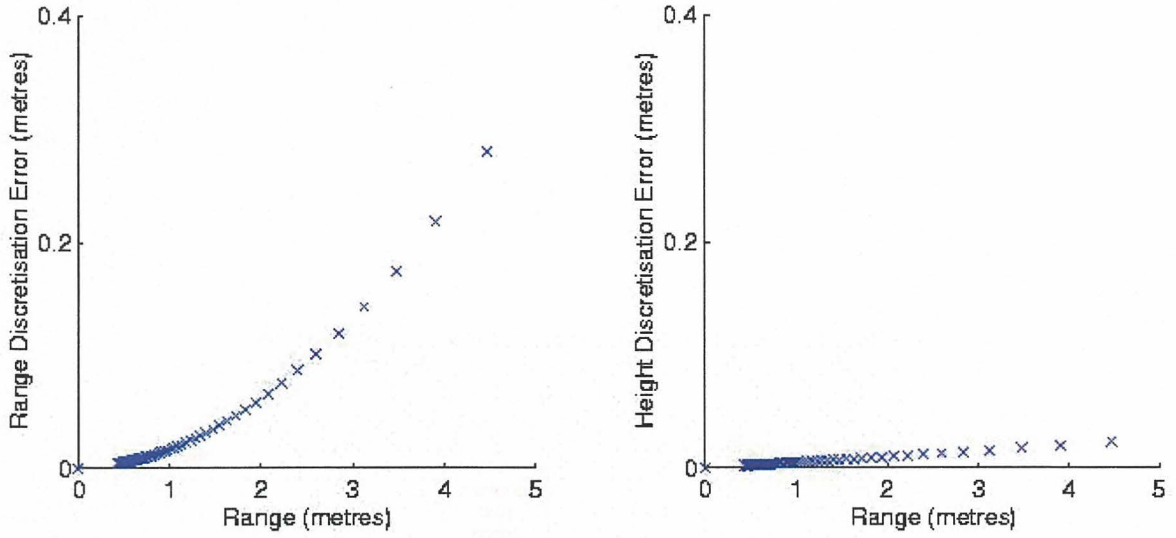


Figure 5.12: Maximum discretisation error of a constant gain stereo panoramic sensor along a ray of light perpendicular to the camera-mirror axis.

mirror, the elevation angle is

$$\phi(\rho) = \alpha\left(\arctan\left(\frac{\rho}{f}\right) - \underline{\theta}\right) + \underline{\phi} \tag{5.2}$$

where f is the focal length in pixels, $\underline{\theta}$ is the minimum radial angle viewed by the camera, and $\underline{\phi}$ is the minimum angle of elevation of a ray of light reflected by the mirror (see Figure 2.13). For a resolution invariant mirror, the elevation angle is calculated as follows:

$$\phi(\rho) = B_\alpha\left(\frac{\rho}{f}\right)^2 + \phi_{\theta=0} \tag{5.3}$$

where B_α is a constant defined by Equation 2.8, and $\phi_{\theta=0}$ is the angle of elevation of the ray of light reflected at the centre of the mirror.

As shown in Figure 5.14, in the case of a constant gain mirror, there is a linear relationship between the elevation angle and the number of pixels available to capture each degree in the azimuth direction. Although the resolution invariant mirror ensures that the image pixel density remains constant in the elevation direction, it is physically impossible to maintain a constant pixel density in the azimuth direction. As a results, there is a redistribution of pixels per azimuth degree. At lower elevation angles, the number of pixels available per azimuth degree increases more rapidly. Then at the edge of the image, the

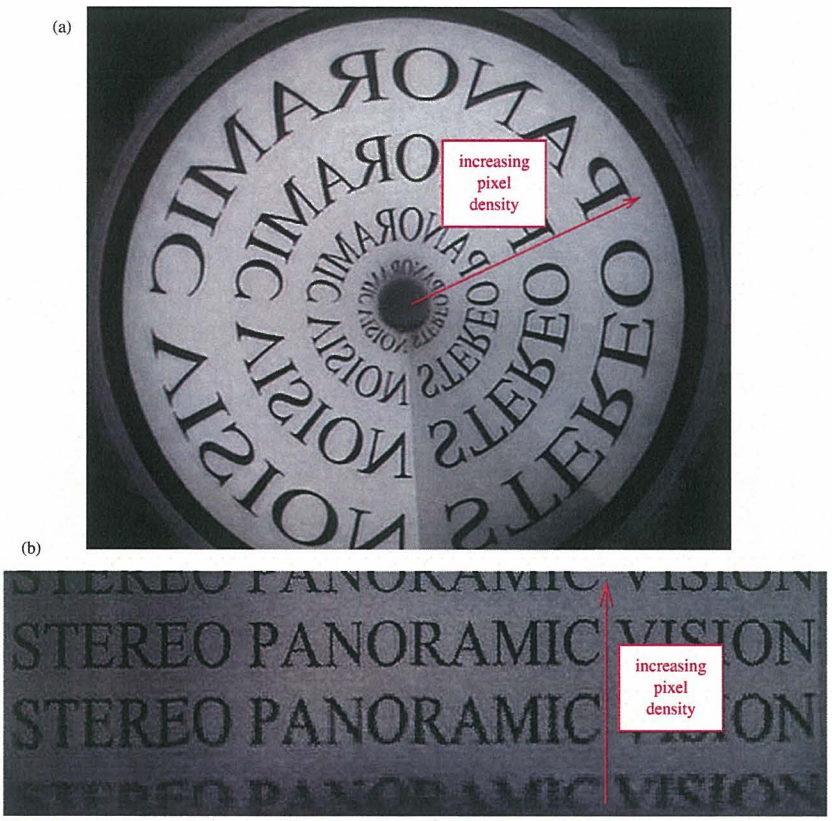


Figure 5.13: Depiction of the direction of increasing pixel density. (a) Raw image. (b) Unwarped image.

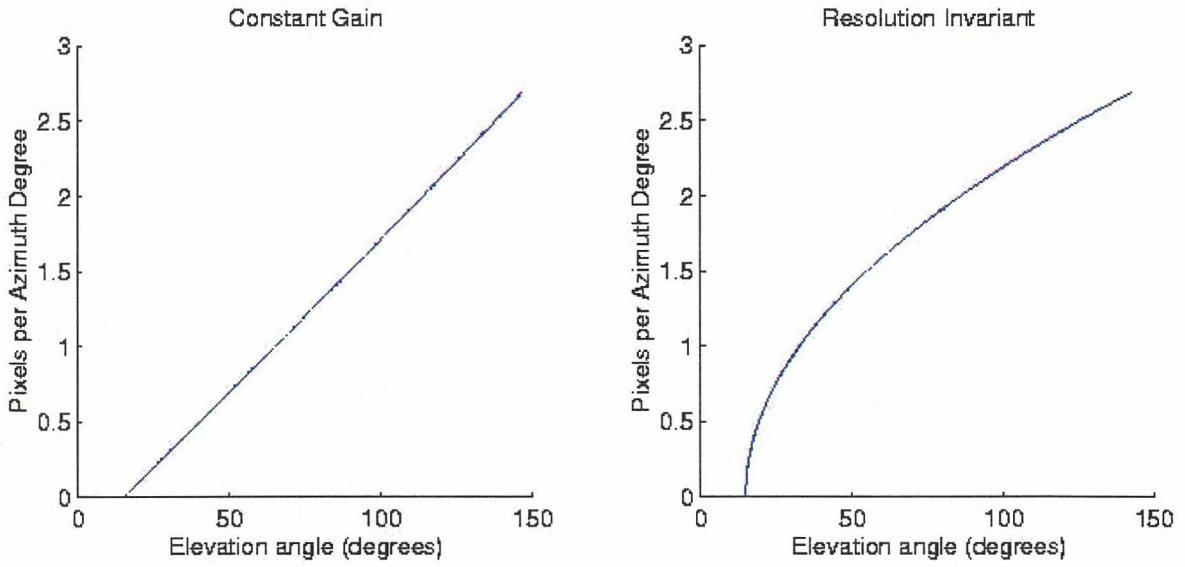


Figure 5.14: Pixels per azimuth degree for various elevation angles for both mirror profiles.

resolution increases at a slower rate.

5.2.2 Baseline

Figure 5.15 shows the estimated range up to a maximum disparity of 64, for various baselines. As with Figures 5.12 or 5.11, the range has been estimated along a ray of light with an elevation angle of 90° to the lower mirror axis. The graph shows the trade offs in sensor characteristics that are involved with selecting a baseline. As mentioned previously, a smaller baseline results in a smaller dead zone. For example, at the maximum disparity of 64 pixels, a 20cm baseline can detect an obstacle as close as 30cm. While a sensor with a 50cm baseline can only begin to detect obstacles at a range greater than 70cm. On the other hand, the advantage of having a larger baseline is the ability to detect objects at larger ranges. The graph shows that a disparity of 10 pixels indicates an obstacle at 2.2m for a 20cm baseline, whereas a 50cm baseline would be detecting an obstacle at 5.4 metres. Therefore we want to select a baseline that will detect obstacles at the the outer edges of the required work volume, while preserving the ability to detect obstacles at close range.

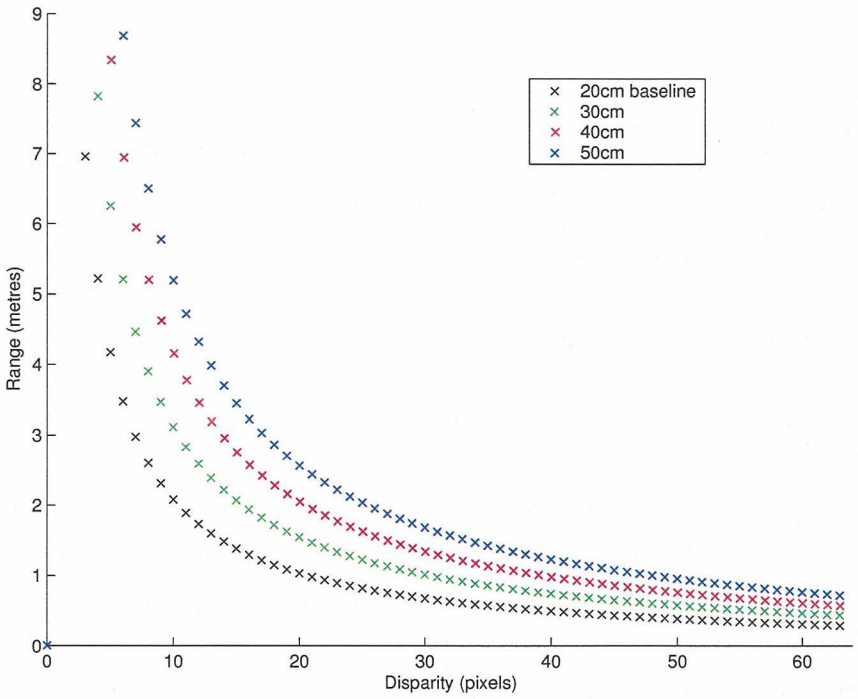


Figure 5.15: The range estimations up to a maximum disparity of 64. Constant gain stereo panoramic sensors with baselines from 20cm to 50cm were used.

5.2.3 Sensor Configuration

The stereo panoramic vision sensor used in this research projecct was designed to be mounted in either a vertical or horizontal configuration to investigate which would be more effective for our application. Figures 5.5 and 5.3 show the images captured in a horizontal configuration. The mirrors used in this project were designed with a minimum viewing angle of 30 degrees and a maximum of 140 degrees. As a result, a sensor in a horiontal configuration will not be able to view objects directly to the left or right, near the rear of the vehicle. However, it will be able to monitor the area directly underneath and above the vehicle.

Conversely, a sensor in a vertical configuration will be able to monitor blind-spots to the right and left of the vehicle, but not directly underneath. Although, this could be improved by mounting the sensor lower on the vehicle. It can also be seen from Figures 5.6 and5.4 that a larger portion of the panoramic image is devoted to the area near the ground plane when using a vertical configuration. Since this is where most obstacles will appear in our application, this is an important consideration.

5.2.4 Summary of Sensor Characteristics

The theoretical limitations of the range finding capabilities of the stereo panoramic vision system have now been presented. A summary of how the various parameters affect the sensor characteristics is can be found in Table 5.1. The minimum accuracy of the range estimation is defined by the depth resolution of the system, which is created by the discretisation of the image space. The next step in evaluating the effectiveness of this system is to carry out a ground truth analysis, to test both the range finding and obstacle detection algorithms.

5.3 Range Estimation

Images were captured of an environment as reflected off a virtual convex mirror surface, using the POV-Ray software package. These were used to simulate an ideal stereo panoramic vision system. The artificial images were unwarped to produce artificial panoramic images, in which the exact location of each object was known. The unwarped images had dimensions of 500 by 200 pixels for a vertical configuration, and 300 by 200 for a horizontal mounting. These were generated from ray traced images that had dimensions of 640 by

Sensor Parameter	Result Characteristic
Increasing the baseline	Increases the size of the dead zone
	Increases the sensing range
	Improves the depth resolution at longer ranges
Increasing the mirror size (requires a larger CCD)	Increases the depth resolution
Resolution invariant mirror profile	Higher discretisation error: 0.2m at a range of 3m
	More pixels per degree in the azimuth direction towards the centre of the image: 1.3 pixels at an elevation angle of 50 degrees
	Smaller dead zone
Constant gain mirror profile	Lower discretisation error: 0.2m at a range of 3.75m
	Less pixels per degree in the azimuth direction towards the centre of the image: 0.75 pixels at an elevation angle of 50 degrees
	Larger dead zone
Increasing maximum disparity	Increases computational burden
	Decreases size of the dead zone
Horizontal configuration	Less image pixels available in working area
	Can monitor area underneath and directly above the rear of the vehicle
Vertical configuration	More image pixels available in working area
	Can monitor area to the immediate left and right
	of the vehicle

Table 5.1: Summary of the sensor characteristics

480 pixels. Corresponding features were selected manually, as shown in Figure 5.16, with the aim of determining the upper limitations on the system's range estimation accuracy. Range was then calculated from these manually selected disparities. Each feature was selected ten times to provide a general indication of the level of human error involved in the manual feature selection. This was done for stereo panoramic systems with various baselines. The following parameters were used in these simulation experiments:

- Baselines of 20cm, 30cm, 40cm and 50cm were used
- A constant gain mirror with a gain of 7.3 and a resolution invariant mirror, with field of views from 30 degrees to 140 degrees
- Both mirror types had outer diameters of 10cm.
- It was decided that a vertical configuration would be used, as this has a larger useful working volume than the horizontal configuration.

The results for range, height and angular estimation for each feature were analysed and graphed. Figures 5.17, 5.20, and 5.21 show the results for constant gain and resolution invariant panoramic systems with a baseline of 30cm. The average error for each feature is plotted on the left hand side, while the standard deviation in the error in measurement is displayed on the right.

As can be seen in Figure 5.17, the average range error for both mirror profiles remains below 0.2m until a range of 3.5m. At this point, the range error for the resolution invariant mirror rises above 0.6m, while the constant gain panoramic system is always below 0.31m. It should be noted that the constant gain mirror meets Specification 4 set out in Chapter 1, that the range sensor should be accurate to within 0.2m up to a range of 4m. The standard deviations increase to 0.76m and 0.85m at a range of 6m for the constant gain and resolution invariant mirrors, respectively.

The increases in error and standard deviation are partially due to human error in the selection of corresponding features. Features on objects that are further away from the sensor can be harder to distinguish and select accurately, which introduces errors in the disparity calculations. For example, if we examine Figure 5.15, a sensor with a baseline of 30cm, a 10 pixel disparity represents an object 3.2m away. While a disparity of 9 pixels representsan object at 3.5m. Therefore, an incorrect selection of a feature with a 1 pixel offset translates to an error of 30cm. Human error is the most likely cause for the spikes in range estimation, if a feature was consistently selected incorrectly over the 10 samples.

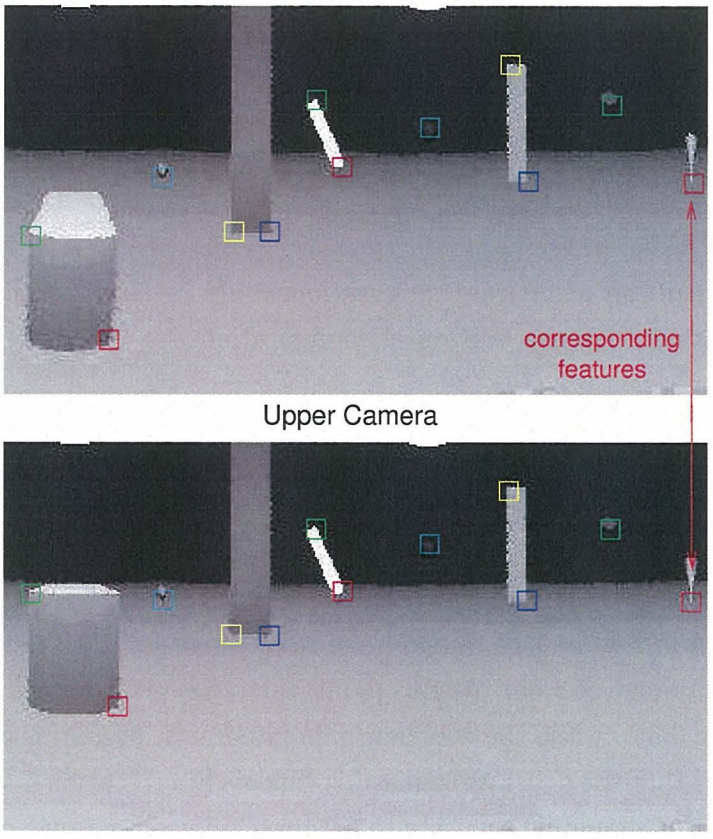


Figure 5.16: Unwarped panoramic images created from the POV-Ray traced images. Corresponding features were selected manually to evaluate the range finding capabilities of the stereo panoramic system.

Another cause of error is the discretisation of three dimensional space was discussed in Section 5.2. This is when an object is positioned at a point inbetween the range estimation points depicted in Figures 5.12 and 5.35.9. The range of the object will be mapped to the closest point on those graphs, and difference between the actual position and the estimated position is shown in Figures 5.12 and 5.11. This accounts for the fact that the error in range estimation is greater for larger distances when the resolution invariant mirror profiles are used, as opposed to the constant gain mirrors. This also explains why the error in range estimation for the resolution invariant mirror increases more rapidly than that for the constant gain profiles, however the majority of this can be attributed to the range discretisation characteristics.

Figures 5.18 and 5.19 show the results of the range estimation as viewed from above the virtual environment. The panoramic sensor is in the centre, at coordinate (0,0), with obstacles at various locations in the surrounding environment.

The average error in height estimation was lower than that for range estimation, as shown in Figure 5.20. For the constant gain mirrors, the error always remained below 0.19m, however the resolution invariant mirrors had errors as high as 0.44m at a range of 3.7m. In both cases, the standard deviations remained below 0.3m. The lower level of precision seen in the results for the resolution invariant mirrors are again due to the higher discretisation errors.

The estimation of the azimuth angular position was particularly impressive, with a maximum error of 3° for the constant gain mirror, and 3.7° for the resolution invariant mirror. The maximum standard deviation was 0.7° for the constant gain profile and 0.55° for the resolution invariant mirror. These results are displayed in Figure 5.21.The high level of precision could be attributed to the fact that there was over one degree per pixel in the unwarped panoramic image, which had a width of 500 pixels.

5.3.1 Comparison with Previous Results

Section 2.5 discussed the previous work to investigate the use of stereo panoramic sensors for range estimation. The theoretical results presented were generated using ray tracing software to simulate the images captured by the panoramic sensors. Matching features were selected manually to remove errors in stereo correspondence computations. (Bunschoten & Kröse 2001) tested hyperboloidal mirrors with a 0.5m baseline placed at the centre of a cube of 6m. The range to the sides of the cube was estimated to within 0.2m, which is similar to the results generated in our simulations.

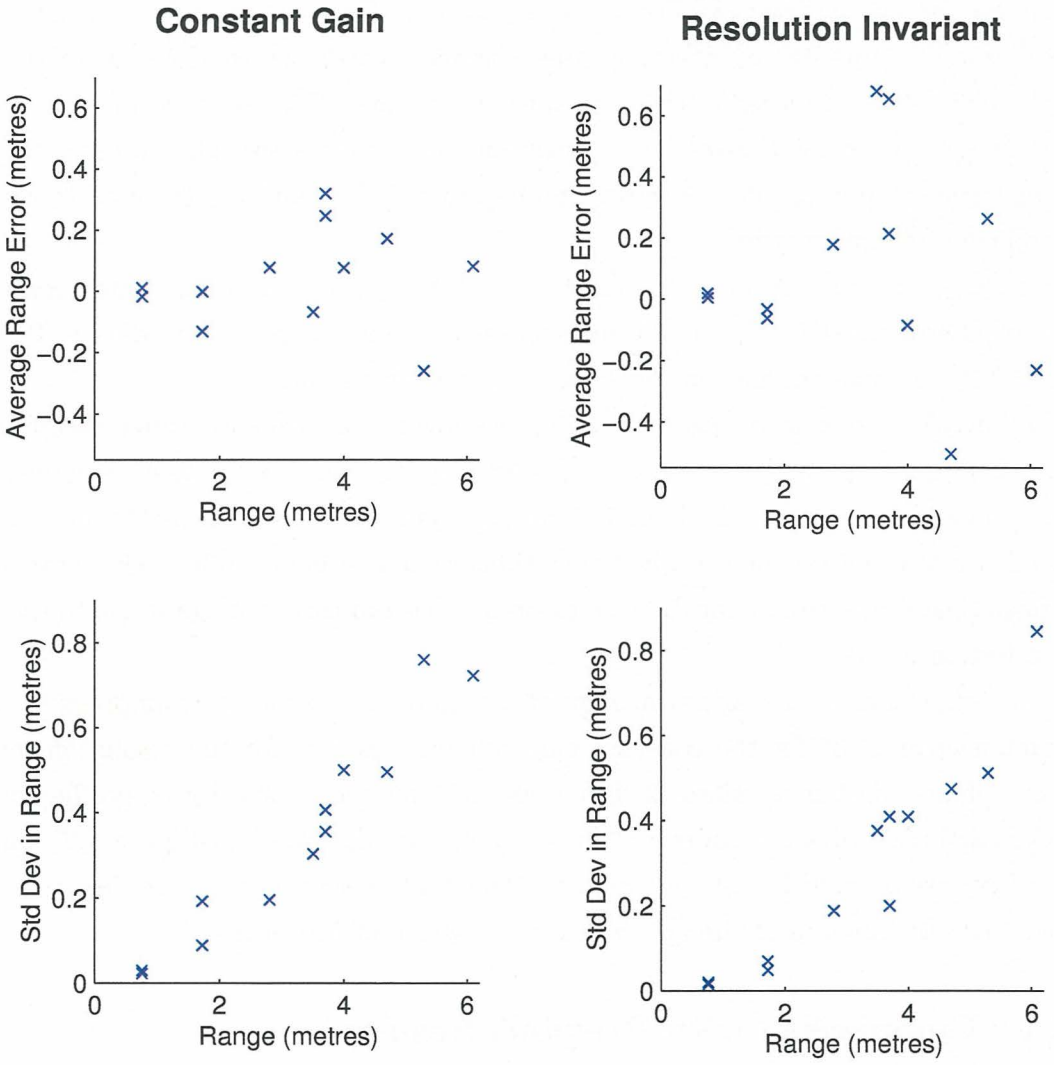


Figure 5.17: The range estimation accuracy calculated from the ground truth experiments, using a baseline of 30cm. The average range error remains below 0.2m until a range of 3.5m for both mirror profiles.

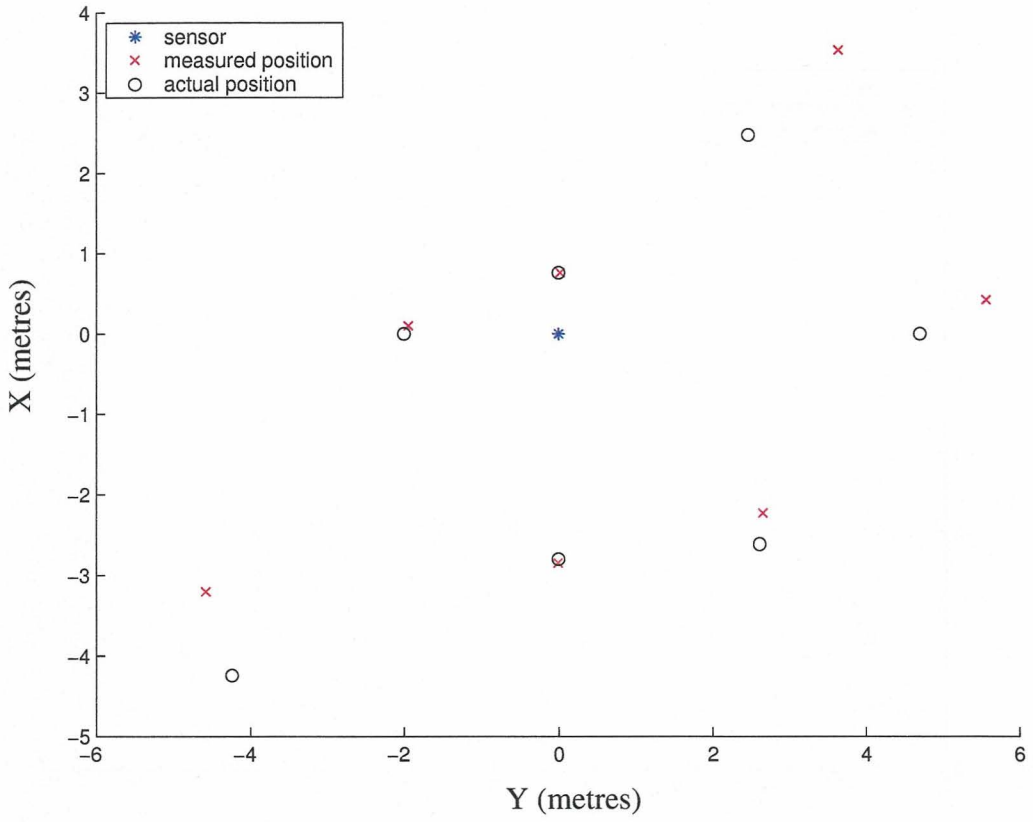


Figure 5.18: A topview of the virtual environment ray-traced for the ground truth experiments. A stereo panoramic sensor using a constant gain mirror is located in the centre at (0,0). The actual position of the objects are shown by the circles, while the positions estimated by the v-disparity algorithm are the red crosses.

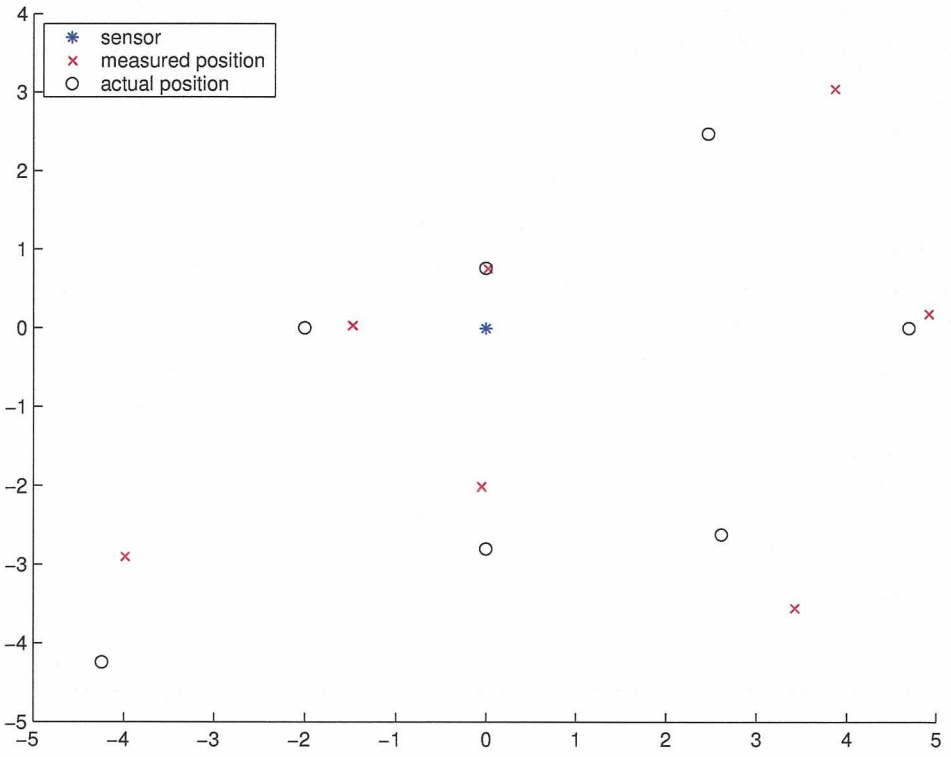


Figure 5.19: A topview of the virtual environment ray-traced for the ground truth experiments. A stereo panoramic sensor using a resolution invariant mirror is located in the centre at (0,0). The actual position of the objects are shown by the circle, while the positions estimated by the v-disparity algorithm are the red crosses.

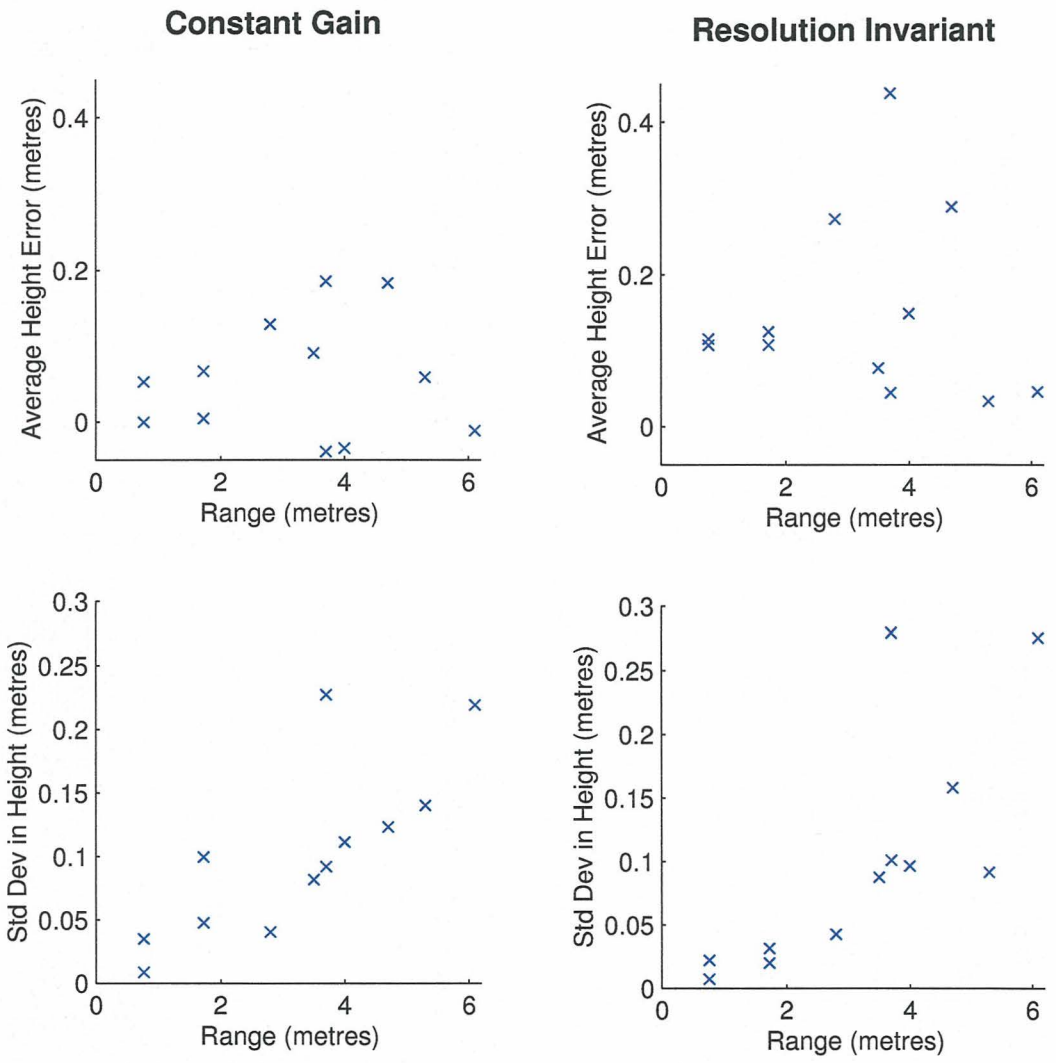


Figure 5.20: The average height estimation error and height error standard deviation calculated from the ground truth experiments. A baseline of 30cm was used.

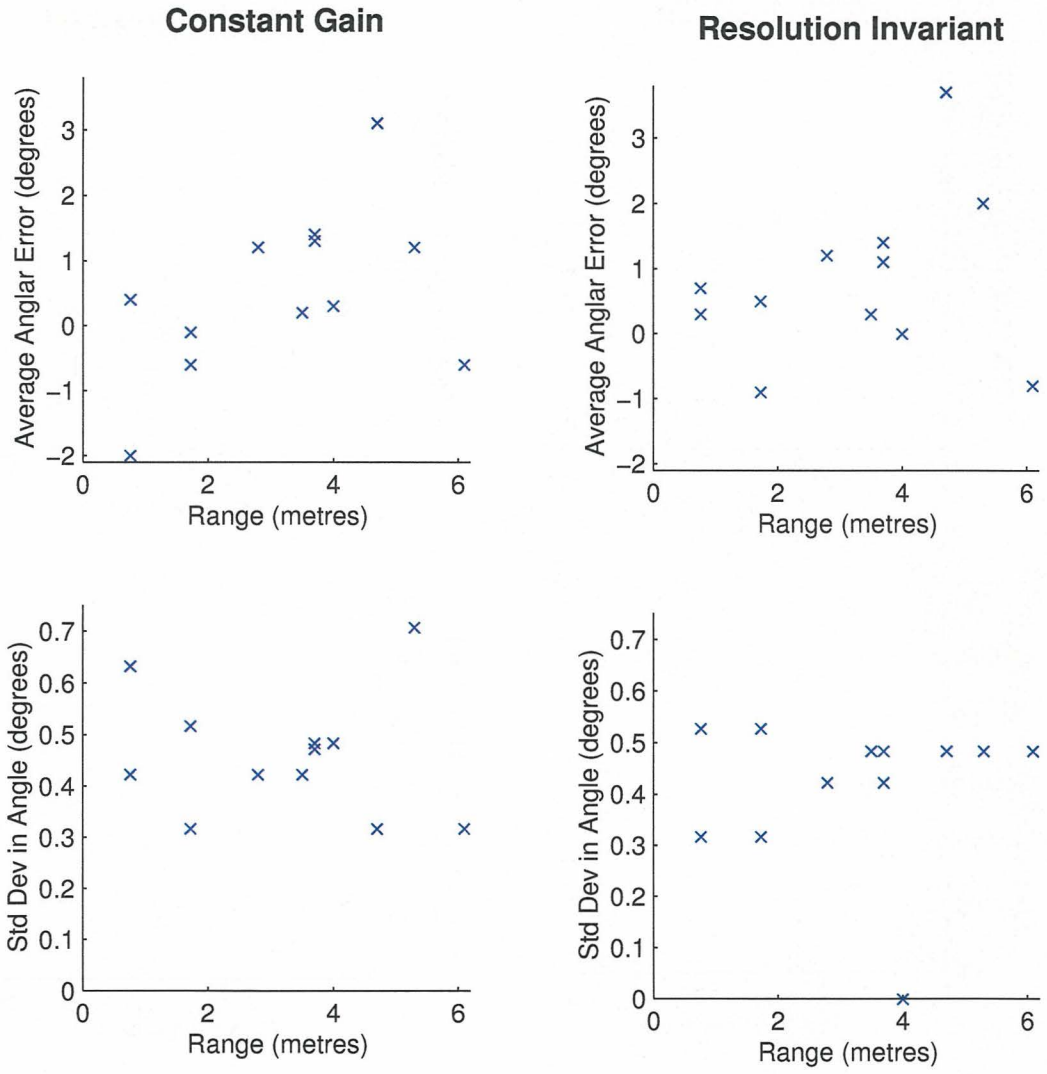


Figure 5.21: The average angular position estimation error, and standard deviation calculated from the ground truth experiments. A baseline of 30cm was used.

(Ollis et al. 1999) used constant gain mirrors with a field of view of 90 degrees in the vertical direction, and a baseline of 0.3m. They reported a measurement error of 0.12m at a range of 5m. This is slightly better than the performance of our constant gain system, which had a range error of 0.2m at the same range. However, this is probably due to the differences in the field of view of each system. Our mirrors have a field of view of 110 degrees in the vertical direction, with the same size CCD (640 x 480 pixels). Therefore our panoramic sensors have a lower resolution than the sensors used by (Ollis et al. 1999).

(Conroy 2000) used a single parabolic mirror that consisted of two resolution invariant profiles to provide disparity information. As a result, this system has small baseline of only a few centimetres, with a maximum range of approximately 1.5m. At this range the maximum measurement error was reported to be 0.45m, which is more than four times larger than the error found in our resolution invariant system, at the same depth.

5.4 Obstacle Detection

The previous section presented the range estimation capabilities of the stereo panoramic vision system. This was done using a human observer to manually determine the corrresponding features, and determine the range to an object without the use of a disparity map. This section evaluates the two different obstacle detection methods suggested in Chapter 4 using the ground truth data generated with the POV-Ray software. The *ground plane subtraction* and *v-disparity* approaches segment regions of interest from the disparity maps. Once this is completed, the range to each object in the surrounding environment can be calculated. The performance of the range estimation of the panoramic system using the disparity maps and obstacle detection algorithms is presented at the end of this section.

5.4.1 Ground Plane Subtraction

As described in Section 4.2.1, the Ground Plane Subtraction approach to obstacle detection assumes a known location and orientation of the ground plane, relative to the stereo sensor. This assumption is used to create a model of the surrounding environment with no obstacles present. From this a theoretical disparity map is calculated, which contains only the ground plane. This is subtracted from a disparity map of the actual environment, leaving behind only non-ground pixel - the obstacles.

Figures 5.22 and 5.23 show the various stages of the Ground Plane Subtraction algorithm. In these figures, (a) shows one of the stereo unwarped images generated from the

ground truth data. The disparity map determined from the unwarped images is shown in (b). The disparity maps have been calculated with respect to the image displayed in (a). Finally, (c) shows the resulting image once the theoretical disparity map has been subtracted from the actual map, which contains several obstacles. It can be seen that the ground plane has been successfully removed from the disparity map, leaving behind the areas of interest.

However, this method suffers from a number of drawbacks which will be discussed further in Chapter 6. In particular, the assumption of a ground plane in a known location decreases the robustness of obstacle detection. Even if the position of the ground plane was determined accurately initially, due to the nature of our application, the orientation of the ground plane may change over time. For example, the vehicle may drive along a road that has a changing slope. In this case, the Ground Plane Subtraction algorithm may no longer successfully segment obstacles, since the assumed model of the environment will differ from the real world scenario. It would be more desirable to be able to detect the ground plane online, and segment objects from around the vehicle accordingly. The results of the ground truth evaluation of such an approach are discussed in the next section.

5.4.2 V-Disparity

Unlike the Ground Plane Subtraction method for segmenting obstacles, the *v-disparity* approach requires no *a priori* knowledge of the exact location and orientation of the ground plane. Instead, the ground plane can be reliably segmented online from each new pair of stereo images, as described in Section 4.2.2.

Figure 5.24 shows the results from applying the *v-disparity* algorithm to a virtual stereo sensor which utilises resolution invariant mirrors. The system has been placed in a vertical configuration. The coloured boxes show areas containing an obstacle, and the colour of the box indicates the range to that object:

- Red is between 0 and 0.8m
- Orange is between 0.8 and 1.8m
- Green is more than 1.8m away from the sensor

Similarly, Figure 5.25 shows the results from a stereo panoramic sensor using constant gain mirrors, in a horizontal position. As can be seen from the figures, the algorithm was able to segment areas of the image containing obstacles. The objects that remain undetected are

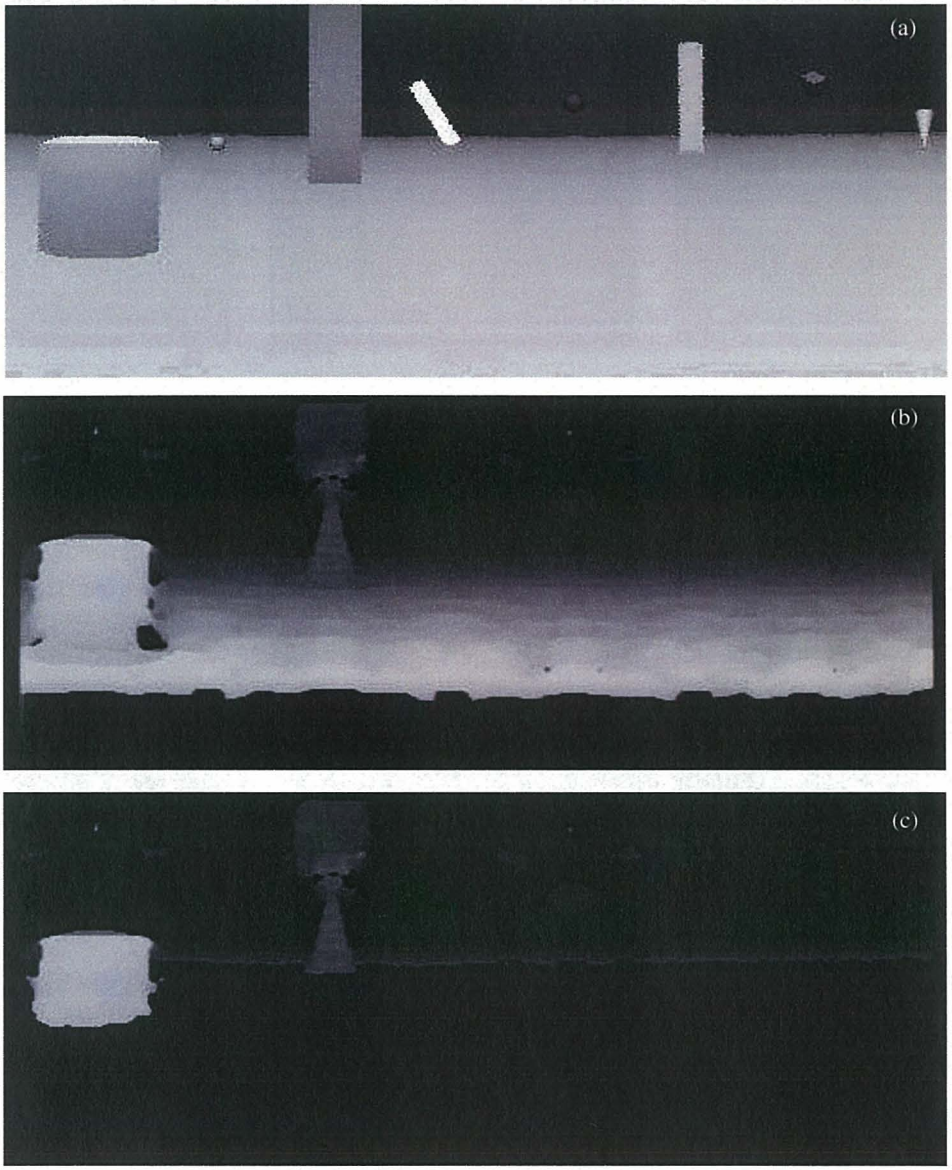
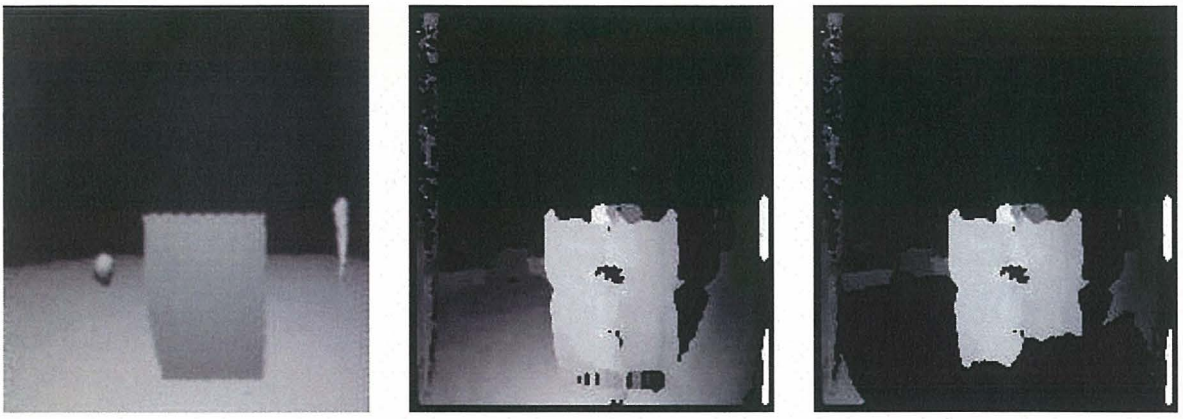


Figure 5.22: The ground plane subtraction algorithm, applied to a stereo panoramic vision system with constant gain mirrors, in a vertical configuration. (a) One of the unwarped stereo images. (b) The disparity map, calculated with respect to image (a). (c) The disparity map resulting from a subtraction between image (b) and the calculated theoretical ground plane.



(a) (b) (c)

Figure 5.23: The ground plane subtraction algorithm, applied to a stereo panoramic vision system with resolution invariant mirrors, in a horizontal configuration. (a) One of the unwarped stereo images. (b) The disparity map, calculated with respect to image (a). (c) The disparity map resulting from a subtraction between image (b) and the calculated theoretical ground plane.

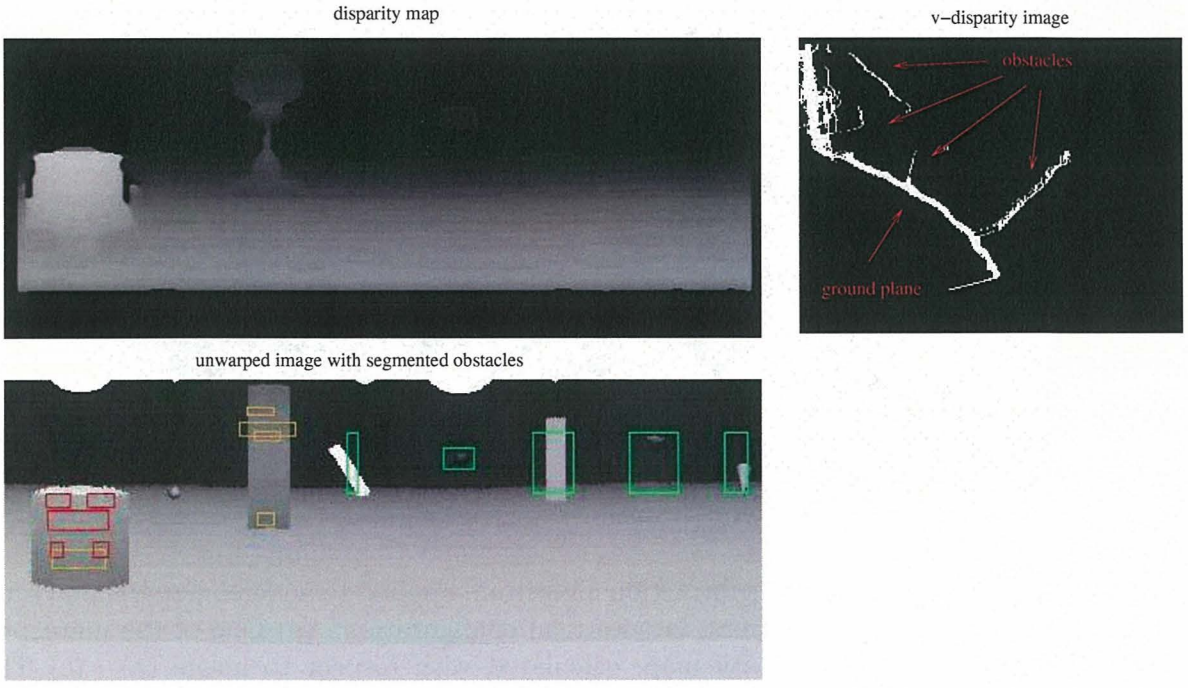


Figure 5.24: The results of the *v-disparity* algorithm, when applied to unwarped images from a resolution invariant mirror system. In this case a 30cm baseline was used with the sensor in a vertical configuration.

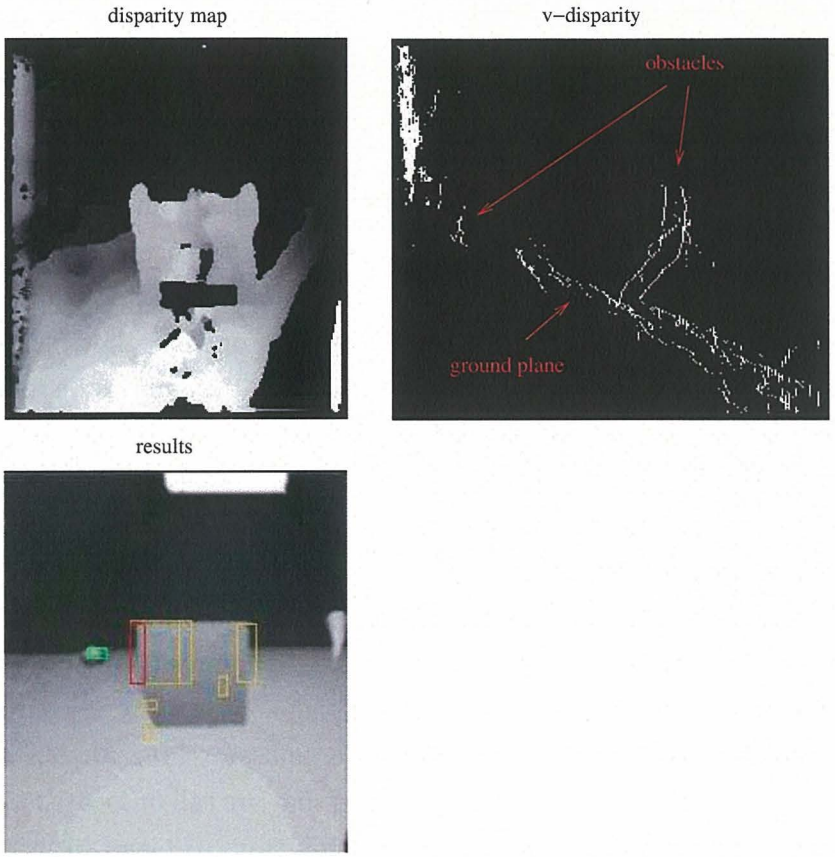


Figure 5.25: The results of the *v-disparity* algorithm, when applied to the unwarped images. In this case constant gain mirrors were used, with the stereo panoramic sensor in a horizontal configuration and a baseline of 30cm.

too far from the range sensor, and begin to merge with the background in the v-disparity images.

The ground plane appears to be more clearly defined in *v-disparity* images produced from a sensor in a vertical configuration (Figure 5.24), when compared to a horizontal configuration (Figure 5.25). This is due to the fact that the ground plane fills a higher proportion of the panoramic image, when the sensor is place vertically. The relative sizes of the unwarped images is also a contributing factor. As a result, extraction of the ground plane with the *v-disparity* algorithm will be more robust when the sensor is in a vertical configuration.

Ground Plane Detection

We have stated that the v-disparity algorithm does not require pre-calibration to determine the orientation of the ground plane. However, the algorithm is more effective when the ground plane is perpendicular to the panoramic sensor axis. While the algorithm can easily handle smaller deviations in the ground plane orientation, the obstacle detection capabilites will begin to deteriorate as the slope of the ground plane increases. Figure 5.26 shows some examples of v-disparity images produced from an environment that contains the ground plane only. The first example shows a ground plane that has no slope and the result is a v-disparity image contatining a downward sloping line representing the ground plane. The remaining examples show the results produced when a slope of 10 degrees is introduced. When the ground plane is sloping either upwards or downwards, the corresponding line in the v-disparity image becomes slightly wider. When the ground plane has a sideways slope, this effect is greater, with a much larger spread of white pixels in the vertical direction.

If we examine Figure 5.27, it is clear that the line representing the ground plane becomes even more spread out as we introduce a larger slope. We have tried to overcome this problem by detecting the upper edge of this line, and searching for obstacles above this edge. This method would be effective for obstacles that are tall or located above the ground plane. However, shorter objects on the ground may start to blend in with the wider line detected in the v-disparity image.

Although these results may be of concern for certain applications, it should also be noted that in this case the panoramic sensor will be located on the rear of a vehicle. It is therefore highly likely that the sensor will be oriented in such a way that the ground will be close to perpendicular relative to the sensor axis. Unless the vehicle is driven on a terrain with a high degree of local curvature, this should not be an issue.

5.4.3 Range Estimation using Automatic Stereo Correspondence

Once the obstacles have been segmented from the disparity map, the range to each object can be determined using a lookup table of precomputed disparities, for increased efficiency. Figure 5.24(a) displays results of the obstacle segmentation for a vertical configuration. The algorithm was able to detect at least sections of all obstacles, except for one sphere, which did not have enough contrast with the background for the stereo matching algorithm to be successful. It should be noted that the Sum of Absolute Differences similarity measure was used as the stereo matching algorithm in the presented results. This enabled the use

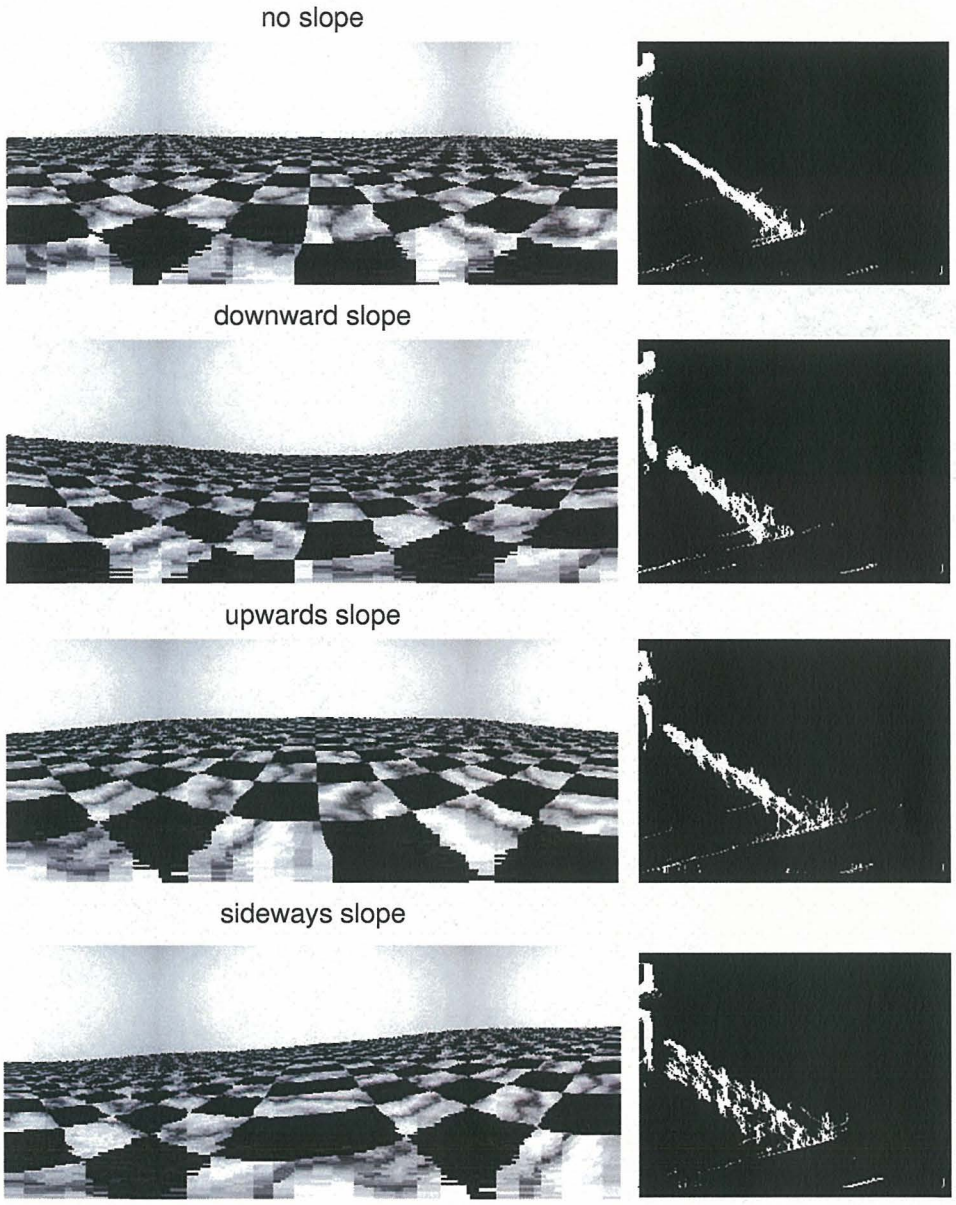


Figure 5.26: A comparison of the effects of a sloping ground plane of 10 degrees in various directions.
The unwarped image is shown in the left column, and the resulting v-disparity image is shown in the right column.

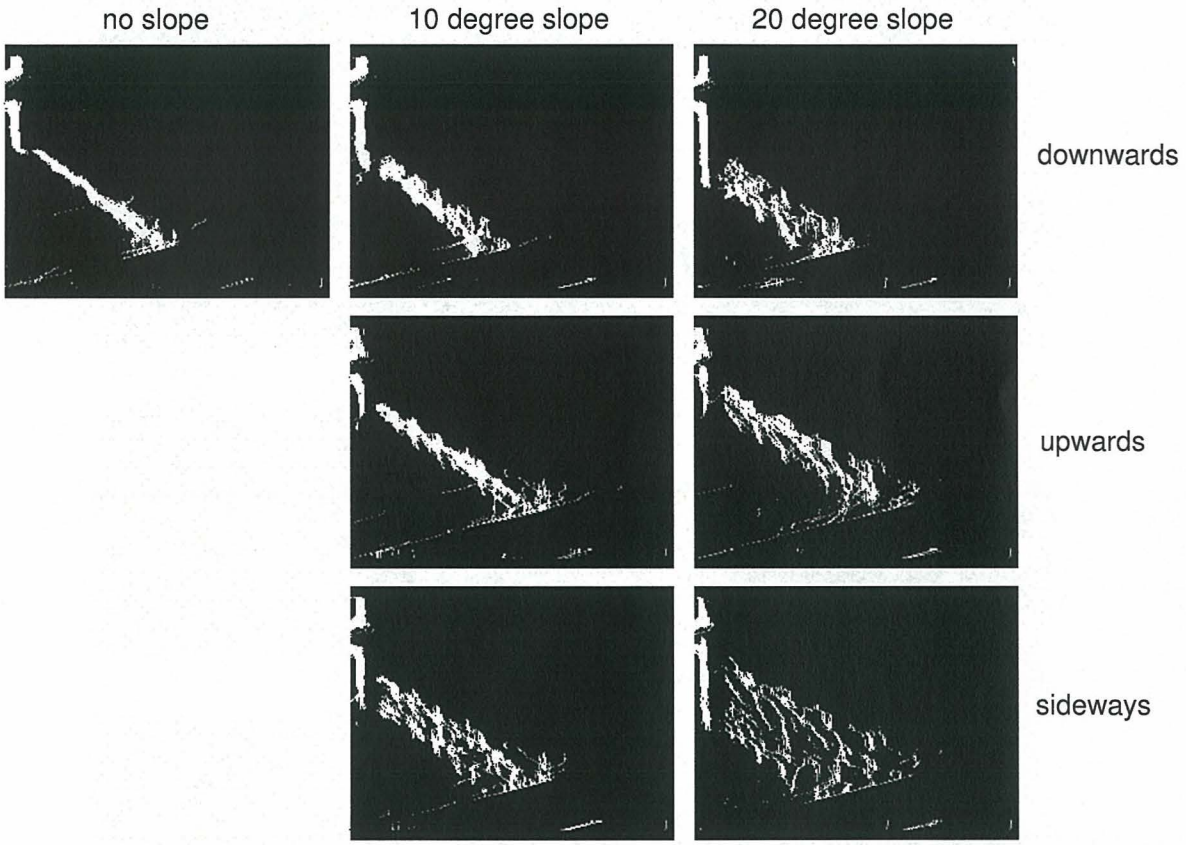


Figure 5.27: The v-disparity images produced from environments containing only ground planes of various slopes.

of a textureless ground plane and obstacles, using only the colour gradient provided by the ambient light in the ray traced images. To ensure that there were no errors introduced because of this, the experiments were repeated using the Normalised Cross-Correlation stereo matching. It was found that there were no significant differences in results.

These experiments were carried out for baselines of 20cm, 30cm, 40cm and 50cm. The error in range estimation for each baseline, and each mirror profile are plotted in Figures 5.28 and 5.29. It can be seen that across all baselines that the constant gain mirrors tend to provide more accuracy in range estimation than the resolution invariant profiles, particularly for larger ranges.

Using the constant gain mirrors with a baseline of 30cm, the system was able to estimate the range to obstacles to within 20cm, up to a range of 4m. The estimation of the azimuth angular position was within 3° of the true value, up to a range of 3.5m. On the other hand, the resolution invariant mirrors provided a lower precision range estimation. The range was estimated to within 20cm, only up to a range of 3m, where the error escalated to above 50cm. The estimation of the azimuth angular position was more reliable, remaining within 2° of the actual value, to a range of 3.5m.

The graphs also show that there is a spike in the error at a range of 4 to 5m, depending on the baseline. Some of this increase can be attributed to discretisation errors. However due to the magnitude and the subsequent decline in error this cannot be the only cause. In Figure 4.12, the background has been highlighted in green, and appears as a thick vertical line on the left of the v-disparity image. There are also several lines to the right of the background, which represent other objects. This is an example showing how objects further away from the sensor can become merged with the background, and other objects at similar distances, due to smaller disparity differences. More examples can be seen in Figures 5.24 and 5.25 . However, as we can see in the graphs in Figures 5.28 and 5.29, as the baseline increases this effect becomes less evident, as it becomes easier to differentiate between objects at larger distances. It should also be noted that the magnitude of the spike in error is not as large for constant gain mirrors in comparison to the resolution invariant profile.

5.5 Selection of Sensor Characteristics

Given the theoretical analysis and simulation experiments that have been presented in this chapter, it is possible to select sensor characteristics that would be most desirable for our

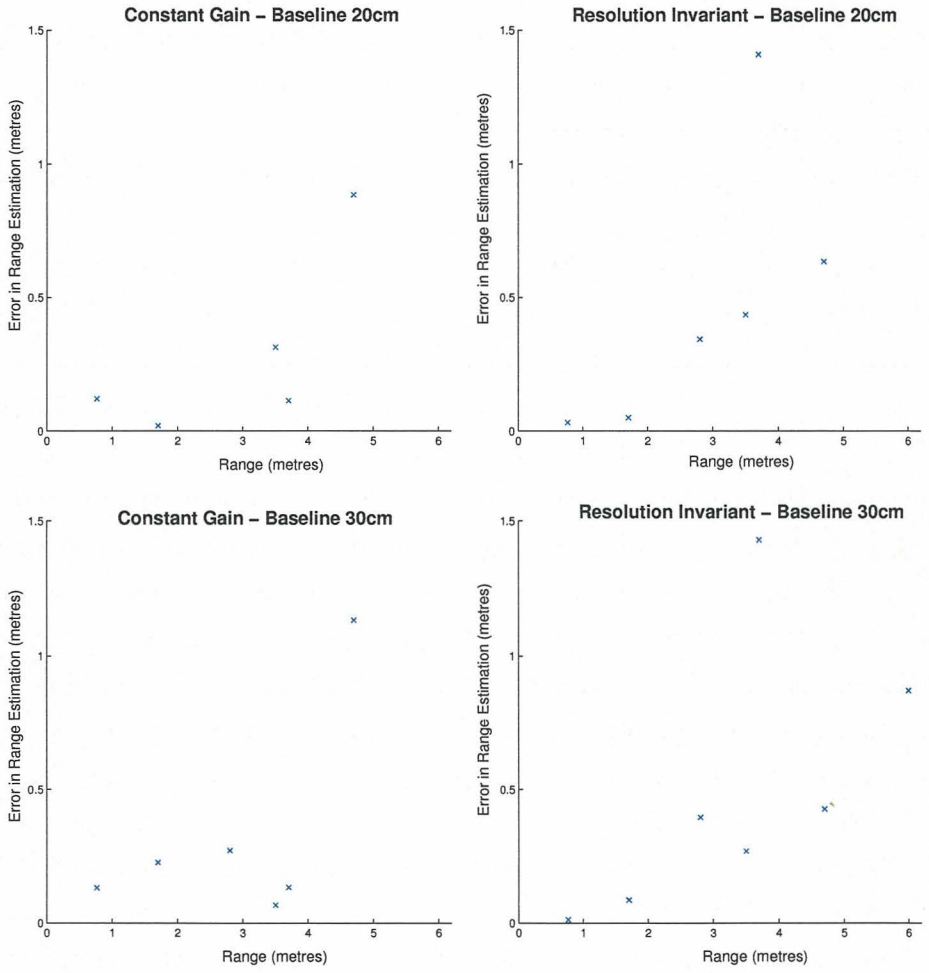


Figure 5.28: Error in range estimation for 20cm and 30cm baselines, using the v-disparity algorithm.

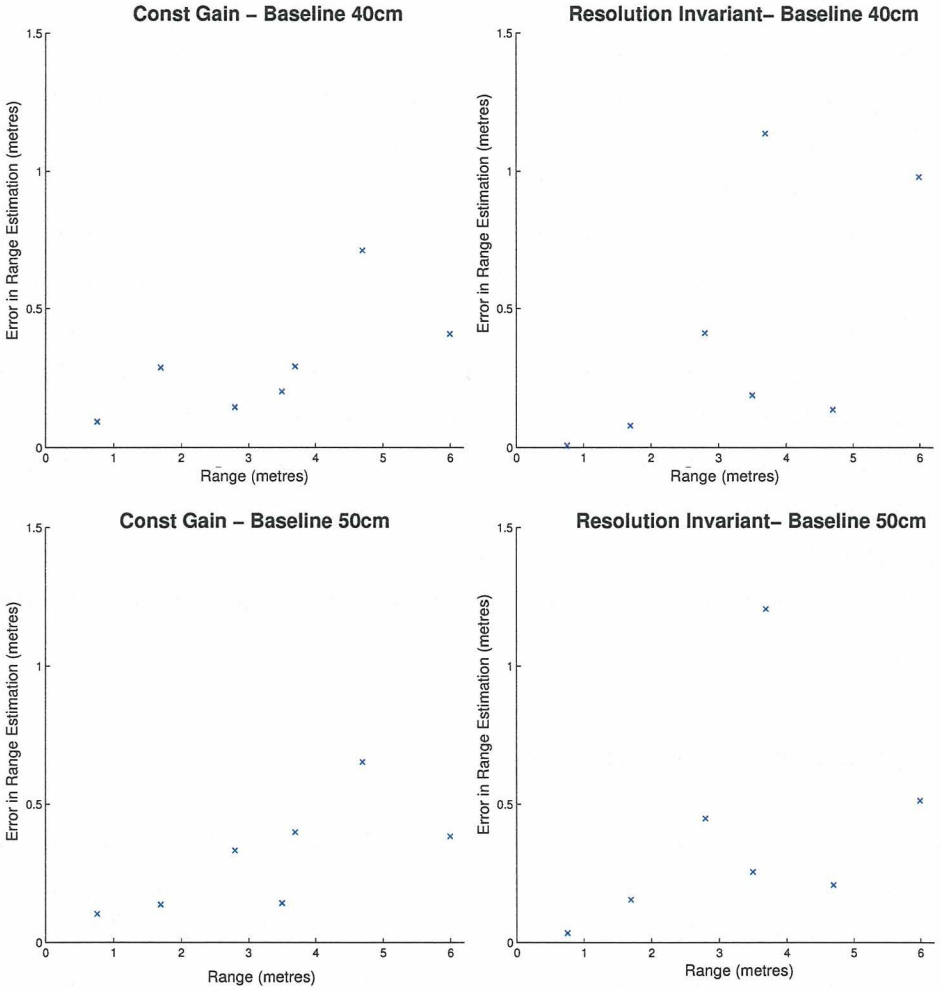


Figure 5.29: Error in range estimation for 40cm and 50cm baselines, using the v-disparity algorithm.

application. A stereo panoramic sensor that utilises constant gain mirrors has a higher accuracy in estimating the range of surrounding obstacles than a system using resolution invariant mirrors.

While larger baselines provide a higher degree of accuracy at further ranges, this causes an increase in the dead zone. Larger baselines also increase the overall size of the sensor, which would make it more difficult to integrate it into a vehicle than a more compact sensor. Therefore, we have chosen to use a sensor with a baseline between 30 and 40cm, which will provide an adequate degree of accuracy for our work volume and a reasonably sized stereo panoramic system.

The horizontal configuration enables the sensor to view the area underneath and directly above the rear of the vehicle. However, this results in a decrease in resolution for the area around the ground plane, where most obstacles will apear. It was therefore decided that a vertical configuration should be used because it has the highest resolution for viewing the desired working area around the sensor. The sensor could be improved to view more of the area directly behind the vehicle by mounting the sensor as low as possible. If the size of the camera were decreased, this would also allow the mirrors to view a lower minimum elevation angle. To further support this decision, after examining the *v-disparity* images for both configurations, it became apparent that ground plane extraction would be more robust when the sensor is mounted vertically.

5.6 Chapter Summary

This chapter presented the results from the theoretical analysis and simulation experiments, with the aim of evaluating the prototype stereo panoramic vision system and selecting suitable sensor characteristics for our application. A theoretical analysis was completed to determine the resolution of the sensor for estimation of range and bearing. Graphs of the sensor resolution in two dimensions were calculated, to investigate the discretisation of space by the sensor. This showed that any point around the sensor will be mapped to a point on the resolution graph, which may result in a slight estimation error. It was found that the constant gain mirror suffered from a lower discretisation error than a panoramic sensor which utilised a resolution invariant mirror. The resolution profiles also revealed the presence of a *dead zone* in which range could not be estimated, due to limitations on the maximum disparity search. This is because any object appearing in this area would have a disparity greater than this value. It should be noted that this phenomenon is found

in most range sensors, due to the time taken to switch from transmitting a ranging signal, to recieving.

A ground truth analysis was performed by generating artificial panoramic images using the POV-Ray software package. This was used to place objects in a virtual environment, at known locations. These images were used to evaluate the range accuracy of the prototype system, and compare various sensor characteristics. It was found that the system was able to estimate the range to an object with an error of 0.2m up to a range of just under 4m, for both mirror profiles, with a baseline of 30cm. This meets the specifications set in Chapter 1 for maximum error. The simulated range estimation performance was found to be comparable or better than the results presented in previous work.

Two obstacle detection methods were evaluated using the ground truth data. The first was *ground plane subtraction.* This algorithm was able to successfully remove the ground plane from the disparity maps, leaving behind the objects of interest. However, it was noted that the assumption of a known ground plane position and orientation would decrease the robustness of such an approach in a real world situation. This is discussed further in the following chapter. The second obstacle detection method tested was the *v-disparity* algorithm. This was also able to successfully detect regions in an image containing objects. Unlike ground plane subraction, this algorithm does not rely on *a priori* knowledge of the ground plane location with respect to the range sensor.

Given the results presented in this chapter, it was decided that the sensor should have the following characteristics:

- constant gain mirrors
- 30-40cm baseline
- vertical configuration

Now that a theoretical evaluation has been completed, we have an indication of the performance of the stereo panoramic vision system. However, the true performance of the system can only be revealed by carrying out real world experiments, which the majority of surveyed work has failed to do.

Chapter 6

Field Experiments

CHAPTER 5 presented the results from a theoretical analysis of the stereo panoramic vision system. The resolution of the sensor was investigated, and ground truth testing was carried out using ray traced images. While this evaluation was invaluable for determining the theoretical limitations of our system, the performance of the sensor under real-world conditions remains to be seen.

One of our main criticisms of previously published work discussed in Chapter 2 was the distinct lack of non-theorectical data. The majority of results presented in the literature have been either purely theoretical, or merely preliminary range estimation (i.e. raw disparity maps) with no evaluation of the accuracy. Only one system (Sogo & Ishiguro 2000) was actually tested in realistic experiments to determine the precision of range estimation. However, the direct implications of these results is not meaningful for our application, as will be discussed further.

This chapter contains one of the main contributions of this body of work - the results of field experiments completed in realistic environments. The range estimation accuracy of our panoramic vision sensor was evaluated, initially without the use of obstacle detection algorithms. This was done using an approach similar to that outlined in Chapter 5. The two obstacle detection algorithms - *ground plane subtraction* and *v-disparity* - were then introduced, and tested in a variety of real-world scenarios. The experiments were completed using different mirror profiles, baselines, and sensor configurations, and the results from this comparison of sensor characteristics are discussed. Finally, the results from the field experiments are reviewed in relation to the work published by (Sogo & Ishiguro 2000).

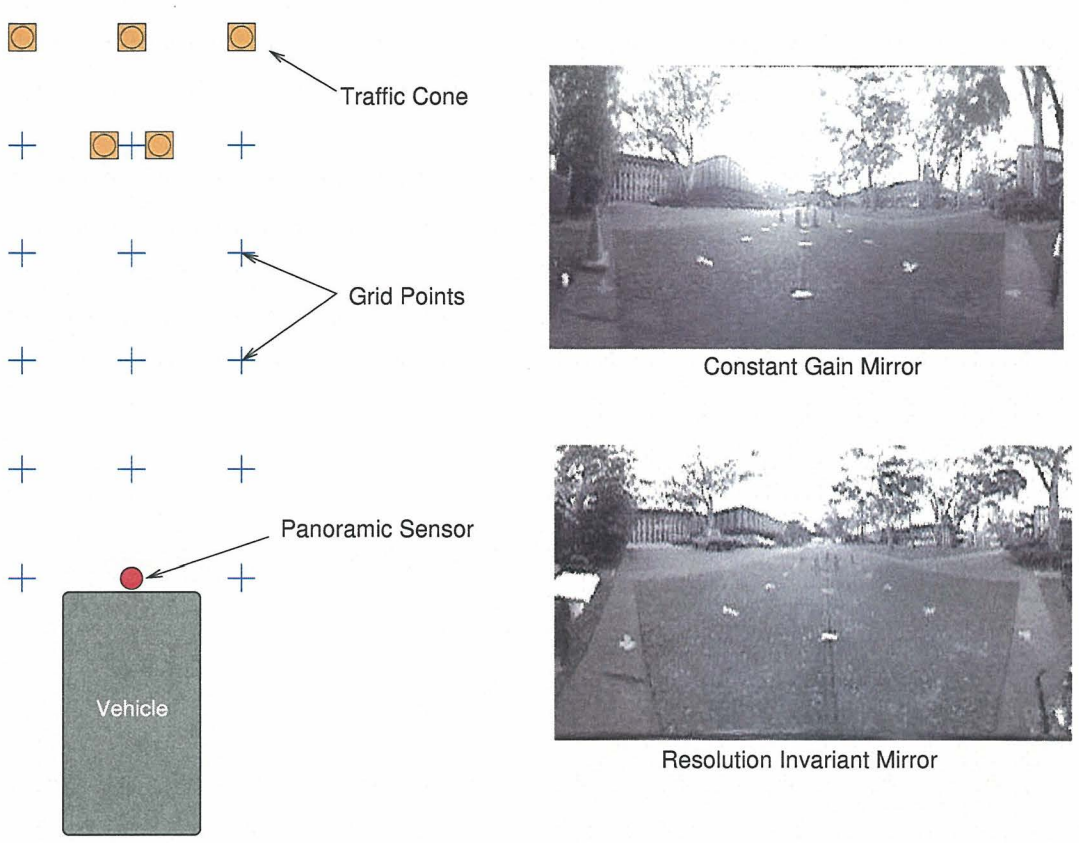


Figure 6.1: A two by five metre grid was used to determine the range accuracy of the stereo panoramic vision system. Left: A top view of the grid. Top right: Image of the grid, captured using a constant gain panoramic sensor. Bottom right: Image of the grid, captured using a resolution invariant mirror.

6.1 Estimation of Object Position

In the previous chapter, we determined the theoretical limits of the range finding capabilities for our system. These ground truth experiments were carried out under ideal conditions. The cameras and mirrors in the virtual stereo sensor were prefectly aligned, and objects could be placed at exact locations in the environment generated by the POV-Ray software. In reality, such perfect situations can be difficult, if not impossible to achieve. Therefore, it is vitally important to perform testing in real-world conditions to gauge the true performance of the system.

The first series of field experiments were carried out without the use of the obstacle detection algorithms. A two by five metre grid of one metre intervals was marked out on the ground, as shown in Figure 6.1. The physical size of the cameras and mirrors prevented a baseline of less than 30cm, while a baseline of greater than 40cm would create a dead zone larger than 60cm (see Figure 5.15). Therefore, images were captured with the sensor using the following parameters:

- baselines of 30cm, 35cm and 40cm
- constant gain and resolution invariant mirror profiles
- vertical configuration (as suggested in Section 5.5)
- Unwarped image dimensions of 500 by 200 pixels

Examples of the unwarped images captured for these experiments are displayed on the right-hand side of Figure 6.1. The complete set of images can be found on the attached CD-ROM. As in Section 5.3, corresponding features (i.e. points on the grid) were selected manually, and the three dimensional position was calculated from these disparities. Each feature was selected ten times to provide a general indication of the level of human error involved in the manual feature selection.

The results for range, height and angular estimation for each feature were analysed and graphed as shown in Figures 6.2, 6.3, and 6.4 respectively. The average error for each feature is plotted the top graphs, while the standard deviations in the error in measurement are displayed below. Each figure contains results for both the resolution invariant and constant gain mirror profiles, as well as results for each of the tested baselines. The results will be discussed in the following sections.

6.1.1 Range Estimation

In Figure 6.2 it can be seen that the error in range measurements was generally larger than for the ground truth analysis (Figure 5.17). In the ground truth experiments, the results were effected by both human error and discretisation issues, as discussed in Section 5.3. However, in the field experiments there are other contributing factors such as the possiblility of slight camera-mirror misalignments, and errors in the placement of the grid.

Both mirror profiles generally estimate the range to objects to within 40cm, as shown in Figure 6.2. However, the constant gain mirrors tend to provide a more accurate range estimation, particularly at larger ranges.

As the sensor baseline increases, the standard deviation in range estimation drops for both mirror profiles. This concurs with the results from the ground truth experiments, and reflects the fact that the discretisation errors decrease as baseline increases.

6.1.2 Height Estimation

There was a significant difference in the error in height estimation for each of the mirror profiles using in the field experiments. Given the emphasis on pixels representing objects closer to the sensor, the resolution invariant mirror provides a much lower degree of accuracy when estimating height at larger ranges. The sensor utilising the constant gain mirror estimates height to within 30cm to ranges greater than 5m, with a baseline greater than 30cm. However, the errror for the resolution invariant profile has a lower accuracy at higher ranges. This agrees with the result for the ground truth data, although the effects of the variation in mirror profile seem more apparent in the field experiments. This is probably due to the nature of the features used. In the artificial images, the corners of objects appeared to be fairly crisp in the resolution invariant panoramic images. On the other hand, the grid was slightly more difficult for the observer to select accurately. The larger the range to a point on the grid, the larger the marker required to enable the point to be seen in the panoramic images. This problem was exacerbated by the comparitively lower resolution provided by the these mirrors. The standard deviation in height estimation was also significantly lower for the constant gain mirrors, in comparison to the resolution invariant profiles.

Once again, we can observe that a larger baseline results in a higher degree of accuracy in height estimation.

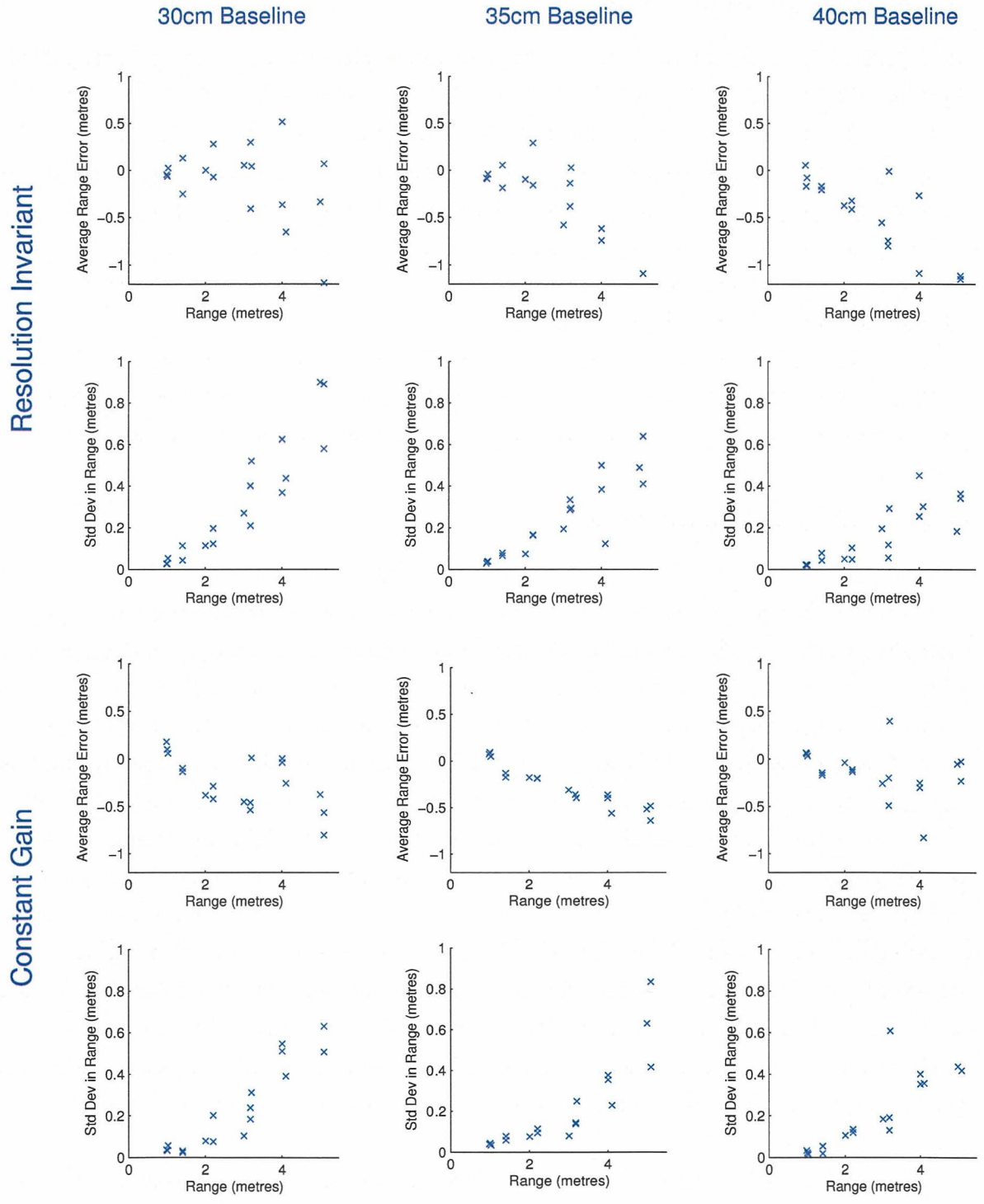


Figure 6.2: The range estimation accuracy and corresponding standard deviations calculated from the field experiments.

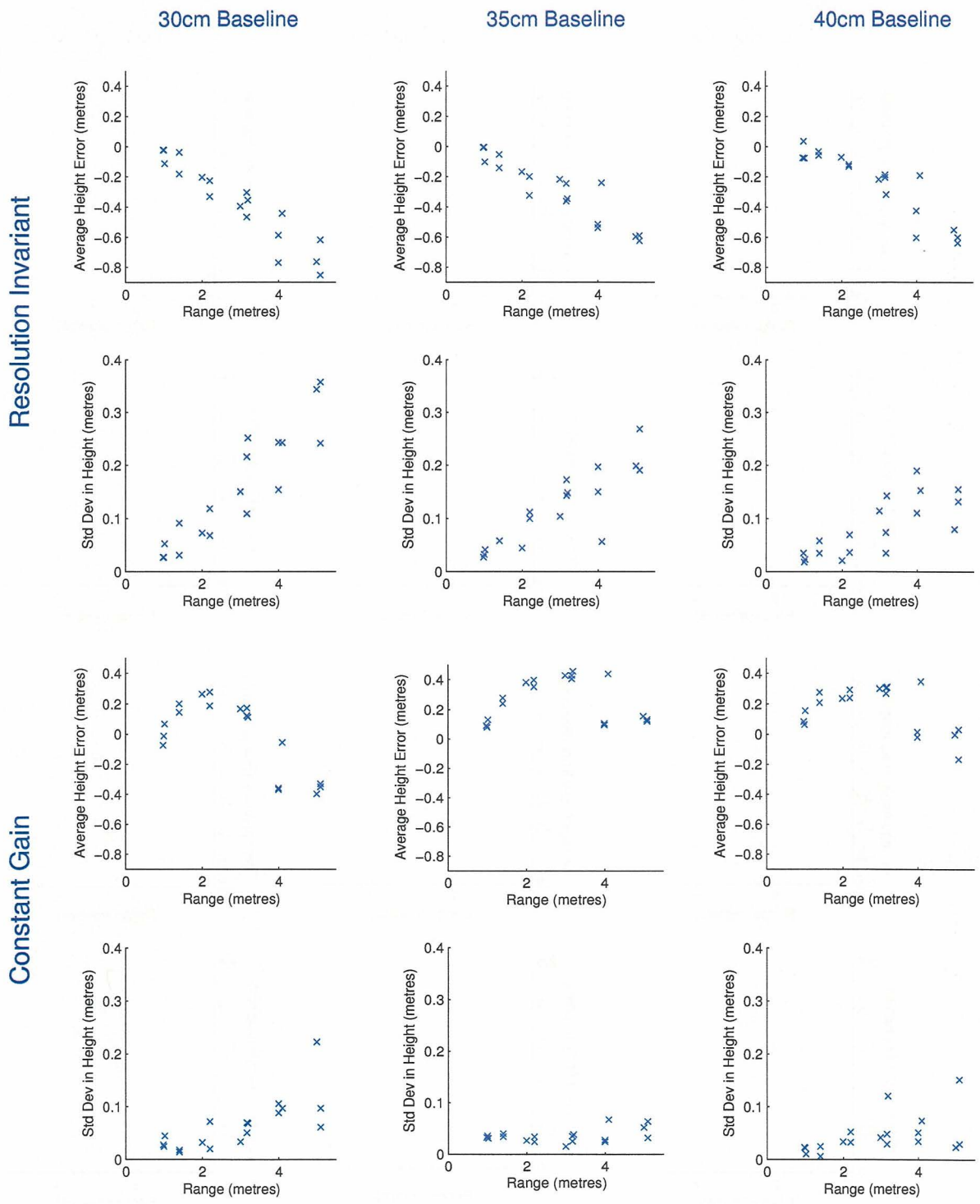


Figure 6.3: The average height estimation error and height error standard deviation calculated from the field experiments.

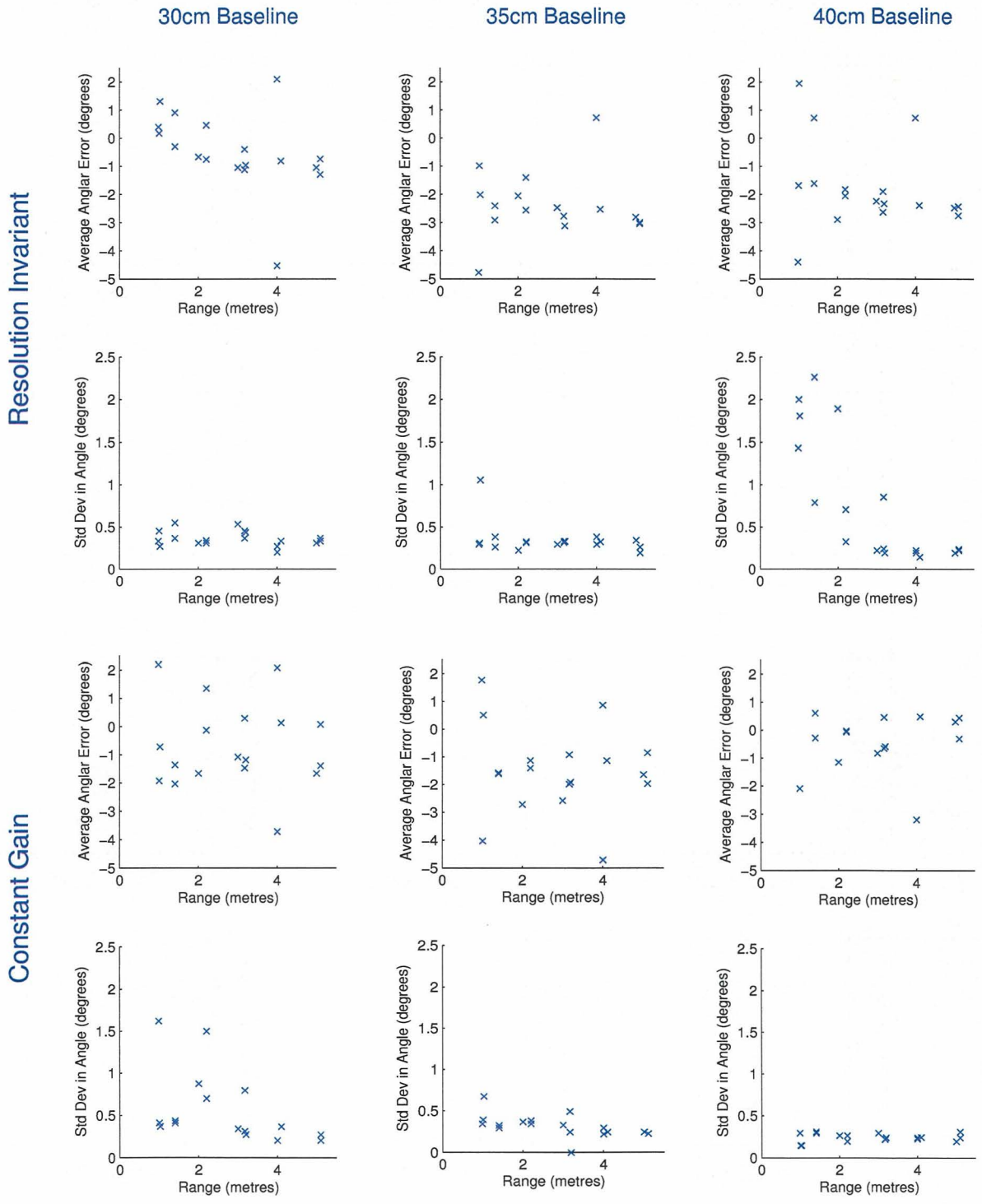


Figure 6.4: The average angle estimation error and standard deviation calculated from the field experiments.

6.1.3 Angular Position Estimation

As in the ground truth experiments, the estimation of the azimuth angular position was quite accurate. The estimated value was generally within 3° of the true value for both mirror profiles. The standard deviation in error was particularly impressive, and remain below 1° for the majority of the collected data. The high level of precision can be attributed to the fact that there was over one degree per pixel in the unwarped panoramic image, which had a width of 500 pixels.

However, in several graphs the standard deviations rise above this value at ranges below 2.5 metres which does not agree with the trend that errors should increase with range. It should be noted that features take up larger number of pixels when they are closer to the sensors. Therefore, there is more likely to be a variation in the point selected by the human observer.

6.1.4 Comparison with Previous Results

As presented in Section 2.5, (Sogo & Ishiguro 2000) used a system that consisted of four panoramic sensors placed around a room at equal height, with the camera axes parallel to one another. The sensors were placed in a rectangular formation with a width and breadth of 1.5m and 2.5m respectively. Obstacles were segmented using the background subtraction method. The aximuth angle of the detected obstacles from each sensor were combined to determine their location. The authors report that they were able to determine range to within 5cm for targets within 3m. However, there were a number of factors that contributed to the higher degree of accuracy that would be inappropriate for our application:

- Multiple sensors - This system requires at least three sensors to estimate range. This would make the system at least 50% more expensive, and there would be a higher degree of difficulty in aligning the extra sensor(s).

- Static background assumption - The obstacle detection is carried out using background subtraction. Since we require the sensor to be mounted to a moving vehicle, this assumption is no longer valid.

- Large baselines - The sensors are mounted at least 1.5m apart, which helps to improve the range estimation accuracy. However, if we were to use a stereo system with a baseline of this magnitude, it would no longer be possible to utilise stereo matching techniques given the large differences in camera view points.

Another disadvantage of the system presented by (Sogo & Ishiguro 2000) is the potential for the detection of false obstacles. Since objects are detected using a set of computed azimuth angles, there are issues with associating the correct subset of angles with a particular object. If there are many obstacles in the environment, then there could be several ways of interpreting the azimuth angles computed from the panoramic sensors. The result could be the detection of a false obstacle, as shown in Figure 2.17. Therefore, this method of obstacle detection would only be reliable when the total number of obstacles was low.

6.2 Obstacle Detection

In the previous section, we presented the results of tests to determine the accuracy of range estimation for the stereo panoramic sensor. These tests were completed using a physical sensor to capture real-world images, as opposed to the ray traced images shown in Chapter 5. However, the final system cannot rely on a human observer to identify objects, and manually select features to calculate range. Instead, one of two obstacle detection methods were employed to segment regions of interest from the disparity maps - *ground plane subtraction* or *v-disparity.* This section discusses the performance of both obstacle detection methods in a realistic environment.

6.2.1 Ground Plane Subtraction

The first step in the obstacle detection algorithm known as Ground Plane Subtraction (GP subtraction), is the creation of a model of the surounding environment without obstacles. This model relies on the assumption that the exact location and orientation of the ground plane is known, relative to the stereo sensor. A theoretical disparity map is generated, which is then subtracted from the disparity maps calculated at run-time to remove any non-obstacle pixels.

In the previous chapter, the GP subtraction algorithm was applied to ray traced images, and it was found to successfully segment obstacles from an environment.The relationship between the ground plane and the cameras was known precisely, as the positions of all objects were specified using the POV-Ray software. The GP subtraction algorithm was then used to locate obstacles with a physical sensor. As shown in Figure 6.5, if the position and orientation of the ground plane has been accurately calibrated, the GP subtraction algorithm will correctly extract obstacles from the surroundings. In this case, the sensor

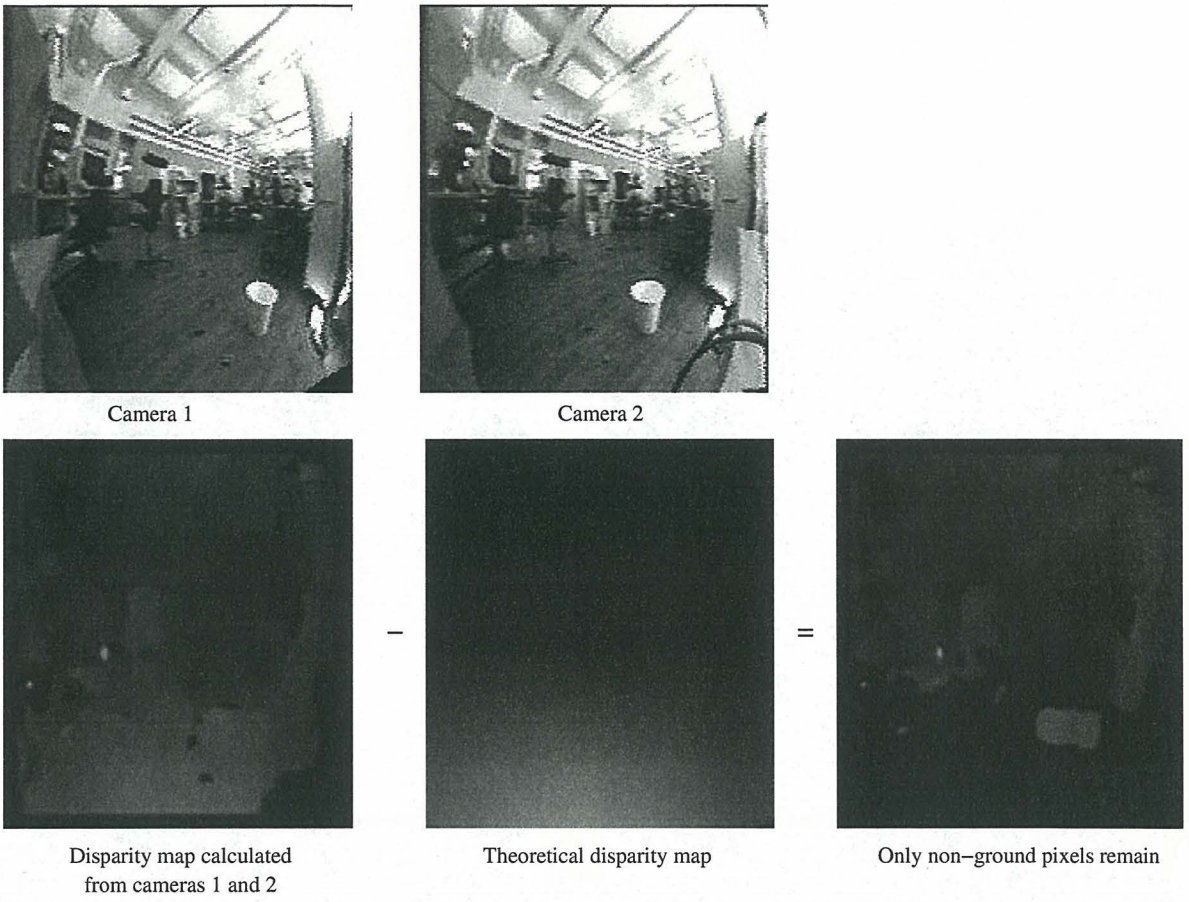


Figure 6.5: The ground plane subtraction algorithm applied to an office scene. The ground pixels have been removed from the final disparity map, leaving behind the obstacles. Only the office furniture, and a rubbish bin in the foreground remain.

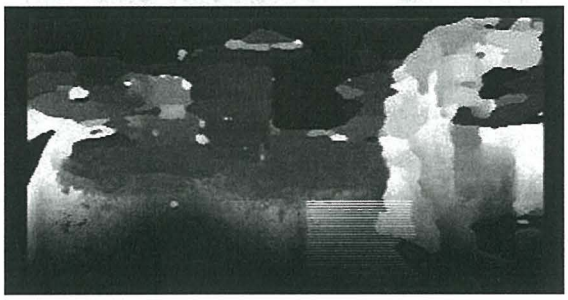
was placed in an office environment, and has removed the ground pixels from the disprity map. Only the office furniture, and a rubbish bin in the foreground remain.

On the other hand, if the ground plane moves from the calibrated position, the GP subtraction algorithm may fail to remove the correct pixels. In our application, with the panoramic sensor mounted to a vehicle, this is a common occurance. The slope of the ground relative to the sensor may change due to a bump in the road, or a change of incline of the ground plane relative to the vehicle. In such situations, pixels on the ground plane may remain in certain areas of the output disparity map, while objects might be erroded, or removed completely. An example of this can be seen in Figure 6.6. In this example the ground plane is removed as shown in (e), leaving the obstacles behind along with a small amount of noise. However, when the sensor is moved from the calibrated position a large

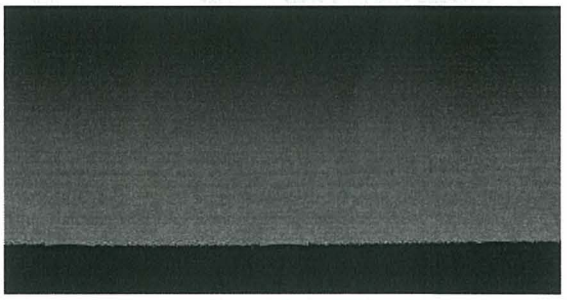
(a) Camera 1 – Calibrated position

(b) Camera 1 – Uncalibrated position

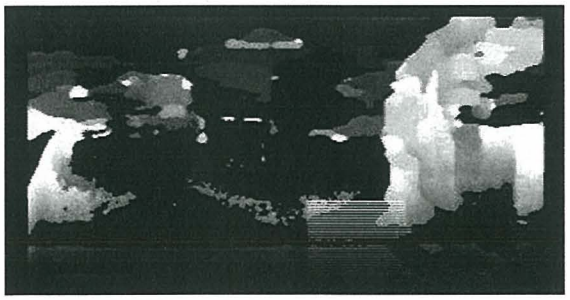
(c) Disparity map – Calibrated position



(d) Precalculated theoretical disparity map



(e) Ground plane subtracted – Calibrated position



(f) Ground plane subtracte – Uncalibrated position

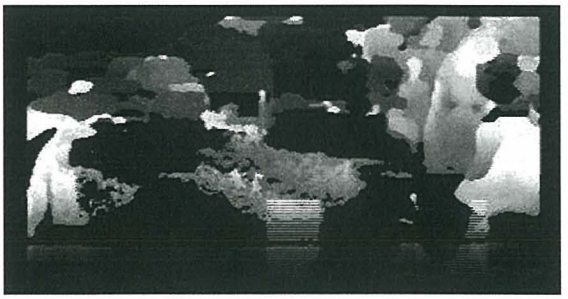


Figure 6.6: If the orientation of the sensor is changed from the calibrated position, the ground plane subtraction will no longer correctly segment obstacles from the environment.

portion of the ground plane remains, while the lower sections of the obstacles have been removed. In our application, the panoramic sensor will be mounted to a moving vehicle, and thus the likelyhood of a change in relative orientation between the ground and the sensor is fairly high. Either the GP subtraction algorithm must be made more robust, or an alternative method of obstacle segmentation employed.

One way to increase the robustness of the GP subtraction algorithm would be to recalibrate for the ground plane in each image. For example, (Molton et al. 1997) suggested a method to dynamically recalibrate for a plane in each image by tracking features on the ground. However, it is still necessary to generate a theoretically disparity map containing the detected plane for each new frame to be processed. This entire process has a much higher computational expense than the initially suggested method, since previously the theoretical disparity map could be calculated off line. Thus it is clear that a less intensive, more robust approach to obstacle detection is required for our real-time application.

Another issue is the high level of noise present in the disparity maps. This is caused by the combination of low resolution panoramic images (in comparison to conventional stereo images), and roughly calibrated panoramic cameras. One way to improve the quality of the disparity maps is to rectify the stereo image pair to further align the epipolar lines. Or, an alternative is to utilise an obstacle segmentation algorithm that is capable of dealing with noisy disparity maps, such as the v-disparity algorithm.

6.2.2 V-disparity

As discussed in the previous section, the Ground Plane Subtraction algorithm relies on an assumption that the orientation of the ground plane with respect to the sensor is know, and will not change. If the relative position changes, the ground plane cannot be successfully removed, and certain objects may also be eroded from the disparity map. There are also problems if there is any noise present in the disparity map. On the other hand, the v-disparity is able to overcome these problems.

The disparity maps produced in the real-world experiments are noisy in comparison to those generated with ground truth data. Despite this, the v-disparity algorithm is able to produce results of a sufficient quality to successfully segment obstacles. The system was not always able to place a bounding box around the entire object, and instead the obstacle was detected in smaller sections. However, the field experiments have shown that the system is able to reliably segment regions of the environment to be avoided. The results of the field experiments will be presented and discussed in Section 6.3.

An example of a noisy disparity map produced in the field experiments is shown in Figure 6.7. The ground plane can be computed, as shown in the v-disparity by the red line. The background is shown in green. Any obstacles will be represented by vertical lines above the ground plane, as shown in yellow. However, since the real-world depth maps are noisier than those computed from the ray traced images, the u-disparity needs to be filtered further to determine the width of the obstacle.

Once an obstacle has been detected in the v-disparity image we now know the depth to the object. We then remove any pixels from u-disparity image that are not at this depth, to determine the width of the object. The filtered u-disparity corresponding to Figure 6.7 is shown in Figure 6.8. The result of the obstacle detection is also shown.

The system currently only runs at approximately 5 Hz, as this is the first software implementation of this algorithm. With improved software design and optimisation, the execution speed would increase significantly. However, optimising the software is not one of the goals of this research project.

6.2.3 Temporal Consistency

Due to the high noise ratio present in the real-world panoramic disparity images, false detection of obstacles became apparent in the image sequences. These generally only occurred in single frames, and as a result many of the false detections were easily filtered out by checking for temporal consistency, as discussed in Section 4.5. The system was modified to only report an object once it had been detected in several consecutive frames, and continued to track the object until it had been lost in the same number of consecutive frames. However, this with solution there is a trade off between the removal of brief false detections, and the responsiveness of the driver assistance system. That is, as we increase the number of frames required before an obstacle is considered to be genuine, the greater the lag in recognising the presence of an actual object. As the current system runs at approximately 5 Hz, it was found that the an object should be accepted once it had been detected in two consecutive frames. Otherwise the delay in the system would be too large. This problem could be improved by increasing the execution speed of the software, either by optimisation or through the use of a faster processor.

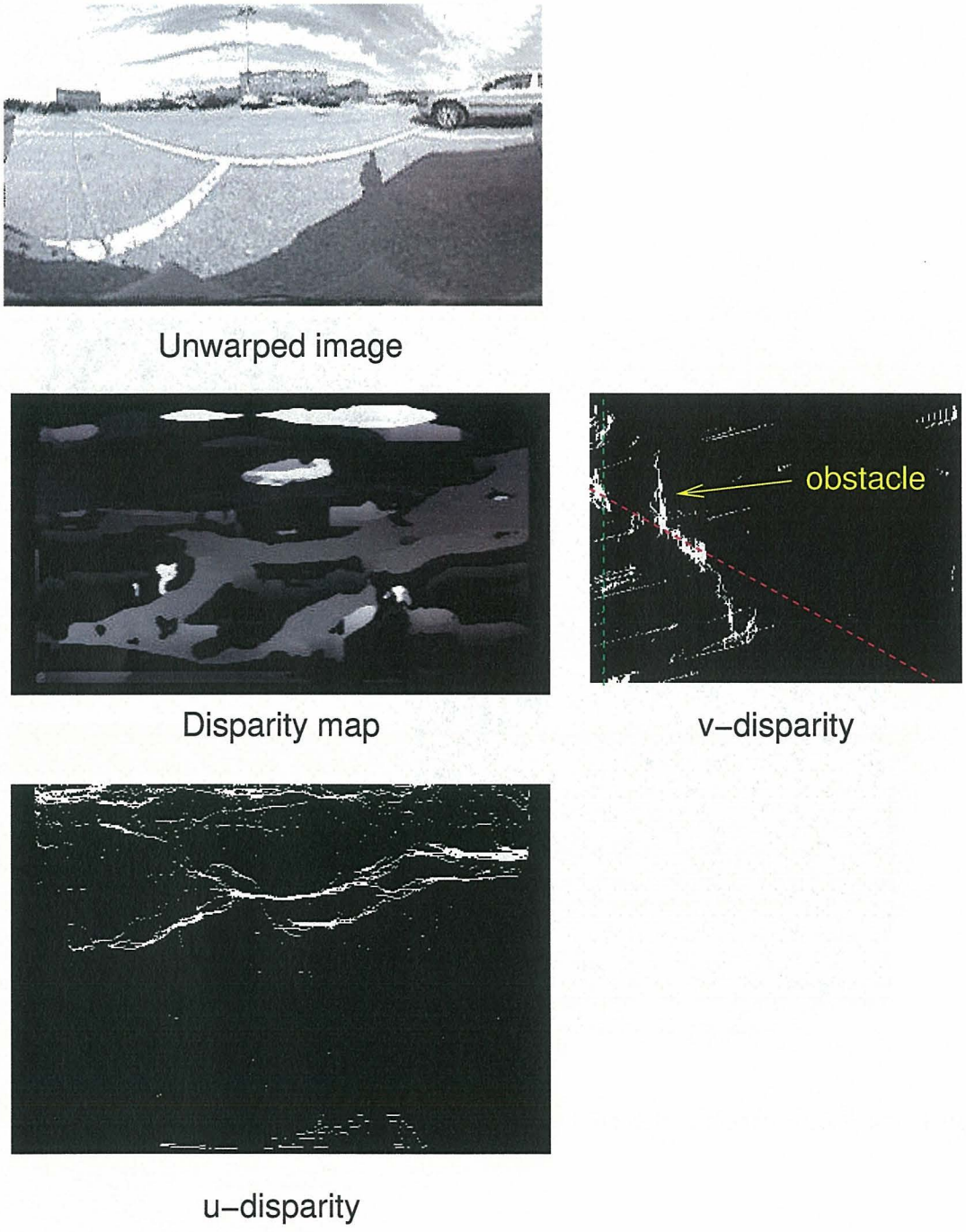


Figure 6.7: An example of a noisy disparity map and the corresponding v- and u-disparities. From the v-disparity map we can determine the ground plane (in red), the background (in green) and the obstacle. However, the u-disparity map needs to be filtered to determine the width of the object.

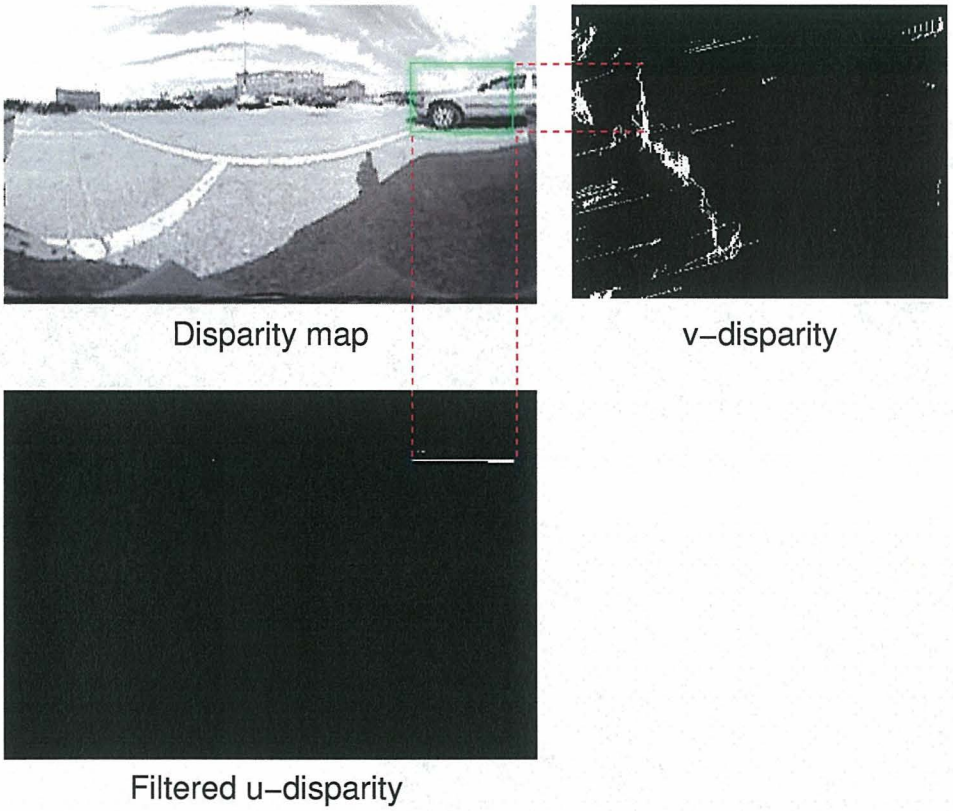


Figure 6.8: The obstacle is detected using the v-disparity and filtered u-disparity images.

6.3 Real World Scenarios

Several sequences were taken of different real-world situations to test the performance of the stereo panoramic vision obstacle detection system. The following is a list of the sequences that were recorded and later analysed:

- **Parking Scenario 1** - Reversing into a parking space, surrounded by other vehicles to the left, right and back of the space.
- **Parking Scenario 2** - Reversing into a parking space with a vehicle located at the rear end of the area
- **Parking Scenario 3 -** Reversing towards a pole surrounded by concrete barriers
- **Reversing into Shed -** Reversing into a shed, passing by several vehicles. This scenario shows the system's ability to detect overhanging objects.

Given the results discussed in Section 6.2, only the v-disparity algorithm was used to segment obstacles from the sequences, due to it's superior performance. Each sequence was analysed frame by frame to visually inspect and record various statistics to quantitatively measure how effectively the system segments objects. The following rates were calculated to provide an indication of the performance of the system:

$$False\ detection\ rate = \frac{\phi}{T} \tag{6.1}$$

where ϕ is the number of false detections and T is the total number of obstacles present in the sequence. A false detection was recorded when the system displayed a rectangle (i.e. a detected object) over an area that did not contain an obstacle.

$$Detection\ rate = \frac{D - \phi}{D - \phi + N} \tag{6.2}$$

where D is the total number of detections and N is the number of objects not detected. An obstacle was considered to be detected when detected in its entirety or in segments which covered the entire object. If no part of the object was detected, or an insignificant portion was displayed, the obstacle was recorded as not detected.

$$Detection\ failure\ rate = \frac{N}{D - \phi + N} \tag{6.3}$$

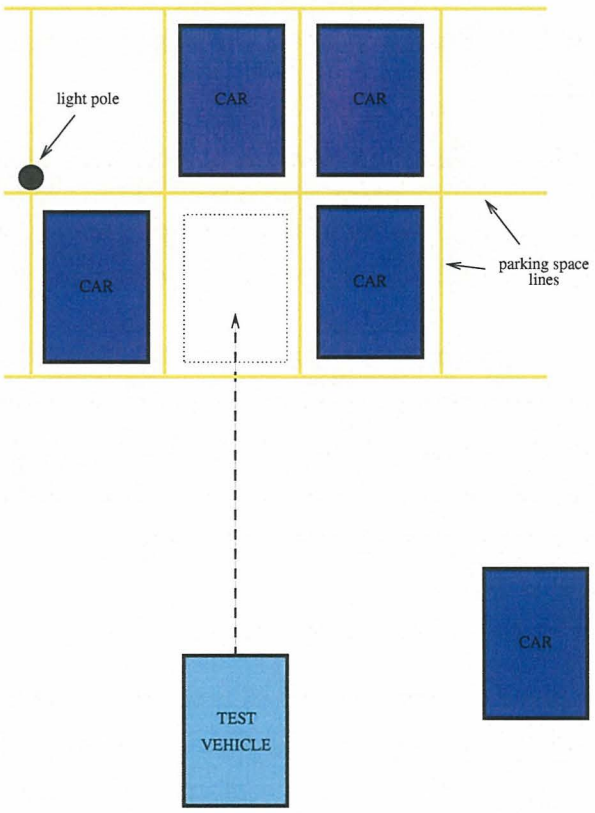


Figure 6.9: Top view diagram of *Parking Scenario 1.*

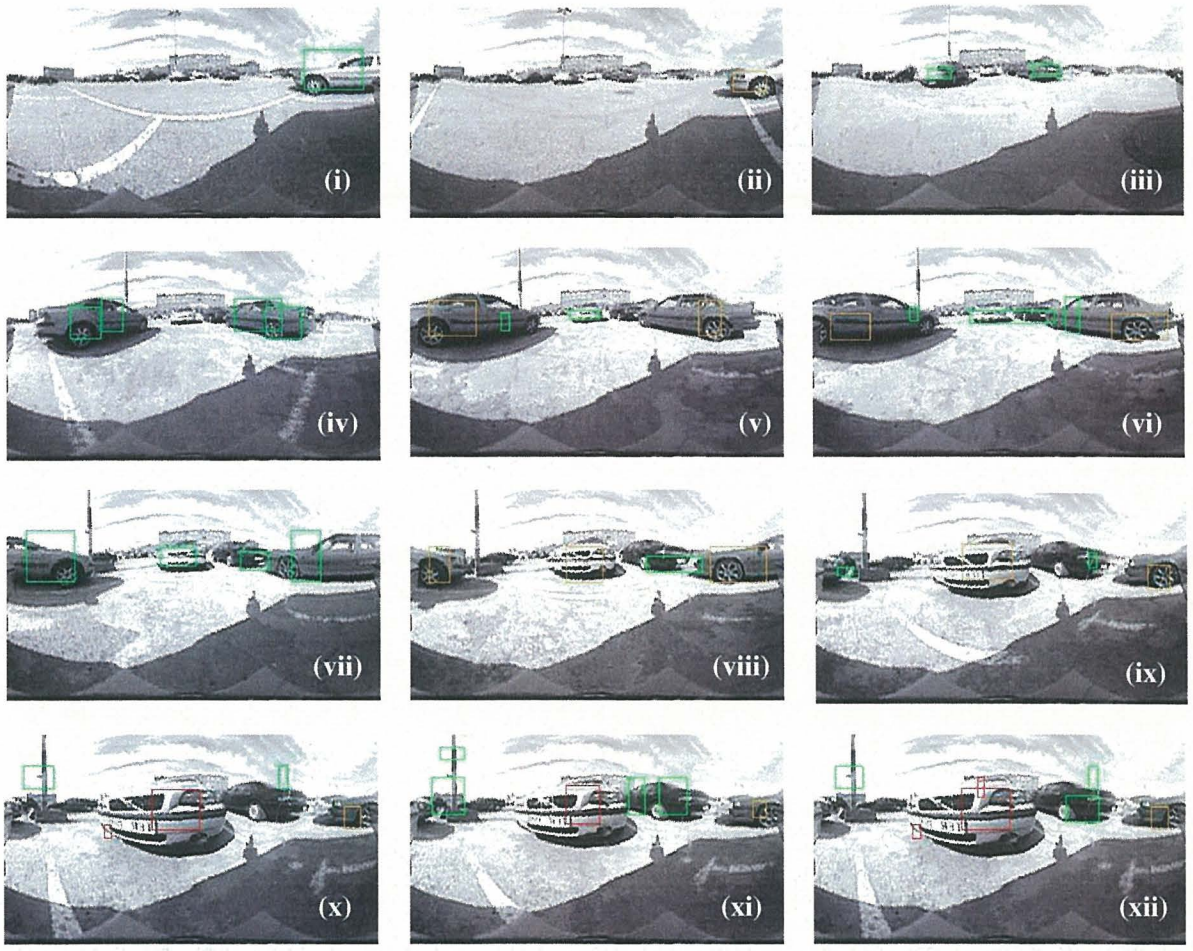


Figure 6.10: Sample frames from *Parking Scenario 1*.

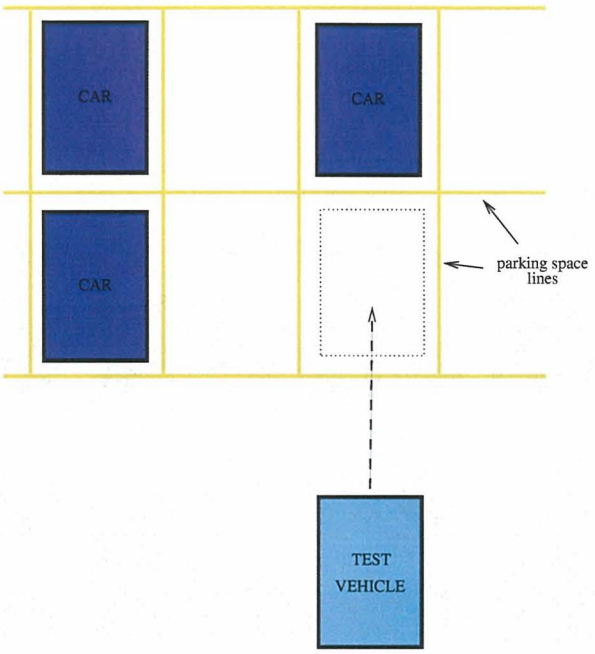


Figure 6.11: Top view diagram of *Parking Scenario 2.*

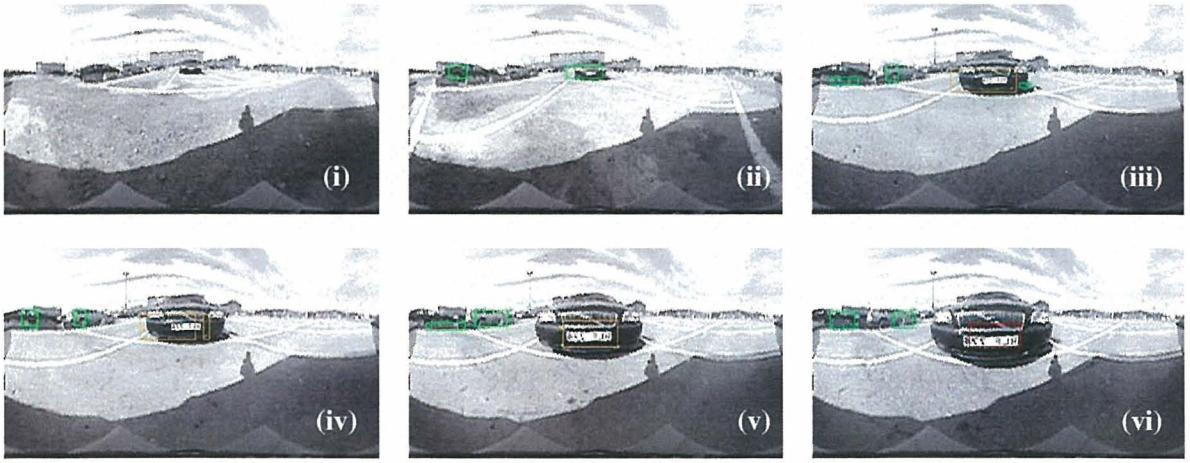


Figure 6.12: Sample frames from *Parking Scenario 2.*

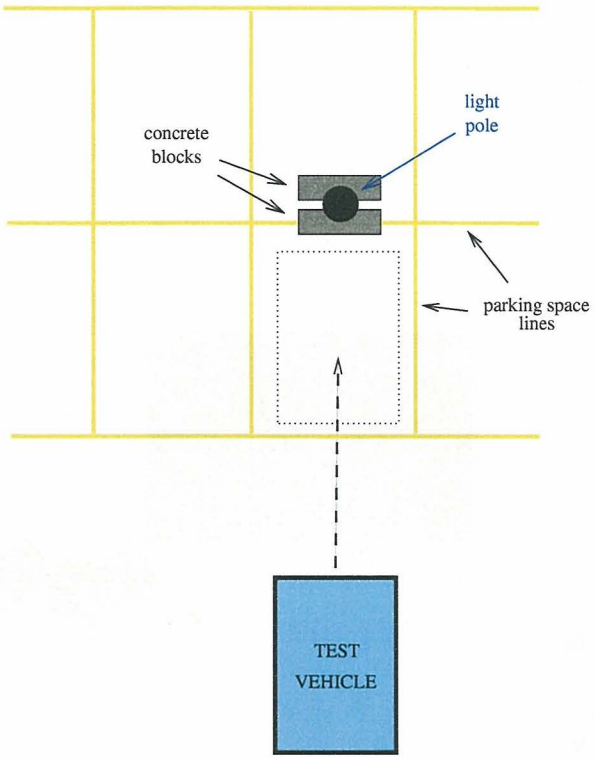


Figure 6.13: Top view diagram of *Parking Scenario 3*.

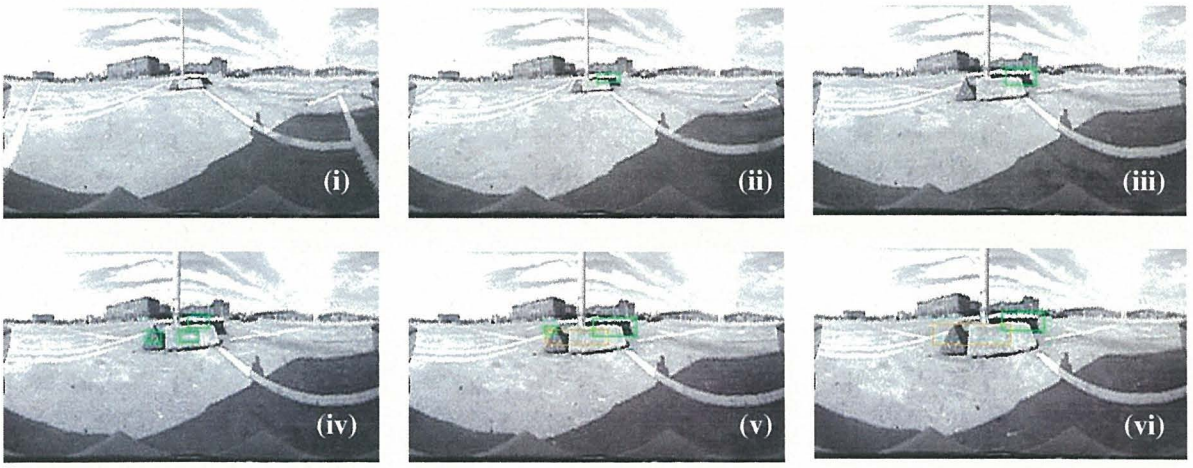


Figure 6.14: Sample frames from *Parking Scenario* 3.

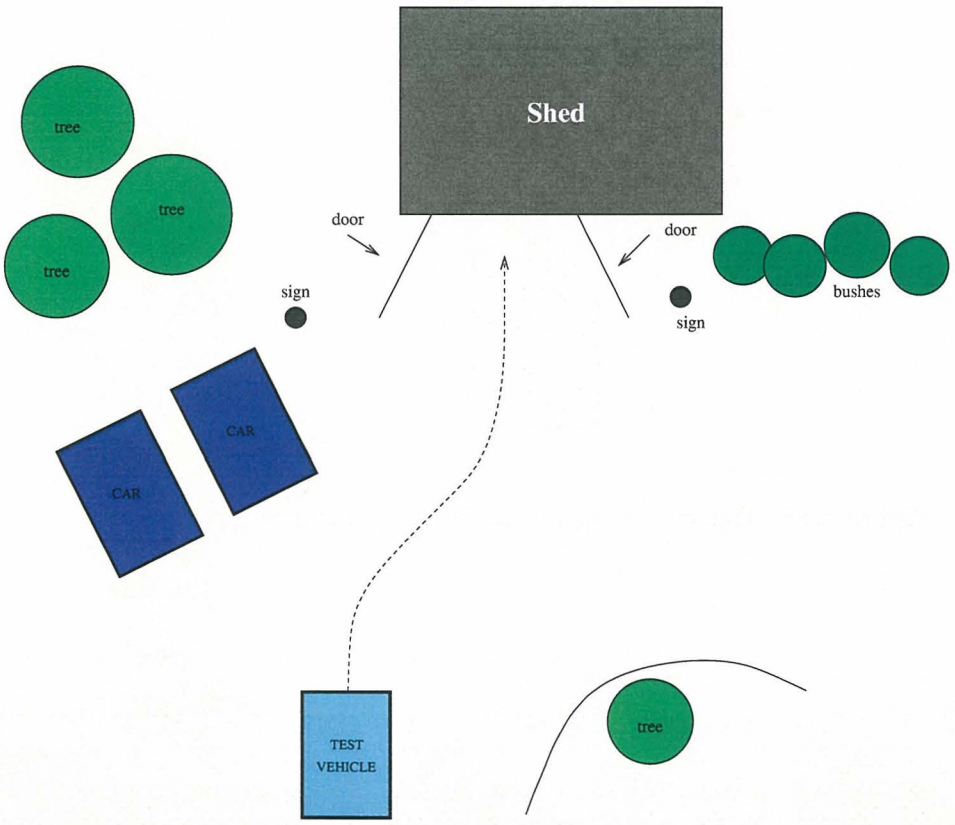


Figure 6.15: Top view diagram of the *Reversing towards a shed* sequence.

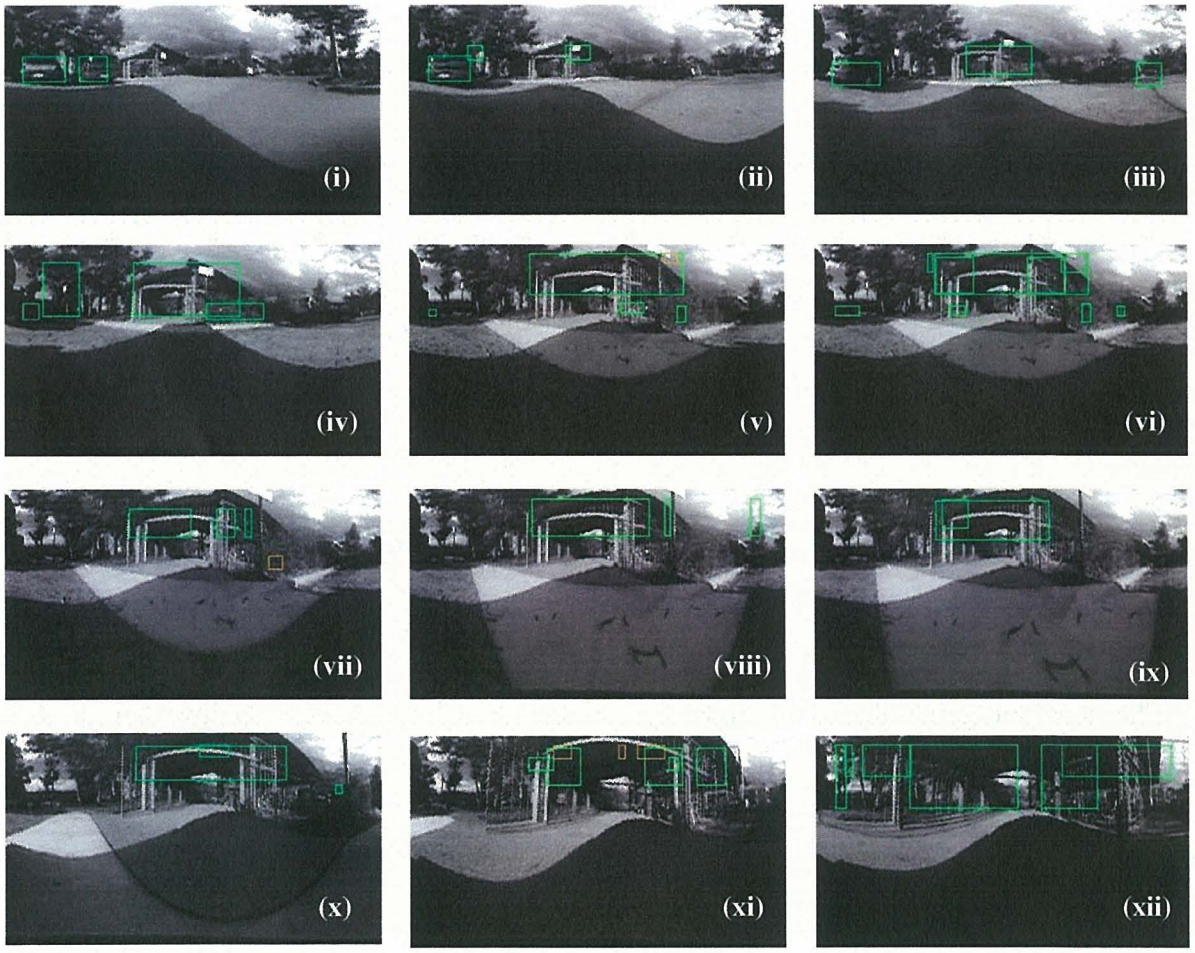


Figure 6.16: Sample frames from the *Reversing towards a shed* sequence.

Scenario	Frames	Objects	False detections	Detection rate	Detection failure
Parking 1	404	1102	1.4%	91.6%	8.4%
Parking 2	319	404	7.7%	94.2%	5.8%
Parking 3	246	92	2.2%	92.8%	7.2%
Into a shed	475	653	19.9%	89.2%	10.8%
Average			7.7%	91.9%	8.1%

Table 6.1: Summary of statistics compiled from the various scenarios

The results of this quantitative analysis have been summarised in Table 6.1. The detection rates for each sequence are fairly similar, with an average detection rate of 91.9%. The false detections were low, with an average of 7.7%. The *Reversing into a shed* sequence had a comparitively high false detection rate of 19.9% which is more than double the average. This can be largely attributed to the harsh lighting conditions, as the sequence was taken in the late afternoon. As a result, there a many long shadows and the low light decreases level of detail that can be viewed in the images[1].

6.4 Selection of Sensor Characteristics

Given the results from both the ground truth, and field experiments we can select characteristics that will provide a sensor with the most desirable attributes for our application. The constant gain mirrors consistently out performed the resolution invariant profiles. While a baseline of 40cm provides a balance in the trade off between a higher level of obstacle position estimation and the increasing size of the deadzone. The vertical sensor configuration was found to be more appropriate than the horizontal configuration, due to the greater coverage of the required working area.

6.5 Chapter Summary

In this chapter we have presented the results of field experiments which have been designed to test the prototype stereo panoramic vision system under real-world conditions. The first set of experiments evaluated the accuracy of the sensor's position estimation, without the use of automated obstacle segmentation. It was found that the range of an obstacle could be estimated to within 40cm, up to a range of 3m for both constant gain and resolution

[1]Videos of the obstacle detection results can be found on the attached CD in the `/Videos` folder.

invariant mirror profiles. The height estimation was significantly more accurate when using the constant gain mirror, with an error of 30cm up to a range of 5m, as opposed to 30cm at a range of 2m. The resolution invariant mirror provided a slightly higher accuracy when estimating the angular position, remaining within 2° of the true value at 4m, while the constant gain mirror had an error of up to 3° at 4m. However, this is more likely to have been caused by human error.

The next step was to introduce the obstacle segmentation algorithms. The *ground plane subtraction* algorithm was able to detect objects fairly easily, if the system was calibrated correctly. However, if the sensor moved slightly from the pre-calibrated position, sections of ground would remain in the disparity map, while portions of obstacles would be erroded. Thus it became apparent that a more robust segmentation method would be required for our application, in which the panoramic sensor is mounted to an automotive vehicle. It was found that the v-disparity algorithm was able to overcome the problems associated with the *ground plane subtraction* method.

Several image sequences were taken of real-world situations that would be faced by the system under normal operation. The majority of which were various parking scenarios containing cars, pedestrians, a garage and other obstacles. These sequences were analysed frame by frame to evaluate the performance of the automated obstacle segmentation. It was found that the system had an average detection rate of around 90%. The average false detection rate (i.e. obstacles that were detected, but did not actually exist) was around 8%, however, this rose as high as 20% under difficult lighting conditions.

Finally, it was found that the most desirable sensor characteristics are attained by using the following parameters:

- Constant gain mirror profile
- Baseline of 40cm
- Vertical sensor configuration

We have now presented the work completed for this project. The following chapter will provide a summary of the entire body of work.

Chapter 7

Conclusion

Statistics show that many accidents occur each year behind vehicles and in blind-spots that results in property damage, injuries and even loss of life. There are two main methods currently available to assist drivers in avoiding these collisions. The first is the installation of devices to increase the driver's field of view, such as extra mirrors and wide-angle lenses. However, these still rely on an alert human observer and do not completely eliminate blind-spot areas. Another approach is to use an automated system which relies on sensors such as sonar or radar to gather range information. Unfortunately, these systems have low angular resolution and limited sensing volumes. Therefore there was an interest in developing a new method of obstacle detection.

In this project we have designed, built and investigated a stereo panoramic vision system for the purpose of monitoring vehicle blind-spots. It was shown that the sensor can be used to generate disparity maps from which obstacles can be segmented. This was done by applying the v-disparity algorithm, which has previously not been utilised in panoramic image processing. We found that this method was very powerful for segmenting objects, even in the case of extremely noisy data. Our results indicate that range can be estimated reliably using a stereo panoramic sensor, with excellent angular accuracy in the azimuth direction. Furthermore, this sensor has the advantage of a much higher angular resolution and larger sensing volume than the driver assistance systems on the market today.

7.1 Summary

Chapter 1 began with a discusion of the motivations behind the research presented in this thesis. Rear and blind-spot collisions occur regularly due to carelessness or areas around the

vehicle obscured from the driver's view point. The safety systems currently employed all have certain drawbacks that reveal the need for the development of a method for avoiding such accidents. The objectives and specifications for a new system were outlined.

Chapter 2 contained a survey of the work related to this topic. The various methods for avoiding rear and blind-spot collisions were presented, and the cons and pros of each system were summarised. We noted that the use of computer vision had not yet been pursued for this application. However, the majority of camera systems have limited fields of view, and would therefore be unsuitable. The different approaches to increasing the field of view of a camera was then presented, and it was decided that panoramic mirrors would be a good solution for our project. One method for estimating range in computer vision is through the use of a stereo sensor. The previous research on stereo panoramic sensors was presented, and this was used to decide the configuration of the sensor to be evaluated in our project.

Chapter 3 provided an overview of the entire system. The hardware and software components as well as the the system calibrations methods were presented. Although the system being evaluated in this project is a prototype to demonstrate the concepts necessary for a commercial product, and therefore requires futher development. However, it was shown that it would be possible to integrate a stereo panoramic vision system into existing vehicle designs through the use of conceptual drawings.

The range estimation and obstacle detection techniques currently employed in computer vision were described in Chapter 4. An explanation of the generation of disparity maps from stereo images was given, followed by a presentation of the two obstacle segmentation algorithms to evaluated. The ground plane subtraction and v-disparity algorithms have thus far only been applied to conventional stereo vision. It was shown that these obstacle detection methods could be modified to suit stereo panoramic system. The calculations required to determine range from the disparity map was then presented. Finally, it was discussed how temporal consistency could be used to help decrease the false detection of obstacles.

Chapter 5 presents the results from the theoretical analysis and simulation experiments that were carried out to evaluate the prototype obstacle detection system. A theoretical analysis was completed to determine the resolution, and working volume of the stereo panoramic vision sensor for range estimation. Ground truth data was generated using a ray-tracing software package (the Persistence of Vison ray tracer, POV-Ray) to create artificial panoramic images. It was found that the sensor was able to estimate range to within 20cm

of the true value, as stipulated in the system specifications in Chapter 1. The error in height remained below 15cm up to a range of 6m. The estimation of azimuth angular position was particularly impressive, with a maximum error of 3°. The constant gain mirrors performed better in the resolution invariant profiles in estimating obstacle position. It was also found that both the ground plane subtraction, and v-disparity algorithms were effective for segmenting obstacles from the panoramic disparity maps.

One of the main criticisms of the published research relating to our topic is the distinct lack of non-theoretical data. Thus, Chapter 6 contains one of the main contributions of this body of work - results from the field experiments carried out in real-world situations. It was found that the physical system tested could estimate the range to an object to within 40cm of the true value, up to a range of over 5m. The constant gain mirror system could estimate height to within 30cm, however, the resolution invariant mirrors performed significantly worse at longer ranges. The angular position was calculated to within 5° for both mirrors. The ground plane subtraction method for obstacle detection was effective if the system had been calibrated correctly. However, if the orientation of the ground plane relative to the cameras changed, obstacles could no longer be segmented correctly. On the other hand, the v-disparity algorithm was found to be robust to such changes, and was still effective when applied to extremely noisy disparity maps. The average successful obstacle detection rate was found to be around 90%, with a false detecion rate of 8%. However, this rose as high as 20% for particularly difficult lighting conditions.

Given the results from both the ground truth and field experiments, it was found that the most desirable sensor characteristics are attained by using the following parameters:

- Constant gain mirror profile
- Baseline of 40cm
- Vertical sensor configuration

7.2 Achievements

- The design and manufacture of a system to detect and locate obstacles using a stereo panoramic vision sensor.
- Implementation, and evaluation of the *ground plane subtraction* algorithm for obstacle segmentation from stereo panoramic images.

- Implementation, and evaluation of the v-disparity algorithm to segment obstacles from stereo panoramic images. It was shown that v-disparity is effective for stereo panoramic vision, as it is robust to changes in ground plane orientation and extremely noisy disparity maps.
- Investigated the characteristics of the stereo panoramic sensor for range estimation and obstacle segmentation. These included the use of different mirror profiles, baselines, and sensor configurations.
- The selection of sensor parameters to achieve most desirable characteristics for our application
- The use of temporal filtering to improve the false detection rate.
- Theoretical evaluation of the system, using Matlab, and ground truth data generated by the POV-Ray raytracing software.
- In the research published to date, there has been a distinct lack of results describing the range estimation performance of panoramic sensors in real-world situations. One of the main contributions of this work is a real-world evaluation of a stereo panoramic vision system. The range accuracy was measured, and a quantitative analysis was carried out to investigate the presented obstacle detection algorithms.

7.3 Further Work

In this thesis we have presented a new solution to the problem of monitoring vehicle blind spots. Although the system meets the goals of this project, the fact still remains that this it is merely a prototype. It was designed to demonstrate and evaluate the main components that would be required for such a driver assistance system, and further development would improve the performance.

The first issue that needs to be addressed is the accuracy of the physical stereo panoramic system. In theory, the sensor should be able to achieve double the precision that was observed in the field experiments. It should be noted that only a rough calibration of the sensor was performed, while the ground truth data was generated such that the panoramic cameras were effectively perfectly aligned. If the panoramic images were rectified to align the epipolar lines, a more complete and accurate disparity map could be calculated. This would greatly improve the segmentation of obstacles, as the dimensions of the obstacles

could be more accurately gauged. Although the current system successfully detects an entire obstacle, it does so in a piecemeal fashion. While it is easy for a human observer to associate individual portions into a single object, the system cannot do the same.

The current implementation of obstacle detection can be unstable. That is, and object detected in one frame may not be found in the next frame, or a different section of the object will be detected. We have also observed false detections, where the system depicts an object that does not exist. Many of these false detections can be filtered out through the use of a temporal consistency check. This means than an object must be detected in several consecutive frames before it is considered to be valid. In the same manner, an object is not believed to be gone until it is no longer detected in several consecutive frames. However, if more accurate disparity maps could be generated, this would be less of a problem and obstacles could be tracked, for example, with a Kalman filter.

The software developed for this project had little optimisation, and therefore operates as slow as 3Hz, depending on the size of the image. The generation of the disparity maps is the most intensive component, and has been implemented using the MMX instruction set by (Fletcher et al. 2001). However, the execution speed could be increased by scaling down the panoramic images at the cost of decreasing the accuracy of the position estimation. If the frame rate of the system was increased, the likelyhood of detecting an obstacle would be greatly increased and the driver can be notified earlier which is particularly important for faster moving objects.

The cameras were chosen for this project because of their flexible characteristics. The cameras have zoom, focus, infra-red and other features that make it configurable for a wide variety of situations. However, this means that they are comparitively large, which propagated in the design of the equipment, resulting in a fairly bulky stereo panoramic system. It would be possible to design a more compact system using smaller, less expensive cameras. The mirrors were all machined from solid blocks of aluminium, which increases the overall weight of the sensor. Instead, the mirrors could be hollowed out, or even made from molded plastic to produce a lighter system.

Bibliography

Badal, S., Ravela, S., Draper, B. & Hanson, A. (1994), A practical obstacle detection and avoidance systems, *in* 'Proceeding of 2nd IEEE Workshop on Applications of Computer Vision', pp. 97–104.

Baker, S. & Nayar, S. (1999), 'A Theory of Single-Viewpoint Catadioptric Image Formation', *International Journal of Computer Vision* **35**(2), 1–22.

Ballard, D. & Brown, C. (1982), *Computer Vision*, Prentice-Hall.

Banks, J. & Corke, P. (2001), 'Quantitative evaluation of matching measures for stereo vision', *International Journal of Robotics Research* **20**(7), 512–532.

Bishop Consulting (2004), Delphi Shows Off Advanced Safety Technologies in D.C. Accessed Nov 2004.
*http://www.ivsource.net/

Bunschoten, R. & Kröse, B. (2001), Range estimation from a pair of omni-directional images, *in* 'Proceedings of the IEEE International Conference on Robotics & Automation', pp. 1174–1179.

Canesta, Inc (2006), Development kits and software. Accessed February 2006.
*http://www.canesta.com/

Cao, Z., Oh, S. & Hall, E. (1985), Dynamic omnidirectional vision for mobile robots, *in* 'SPIE Intelligent Robots and Computer Vision', Vol. 579, pp. 405–414.

Chahl, J. & Srinivasan, M. (1997), 'Reflective surfaces for panoramic imaging', *Applied Optics* **36**(31).

Chandler, D. D. (1997), 'Visual perception 1', Lecture notes. University of Wales, Aberstwyth.

Conroy, T. (2000), Topics in Nonlinear Signal Processing, PhD thesis, Australian National University.

Conroy, T. L. & Moore, J. B. (1999), Resolution invariant surfaces for panoramic vision systems., *in* 'Proceedings of the International Conference on Computer Vision', Kerkyra, Corfu, Greece, pp. 392–397.

Department of Motor Vehicles, California (2001), Mobility and gridlock - accident avoidance. Accessed July 2003.
*http://www.dmv.ca.gov/pubs/hdbk/pgs62thru64.htm

Driscoll, T., Mitchell, R., Mandryk, J., Healey, S. & Hendrie, L. (1998), Work-related traumatic fatalities in australia, 1989 to 1992, Technical report, National Occupational Health and Safety Commission of Australia.

Dudek, G., Jenkin, M., Milios, E. & Wilkes, D. (1996), Reflections on modelling a sonar range sensor, Technical report, McGill Research Centre for Intelligent Machines.

Elkins, R. & Hall, E. (1994), Three-dimensional line following using omnidirectional vision, *in* 'SPIE Intelligent Robots and Computer Vision XIII: 3D Vision, Product Inspection, and Active Vision', Vol. 2354, pp. 130–144.

European Commission (2003), Road safety - vehicles and vehicle equipment.
*http://europa.eu.int/comm/transport/road/index_en.htm

FastLane Communications (2003), Safety first - in reverse. Accessed September 2003.
*http://www.fastlane.com.au/4WD-OffRoad/Bushranger_widelens.htm

Faugeras, O., Hots, B., Mathieu, H. & *et all* (1993), Real time correlation-based stereo: algorithm, implementations and applications, Technical Report 2013, INRIA.

Fletcher, L., Apostoloff, N., Chen, J. & Zelinsky, A. (2001), Computer Vision for Vehicle Monitoring and Control, *in* 'Australasian Conference on Robotics and Automation'.

Gandhi, T. & Trivedi, M. (2004), Motion based vehicle surround analysis using an omnidirectional camera, *in* 'IEEE Intelligent Vehicles Symposium', Parma, Italy, pp. 560–565.

Gaspar, J., Winters, N. & Santos-Victor, J. (2000), Vision-based navigation and environmental representations with an omnidirectional camera, *in* 'IEEE Transactions on Robotics and Automation', Vol. 16, pp. 890–898.

Gluckman, J., Nayar, S. & Thoresz, K. (1998), Real-time omnidirectional and panoramic stereo, *in* 'Proceedings of the DARPA Image Understanding Workshop', Monterey, USA.

Gokturk, S. B., Yalcin, H. & Bamji, C. (2004), A time-of-flight depth sensor - system description, issues and solutions, *in* 'Conference on Computer Vision and Pattern Recognition Workshop', Vol. 3, p. 35.

Heikkila, J. & Silven, O. (1997), A four-step camera calibration procedure with implicit image correction, *in* 'Computer Vision and Pattern Recognition'.

Hicks, A. & Bajcsy, R. (1999), Reflective Surfaces as Computational Sensors, *in* 'Second IEEE Workshop on Perception for Mobile Agents '.

Holden (2003), 'Commodore Series II', Pamphlet.

Holland, A., Liang, R., Singh, S., Schell, D., Ross, F. & Cass, D. (2000), '"Driveway motor vehicle injuries in children"', *Medical Journal of Australia* **173**, 192–195.

IBEO Automobile Sensor (2004), Products - data & prices. Accessed September 2004. *http://www.ibeo-as.de/html/prod/prod_dataprices.html

Kagami, S., Okada, K., Inaba, M. & Inoue, H. (2000), Design and implementation of onbody real-time depthmap generation system, *in* 'Proceedings of the IEEE International Conference on Robotics & Automation'.

Koeppen, S. (2003), Kids 'n cars promotes blind-spot safety. Accessed July 2003. *http://www.thepittsburghchannel.com/call4action/2041787/detail.html

Koyasu, H., Miura, J. & Shirai, Y. (2003), Mobile robot navigation in dynamic environments using omnidirectional stereo, *in* 'IEEE International Conference on Robotics and Automation', Taipei, Taiwan, pp. 893–898.

Labayrade, R. & Aubert, D. (2003), In-vehicle obstacles detection and characterization by stereovision, *in* 'Proceedings the 1st International Workshop on In-Vehicle Cognitive Computer Vision Systems'.

Labayrade, R., Aubert, D. & Tarel, J. P. (2002), Real time obstacle detection on non flat road geometry through v-disparity representation, *in* 'IEEE Intelligent Vehicles Symposium, Versailes'.

Lacroix, S. & Dudek, G. (1997), On the identification of sonar features, *in* 'IEEE/RSJ International Conference on Intelligent Robots and Systems'.

Matsumoto, Y., Ikeda, K., Inaba, M. & Inoue, H. (1999), Visual navigation using omnidirectional view sequence, *in* 'Proceedings of 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems', pp. 317–322.

Maycock, G. (1996), 'Sleepiness and Driving: The Experience of UK Car Drivers', *Journal of Sleep Research* **5**, 229–237.

Mazokhin-Porshnyakov, G. (1969), *Insect Vision*, Plenum Press.

McKerrow, P. J. & Zhu, S. (1996), Modelling multiple reflection paths in ultrasonic sensing, *in* 'Proceedings of the International Conference on Intelligent Robots and Systems'.

Molton, N., Se, S., Lee, D., Probert, P. & Brady, M. (1997), Robotic sensing for the guidance of the visually impaired, *in* 'International Conference on Field and Service Robotics'.

Ng, K., Trivedi, M. & Ishiguro, H. (1998), 3d ranging and virtual view generation using omni-view cameras, *in* 'Multimedia Systems and Applications, SPIE', Vol. 3528.

Ollis, M., Herman, H. & Singh, S. (1999), Analysis and design of panoramic stereo vision using equi-angular pixel cameras, Technical report, Carnegie Mellon University Robotics Institute.

Paine, M. & Henderson, M. (2001), Devices to reduce the risk to young pedestrians from reversing motor vehicles, Technical report, Motor Accidents Authority of NSW.

Panerai, F., Capurro, C. & Sandini, G. (1995), Space variant vision for an active camera mount, *in* 'SPIE AeroSense95', Florida, USA.

Peer, P. & Solina, F. (2002), 'Panoramic depth imaging: Single standard camera approach', *International Jornal of Computer Vision* **47**, 149–160.

Peleg, S. & Ben-Ezra, M. (1999), Stereo panorama with a single camera, *in* 'Computer Vision and Pattern Recognition', pp. 395–401.

Polaroid OEM Components Group (1999), 'Piezo Transducers - K Series', Technical Specification Documents.

Sensing and Control, Honeywell Inc. (n.d.), 'Ultrasonic sensor', Product Catalog.

Sogo, T. & Ishiguro, H. (2000), Real-time target localization and tracking by n-ocular stereo, *in* 'IEEE Workshop on Omnidirectional Vision'.

Szeliski, R. (1994), Image mosaicing for tele-reality applications, *in* 'Workshop on Application of Computer Vision', IEEE, pp. 44–53.

Thompson, E. (2002), 'Bus and vehicle collision prevention', *News & Views* **6**. Texas Association of School Boards Risk Management Fund.

Trilogix Electronic Systems Inc. (2003), Donnelly's VideoMirror. Accessed July 2003. *http://www.12voltdirect.com/

Trucco, E. & Verri, A. (1998), *Introductory Techniques for 3-D Computer Vision*, Prentice Hall.

Wijk, O. (2001), Triangulation Based Fusion of Sonar Data with Application in Mobile Robot Mapping and Localization, PhD thesis, Royal Institute of Technology (KTH).

Yagi, Y. (1999), Omnidirectional sensing and its applications, *in* 'IEICE Transactions on Information and Systems', Vol. E82D, pp. 568–579.

Yagi, Y., Kawato, S. & Tsuji, S. (1994), Real-time omnidirectional image sensor (COPIS) for vision-guided navigation, *in* 'IEEE Transactions on Robotics and Automation', Vol. 10, pp. 11–21.

Zhang, Z. (1999), Flexible camera calibration by viewing a plane from unknown orientations, *in* 'International Conference on Computer Vision'.

Zhu, Z., Rajasekar, K. D., Risemane, E. & Hanson, A. (2000), Panoramic virtual stereo vision of cooperative mobile robots for localizing 3d moving objects, *in* 'IEEE Workshop on Omnidirectional Vision', pp. 29–36.

Appendix A

To increase the speed of unwarping images, a lookup table can be precomputed. Therefore, instead of having to calculate the mapping of pixels while the system is online, each pixel can be mapped according to the lookup table. An example can be seen in Figure 7.1. The lookup table in the figure provides a mapping to flip a 640x480 pixel image up-side-down. For example, to determine the mapping for pixel (0,0) in the unwarped image, we use cell (0,0) from the lookup table. This shows that pixel (0,0) from the unwarped image should be set to the same value as pixel (0,479) in the raw image. This is done for each pixel in the unwarped image.

	0	1	2	3	4	
0	(0, 479)	(1, 479)	(2, 479)	(3, 479)	(4, 479)	
1	(0, 478)	(1, 478)	(2, 478)	(3, 478)	(4, 478)	
2	(0, 477)	(1, 477)	(2, 477)	(3, 477)	(4, 477)	
3	(0, 476)	(1, 476)	(2, 476)	(3, 476)	(4, 476)	

Figure 7.1: An example lookup table showing the mapping from the unwarped to the warped (raw) image. In this example, pixel (0, 2) of the unwarped image corresponds to pixel (0, 477) of the raw image.

Appendix B

A CD-ROM is attached to this thesis and contains electronic of the various documentation produced for this body of work, as well as all collected data, and code. It contains the following folders:

1. **C++ code** - contains a zip file of the C++ code written for this research project.
2. **Documentation** - contains electronic copies of the final seminar presentation, the Intelligent Vehicle Symposium paper, the progress report, this thesis and the user guide for setting up and using the developed hardware and software.
3. **Matlab code** - contains the code for analysing the results from the test sequences, camera calibration and the initial ground plane subtraction algorithm.
4. **Mirror design** - contains the Matlab code and ProEngineer files used during the design of the mirrors and camera mounts, including the technical drawings.
5. **Ray tracing** - contains the POVray code for generating the ground truth images.
6. **Videos** - contains videos of the obstacle detection results using the test sequences.