

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Surface reconstruction of a blast plate using stereo vision

Prepared by:

Gladys Nabagala Kigozi

Submitted to the Department of Electrical Engineering,
University of Cape Town, in fulfilment of the requirements for the
degree of Master of Science in Engineering.

February 2007

University of Cape Town

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in Engineering in the University of Cape Town. It has not been submitted before for any degree or examination in any other university.

Signature of Author

Cape Town February 2007

Signed by candidate

Abstract

This thesis presents method for reconstructing and measuring the profile of a blast metal plate. Among the many methods in computer vision, stereo vision using two cameras is chosen as the range finding method in this thesis. This is because it is a non-contact method and hence eliminates the need to calibrate moving parts. A stereo-rig consists of two calibrated cameras and hence gives two view geometry. Stereoscopic reconstruction relies on epipolar geometry to constrain the relationship between the views. The 3-D point is then estimated using triangulation of the corresponding points from the two views.

The blast plates that are reconstructed have highly reflective surfaces. This causes a problem due to specular reflection. This thesis further studies the reflective properties of the metal plate surface.

Different methods of scanning the plate using the stereo-rig are investigated. The reconstructions obtained from these methods are analyzed for accuracy and consistency. Since low cost cameras are used in constructing the stereo-rig, the point cloud data obtained is further investigated for consistency by aligning different instances of the reconstruction. This is done using the Iterative Closest Programme (ICP) algorithm which tries to align two sets of data iteratively.

Contents

1	Introduction	1
1.1	Problem description	1
1.1.1	Deformation and tearing of thin plates subjected to impulsive loading	1
1.1.2	Blasting process	2
1.1.3	Blasting experimental discussion	3
1.2	Background to thesis	4
1.3	Purpose and limits of thesis	5
1.4	Outline of thesis	6
2	Reflective surfaces and their geometric properties	8
2.1	Types of radiances from an object	8
2.1.1	Reflection due to the surface of the object	8
2.1.2	Reflection due to the body of the object	9
2.2	Geometric properties of highly reflective surfaces	10
2.3	Implications of intensity of a point and inaccuracy in stereo analysis	11
2.4	Types of surfaces	13
2.4.1	Rough surfaces	13
2.4.2	Lambertian surfaces	14
2.5	Surface related problems encountered in range sensing	15
2.6	Related works on surface reconstruction and measurement of specular objects	15
3	Different methods used to obtain range data of an object for measurements	16
3.1	Direct and active methods	16
3.2	Passive monocular image-based methods	16
3.3	Motion and multiple relative positions of camera scene	17
3.3.1	Setup comprised of calibrated camera and movable laser plane	17
3.3.2	Setup comprised of two calibrated cameras	18

4	Camera modeling and epipolar geometry	20
4.1	Perspective projection of a Pinhole Camera	20
4.2	Reference frames	21
4.2.1	World reference frames	21
4.2.2	Image reference frame	22
4.2.3	Camera reference frame	22
4.3	Epipolar geometry	23
4.4	The fundamental matrix F	24
4.5	The Essential Matrix	25
4.6	Obtaining the camera matrices P and P'	25
4.7	Triangulation of the corresponding points	26
5	Camera calibration	27
5.1	Intrinsic camera parameters	27
5.2	Extrinsic camera parameters	28
5.3	Lens distortion in cameras	28
5.4	Distortion models	29
6	Registration of data using Iterative Closest Point (ICP) algorithm	30
6.1	Introduction	30
6.2	The ICP algorithm	31
6.3	Uses of ICP in image processing	31
6.4	The working of the ICP algorithm	31
6.4.1	Mathematical preliminaries	31
6.4.2	The ICP algorithm is stated	32
6.5	Disadvantages of using the ICP algorithm	33
6.6	Advantages of ICP	34
7	System design and functioning	35

7.1	Objectives of scanners	35
7.2	Calibrating the rig using Matlab Toolbox	35
7.2.1	Estimation of the fundamental matrix	37
7.2.2	Validating the fundamental matrix	37
7.3	Designing of the stereo-rig	37
7.4	Factors affecting the laser scanner design	39
7.4.1	Camera positioning	39
7.4.2	Laser line/stripe	39
7.4.3	Resolution of scanner	39
7.5	Scanning procedure	39
7.6	Data processing	40
7.6.1	Pre-processing data	40
7.6.2	Processing data	41
8	Results, analysis and discussions	43
8.1	Introduction	43
8.2	3-D reconstruction of the plates using laser line	43
8.2.1	Projective reconstructions as a result of using a laser line	44
8.2.2	Effects of visibility of the laser line images	44
8.3	3-D reconstruction using a laser spot	47
8.3.1	Projective reconstructions of the plate as a result of using a laser spot	47
8.4	Estimating plane of best-fit	48
8.5	Results analysis	49
8.5.1	Analysis of the best-fit plane	50
8.5.2	Depth Error of the height of the dome	50
9	System accuracy testing	52
9.1	Comparison with simple object	52

9.2	Reprojection error	54
10	Capabilities of the ICP registration algorithm	55
10.1	Applying ICP on randomly generated data	56
10.1.1	Point set matching on 2-D data	56
10.1.2	Point set matching on 3-D	58
10.1.3	Analysis of the randomly simulated data	58
10.1.4	Conclusion on the synthetically generated data	59
10.2	Applying ICP on measured data set	59
10.3	Analyzing results of registered point cloud data of the blast plate	62
10.3.1	Analyzing choice of local minima	62
10.3.2	Analyzing computational approaches and time taken	62
10.4	ICP registration on surface reconstruction of other objects	63
10.4.1	Registration of cloud point data of a tennis ball	63
10.4.2	Analyzing the registration	64
10.5	Conclusions made on the ICP algorithm alignment	64
11	Conclusions	66

List of Figures

1	Ballistic pendulum experimental arrangement [25].	2
2	Schematic layout of ballistic pendulum test rig showing circular holding plate [23].	3
3	Laser range scanner set up	4
4	Polar plots of the three reflection components as a function of the viewing angle for a fixed source direction [22].	9
5	The geometry needed to define radiometric terms [22].	10
6	Rays reflected off (a) a smooth surface and (b) a rough surface.	11
7	An example to show that specular surfaces can cause existing shape recovery algorithms to produce results that do not lie close to the desired result [21].	12
8	Camera images of a laser when projected on three different objects. (A) Object with diffuse surface, (B) and (C) objects with specular surfaces [14].	13
9	Image showing laser light directed on the blast plate [22].	13
10	Diffuse reflection resulting from multiple reflection and internal scattering mechanism [22].	14
11	Diagram showing perception of depth with change in gradient [30].	17
12	Single camera and laser plane [16].	18
13	Systematic diagram of stereo using multiple camera setup [9].	18
14	Setup for deriving the equations of the pinhole camera perspective projection [5].	20
15	Image and camera coordinate system [18].	22
16	Epipolar geometry of two cameras with optical centers O and O' . Images x and x' of X lie in the same plane (shaded region) [5].	23
17	Illustration of lens distortion. The original grid (left), and the effect of barrel (center), and pincushion distortion (right) [17].	28
18	Image of a rabbit reconstructed by the Stanford research group with high-resolution cameras [15].	30
19	Image of the stereo setup.	35
20	Checkerboard images used in the calibration of the two camera.	36

21	Figure showing extracted corner on one of the checkerboard images used in calibration.	36
22	Figures showing corresponding points on a checkerboard.	38
23	Blasted plate setup	38
24	Figure showing gaussian property of light.	41
25	Images of the blast metal plates to be reconstructed.	43
26	The dimensions of the blasted plate to be reconstructed.	44
27	3-D Reconstruction of the painted plate from different views.	45
28	3-D Reconstruction of the unpainted plate using a laser line from different views.	45
29	Figure (a) shows a laser line image captured by the left-hand image. Figure (b) shows the laser line image after filtering.	46
30	Figure (a) shows a laser line image captured by the right-hand image. Figure (b) shows the laser line image after filtering.	46
31	Point cloud data of the surface reconstruction using a laser spot.	47
32	Point cloud data in the x-y direction to show patchiness as a result of using a manually positioned laser spot.	48
33	Image of the reconstructed surface in the form of a surf.	48
34	Images showing plane of best fit for the flat section of the plate.	49
35	Close up of estimated plane	49
36	Image of a box used for accuracy check.	52
37	Figure showing point cloud data of reconstructed faces of the wooden box from different views	53
38	Figure showing estimated planes best-fit used to calculate the angle between the reconstructed faces.	53
39	Re-projection back onto the checkerboard	55
40	Image showing re-projection of reconstructed points of a box back to original points	55
41	Re-projection of the reconstruction of the blast plate to its original image points	55
42	Figure(a) shows two sets of randomly generated data in 2-D, before alignment, (b) shows how the two data sets align after being registered using ICP algorithm.	57

43	(a) shows two sets of randomly generated data in 3-D before alignment, (b) shows how the two data sets align after being registered using ICP algorithm .	58
44	(a) shows a larger set of randomly generated data in 3-D before alignment, (b) shows how the two data sets do not register after being registered using ICP algorithm	59
45	Real complete points cloud data of set X and X' before ICP.	60
46	500 randomly chosen points from both set X and X'	60
47	Figure showing process of the aligning subsets of X and X'	61
48	Alignment of set X and X' after transformation obtained from using 500 points.	61
49	Image of the tennis ball that is introduced for comparison.	63
50	3-D point cloud data obtained from scanning a tennis ball at different orientations. (The red points show the first scan in one view and the blue scan show the point cloud data of another view).	63
51	Image showing the results obtained from aligning point set-one (red points) with point set-two (blue points) of different surface reconstructions.	64
52	Time taken to perform ICP algorithm using different number of points chosen as subsets	65

1 Introduction

This thesis falls under an area of surface reconstruction in computer vision. Computer vision is a combination of artificial intelligence and image processing and is used to generate 3-D models of real world objects from images on a computer [6]. Vision techniques have improved and been adapted to suit specific applications. They are used extensively in a wide range of industries. Practical machine vision applications include surface reconstruction, auto navigation and bio-medical presentation of anatomical parts, eye/gaze tracking, lip reading, and the tracking and recognition of human gestures [5]. The theory of surface reconstruction is becoming common in the research field of computer vision, because it makes for easier and faster study of structures.

Over the years, new image processing methods have been researched and developed and these are being used for surface reconstruction and object measurements. One of these methods is stereo-vision using a laser triangulation scanner. This is a non-contact method; its popularity is based on the robustness and on the level of precision attainable [11]. Hence this is the method chosen to create surface reconstructions of specified objects in this thesis. The accuracy of the laser scanner is largely dependent on the cost of its components, i.e. the higher the cost, the better the quality of the image, and thus the better the reconstruction [11]. Even if the method is currently only being used on a small scale in the laboratory, it can be optimized for larger works since the concept remains the same.

This thesis therefore present both the process in general and the specific steps that were carried out to generate surface reconstructions of blasted metal plates using stereo-vision. From the results generated, relevant conclusions are drawn, and recommendations are made with regard to future research in this field.

1.1 Problem description

The blasted metal plates that were used in the reconstruction were obtained from the Blast Impact and Survivability Research Unit (BISRU) of University of Cape Town. The BISRU research involves reducing the risk of injury caused by blast events and equipment failure. They are also involved in improving survivability through investigation of fundamental principles of science and engineering, using experimental, analytical and computational tools to understand the mechanics and dynamics of blast and impact loads.

1.1.1 Deformation and tearing of thin plates subjected to impulsive loading

The deformation and tearing of materials is carried out in a blasting chamber on the UCT grounds for experimental analysis. The experiments are carried out to analyze how different

materials deform when subjected to different sizes of impulsive loading. One set of experiments studied the effect of underwater explosive forming on fully clamped circular discs of various thickness and different materials. Another investigation using several plate materials and thickness studied the relationship between deformation and explosive parameters such as charge mass and stand off distance [24]. The plates used were about 0.6mm thickness and made of mild steel.

There are number of factors that were observed while carrying out the above experiments. These include:

- Observing the dynamic behaviours of the various structures; in this case metal plates. The dynamic response of the structure (beams, plates) was measured using various techniques.
- Analysing the final deformed shape and logging the deflection-time history. These measurements provide a useful guideline for predicting maximum central deflection of impulsive loading and hence predicting when maximum tearing will occur.

1.1.2 Blasting process

Methods of impulsive loading used include: air pressure waves created by explosives devices; direct impulsive loading using plastic sheet explosives and ballistic pendulum; and spring loaded arms [23].

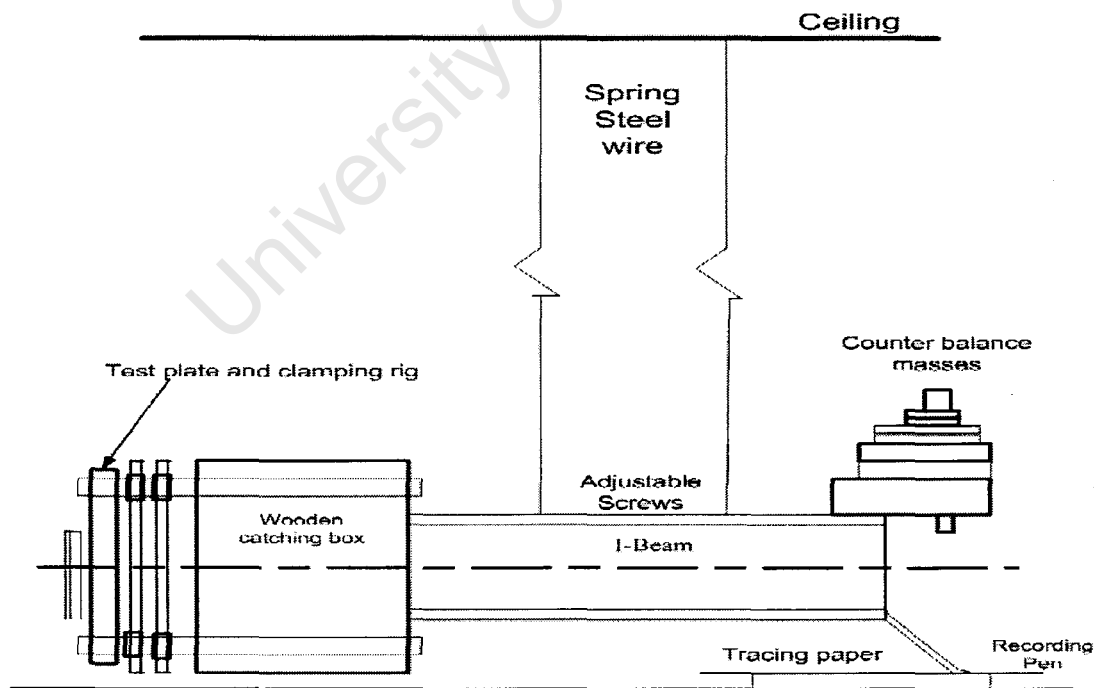


Figure 1: Ballistic pendulum experimental arrangement [25].

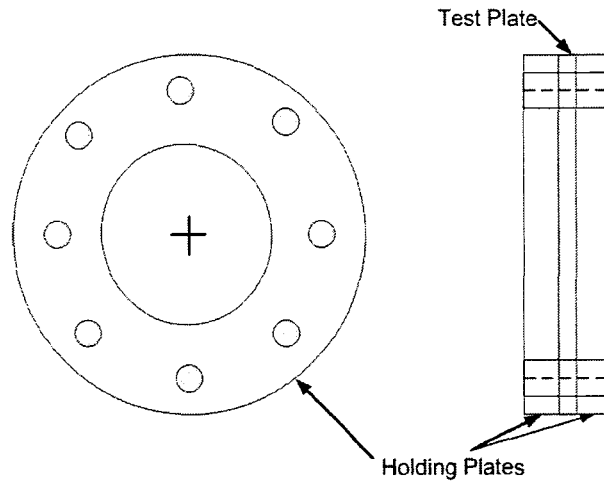


Figure 2: Schematic layout of ballistic pendulum test rig showing circular holding plate [23].

- **Ballistic pendulum blasting:** The ballistic pendulum is structured as shown in Figure 1. The experimental plate is rigidly clamped on the front of the pendulum as shown in Figure 2. Layers of Dupont sheet explosives of thickness 0.4mm each are applied to the clamped beam [23].

The explosion is initiated at the center of the test specimen by means of a pigtail leading back to a standard detonator blast. The recorded pendulum swing gives a direct indication of the maximum potential energy of the system [23]. The potential energy is used to calculate the maximum velocity of the whole pendulum system and hence gives a measure of the applied impulse. The beam is protected by a layer of Neoprene.

1.1.3 Blasting experimental discussion

From tests carried out the following analysis were made:

The magnitude and shape of the deformed plate depends on the size and form of the impulsive loading. From the many experiments carried out, the deformation of a structure can be predicted. Through repeatability with more than 200 tests a conclusion on deformed shapes can be made. This result can be used to predict deformation of materials before blasting so as to know which material can be reliable in a given working environment.

In industrial or in laboratory experiments, when structures are subjected to blast loads, they undergo deformations that depend on the type of material used and its characteristics. In blasting experiments, metric 3-D measurements of the structures have to be taken both before and after the blast event. It would be convenient to be able to take these measurements on site, as this would reduce delays and the difficulties involved in transporting the material elsewhere. Laser scanners could be used to take these measurements, but off-the-shelf laser scanners are expensive. The BISRU group thus requested a system that could take the requisite metric

measurements on site. Furthermore this system has to be able to make reconstructions of their objects.

1.2 Background to thesis

The reconstructed object chosen for the purposes of this research was a blast metal plate. It was found that the deformed section of the plate has high specular reflection when swept over with a laser line. This property makes it difficult to use a laser scanner for range imaging of metal objects. Sometimes the difficulty can be overcome or at least bypassed by painting or spraying the metal surface. However in industry, it is often important to retain the original appearance of the metal surface, and therefore it is not possible to paint it with a matte colour. This was not the case for this thesis though, and consequently the blast plate was painted so as to obtain a 3-D model of the plate. In section 2 an explanation on how the direction of incident light on a surface affects the reflected ray is given. Scanning an unpainted metal plate using a laser spot while changing the angle of incident of the laser is investigated to see the type of reconstruction that can be obtained.

The main task in this project is to create an accurate surface reconstruction of the blast metal plate in three dimensions using a stereo-rig (multiple view geometry). The prototype used is made up of two calibrated charged-couple device (CCD) cameras and a laser line generator that is not calibrated into the system. A figure of the setup is shown in Figure 3.

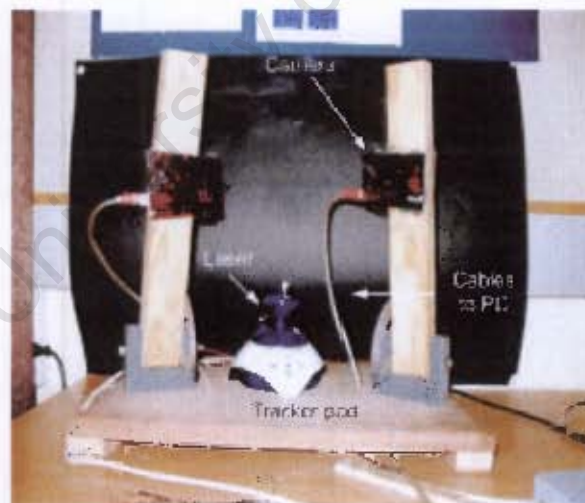


Figure 3: Laser range scanner set up

The laser range scanner relies on a known calibrated camera geometry in order to reconstruct depth. The process used in calibrating the stereo-rig using the Matlab camera calibration toolbox is explained in section 7.2. The laser line generator sits on a trackerpod that moves the laser in a horizontal direction. The two cameras face the blast plate and the laser line is swept across the plate's surface at intervals. For every interval the cameras capture the

laser line simultaneously. The images captured are processed to extract the co-ordinates of the laser line image and the data points extracted are used to calculate the 3-D position of the laser line. Although it is easy for the human eyes to judge stereoscopic vision, this is not simple or straightforward for a computer. A computer needs to calculate the depth of a world point by using images from two cameras. This involves knowing the position of the first camera relative to the second through calibration of the cameras. Hence for an accurate 3-D reconstruction using stereo-vision, accurate calibration of the stereo-rig is required.

The 3-D model acquired from a two view geometry reconstruction is projective. Projective cameras P and P' map world point X to image points x and x' according to $x = PX$ and $x' = P'X$ [9]. A matrix describes the projective structure between the two images. This matrix is known as the fundamental matrix and its mapping between corresponding points $x \leftrightarrow x'$ is $x'^T F x = 0$. This matrix essentially describes the intrinsic projective geometry between two views. Robust methods for determining the fundamental matrix are especially important when dealing with real image data. The two sets of image data used to estimate the fundamental matrix consist of a corner points of a checkerboard used from the calibration process.

The blast plate is thus scanned after calibrating the stereo-rig. The images captured are then processed to detect the co-ordinates of the laser line in the image. The epipolar theory of the mapping $x \mapsto l'$ [9] is used to obtain corresponding points from the two images of the line. These corresponding points are then used in the triangulation process when estimating the 3-D model of the plate.

1.3 Purpose and limits of thesis

The aim of this thesis was to investigate stereo-vision as a technique to obtain 3-D surface reconstructions of objects specified by the BISRU research group. Software to implement the stereo-vision technique was required. Furthermore, it was necessary to determine the level of accuracy that could be achieved by using the suggested technique and the developed prototype. However, a number of problems were encountered during the attempts to achieve the desired outcomes, namely:

- High resolution equipment is very expensive, and therefore low quality CCD cameras were used to construct the prototype. These create low resolution images of 480×640 pixels. Poor quality images were thus produced during the scanning process, which caused difficulties in the detection of the laser stripe image.
- The specified object (blast plate) that was reconstructed had a poor reflective index because of its surface properties. On certain parts of the blast plate, no visible laser image could be captured. This led to loss of data during the scanning process. In order to avoid this problem, the plate was thinly painted with a matte colour.
- Scanning of the plate was done in a controlled light environment. Daylight was blocked out using blinds whereas the other light sources were controlled using filters. This means that the scanner can only be used in limited areas.

- Stereo-vision depends on the scope of view of two cameras. The cameras thus have to be placed in a position such that maximum view of the object is seen by both cameras. Wider surface reconstructions are at times hard to obtain, depending on the curvature/occlusion of the object. If a larger view of the object is needed, then the object is moved further away from the scanner. However, this reduces the resolution of the 3-D model, thus reducing its quality.

1.4 Outline of thesis

The layout of this thesis follows the systematic steps that were taken to arrive at the final goal of this project. One of the difficulties experienced when scanning a metal surface is the specular reflection that occurs when a laser light is directed towards the surface. Surface properties of objects affect their reflection properties. Chapter 2 gives some insight into these surface properties and how an incident beam behaves when it strikes a surface. From the surface properties an object is grouped as either having specular reflections or Lambertian reflections. The surface properties determine the type of range finding method that can be used on the object.

Different methods that can be used to generate a 3-D model of an object are discussed in Chapter 3. These methods are characterized according to their properties and steps needed to carry out the experiments. Some of these methods include passive and active range finding. The advantages and disadvantages of the methods are given to show why the method was chosen for this research.

In Chapter 4 the working of a pinhole camera is presented. The chapter further explains the relationship between what is observed in the camera images and what is seen in the world. Further the concept used to mathematically relate two corresponding points (which are used in the linear triangulation) in order to calculate a world point is explained. This is defined as epipolar geometry.

Chapter 5 describes the process involved in calibrating the stereo-rig using the Matlab camera calibration toolbox [4]. The intrinsic parameters obtained from calibration are defined and these are further explained to show how they are used in modeling in the camera. This chapter further explains how to obtain the orientation, i.e. rotation and translation, of a view point in multiple view geometry with reference to another point. Problems experienced and solutions found in using real cameras are discussed. These include reducing lens distortion.

The Iterative Closest Program (ICP) algorithm was used to align two different reconstructions to show that two reconstructions are consistent. The literature and methodology relating to this algorithm is discussed in Chapter 6.1. The ICP algorithm aligns two sets of data by minimizing the overall mean-squared distance metric between the two sets of data points. The transformation required to align the two sets is applied to one set and the mean-squared distance between data points is calculated. Another transformation is estimated from the information of the mean-squared distance and applied to the data again. This is repeated until the mean-squared distance between the data points is very small, i.e. when the algorithm

reaches its local minimum. Finally, the advantages and disadvantages of the algorithm are given.

The system's design and setup is explained in Chapter 7. The steps taken in calibrating the system are explained in detail. Since a rigid rig is constructed for the scanning, initial positioning of the cameras is crucial before calibration. The steps taken in operating the rig are given as well as the procedure taken in processing the captured images in order to extract the coordinates of the laser stripe.

Chapter 8.1 presents the results obtained from the reconstruction and the analysis of the findings. These are presented in the form of images, and the measurement taken are tabulated. The error between expected height and the measured height of the blast plate profile is analysed and discussed. The possible causes of the errors obtained are listed and solutions are given.

Checking of the system's accuracy is presented in Chapter 9. Different measurement tests were done on a wooden box. The investigations done on the reconstructions were to check that the correct length of an object can be obtained and that the angle between planes hold. The re-projection error was also calculated for the different reconstructions and is discussed.

In Chapter 10 the ICP algorithm is used to further analyze the consistency of the reconstructions obtained, by checking how close two sets of reconstruction align. In order to carry this out the ICP algorithm was tested on 2-D randomly generated data. This was then further upgraded to 3-D randomly generated data. The algorithm is also used on the measured data set, in this case the point cloud data of the metal plate. The results obtained in terms of average RMS distance error, number of iterations used and total time taken to carry out the algorithm are tabulated, analysed and conclusions are made. Finally, Chapter 11 draws conclusions on the results obtained.

2 Reflective surfaces and their geometric properties

The aim of this section is to qualitatively describe the effect of specular reflections in stereo-vision. Computer vision can benefit from the understanding of surface behavior of different materials. In [22] it is explained that many of the machine vision problems involve analysis of an image and are a result of poor light reflection. The visibility of a point in an image depends on its ability to reflect incident light in the direction of the camera/sensor. This is known as the point's reflectance [22].

Optical properties of the surface determine the quality of the surface reconstruction. An ideal surface is a Lambertian surface with a high reflective index [12]. Laser scanning of a surface with a mirror-like reflection is a challenging task and many metal objects have this property. Although the physics and geometry governing specular reflection can be accounted for as explained in [20], the visual interpretation of specular surfaces is still an ongoing research topic [21]. Some researchers eliminate surface related problems by painting the object so as to ensure that the surfaces reflect the light source diffusely [14], but this is not always desirable.

Many objects of importance in industry have specular reflecting surfaces. In this project the prototype that has been developed to make a surface reconstruction of a blast plate is a laser scanner. It has proven difficult to scan an unpainted blast plate because of the plate's surface properties.

2.1 Types of radiances from an object

This section looks at how light behaves once it hits a surface, i.e. through reflection. The surface can be decomposed into three primary reflection components, namely a diffuse lobe, a specular lobe and a specular spike [20].

A presentation of this explanation can be viewed in Figure 4. The sum of the three lobes determines the surface radiance detected by the viewer for a specific position. For smooth surfaces the specular spike component is larger than the specular lobe component. As the surface becomes rougher the specular spike reduces and the specular lobe increases. An object's surface is able to reflect light due to two factors: reflection due to the object's body, and reflection due to its surface. These are further discussed below as read from Nayar et al. [22].

2.1.1 Reflection due to the surface of the object

Reflection due to surface of the object occurs at the boundaries between materials such as two dielectrics. Surface reflection can be broken down further into two components. These are coherent and incoherent components depending on the surface smoothness or roughness.

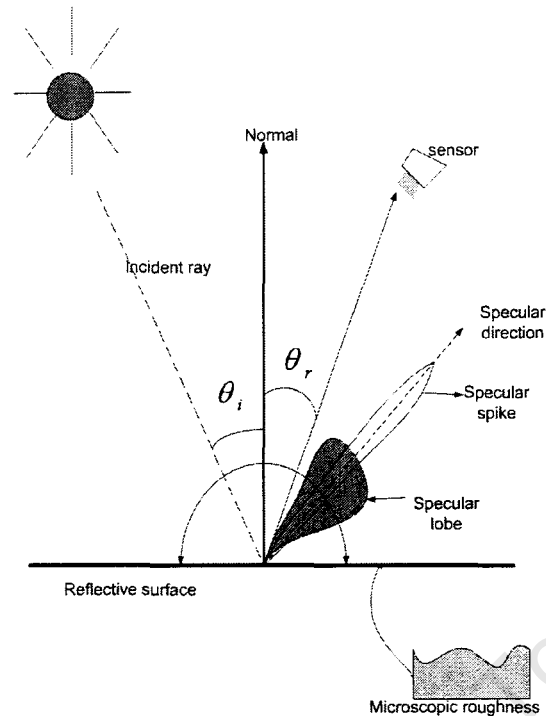


Figure 4: Polar plots of the three reflection components as a function of the viewing angle for a fixed source direction [22].

- Coherent¹ reflection obeys the mirror-like laws of reflection. This component is also defined as the specular spike in Nayar et al. [22]. In this reflection the surface variations are smaller than the wavelength of light hence coherent reflection occurs and predominates.
- As the wavelengths become smaller compared to the variations of the surface (i.e. the surface becoming rougher), the coherent component weakens and the incoherent components dominate. The incoherent component spreads in directions other than the specular spike. The spreading depends on the roughness of the surface. This component is also termed the specular globe, but in this report it is referred to as specular reflection for simplicity [22].

2.1.2 Reflection due to the body of the object

Body reflection is due to internal scattering of light [2]. This is where rays enter the surface and due to the surface in-homogeneties, it undergoes multiple reflections and refraction. A fraction of the rays emerge from the surface. Because of the non directional nature of the rays this is known as diffuse reflection. Diffuse reflections are sometimes assumed to be

¹Coherence refers to the predictability of the phase of reflected light waves with respect to the incident light wave. If the phase of the reflected wave is well defined, then it is called coherent and if it is random it is called incoherent [2].

Lambertian, but Oren and Nayar [21] developed a reflectance model to show that this is an incorrect assumption when considering rough surfaces.

2.2 Geometric properties of highly reflective surfaces

Depending on the incident ray, the reflected ray will either be intense (i.e. easily noticeable) or weak. This section presents some of the terms used to account for the reflected ray's intensity. Radiometric terms that are related and used to define properties of reflective surfaces are presented below, as discussed in [22]. Detailed derivations and descriptions of these terms are given by Nicodemus et al. [7].

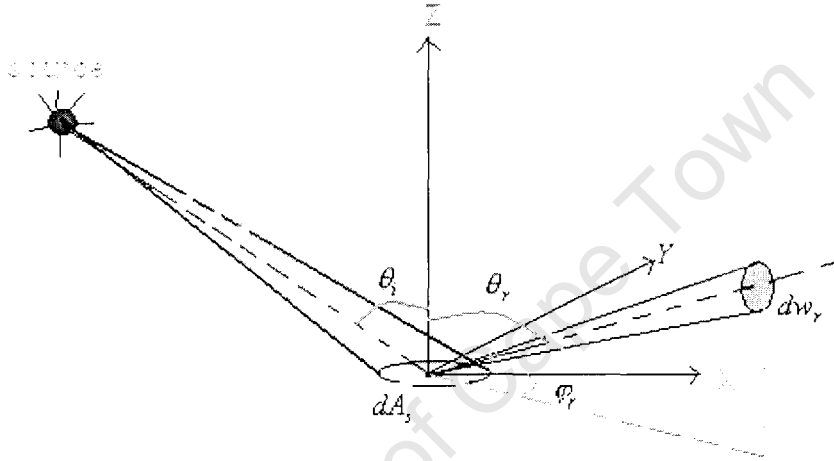


Figure 5: The geometry needed to define radiometric terms [22].

In Figure 5, the source of light is assumed to lie in the $x - z$ plane and it is defined by the angle θ_i . The monochromatic flux $d\Phi_i$ is incident on the surface area dA_s from the direction θ_i . Some of the flux $d^2\Phi_r$ is reflected in the direction (θ_r, Φ_r) [22]. The irradiance (the density of radiation incident on a given surface) I_s of the surface is defined by the incident flux density (Φ_i) over the incident area (dA_s),

$$I_s = \frac{d\Phi_i}{dA_s}. \quad (1)$$

The radiance, L_r , is the flux density of emitted energy per unit solid angle and per unit projected area of the radiating surface. In the direction (θ_r, Φ_r) it is defined as

$$L_s = \frac{d^2\Phi_r}{dA_s \cos \theta_r d\omega_r}. \quad (2)$$

The bi-directional reflectance distribution function f_r of a surface is the measure of how bright the surface appears when it is illuminated from one direction and viewed from another. The

measure is defined as

$$f_r = \frac{L_r}{I_s}. \quad (3)$$

As explained in [22], the way in which light is reflected off a surface depends on the microscopic shape characteristics of the surface. A smooth surface tends to reflect incident light in a single direction while a rough surface will scatter the incident light in all directions as shown in Figure 6. In Figure 6(b), the reflected ray from the rough surface is dispersed in multiple

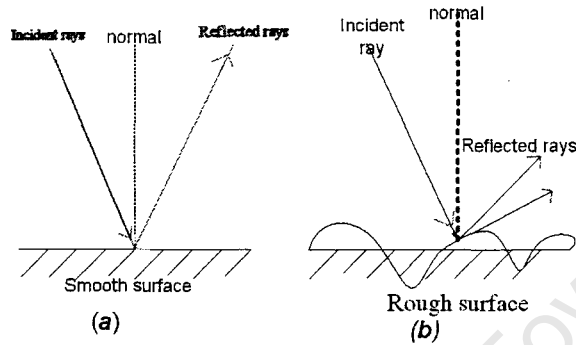


Figure 6: Rays reflected off (a) a smooth surface and (b) a rough surface.

directions. This means that for an image ray to be visible in a sensor the model of the surface has to be known so as to predict where to place the sensor.

2.3 Implications of intensity of a point and inaccuracy in stereo analysis

The amount of light reflected off a surface depends on the incident beam and the type of material of the surface. An observer will see a percentage of the reflected beam depending on the amount of light that was absorbed, transmitted, scattered off the material, or the angle at which the observer is from the surface.

The total intensity at a point in a scene is a sum of the intensity due to diffuse and specular reflections [21]. Hence, given a pair of corresponding images, the intensity at corresponding points is different because each component varies due to the change in viewing direction. This is sometimes the reason why an image point is easily viewable in one sensor but not in the other. If the detection of a point is incorrect, this can lead to false calculations of depth, consequently leading to inaccurate results. An example of such a result is shown in Figure 7.

In Figure 7, it is shown that the virtual features of a point (when matched) produce the wrong depth estimation of that point on the sphere. The estimated depth is at point X. Depending on whether a surface is concave/convex the estimated depth will either be below/above the surface [21]. Distinguishing between a virtual feature and a real one is not an easy task since the photometric properties, e.g. brightness, colour and polarization of both features could be identical [14]. Several techniques have been developed to recover the shape of a specular surface using illumination/structured light. Some of these include multiple viewpoint images

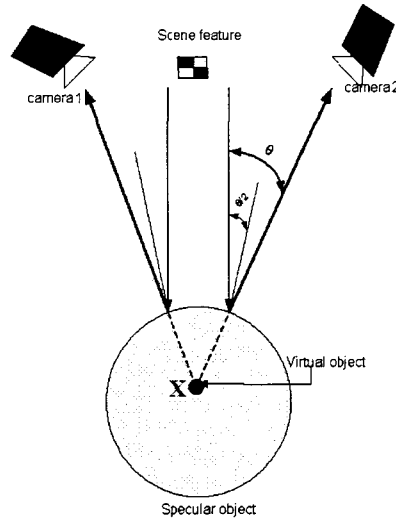


Figure 7: An example to show that specular surfaces can cause existing shape recovery algorithms to produce results that do not lie close to the desired result [21].

of a specular surface to iteratively recover shape, and developing an irradiance map of the illumination to compute the surface that will minimize discrepancies between rendered and measured images [13].

Methods developed in Ikeuchi [10] show that surface reconstruction of specular objects using photometric stereo are possible. This method different from the methods used in the prototype that has been developed for this research (i.e. stereo-vision using a laser scanner). Ikeuchi [10] describes photometric stereo as the process of determining the local surface orientation of specular materials from intensity information under different illumination but from the same view position. This is further explained as:

A pair of images of the same object are obtained from the same viewpoint while varying the direction of incident illumination. Different pairs of reflectance maps are obtained, because the reflectance map depends on the direction of the light source. Hence surface orientation is determined locally with the help of intensity pairs recorded at each image point as an intersection of constant brightness lines on reflectance maps [10].

Figure 8 shows images of a laser line when projected on three different surfaces. Image A is an ideal situation, where the image of the laser reflection is clear for each laser scan. Image B and C are laser images of reflection obtained from specular surfaces. In images B and C, multiple peaks (laser reflections) are observed in the same camera scan. This makes it difficult choosing the highest peak value in the scan thus making it difficult to process.

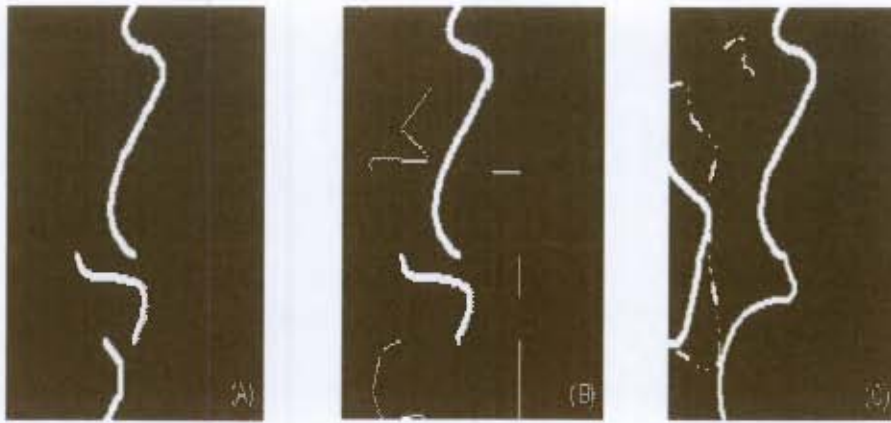


Figure 8: Camera images of a laser when projected on three different objects. (A) Object with diffuse surface. (B) and (C) objects with specular surfaces [14].

2.4 Types of surfaces

Reflections off a surface depend on the material properties of the surface. Below is a brief description of the surface types of that can be encountered.

2.4.1 Rough surfaces

When defining roughness of a surface, surface irregularities are related to the wavelength and the angle of incidence of the light beam. The roughness of a surface is analyzed by looking at how the incident light for a given wavelength scatters (in different directions). If the irregularities on the surface are smaller than the wavelength of the incident light, most of the light will be reflected specularly in a single direction [22]. If the surface regularities are

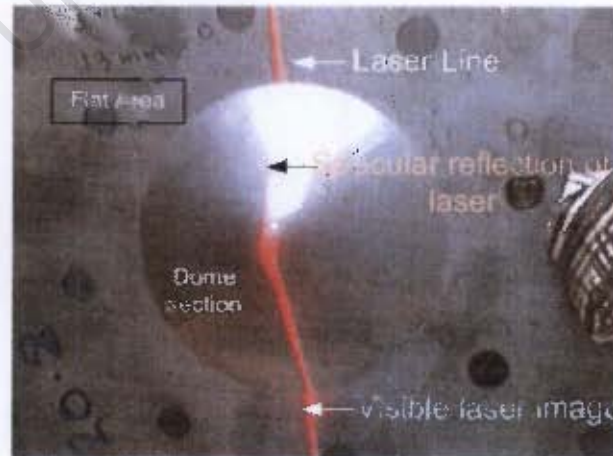


Figure 9: Image showing laser light directed on the blast plate [22].

larger compared to the wavelength, the incident light will be scattered in all directions. A surface can appear rough or smooth depending on the wavelength of the incident light, or the angle at which it hits the surface. An example of this case is shown in Figure 9. The laser beam is directed towards the blast plate. The reflected beam is clearer on the flat area of the plate and the lower part of the dome. When it hits the upper section of the dome at a different angle, the reflected beam is scattered and is seen as white light.

2.4.2 Lambertian surfaces

A Lambertian surface [22] appears equally bright from all directions, which means its appearance is independent of the viewing direction. As explained in [22] there are two mechanisms for producing a Lambertian surface. These are explained below:

- In one mechanism the incident light is reflected many times by the surface's irregularities before it is scattered into space. The multiple reflections occur in a random manner resulting in scattering of light in all directions as shown in Figure 10(a).
- The other mechanism occurs when the incident light encounters microscopic non-homogeneities once it has penetrated the surface. These irregularities cause internal reflection as illustrated in Figure 10(b). The incident light ray is repeatedly reflected and refracted at the different boundaries between the differing refractive indices/regions. Some of the scattered rays find their way to the surface with varying directions. This result is diffuse (scattered) reflection.

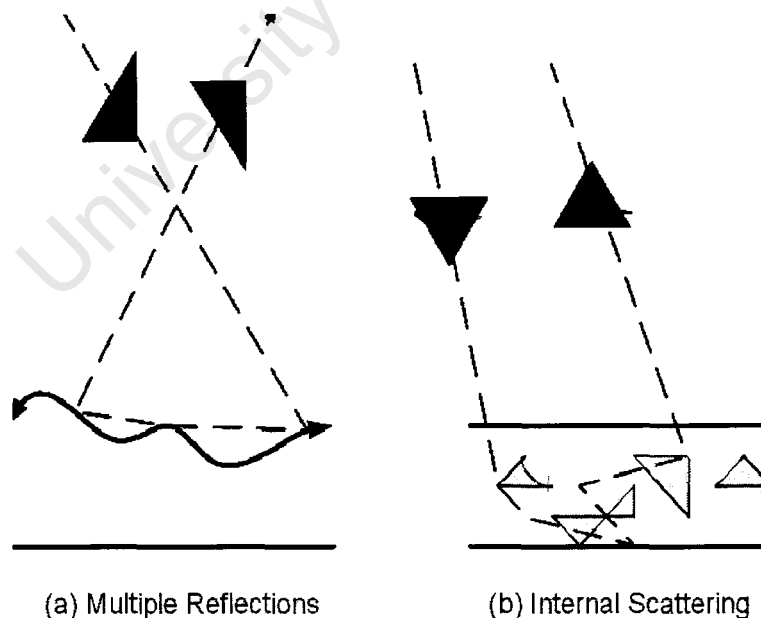


Figure 10: Diffuse reflection resulting from multiple reflection and internal scattering mechanism [22].

2.5 Surface related problems encountered in range sensing

In [21] Oren presents some of the issues associated with surface-related properties in range sensing. Of the various surface properties, specularity is the most problematic. This poses the question as to how accurately the image in the sensor corresponds to the actual feature of the object. The ambiguity that specular surfaces introduce in image analysis is whether an image feature is real or virtual. A real image feature corresponds to a physical scene point such as a surface marking, while a virtual image is the reflection, by the specular surface, of another physical point [21].

Specular surfaces can cause inaccuracy in image processing. For example, if a prototype to track an object is built for a particular ecological system, and if there are any specular surfaces in the vicinity, their reflections move with the change in light even though they themselves are stationary. Hence with the change in angle of observation there is change in the position of a stationary object according to the prototype. This will lead to inaccurate tracking.

2.6 Related works on surface reconstruction and measurement of specular objects

Several techniques have been developed to recover the shape of a specular surface using illumination/structured light. Some of these include multiple viewpoint images of specular surface to iteratively recover shape [13], developing an irradiance map of the illumination to compute the surface that will minimize discrepancies between rendered and measured images [29] and binocular stereo [2]. Methods in Ikeuchi [10] show that surface reconstruction of specular objects using photometric stereo is possible.

From the analyses of radiance behavior of types of surfaces given in literature an approach of scanning the unpainted plate using a laser spot while changing the angle of incident of the laser is investigated. The results were analyzed to check the concept's credibility (i.e. scanning while changing direction of light). The next chapter presents three methods that can be used to obtain a surface reconstruction of an object.

3 Different methods used to obtain range data of an object for measurements

Over the years, research in computer vision and robotics has investigated how to obtain a 3-D model of an object using non-contact methods. The following section presents some of these methods. The choice depends on the type of surface of the object, or on its accessibility.

In order to obtain a 3-D model of an object a range point for every accessible point on the object is measured using a rangefinder device [26]. These usually use a set of known distances or target sizes to make the measurements. Range-finders are used extensively in 3-D object recognition, 3-D object modeling, and a wide variety of computer vision related fields. Below are different methods of obtaining the range of a point from a reference point in the world.

3.1 Direct and active methods

Active range acquisition methods involve a controlled energy beam and reflected energy beam. These include ultrasonic, radar, laser and light time of flight estimation and triangulation [16]. An example is given below.

A laser range-finder is a device which uses a laser beam to determine the distance from a reflective object. Based on the information received from the echoed laser beam, one can determine the position of the reflective surface relative to the laser source [16]. The most common form of laser range-finder operates on the time-of-flight principle by sending a laser pulse in a narrow beam towards the object and measuring the time taken by the pulse to be reflected off the target and return to the sender.

One of the advantages that laser range-finders offer over other methods of computer vision is that the computer does not need to correlate features from two images to determine depth information as in stereoscopic methods [26].

3.2 Passive monocular image-based methods

Passive monocular image-based methods include:

- Texture gradient analysis [26]: A texture gradient is a surface pattern that provides information about the distance of an object [30]. This is another way of perceiving depth. When looking at an image the resolution of the details/texture that is seen decreases with increasing distance from the observer. This can be seen in Figure 11.
- Photometric stereo [10]: This is where surface normal vectors from reflectance are analyzed to determine distance. In cases where the analyzed surface is specular/Lambertian, various steps that best suit the degree of reflection have been developed and researched. In [10] Katsushi explains that photometric stereo is the determination of local surface

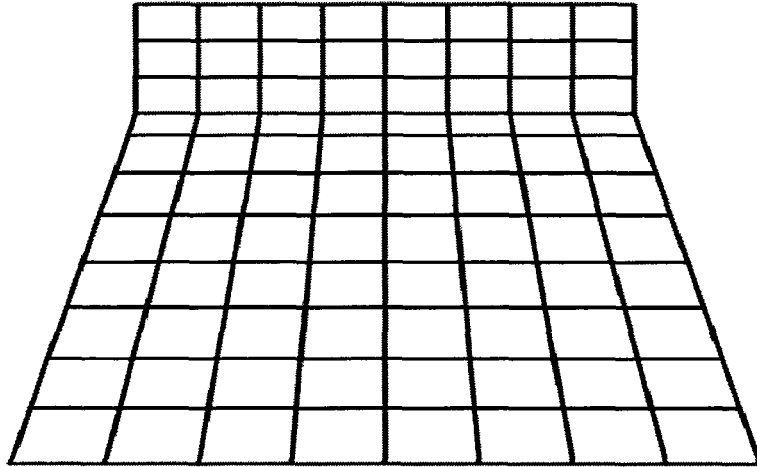


Figure 11: Diagram showing perception of depth with change in gradient [30].

orientation of specular materials from the intensity information under different illumination, but from the same viewing point. This method depends on reflectance maps which were obtained from previous experiments. Reflectance maps represent the relationship between surface orientation and image intensity in gradient space.

3.3 Motion and multiple relative positions of camera scene

The above method includes reconstruction from multiple views, stereo disparity and retinal flow [26]. Stereo-vision consists of using multiple views of the same point object and geometry of calibrated cameras to estimate the depth of that point object. The method of obtaining the multiple views depends on the type of surface being analyzed. In situations where the surface is matte (not shiny), structured light is used. Laser scanners using triangulation are a popular methods for acquiring 3-D geometry due to their accuracy, simplicity and robustness [11]. There are different setups of laser scanners. These are listed below.

3.3.1 Setup comprised of calibrated camera and movable laser plane

The traditional laser scanner consists of a movable laser and a camera. This is shown in Figure 12. A calibrated geometry exists between the plane of the laser light and the camera. By using triangulation for the observed laser image and the known laser plane, a single strip of depth can be recovered for every scan [11]. The calibration involves static geometry and optics as well as actuated components. The relationship between the laser stripe, the shape of the laser stripe, and the repeatability and motion of any motorized component must be completely characterized in order to produce an accurate depth image [11]. The components must be characterized in terms of geometry, velocity and acceleration.

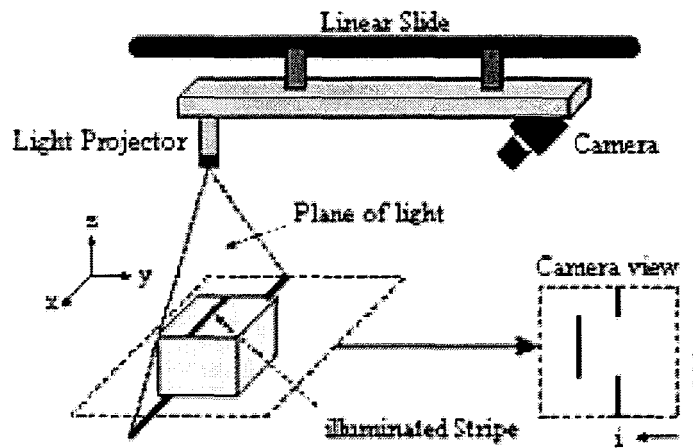


Figure 12: Single camera and laser plane [16].

3.3.2 Setup comprised of two calibrated cameras

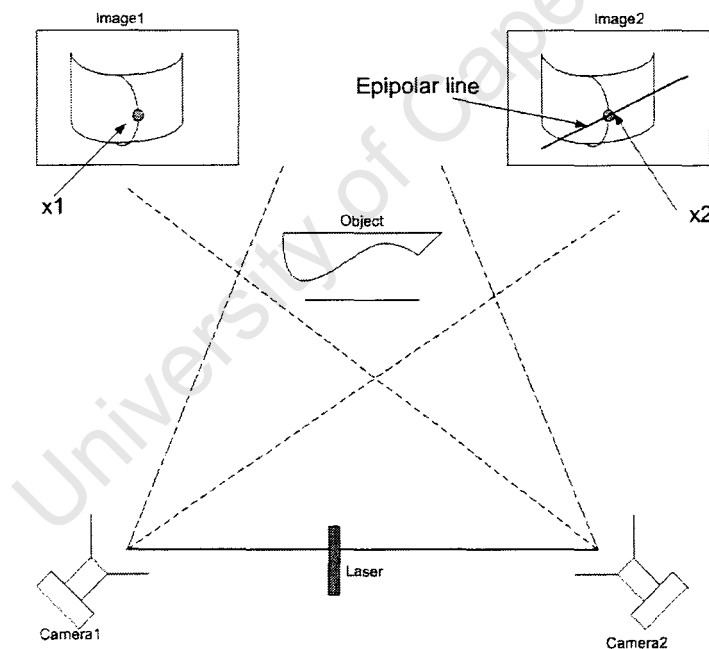


Figure 13: Systematic diagram of stereo using multiple camera setup [9].

A setup consisting of two calibrated static cameras can be used to locate a laser stripe and recover the depth of the un-calibrated stripe swept over the object [11]. Depth is recovered through triangulation between the two known viewpoints. This setup minimizes complexity as it removes actuated components from the calibration geometry. An image of this setup is shown in Figure 13. Since it is hard to ensure absolute repeatability of moving parts, this design has only static components, excluding the laser which is not calibrated into the system.

This setup is easier to use the one presented in Figure 12 as it does not includes movable parts which are subjected to wear.

In [20] Levoy reports that calibration is the most difficult and ongoing problem. The movable components create problems with repeatability and error modeling. The inaccuracy in measurements are mainly caused by inaccurate calibration parameters.

For this thesis the two calibrated camera setup was used to obtain the range of the metal plate. Further explanation on the epipolar geometry used to estimate depth is explained in Chapter 4.

University of Cape Town

4 Camera modeling and epipolar geometry

This section describes how an object in real world relates to a 2-D image produced by a camera. The camera's geometry can be modeled and expressed as a mathematical relation. By estimating parameters in these relationships a camera can be calibrated from a reference point.

4.1 Perspective projection of a Pinhole Camera

In this section the relationship between an object in 3-D and its camera image in 2-D is discussed using a simple model of pinhole perspective camera. The geometry of a pinhole camera is presented in Figure 14.

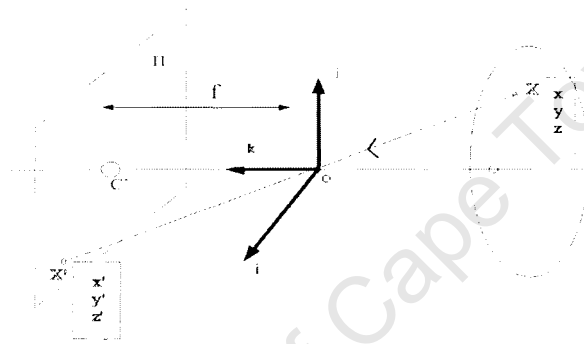


Figure 14: Setup for deriving the equations of the pinhole camera perspective projection [5].

In Figure 14, consider the co-ordinate system (O, i, j, k) to be a point attached to a pinhole camera. The vectors j and i form the basis of a vector plane parallel to the image plane π , which is located a distance f from the origin of the pinhole camera. The line perpendicular to π and passing through the pinhole is known as the optical axis. Point C' on π is the image center. This is used as the origin of the image plane coordinate frame and it plays an important role in camera calibration [5].

In Figure 14, X represents a scene and X' denotes its image with coordinates (x, y, z) and (x', y', z') respectively. Forsyth [5], shows that if X' lies in the same image plane as C' , then $z' = f$. Since the three points X, O, X' are collinear, $OX' = \lambda OX$ where λ is a constant. From this analysis the geometry in the imaging process is:

$$\begin{aligned} x' &= \lambda x \\ y' &= \lambda y \Leftrightarrow \lambda = \frac{x'}{x} = \frac{y'}{y} = \frac{f}{z} \\ f &= \lambda z \end{aligned} \tag{4}$$

Therefore the perspective projection equations of a point from 3-D to 2-D of a camera of focal

length f are:

$$\begin{aligned} x' &= f \frac{x}{z} \\ \text{and} \\ y' &= f \frac{y}{z}. \end{aligned} \tag{5}$$

However this assumes that the origin of the image is at the principal point. This is generally not true in practice. In [9] a practical mapping of 3-D to 2-D is considered to be:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} f \frac{x}{z} + c_x \\ f \frac{y}{z} + c_y \end{pmatrix} \tag{6}$$

where the principal point with coordinates c_x and c_y is introduced. This can further be expressed in homogeneous coordinates as:

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} f \frac{x}{z} + c_x \\ f \frac{y}{z} + c_y \\ 1 \end{pmatrix} = \begin{bmatrix} f & \alpha & c_x & 0 \\ & f & c_y & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{7}$$

where α is a skew factor, which is normally zero. The parameter α is further explained in section 5. From the above mapping the camera calibration matrix \mathbf{K} is then expressed as:

$$\mathbf{K} = \begin{bmatrix} f & \alpha & c_x \\ & f & c_y \\ & & 1 \end{bmatrix}. \tag{8}$$

The relationship between the image coordinates \mathbf{x}_{img} and the camera point \mathbf{X}_{cam} can then be expressed as $\mathbf{x}_{img} = \mathbf{K}[I|0]\mathbf{X}_{cam}$ [9]. Explanation from \mathbf{X}_{wld} to \mathbf{X}_{cam} is explained in section 4.2. Explanation of the camera matrix \mathbf{K} is given in Chapter 5.

4.2 Reference frames

There are different spaces and corresponding reference frames involved in mapping points from the world to images. These are explained below.

4.2.1 World reference frames

Objects in the real world are known to be in three-dimensions, but the human eye sees them in two dimensions. The human vision system helps put the two-dimensions into three-dimension. This enables a person to know how far an object is from a reference point. The world reference frame is used to associate the position of the object in space [18]. The world

reference frame is a right-handed Cartesian coordinate frame consisting of three orthogonal vectors of unit length. A point $\hat{\mathbf{X}}$ in the world is specified by an ordered triple of numbers using the following notation:

$$\hat{\mathbf{X}} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (9)$$

A world point $\hat{\mathbf{X}}$ represented in homogenous co-ordinates by \mathbf{X}_{wld} , is mapped to its camera point \mathbf{x}_{cam} through the camera matrix \mathbf{P} with the following equation;

$$\mathbf{x}_{\text{cam}} = \mathbf{P}\mathbf{X}_{\text{wld}} \quad (10)$$

where both \mathbf{X}_{wld} and \mathbf{x}_{cam} are both homogenous vectors.

4.2.2 Image reference frame

There are two image planes associated with a camera. One is the normalized image plane when f , the focal distance, is a unit length. The other is the physical image plane associated with the camera (i.e. virtual image, image standing upright). This image plane is described by the intrinsic parameters of the camera found in \mathbf{K} [18]. The intrinsic parameters are obtained from the calibration of two cameras. A presentation of this can be found in Figure 15.

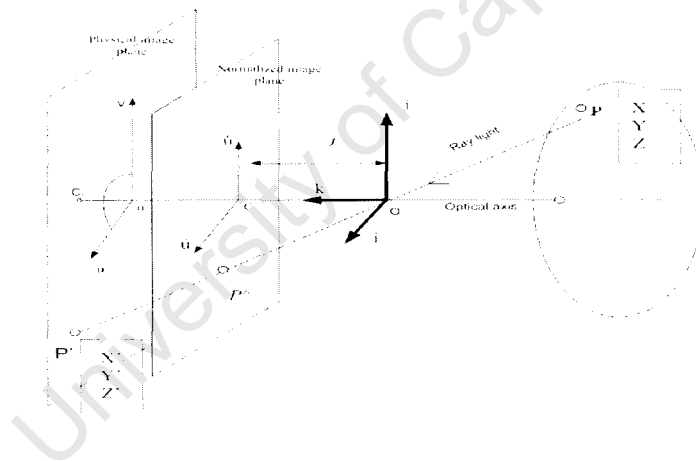


Figure 15: Image and camera coordinate system [18].

4.2.3 Camera reference frame

For a point in space, a reference frame can be defined to coincide with the pinhole. This reference frame is known as the camera reference frame. The camera's principal axis coincides with the basis vector normal to the image plane [9]. The world reference and camera reference frame will not coincide because they are separated by transformation brought about by a rotation and translation $[\mathbf{R}|\mathbf{t}]$.

4.3 Epipolar geometry

This section defines epipolar or two view geometry. In order to understand 3-D reconstruction, an understanding of the camera's geometry is required. In [5] Forsyth discusses that it is impossible to obtain the depth of a point from a single picture along the corresponding projection ray. However with two or more views, it is possible through triangulation. A discussion on what factors to consider in finding the relationship between corresponding image points $\mathbf{x} \Leftrightarrow \mathbf{x}'$ and their world point \mathbf{X} is presented below (as adapted from [5]).

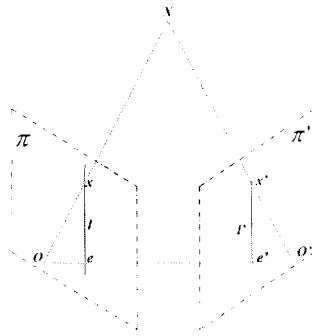


Figure 16: Epipolar geometry of two cameras with optical centers O and O' . Images \mathbf{x} and \mathbf{x}' of \mathbf{X} lie in the same plane (shaded region) [5].

Consider an instance of a world point \mathbf{X} with point images \mathbf{x} and \mathbf{x}' obtained from cameras with camera centers O and O' . The point of intersection of the line joining the camera centers O and O' with the image plane Π' , is the epipole e' . This is shown in Figure 16. The images \mathbf{x} and \mathbf{x}' of point object \mathbf{X} observed by the two cameras belong to the epipolar plane defined by the two intersecting rays $O\mathbf{X}$ and $O'\mathbf{X}$. If \mathbf{x} and \mathbf{x}' are images of the same point, then \mathbf{x}' must lie on the epipolar line associated with \mathbf{x} , namely l' [9]. This epipolar constraint plays a fundamental role in stereo-vision. The most difficult part in stereo data analysis is establishing correspondence between points in two images [5].

Assuming that the intrinsic and extrinsic parameters of the camera are known from calibration and using the epipolar geometry the following equation is obtained;

$$\mathbf{Ox} \cdot [\mathbf{OO}' \times \mathbf{O}'\mathbf{x}'] = 0. \quad (11)$$

This coordinate-independent equation can be written in association to the first camera as

$$\mathbf{x} \cdot [\mathbf{t} \times \mathbf{R}\mathbf{x}'] = 0. \quad (12)$$

$\mathbf{x} = [u, v, 1]^T$ and $\mathbf{x}' = [u', v', 1]^T$ denote the homogenous image coordinates of \mathbf{x} and \mathbf{x}' [5]. The vector \mathbf{t} is the coordinate vector of the translation \mathbf{OO}' , that separates the two coordinate system in the two cameras. \mathbf{R} is the Rotation matrix such that a vector \mathbf{v} in the second camera corresponds to $\mathbf{R}\mathbf{v}$ in the first. With further analysis equation 12 can be written as:

$$\mathbf{x}^T \mathbf{F} \mathbf{x}' = 0 \quad (13)$$

where \mathbf{F} is the fundamental matrix. \mathbf{F} can also be used to define the relationship between the two camera calibration matrices as:

$$\mathbf{F} = \mathbf{K}^{-T} \mathbf{E} \mathbf{K}'^{-T}. \quad (14)$$

For normalized cameras equation 13 can be written as:

$$\hat{\mathbf{x}}^T \mathbf{E} \hat{\mathbf{x}}' = 0. \quad (15)$$

\mathbf{E} is the essential matrix, and $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ are the normalized image coordinate vectors. Normalization is a situation where points are referred to the normalized image plane [9]. The relationship between normalized points and points that have not been normalized is shown below:

$$\mathbf{x} = \mathbf{K} \hat{\mathbf{x}} \quad \text{and} \quad \mathbf{x}' = \mathbf{K}' \hat{\mathbf{x}}'. \quad (16)$$

For a case with more than two views e.g. a three-view situation, each camera pair is defined by an epipolar constraint:

$$\begin{aligned} \hat{\mathbf{x}}_1^T \mathbf{E}_{12} \hat{\mathbf{x}}_2 &= 0 \\ \hat{\mathbf{x}}_2^T \mathbf{E}_{23} \hat{\mathbf{x}}_3 &= 0 \\ \hat{\mathbf{x}}_3^T \mathbf{E}_{31} \hat{\mathbf{x}}_1 &= 0. \end{aligned} \quad (17)$$

Here \mathbf{E}_{ij} denotes the essential matrix associated with the different images. These constraints are not independent of each other since $\mathbf{e}_{31}^T \mathbf{E}_{12} \mathbf{e}_{32} = \mathbf{e}_{12}^T \mathbf{E}_{31} \mathbf{e}_{21} = \mathbf{e}_{23}^T \mathbf{E}_{31} \mathbf{e}_{21} = 0$. Further explanations can be found in [5]. Further properties of \mathbf{F} and \mathbf{E} and how they relate to 3-D reconstruction are explained in Section 4.4.

4.4 The fundamental matrix \mathbf{F}

The fundamental matrix is the algebraic representation of the epipolar geometry [9]. For a pair of matching points in two images \mathbf{F} is defined by the equation [9]

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0. \quad (18)$$

At least seven pairs of corresponding points are required to compute the fundamental matrix. This is because for a given two-view geometry, the fundamental matrix \mathbf{F} has to be a homogeneous matrix of size 3×3 , which has nine elements. However, the common scaling of the elements is not significant, which eliminates one degree of freedom. A further constraint is that the determinant of \mathbf{F} is zero, which means that \mathbf{F} is not invertible. This eliminates a second degree of freedom. From this analysis the fundamental matrix \mathbf{F} has $9 - 2 = 7$ degrees of freedom. Hence a minimum of seven pairs of corresponding points is needed to estimate the fundamental matrix.

In [9] it is further explained that with seven corresponding points a non-linear solution of \mathbf{F} is obtained. To be able to solve the equation linearly at least eight points are needed.

This point correspondence enables the equation to be resolved using a least-squares solution (SVD)². With exactly eight points it is possible to get a unique solution. The 8-point algorithm from [9] is used in this project to compute the fundamental matrix. In order to have a proper working method, the input data (corresponding points \mathbf{x} and \mathbf{x}') should be properly normalized before constructing the Equation 19 [9].

Since the fundamental matrix is the algebraic presentation of the epipolar geometry it is used in point correspondence. Given two corresponding images, for an image point \mathbf{x} in Image1, there exists a corresponding epipolar line l' in the other image. On this line l' lies the corresponding point \mathbf{x}' . The mapping defining this constraint is $\mathbf{x} \mapsto l'$ and its mathematical equation is [9]:

$$l' = F\mathbf{x}. \quad (19)$$

4.5 The Essential Matrix

The essential matrix is the specialization of the fundamental matrix of the case of normalized image coordinates $\hat{\mathbf{x}}'$ and $\hat{\mathbf{x}}$. The fundamental matrix can further be thought of as the essential matrix in which the assumption of calibrated camera matrix (\mathbf{K}) is removed. In [9], Hartley and Zisserman show that normalized points are image coordinates, obtained from the camera matrix, where the effect of the calibration matrix \mathbf{K} has been removed. The results from this are that the camera matrix \mathbf{K} becomes $\mathbf{K} = \mathbf{I}$ (identity matrix). The mathematical presentation of the normalized points become $\hat{\mathbf{x}} = [\mathbf{R}|\mathbf{t}]\mathbf{X}$. The nine coefficients of the essential matrix are parameterized by the three degrees of freedom of the rotation matrix \mathbf{R} and the two degrees of freedom defining the direction of the translation vector \mathbf{t} . The defining equation for corresponding normalized coordinates $\mathbf{x} \leftrightarrow \mathbf{x}'$ is

$$\hat{\mathbf{x}}'^T \mathbf{E} \hat{\mathbf{x}} = 0. \quad (20)$$

Compared to the fundamental matrix, the essential matrix has fewer degrees of freedom, namely five, and more constraints [9]. One of the constraints of an essential matrix is that two of its singular values are equal and the third is zero. This contributes to the reduction in its degrees of freedom.

4.6 Obtaining the camera matrices \mathbf{P} and \mathbf{P}'

The camera matrix is a 3×4 projective matrix. The method used to compute a camera matrix is known as resectioning [9]. The camera matrices needed in a 3-D reconstruction can be estimated from point correspondence of 2-D image points to 3-D points. It is usually advisable to work in a normalized coordinate frame. Normalized camera matrices can be estimated from the essential matrix \mathbf{E} . These are called normalized because the effect of the calibration matrix \mathbf{K} is removed.

²SVD : Singular Value Decomposition: Detailed explanation of SVD can be obtained in [9].

For a given set of data from two views, the first normalized camera matrix \mathbf{P} is positioned at the origin and it is symbolized by the identity matrix, $\mathbf{P} = [\mathbf{I}|0]$. The second normalized camera matrix $\mathbf{P}' = [\mathbf{R}|\mathbf{t}]$, is obtained from the decomposition of the essential matrix, \mathbf{E} [9]. Singular value decomposition (SVD) is used in the decomposition of \mathbf{E} . From this decomposition the results of four possible matrix solutions of \mathbf{P}' are obtained. The solutions obtained from the decomposition can place the 3-D point of $\mathbf{x} = \mathbf{P}'\mathbf{X}$ in four possible positions: behind both cameras, in front of both cameras, behind camera one but in front of camera two, or behind camera two and in front of camera one. The correct matrix is the one that corresponds to a positive depth of corresponding points (i.e. in front of both cameras).

4.7 Triangulation of the corresponding points

After obtaining corresponding points, their depth is estimated using triangulation. Consider Figure 16: from the epipolar constraint \mathbf{x}' corresponds to the epipolar line l , hence $\mathbf{x} \leftrightarrow \mathbf{x}'$ are corresponding points. For the two rays back-projected from images, \mathbf{x} and \mathbf{x}' lie in a common epipolar plane and intersect at point \mathbf{X} on the plane. Because of the error found in estimating \mathbf{x} and \mathbf{x}' , the back-projected rays may not intersect. A method that finds a point in the world that projects back into the image closely as possible to the original image points is needed. A representation of this is shown in Figure 17.

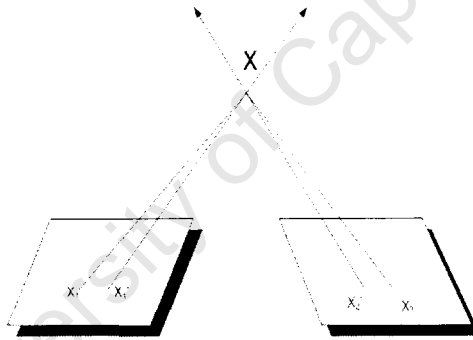


Figure 17: Image showing triangulation error.

The method must also be invariant to transformations of the appropriate class [9], i.e.

$$\tau(\mathbf{x}, \mathbf{x}', \mathbf{P}, \mathbf{P}') = \mathbf{H}^{-1}\tau(\mathbf{x}, \mathbf{x}', \mathbf{P}\mathbf{H}^{-1}, \mathbf{P}'\mathbf{H}^{-1}) \quad (21)$$

where τ is the triangulation method and H is the transformation. Such a triangulation process is projective-invariant. This method minimizes the image distances $(x_1 - x'_1)$ and $(x_2 - x'_2)$. There are a number of methods used to obtain point \mathbf{X} . The method used in this project is linear homogeneous triangulation.

5 Camera calibration

Camera calibration in the context of three dimensional reconstruction is the process of determining the internal camera geometric and optical characteristics (intrinsic parameters) and/or the 3D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters) [28]. When cameras are calibrated this makes the task of recovering a scene simpler. Cameras are said to be calibrated if the mapping between image coordinates and directions relative to the camera center are known.

The calibration process allows calculation of both the intrinsic and extrinsic parameters i.e. focal distance coordinates, principle points, radial and tangential distortions. The extrinsic values include positions and orientations with respect to an absolute reference system.

In order to obtain the intrinsic and extrinsic parameters a readily available Matlab camera calibration toolbox was downloaded from [4].

5.1 Intrinsic camera parameters

As discussed in section 4, in an ideal pinhole camera a representation of a perspective image is characterized by parameters called the intrinsic parameters. These are listed below as obtained in [3]:

- The x co-ordinate of the center of the camera, in pixels (c_x)
- The y co-ordinate of the center of the camera, in pixels (c_y)
- The focal length, in pixels (f)
- The skew coefficient defining the angle between the x and y pixel axes, α
- The aspect ratio(s)

The presentation of the intrinsic parameters is in the form of a matrix and this is shown below.

$$\mathbf{K} = \begin{bmatrix} f & \alpha & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (22)$$

The skew parameter α is assumed to be zero. It is non zero when there is skewing of the pixel elements in the CCD array, such that the x - and the y - axes are not perpendicular. The aspect ratio s is assumed to be 1 [4]. When using the Matlab calibration toolbox, the intrinsic camera matrix is provided as a variable KK , after calibration. This matrix is used to convert between the retinal (normalized) plane and the physical plane. The pixel co-ordinate vector \mathbf{x}

relates a point on the normalized image plane as

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & \alpha & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x_d \\ y_d \\ 1 \end{pmatrix}. \quad (23)$$

5.2 Extrinsic camera parameters

The extrinsic camera parameters consist of the rotation matrix \mathbf{R} and the translation vector \mathbf{t} [3]. The camera extrinsic parameter rotation matrix, \mathbf{R} (3×3) and translation vector, \mathbf{t} relate a point in a world coordinate system to a point in the camera coordinate system. Given a point in space $\mathbf{X} = (x; y; z)$ and a point on the camera reference frame $\mathbf{X}_c = (x_c; y_c; z_c)$, the relation between the two points is [3]:

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \mathbf{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \mathbf{t}. \quad (24)$$

Real cameras deviate from the pinhole camera in several ways. One of these is that in terms of camera calibration, each image point corresponds not to a single ray from the camera center, but to a set of rays from across the front of the lens all converging on a particular point on the focal plane. The effects of this can be minimized by making the camera aperture small. Another difference is that due to various constraints in manufacturing, straight lines in the world imaged through a real lens become curved on the image plane. This is known as lens distortion.

5.3 Lens distortion in cameras

Distortion exists in an image but cannot generally be seen by the naked eye. In cameras radial distortion becomes prominent with a decrease in focal length. There are mainly two kinds of radial distortion: barrel distortion and pincushion distortion [28]. Barrel distortion is due to the fact that many wide angle lenses suffer from higher magnification in the center of the image than at its edges. This causes the image to expand around the center and shrink around the edges, forming the shape of a barrel. This is illustrated in Figure 18. The pincushion is the inverse effect of the barrel, i.e. the edges are magnified more than the shape.

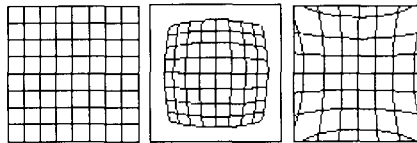


Figure 18: Illustration of lens distortion. The original grid (left), and the effect of barrel (center), and pincushion distortion (right) [17].

5.4 Distortion models

Radial distortion is one of the many aberrations altering the geometry of the projected image [9]. In the Matlab calibration toolbox, the tangential and radial distortions are estimated and stored as a (5×1) vector kc . In camera geometry an accurate model is one whose model is linear which means the world points, image points and optical center are all collinear.

In [9], Hartley et. al. explain that an ideal (undistorted) point (x_{und}, y_{und}) is related to an actual image point after radial distortion (x_d, y_d) by a radial distortion factor. This radial distortion is modeled as:

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = L(r) \begin{pmatrix} x_{und} \\ y_{und} \end{pmatrix}. \quad (25)$$

where (x_{und}, y_{und}) is the ideal image position (which obeys linear projection). For a measured point (x, y) with a center of radial distortion as (x_c, y_c) , the radial distance r can be estimated as

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2}. \quad (26)$$

$L(r)$ is the distortion factor which is a function of the radius r :

$$L(r) = 1 + k_1 \times r + k_2 \times r^2 + k_3 \times r^3 + \dots \quad (27)$$

From the above equations for a measured set of points (x, y) , its corrected coordinates (x_{cc}, y_{cc}) , can be calculated using

$$\begin{aligned} x_{cc} &= x_c + L(r)(x - x_c) \\ y_{cc} &= y_c + L(r)(y - y_c). \end{aligned} \quad (28)$$

After looking at the calibration parameters and their representations the steps taken in calibrating the stereo-rig are presented in Section 7. Section 7 further explains the design of the system and some of the steps taken in processing the captured data.

6 Registration of data using Iterative Closest Point (ICP) algorithm

6.1 Introduction

Registration of data is the process of estimating a transformation Q , that is applied to two sets of data so as to bring the pair to the same co-ordinate system [19]. The registration technique used to carry this out is the ICP algorithm developed by Besl and McKay [1].

The algorithm is used to align data or find proper relative positioning of digitized data from un-fixtured rigid objects with an idealized geometric model prior to shape inspection. In this algorithm neither extracted features nor curve or surface derivatives are required and the data does not have to be preprocessed except for removal of statistical outliers [1].

ICP algorithm is widely used in alignment of three-dimensional point sets. The ICP mechanism is used when point correspondence between two data sets is required and it is unknown. After establishing point correspondence between two point data, alignment of the two sets possible. The quality of the alignment depends on pairing of good corresponding points from the two sets of data. If an object is highly detailed, like the model of the rabbit in Figure 18, a high resolution camera is needed to obtain a high quality image of the rabbit so as to eventually obtain a good reconstruction.



Figure 18: Image of a rabbit reconstructed by the Stanford research group with high-resolution cameras [15].

A single reconstruction (from one view) of an object can only contain as many points of the scene as it is able to process. In order to obtain a complete scene model, multiple reconstructions from different views have to be merged. Registration and integration of the data is required to obtain the merge [19]. In other cases registration is used to validate two or more surface reconstructions of the same object from different views.

6.2 The ICP algorithm

The ICP algorithm involves first solving the closest point problem for a pair of data and then determining the transformation Q required to align the two data sets. The transformation is applied to one point set and the procedure iterated until convergence is reached. In [19] it is explained that one set is assumed to be a subset of the other data. If this condition is not met, wrong matches are made which negatively influence the convergence of the ICP algorithm.

The ICP algorithm tries to converge to a minimum with respect to the mean-square distance objective function [19]. At times even if a good initial approximation of the transformation matrix is provided, ICP does not converge to the right solution due to the closest point search method. The ICP algorithm has no reliance on feature extraction to establish correspondence between views [19]. The method instead relies on establishing closest point pairs.

6.3 Uses of ICP in image processing

As [1] already explains, ICP is a method of registering 3-D shapes. Examples of the registrations are listed below:

- Given a 'model shape' and a 'section data' that represents a certain portion of the model shape, the point registration (or common ground) can be obtained with an initial translation and rotation. ICP can be used to obtain the initial translation.
- ICP is used to decide the congruence (shape equivalence) of different geometric representations.
- ICP is also used to estimate motion between point sets where the correspondences are not known.

6.4 The working of the ICP algorithm

The registration starts with two sets of data from different views and an initial estimate of the aligning rigid-body transformation [27]. It then iteratively refines the transformation by alternately choosing corresponding points in the two data sets and finding the best translation and rotation that minimizes an error metric based on the distance between the corresponding points. The numerical step by step operation of the algorithm is explained below as adapted from [1].

6.4.1 Mathematical preliminaries

Mathematical preliminaries of computing the closest point on a shape to a given point are explained below as adapted from [1] Besl and McKay. Consider two sets of data X and X' .

The task in this algorithm is to find the transformation Q that is going to be applied to source data X . Q is used to align the data in X' with data in X so that $Q\mathbf{x} = \mathbf{x}'$ is satisfied where $\mathbf{x} \in X$ and $\mathbf{x}' \in X'$ [19]. In [19] X and X' are considered as measured and model data respectively. Myers [19] explains that $Q\mathbf{x} = \mathbf{x}'$ is only taken into account when considering an ideal situation, i.e. for real data it will not work. This is because factors such as noise that affect real data have not been taken into account.

For real data $\|Q\mathbf{x} - \mathbf{x}'\|^2$ is instead considered. The distance between the points in X and the points in X' is minimized by finding the corresponding closest points in data set X for X' . A method that uses quaternions to represent the rotation is used. A method based on singular value decomposition can also be used to compute the transformation in data with dimensions greater than three.

Calculation of the closest corresponding point is one of the concepts behind ICP algorithm. In order to do this the Euclidean distance between two points from different sets has to be known. Let the number of points in X and X' be N_x and $N_{x'}$ respectively. The Euclidean distance between any two points \mathbf{x} and \mathbf{x}' in each set is $d(\mathbf{x}, \mathbf{x}')$ where:

$$\mathbf{x} = (x_1, y_1, z_1) \quad \text{and} \quad \mathbf{x}' = (x_2, y_2, z_2) \quad (29)$$

$$d(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}. \quad (30)$$

Hence the closest distance between points in set X and X' is:

$$\min_{i=1 \dots N} d(\mathbf{x}_i, \mathbf{x}_i'). \quad (31)$$

6.4.2 The ICP algorithm is stated

In the following section an abstract on how the algorithm is carried out is given as used in [1]. The input data is two sets of digitized points obtained from surface reconstruction using a laser scanner. The objective of the algorithm is to minimize the Euclidean distances between corresponding points using least squares. The "measured" data set X is moved into best alignment with "model" data set X' . The metric distance d between a single data point \mathbf{x} and a model set X' is denoted by:

$$d(\mathbf{x}, X') = \min_{\mathbf{x}' \in X'} \|\mathbf{x}' - \mathbf{x}\|. \quad (32)$$

Let y denote the closest point in X' such that the minimum distance obtained from \mathbf{x} and a point from X' is:

$$d(\mathbf{x}, y) = d(\mathbf{x}, X'). \quad (33)$$

Let the function to denote the computation of the closest point be $O(N_x)$ and Y denotes the resulting set of closest points. C is the closest point operator. Given the resulting corresponding point set Y , the least squares registration is computed as $Y = C(X, X')$. The "measured" data is updated as $X = q(X)$ where $q(X)$ is used to denote the data set X after transformation by the registration vector q . The ICP algorithm can be outlined as:

- Point set X and X' are given.
- An initial rotation matrix and translation is defined relative to the data set X and X' , so that the final complete transformation can be estimated.
- The iteration is initialized and the closest points are computed.
- The registration needed for repositioning of data set X is computed.
- Registration is applied to data set X and the closest points are calculated again.
- The iteration is terminated when the change in mean square error of the Euclidean distance falls below a given threshold τ , or when it is stuck in a local minima for up to 6 iterations.

6.5 Disadvantages of using the ICP algorithm

In [1] it is explained that since ICP is a non-linear local search algorithm it suffers from many problems. These include:

1. Slow convergence due to shallow error landscape, and the tendency to fall in the wrong local minima. The way the corresponding points are selected and the way the metric error is minimized play a big role in the rate of convergence and the accuracy of the resulting pose.
2. If too many points are used from a featureless region of the data, the algorithm converges slowly, could find the wrong pose, or could even even diverge especially in the presence of noise [8].
3. Poor alignment between points can also be caused by noise. Noise in the input data can cause the ICP algorithm to converge to a wrong local minima [8]. The frequency of the local minima in the error landscape depends on the input geometry and on the minimized distance metric.
4. ICP algorithm is susceptible to gross statistical outliers unless a statistically robust method is substituted at some point. In this case the outliers are removed or reduced before the registration.
5. From a computational expense point of view ICP is highly time consuming because of the exhaustive search for the nearest point [19].
6. In order to get a near to zero error when using ICP algorithm, every point on one surface has to have a corresponding point on the other surface.

6.6 Advantages of ICP

Experiments in [1] were carried out aligning different forms of data from previous work; i.e. data points, curves, surfaces. From these the following conclusions were drawn.

- ICP algorithm can handle the full six degrees of freedom in a similar transformation.
- ICP algorithm is independent of shape representation.
- It does not generally require smoothing of data as long as the outliers are near zero. This is usually the case with non contact sensors used for inspection purposes.
- If data of dimension greater than three is being dealt with, the quaternion method can be substituted for by the SVD algorithm.
- The cost of aligning a data set can be predicted from complexity in shape of the object. If a simple object is being dealt with, there is a better guarantee that the data will settle at the local minima compared to if a complex object with occlusions were used.

7 System design and functioning

A laser rig was designed in the laboratory to make surface reconstructions and take measurements of the blast plate. An image of the stereo-rig is shown in Figure 19. The laser range scanner relies on a known calibrated camera geometry in order to reconstruct depth. The process used in calibrating the stereo-rig using the Matlab toolbox is explained in Section 7.2. The triangulation process requires at least two viewpoints. The arrangement as seen in Figure 16 allows triangulation between the two view rays, thus allowing for the reconstruction of a single depth. This is better explained in section 4.3.

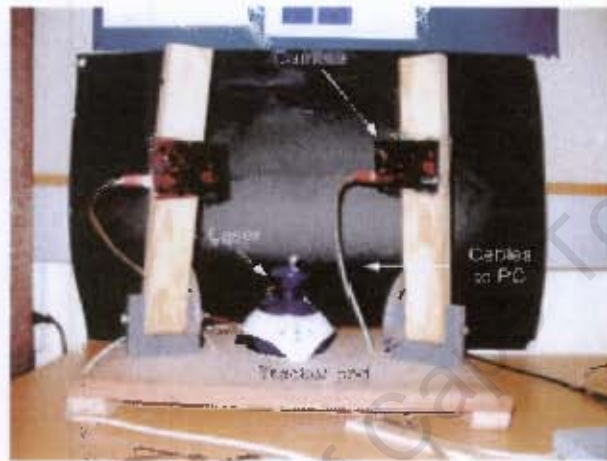


Figure 19: Image of the stereo setup.

7.1 Objectives of scanners

Maximizing scanner accuracy while minimizing engineering costs is a major target when designing scanners. The building costs arise from the expensive components (e.g. cameras, laser setup) and the difficult calibration requirements needed for the setup.

A choice has to be made between accuracy of the reconstruction and expense of the setup. Since this project is based on reconstruction of surfaces using low cost cameras, other ways of ensuring accuracy of the reconstruction are considered. Two of these aspects include proper calibration of the rig, and accurate extraction and detection of the laser stripe. When calibrating the system the two major quantities carefully considered were the re-projection error and the distortion parameter.

7.2 Calibrating the rig using Matlab Toolbox

The Matlab camera calibration toolbox requires images of a checkerboard that have been taken in different positions in front of the two cameras for calibration. Each position of the

checkerboard should be taken simultaneously by both cameras. This is because point matching is required to obtain the rotation and translation that relates the two cameras. When taking images of the checkerboard, as much space as possible in the image plane is covered, so as to consider most of the plane's edges and center when calculating distortion. Figure 21 shows an example of images used in calibration of one of the cameras.

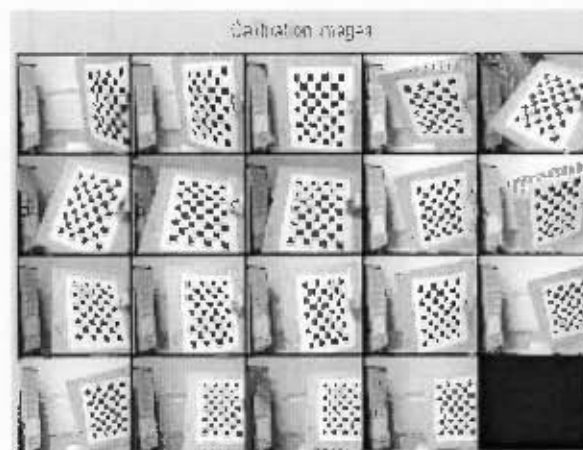


Figure 21: Checkerboard images used in the calibration of the two camera.

Each camera was calibrated separately. All of the images for each individual camera were read in the toolbox. Then corner extraction for each checkerboard image is performed using the corner extraction engine provided by the toolbox. When selecting the outline of the checkerboard, a consistent pattern should be used, i.e. grid points need to correspond across the different camera views.

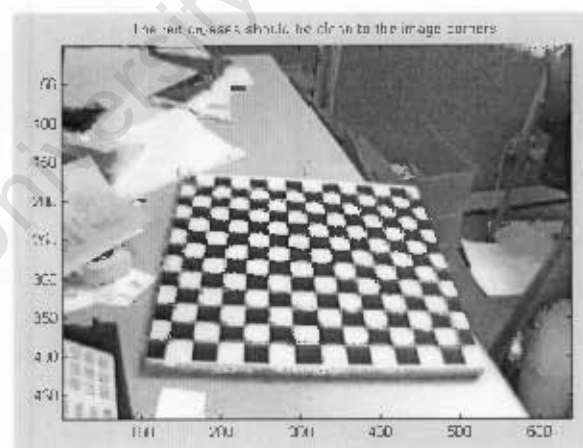


Figure 22: Figure showing extracted corner on one of the checkerboard images used in calibration.

After corner extraction the main camera calibration procedure in the toolbox was performed. After calibration, a data set of the corresponding pixel coordinates of the corners of the squares

is stored in the toolbox. These are used in estimating the fundamental matrix \mathbf{F} using the 8-point algorithm found in [9]. From the rig calibration using the matlab calibration toolbox, the intrinsic camera matrices \mathbf{K} and \mathbf{K}' were obtained.

7.2.1 Estimation of the fundamental matrix

Using the corner coordinates and knowing they are stored according to their corresponding points, the fundamental matrix \mathbf{F} , which is a representation of the epipolar geometry, is estimated as discussed in Section 4.3. The camera matrices of two views can be derived from the fundamental matrix. This is because \mathbf{F} depends on the projective properties of the camera matrices \mathbf{P} and \mathbf{P}' as discussed earlier. Since the essential matrix is $\mathbf{E} = \mathbf{K}'\mathbf{F}\mathbf{K}$, \mathbf{E} was derived from \mathbf{F} and used instead. This is because the camera matrices can then be computed up to scale. From the estimated essential matrix \mathbf{E} , there were four possible solutions for the second camera matrix \mathbf{P}' given that the first camera is

$$\mathbf{P} = [\mathbf{I}|0]. \quad (34)$$

Using a pair combination of \mathbf{P} and one of the four solutions of \mathbf{P}' , world points \mathbf{X} were calculated using triangulation. The distance from the cameras to \mathbf{X} were then calculated. The \mathbf{P}' of the four possible \mathbf{P}' 's whose depth is positive, is the correct solution. After choosing the correct camera matrix \mathbf{P}' for the reconstructed points, 3-D point cloud can be reconstructed for the square corner data.

7.2.2 Validating the fundamental matrix

In order to check that the estimated fundamental matrix is correct, images of the checkerboard were retaken by the calibrated cameras. The two images were then explored using the mapping

$$\mathbf{x} \mapsto \mathbf{l}'.$$

For a given point in the first image there is a corresponding epipole line \mathbf{l}' in the other image. \mathbf{F} was used to provide this projective mapping from points to lines using the following equation.

$$\mathbf{l}' = \mathbf{F}\mathbf{x} \quad (35)$$

An illustration of this is shown in Figure 23. For a point A in the left-hand image the corresponding epipolar line \mathbf{l}' passes through the corresponding point A' . This is the case for point $B \Leftrightarrow B'$ and $C \Leftrightarrow C'$. Throughout the project this test was used to check that the parameters obtained from the calibration were still accurate, i.e. the cameras had not been moved or bumped. After calibrating the system the blast plate and other objects used in the project were scanned and reconstructed.

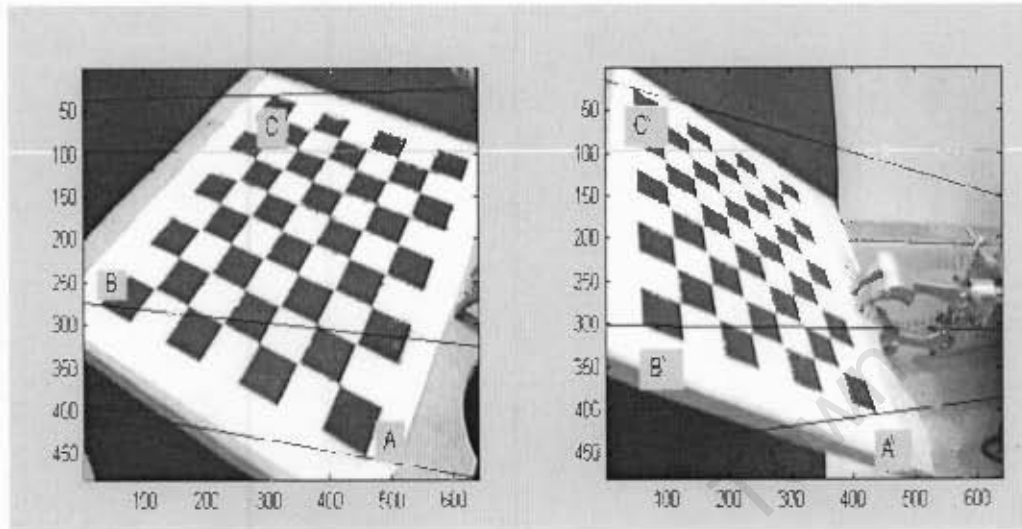


Figure 23: Figures showing corresponding points on a checkerboard.

7.3 Designing of the stereo-rig

The stereo-rig system was composed of a pair of CCD cameras, a laser line generator and software to control the hardware and process the images as shown in Figure 20. A small robotic tripod which was controlled by a computer using a USB port was used to move the laser in order to sweep the object. This tripod is called a trackerpod and its movements involve panning and tilting which are controlled through programming. The measurement process does not require calibration of the laser light source. The two CCD cameras were made to face the object. The trackerpod was placed between the two cameras and the laser was mounted on it. The trackerpod was programmed such that it moved a vertical laser line horizontally in small steps across the object.

For this project the main object analyzed was a blast metal plate with a dome in the middle. An image of the blast plate is shown in Figure 24. The dome caused by the blasting was a source of occlusion as the plate was generally flat. The full surface reconstruction was made up of aligned depth laser stripes.

7.4 Factors affecting the laser scanner design

The following section discusses the components of the stereo-rig setup in detail.

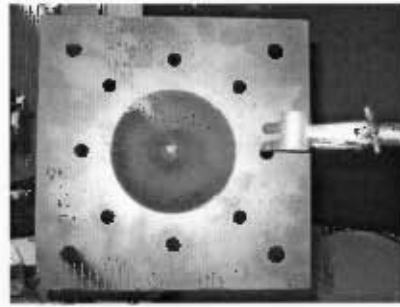


Figure 24: Blasted plate setup

7.4.1 Camera positioning

Simultaneous observation of the laser line from both views was required in order to obtain corresponding laser line image coordinates. In order to observe a point simultaneously the cameras were stationed at an angle to each other depending on the size and formation of the object being scanned. The cameras were moved until an image of the blast plate was observed and was central in the field.

7.4.2 Laser line/stripe

A laser diode of wavelength 650nm and output power less than 5mW was used. The laser light is a beam of circular cross section with a Gaussian intensity profile. The beam was spread by a cylindrical lens into a thin line. This still makes the intensity profile in the image planes of the camera Gaussian along the stripes. For this project a laser that produced a line was readily available.

In the traditional method (using a calibrated camera and laser), the accuracy of the system is dependent on the laser sheet being perfectly planar. This implies the laser optics have to be carefully chosen and precisely mounted with respect to the laser. However, in our case the laser plane is not calibrated into the system hence this does not affect the accuracy. An inexpensive laser was therefore used. A thin laser line is desirable in these experiments. In this case the line was about 3 pixels wide.

7.4.3 Resolution of scanner

The resolution obtainable by the scanner design is dependent on both the working volume and the camera resolution. The camera that was used has a resolution of 640×480 pixels in each captured image. The working volume dimensions were *width* = 20 cm, *height* = 30cm and *depth* = 50cm. From these details a resolution was estimated for our scanner. For better scanner resolution it is preferable if the working volume is made smaller and higher resolution cameras are obtained.

7.5 Scanning procedure

A painted metal plate was placed in front of the calibrated cameras. The motion of the laser line and capturing of the images was synchronized such that for every movement of the laser an image was acquired. A programme in Matlab was created to control this step. Irregular motions of the laser merely resulted in irregular sample spacing of the laser line, rather than incorrect or noisy depth measurements. Even if the laser plane was swept over the object by hand, inaccurate measurements would not result. Irregular stripes would appear because of irregular sampling.

Multiple surface reconstructions of the object are processed for both sides of the metal blast plate. In Chapter 10 these are then aligned in order to check for consistency of the reconstructions. Registration of these views was done in pairs.

7.6 Data processing

After scanning and capturing the laser line images, they are converted from a 3 band RGB image to a greyscale image with 256 grey levels, where an intensity of $I(x, y) = 0$ means it is black and $I(x, y) = 255$ white.

7.6.1 Pre-processing data

Before epipolar geometry is applied to the image data in order to extract information, it is necessary to process the data to ensure that it satisfies certain assumptions implied by the theory. These include:

1. **Background subtraction:** Before scanning the blast plate, 20 images of the blast plate without laser lighting are taken. These are then averaged out using image averaging and stored as background noise. For every captured laser image an image subtraction of the average background noise is done. This is to ensure that sensor noise does not introduce false information.
2. **Peak filtering:** Lighting from the laser is isolated from ambient background lighting with a medium red filter. The laser stripe produces an approximately Gaussian intensity distribution as shown in Figure 25 whose intensity is higher than the surrounding one.

During scanning surrounding highly reflective objects reflect off rays from either surrounding light or the laser onto the plate. When capturing images of the plates these are seen on the images as equally high intensity spots as the laser line. The detection of the laser stripe is disturbed by the presence of these undesirable high intensities in the scene. Since these undesirable intensities occur in spots of not more than 4×4 pixel a window filter of 12×12 pixels is swept through the image. In areas on the image where high intensities occupy an area of less than 5×5 pixels, that area is changed to

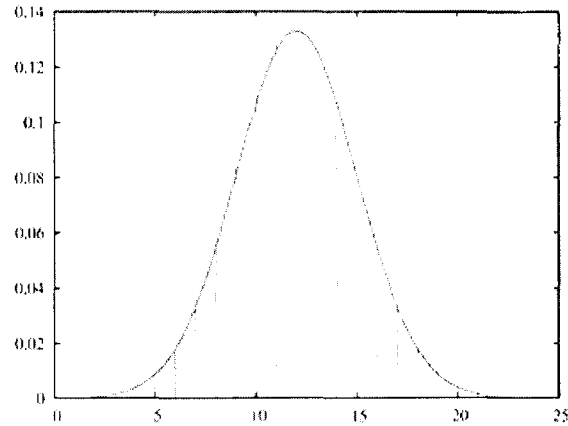


Figure 25: Figure showing gaussian property of light.

an intensity of $I(x, y) = 0$. A dark background cardboard is also used to reduce noise from lighting of false reflections.

3. **Laser stripe extraction:** The image was changed to greyscale using Matlab's `im2grey` which converts the image to HSV and uses the value band for the grey image. The image frame at this point is in greyscale. The co-ordinates of the pixels with the intensity above the threshold for each row of the image are extracted, assuming that the laser plane runs vertically in the image.

7.6.2 Processing data

After extracting the coordinates of the laser line, the epipolar geometry was used establish point correspondence between the data sets from the two images, and triangulation was used to obtain the 3-D position of the laser line in the world reference plane. The steps used in this process are explained below:

1. **Stripe processing:** After performing the above preprocessing operations, point correspondences are established between the corresponding images. To find corresponding points, pairs of images are considered individually. For every point on the laser stripe (\mathbf{x}) there is a corresponding point (\mathbf{x}') in the other image. In order to find the corresponding point (\mathbf{x}'), the epipolar line l' in the other image corresponding to \mathbf{x} is searched for the position where it crosses the laser line. This point's pixel is checked for the required intensity. If the intensity is above the required threshold, the pixel is taken as the corresponding point \mathbf{x}' . These corresponding points are used to calculate a depth value in 3D.

In order to ensure that the laser line in the right image crosses the epipolar line only once, careful positioning of the laser plane is required. When placing the laser on the trackerpod, it was ensured that the laser plane always hits the object vertically.

2. **High-level processing:** At this stage the information has been reduced to a small set of threshold intensity value. The data is then used to obtain the depth measurement for each point pair. The pixel co-ordinates of the corresponding points are obtained from the physical plane. These are transferred to the normalized image plane. The relationship between the physical and normalized image plane is

$$\mathbf{x}_n = \begin{bmatrix} x/z \\ y/z \end{bmatrix} = \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}. \quad (36)$$

where (\hat{x}, \hat{y}) are the normalized coordinates and (x, y) are the points from the physical plane. Distortion was taken into consideration and the corrected points were calculated using equation 28 where (x_{cc}, y_{cc}) were the corrected points. From here the normalized points were obtained as;

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ 1 \end{bmatrix} = \begin{bmatrix} f & \alpha & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{cc} \\ y_{cc} \\ 1 \end{bmatrix}. \quad (37)$$

Code from the Matlab camera calibration toolbox was used to convert the physical pixel co-ordinates to normalized pixel co-ordinates. After the points were normalized, these were converted to the world points using the transformation $\mathbf{x} = \mathbf{P}\mathbf{X}$. The depth of the laser line was calculated using stereo triangulation between two views.

The results of the reconstruction are presented in Chapter 8.1.

8 Results, analysis and discussions

8.1 Introduction

This chapter presents and discusses the results obtained in a 3-D reconstruction of a basic scene (blast plate). There were two types of reconstruction done. The first reconstruction

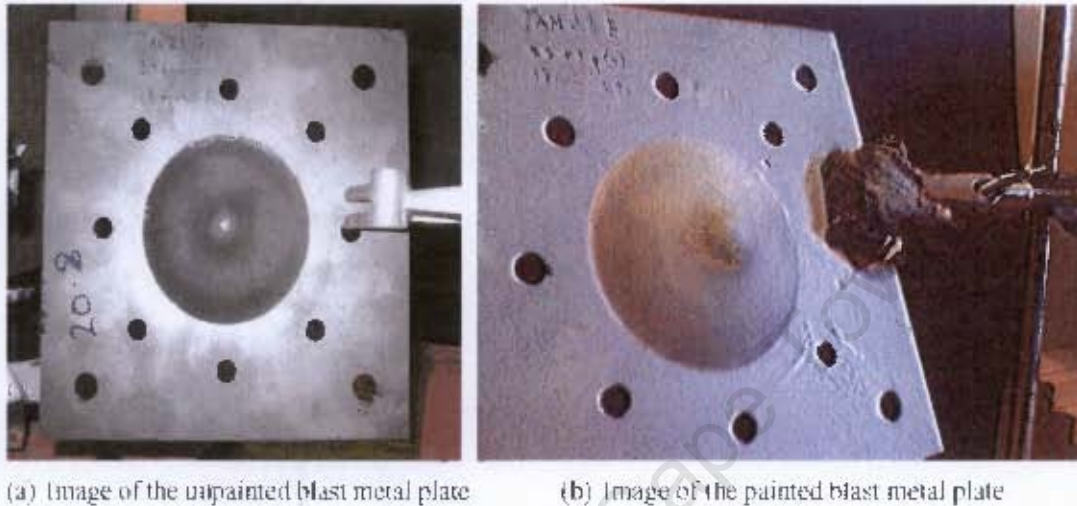


Figure 25: Images of the blast metal plates to be reconstructed.

was of the unpainted plate in Figure 25(a) and the second reconstruction was of the painted blast plate in Figure 25(b). The reconstruction of the plates using a laser spot is discussed in Section 8.3, and reconstruction using a laser line is discussed below in Section 8.2.

8.2 3-D reconstruction of the plates using laser line

Since the blast plate has poor reflection properties it was painted to obtain a good surface reconstruction. The plate was painted with a matte grey colour. The thickness of the plate was an average 2mm unpainted and an average of 2.006mm when painted. Since the paint was thinly layered over the plate, it did not cause a drastic deviation in the readings.

Figures 25(a) and 25(b) show the plates that were reconstructed. The scene was a blast plate held up by a stand (clamp). The dimensions of the square metal plate are shown in Figure 26. Reconstructions of both painted and unpainted plates using a laser line were obtained. The method used to scan the object and locate the position of the laser was simple yet effective. However, a number of difficulties were encountered in this experiment.

Disadvantages of the laser line finding method:

- It fails if the colour of the laser is similar to that of the scene.

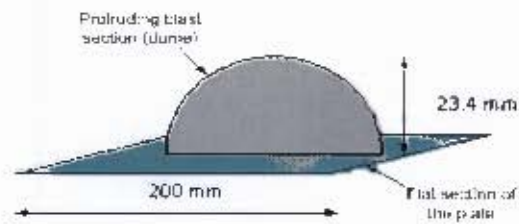


Figure 27: The dimensions of the blasted plate to be reconstructed.

- The laser scanning method was **only** used on stationary scenes as movement can cause **blurring** of the images. This was because the camera was set to a frame rate of one frame per trigger.

The total scanning time of the plate took just under 5 minutes. A high-speed camera can accelerate the process, and the laser line interval sizes can be made smaller, thus getting a better scan with more information. Processing of the data to get the final reconstruction took about 10 minutes.

8.2.1 Reconstructions as a result of using a laser line

Using the calculated matrices P and P' a point cloud data of the scene was obtained. Figures 28(a) and 28(b) show the reconstructions of the painted plate from two different angles in form of point cloud data. Figure 28(c) show surface presentation of the reconstruction. Reconstruction of the unpainted plate are shown in Figure 29(a), 29(b) and 29(c). For a pro-

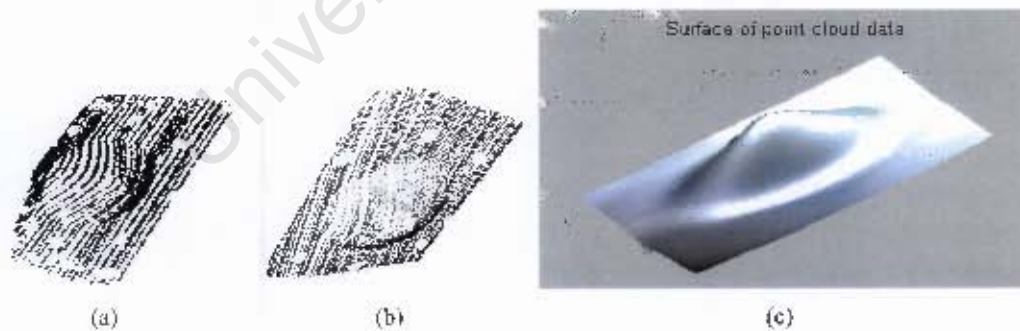


Figure 28: 3-D Reconstruction of the painted plate from different views.

jective reconstruction, geometric information such as orthogonality of lines and dimensions and ratios of length are not maintained [9]. However in the results obtained straight lines intersect at the correct places and perpendicular planes are perpendicular. Hence the reconstruction obtained is metric because it is correct to a similar transformation of the origin. The

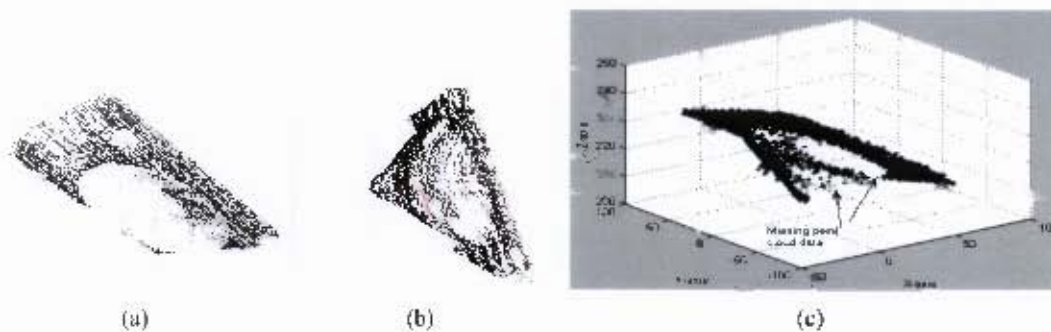


Figure 29: 3-D Reconstruction of the unpainted plate using a laser line from different views.

reconstruction obtained is metric because the camera matrices estimated as explained earlier are as a result of the fundamental matrix and the camera matrices. In the process of estimating these parameters any form of distortion is removed, Hence this enable a metric reconstruction to be obtained. Reconstruction of other objects were made so as to analyze the accuracy and consistency of the reconstructions obtained from this setup. These will be discussed in Chapter 9 and 10. These objects consisted of a tennis ball and a wooden box. Each scan resulted in a 3D point cloud data.

8.2.2 Effects of visibility of the laser line images

For a reconstruction of an unpainted plate using a laser line, it was possible to obtain a reconstruction of the flat section of the blast plate compared to the deformed section of the plate (dome) as shown in Figure 29. This was because of the behavior of light as it was reflected off different surface sections of the plate. On the dome, for most of the scanning, there was no laser image captured by the two camera sensors simultaneously. This was because of the reflection property of that area of the plate. The different amount of 3-D point cloud data obtained on each section was due to the different surface properties of the plate. When the plate went through a deformation from the blast, material properties changed and therefore its surface properties also changed. Because of this the reflected beam acted differently on the various sections of the plate. The surface scan therefore resulted in different amounts of point cloud data being collected from the different sections of the plate.

Figures 30(a),30(b) and Figures 31(a), 31(b) show laser images obtained from the scanning an unpainted plate. The image captured in the right hand camera is more visible than that of the left. This is because of the different viewing direction of light from the plate. As there is missing data during scanning, this carries through to the point correspondence process. Since most of the data in the left camera is missing there will be many points in the right camera that do not have matching points. Hence fewer 3-D points are obtained in the triangulation process.

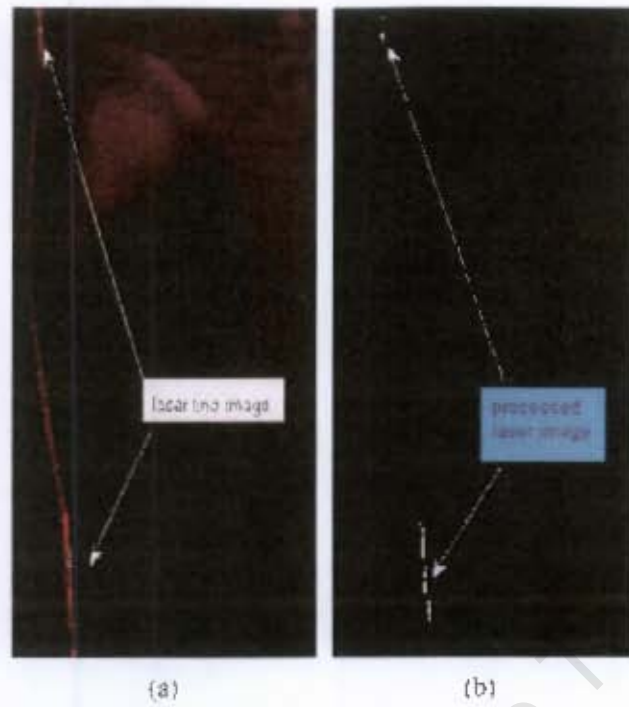


Figure 29: Figure (a) shows a laser line image captured by the left-hand image. Figure (b) shows the laser line image after filtering.

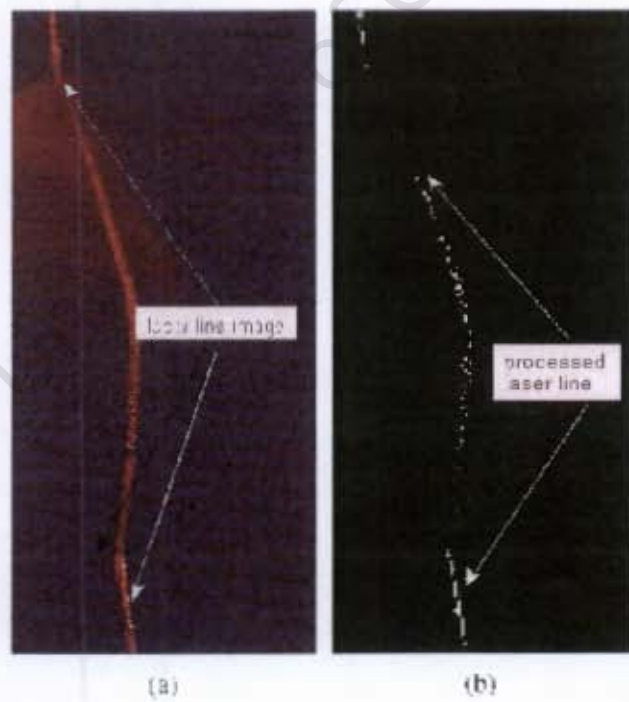


Figure 30: Figure (a) shows a laser line image captured by the right-hand image. Figure (b) shows the laser line image after filtering.

8.3 3-D reconstruction using a laser spot

One of the problems encountered in machine vision, is analysing an image which is captured under poor conditions, in this case with high levels of specular reflection. This was one of the problems experienced when scanning the unpainted metal plates in this project. When the unpainted blast plate was scanned using a laser line, an incomplete reconstruction of the plate was obtained because of the surface's specular reflection. This made the laser line not visible in the camera. However this was not always the case during the experiments carried out. Another scanning method was proposed. This involved replacing the laser line with a laser spot which was controlled manually so as to change the direction of the light according to the shape of the surface.

As discussed in Chapter 2 the visibility of a laser point depends on the material properties and the angle of incidence. The material texture of the dome section on the blast plate was different to that of the flat section see Figure 9. This was because it was stretched during the blast. This made visibility of the laser beam on the flat surface clearer than that of the deformed part. A laser spot was directed towards the blast plate and a visible image of the spot was obtained in both cameras. This was because of the intensity and angle of incidence the laser spot had in each instance.

The unpainted plate was scanned using a laser spot while changing the angle of incidence so as to get a reflection in the cameras.

8.3.1 Projective reconstructions of the plate as a result of using a laser spot



Figure 31: Point cloud data of the surface reconstruction using a laser spot.

Similarly using the calculated matrices P and P' a projective point cloud data of the unpainted plate was obtained. The reconstructions obtained from scanning using a laser spot are shown in Figures 31 and 33.

Figure 31 shows point cloud data of the reconstructed blast plate obtained by changing angle of incidence of the laser spot. This shows that a more detailed surface reconstruction is obtained by varying the angle of incidence with a laser spot as compared to using a laser line

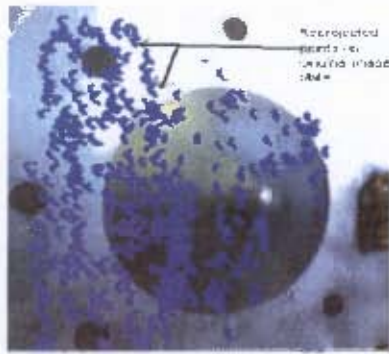


Figure 32: Point cloud data in the x-y direction to show patchiness as a result of using a manually positioned laser spot.

at a constant tilt. The disadvantage of using a hand held laser spot to scan is that the surface coverage is inconsistent as shown in Figure 32. This is as a result of the inconsistent scanning intervals that were performed manually. In order to get a better presentation of the surface, the point cloud data was converted to a mesh as shown in Figure 33.



Figure 33: Image of the reconstructed surface in the form of a surf.

8.4 Estimating plane of best-fit

Figures 27(a), 27(b) and 27(c) show the surface reconstruction of the painted blast plate. The gaps between each stripe in the reconstruction is due to the varying sample interval caused by the inconsistent movement of the trackerpod. Compared to the unpainted plate, the scanning of the painted plate produces a full reconstruction of the scan.

A number of steps were taken to analyze the accuracy of the reconstruction. One of these involved estimating a plane of best fit for the flat section of the plate, from which the highest depth projection of the dome section can be inferred.

The plate that was reconstructed has a section that is flat (namely the section that was not

deformed). The points in this section were extracted manually and a plane of best fit was estimated for the plate. This is the same plane from which the highest point of the dome is referenced. An example of this is shown in Figure 34. A more detailed view of the plane

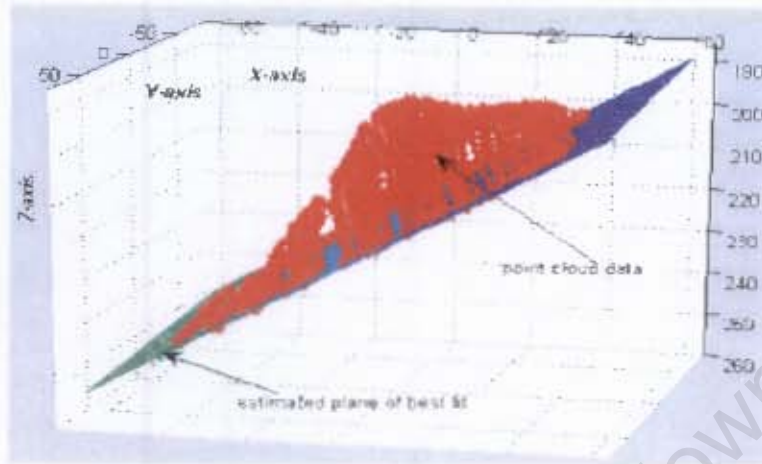


Figure 34: Images showing plane of best fit for the flat section of the plate.

is given in Figure 35. This figure shows the plane in the form of a best-fit line with the 3-D

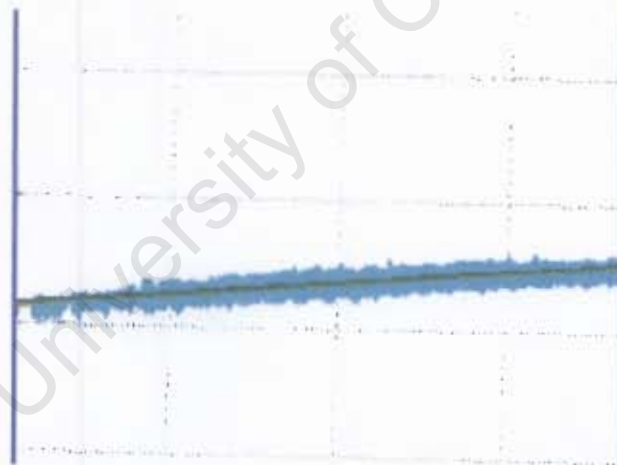


Figure 35: Close up of estimated plane

points scattered around the line. Measurements were carried out to show how the estimated plane relates to the cloud point data. These are shown in Table 1.

8.5 Results analysis

This section compares a metal blast plate of depth 23.4mm to its reconstruction results. First the plane of best-fit is considered.

Table 1: Results measurements of distances taken from best fit plane in different reconstruction instances of the painted plate.

Diff scan data	Ave dist of Pts from best fit plane (mm)	RMS error from best fit plane (mm)
Plate 1	0.473	0.537
Plate 2	0.475	0.324
Plate 3	0.369	0.272

8.5.1 Analysis of the best-fit plane

The estimated flat plane is assumed to be one pixel thick. The world points \mathbf{X} in space, which are as a result of back projecting the rays from the corresponding points, are as a result of minimised error distance to hold the following equation $\mathbf{x}^T \mathbf{F} \mathbf{x} = 0$.

Since real points have errors in them from detecting them, the constraint $\mathbf{x}^T \mathbf{F} \mathbf{x} = 0$ does not hold. In order to estimate a depth for two corresponding points the distance d where the back projected rays are closest is minimised. Since the depth for the different corresponding points is an estimate of the midpoint where the rays are closest this will not enable points on a plane in the real world to all lie on the same plane. The other points will lie in the uncertainty region. Hence when estimating a flat plane the reconstructed points will be scattered about and on the ideal position of the flat plane.

The dispersal of the cloud points from the plane of best fit is considered. For each reconstruction, the RMS error of the point cloud data (on the flat section) from the plane of best fit is calculated. The results obtained are shown in Table 1. Different RMS error values are obtained for the different reconstruction data. A reason for this variation is inclusion of outliers in the result data that are lying too close to the correct data to be removed. A close-up distribution of the distance of the points from the plane is seen in Figure 35.

It was found that the worst deviation error distance of the points from the best-fit plane was about 2mm. From Table 2 an average of 0.378mm RMS error was obtained. This estimated plane was be used as a reference from which the highest point of the dome can be calculated.

8.5.2 Depth Error of the height of the dome

In Table 2, for the different reconstructions instances, the maximum point of the dome from plane of best-fit is obtained. For the painted plate in Table 2 the maximum point has an average value of 24.56mm hence giving it an average error of 1.16mm. The error in the measurements can be accounted for by:

- The maximum point is referenced perpendicularly from the plane of best fit. Since this plane is an estimate itself; it brings an error in the calculations.

Table 2: Results of measurements of depth taken from 3 different reconstruction instances of the painted plate.

Diff scan data	Depth of dome (mm)	Depth error (mm)	re-Projection err of image points (pixels)
Plate 1	24.89	1.49	6.57e-4
Plate 2	24.27	0.87	6.68e-4
Plate 3	24.52	1.12	6.60e-4

- When extracting the laser image in the pixel co-ordinates it is assumed that the width of the laser image is one pixel, but this is not the case. In data processing for a wider laser stripe, the left most pixel of the stripe is considered as the required coordinate which might not be true. Its corresponding point in the second image is obtained using the epipolar constraint. The assumption of using the left most pixel incorporates error in the horizontal direction of the image plane.
- Another cause of inaccuracy in measurements is the outliers that still exist in the reconstruction even after outside noise has been removed. Some of these outliers are a cause of false reflections as a result of the poor paint colour of the plate.

For the laser line reconstruction obtained for the unpainted plate, it was difficult to estimate the depth of the dome as there was missing data of the surface. For the reconstruction obtained from laser spot scanning, the height of the highest 3-D point was calculated. This was found to be 23.23mm as opposed to the actual value of 20.50mm for the unpainted plate. This gives geometric error of 2.73mm. The error of the calculated height was based on a number of factors. One of these is multiple and false reflections off the plate since it has specular reflections. Another cause of such an error would be noise due to speckle caused by the laser as it is directed towards the plate.

9 System accuracy testing

9.1 Comparison with simple object

In order to determine the system's accuracy several tests were carried out. These included metric measurements of an object (box), for example measuring height of one of the box's faces and the angle between the planes. A simple object was used to assess accuracy, so that the reconstruction results were straight forward and could be compared for accuracy analysis. A box of height 44mm, width 44mm and length 78mm was chosen to investigate the accuracy.

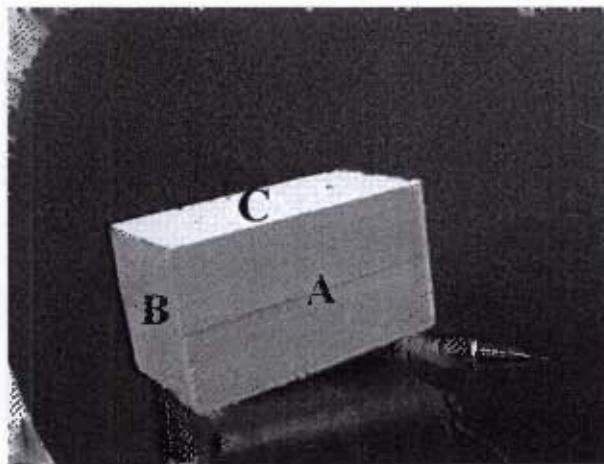


Figure 36: Image of a box used for accuracy check.

The wooden box is shown in Figure 36. The point cloud data of the faces obtained are shown in Figure 37. Due to the angle of the cameras, face *A* and face *B* were the only faces that were reconstructed properly. Face *C* has a few 3-D points. This is because the laser intensity arriving and leaving face *C* was not as strong as that of face *A* and *B*. This can be explained using radiometry (Chapter 2). A surface reflects more energy collected while viewing a source frontally compared to energy received from a source viewed tangentially. A number of height measurements of the faces were taken and tabulated in order to calculate the average reading. Manual selection of the correct points used in calculating the height of the faces was done in the Matlab's figure window. Manual selection of the 3-D points for plane estimation was also done in the Matlab's figure window. Planes of best-fit from the reconstructed data were estimated for face *A* and *B*. Figure 38 shows the estimated planes of two faces and how they intersect with one another. From the plane estimation process an equation of the plane is obtained: $Ax + By + Cz = D$. The normals $[A_1; B_1; C_1]$ and $[A_2; B_2; C_2]$ obtained from the two plane equations are used to calculate the angle between the planes. The measured results obtained for different reconstruction instances are shown in Table 3. The calculated results are rounded off to the nearest decimal place. The experimental points have been compared with planes of best-fit of the surface faces. The analysis of the measurements yielded an average perpendicular distance of 0.77mm from the plane. An estimate of the height was obtained with an average error of 0.30mm. The angle between the two faces had an average error of

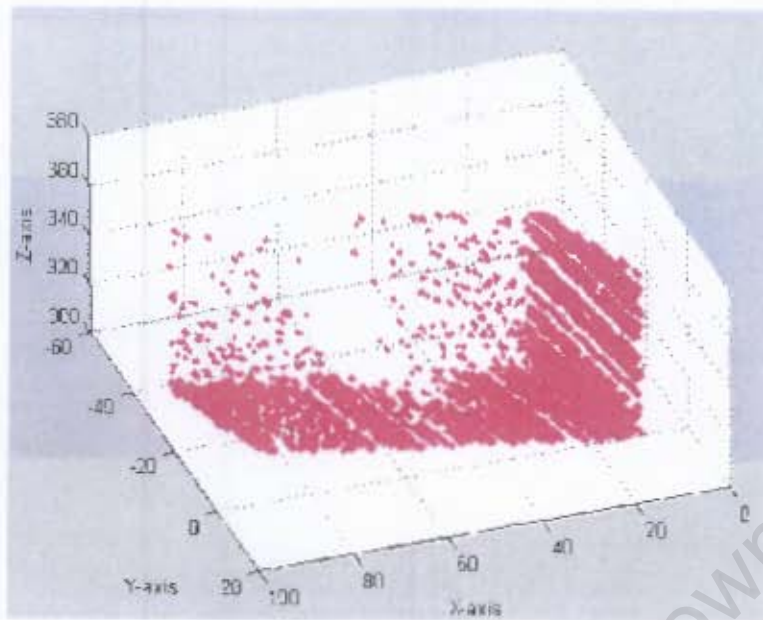


Figure 37: Figure showing point cloud data of reconstructed faces of the wooden box from different views

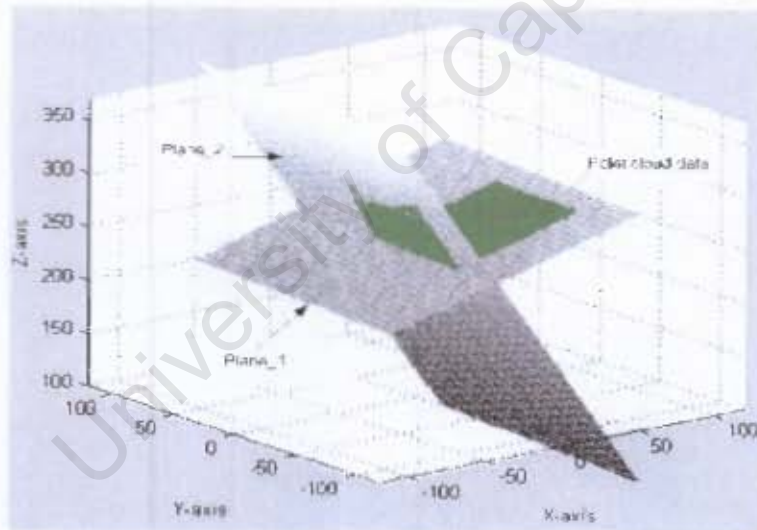


Figure 38: Figure showing estimated planes best-fit used to calculate the angle between the reconstructed faces.

0.34degrees. Note that the results of the reconstructed box have a smaller error than those obtained for the blasted metal plate. This is probably due to the better surface quality of the wooden box, which had negligible reflection problems.

Table 3: Table showing measurements obtained from the reconstructed box.

Instances	Height (mm)	Height error (mm)	Angle (deg)	Angle error (deg)
1st instance	44.3	0.3	90.7	0.7
2nd instance	44.0	0.0	90.4	0.4
3rd instance	44.6	0.6	90.6	0.6

9.2 Reprojection error

The re-projection error is a measure of what is needed to obtain perfectly matching points that agree with the following equation $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$. This is done by minimizing the distance between the perfectly matched points $\hat{\mathbf{x}}, \hat{\mathbf{x}}'$ and the measured points \mathbf{x}, \mathbf{x}' by use of an invariant homography \mathbf{H} in $\mathbf{x}' = \mathbf{H}\hat{\mathbf{x}}$. In order to look at the reprojection error, the 3-D points were taken into account. The point cloud data was transferred back to the normalized image plane using the following mapping; $\hat{\mathbf{x}} = \mathbf{P}\mathbf{X}$ [9].

The results obtained were then mapped back onto the physical image plane of the camera. These are the 2-D reprojected data sets. The sets were then compared to the original 2-D data extracted from the physical images. The error between the original image data and the reprojected data was calculated and is presented in Table 4. Corner points from one of the

Table 4: Table showing re-projection errors obtained from the different reconstructions.

	error in left image (pixel)	error in right image(pixel)
Corner pts of Checkerboard	$0.4817 e^{-04}$	$0.4047 e^{-04}$
data pts of a box	$6.6032 e^{-04}$	$5.8848 e^{-04}$
data pts of the metal plate	$9.0562 e^{-04}$	$9.0431 e^{-04}$

checkerboard images were used to make a representation of the re-projection error. This is shown in Figure 40, where the cross 'x' is the original point image and the circle 'o' is the estimated image point from re-projection. Figures 41 and 42 give a view of how the original image points and re-projected points align. The green is the original data, and the black is the re-projected point. The average re-projection error in the camera image frame was $0.40473e^{-04}$ for the right handed side and $0.4818e^{-04}$ for the left handed side of the box. This was caused by the error obtained when calibrating the rig and the distortion not accounted for in the images.

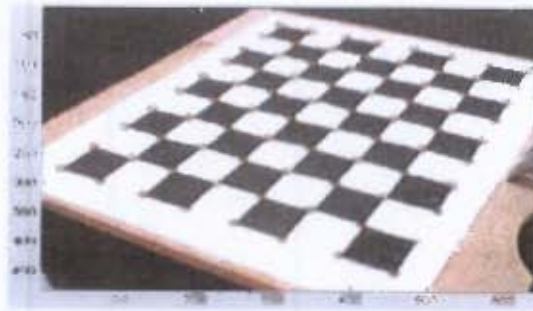


Figure 39: Re-projection back onto the checkerboard

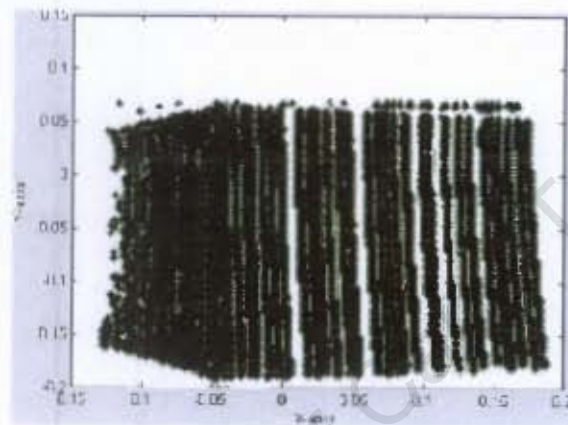


Figure 40: Image showing re-projection of reconstructed points of a box back to original points

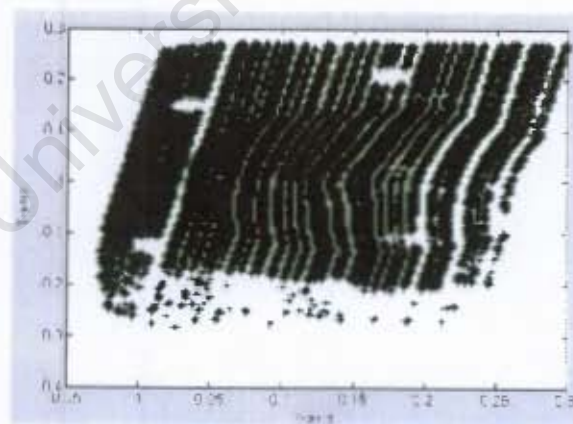


Figure 41: Re-projection of the reconstruction of the blast plate to its original image points

10 Capabilities of the ICP registration algorithm

This section discusses the results obtained from implementing the ICP algorithm on randomly generated data sets in 2-D and then in 3-D. It then shows how the ICP algorithm works on the

3-D point cloud data obtained from reconstructing the blasted metal plate. Since the project was mainly run using Matlab, this section of the project was also done in Matlab. The approximate times taken to run ICP algorithm on different sets of data are given for execution on a single-processor Intel Pentium computer with 512 Mega bytes of RAM 800MHZ.

After the reconstruction of the blast plates the different data obtained were aligned using ICP algorithm to check for consistency. First two sets of data were generated randomly in 2-D and 3-D to test the algorithm. The algorithm was then applied on the 3-D points of the blast plate. Results of the alignment are presented below.

10.1 Applying ICP on randomly generated data

In order to check that the ICP algorithm was functioning properly, it was first implemented on 2-D data assuming known rotation and translation. For the two sets of data an initial rotation matrix and translation vector was estimated. In this project the initial rotation matrix is chosen randomly, and the difference in position of the center of mass for the two sets of data is chosen as the initial translation matrix. The "center of mass" μ_X of the set X and "center of mass" $\mu_{X'}$ of the set X' were estimated as shown below:

$$\vec{\mu}_X = \frac{1}{N_X} \sum_{i=1}^{N_X} \mathbf{X}_i \text{ and } \vec{\mu}_{X'} = \frac{1}{N_{X'}} \sum_{i=1}^{N_{X'}} \mathbf{X}'_i. \quad (38)$$

The transformation required to align the set of data points X with the set X' was calculated. This required choosing parameters such that they move points x' in X' to points x in X , such that the relationship $x' = sR_x + t$ is true between x and x' . The scalar s is assumed to be positive, so as to prevent reflections which are not permitted in this algorithm.

10.1.1 Point set matching on 2-D data

This section demonstrates the use of ICP on 2-D generated data. A set of 50 corresponding points was generated randomly using Matlab's function 'rand'. The values obtained were then multiplied by a 1000 to give a bigger working space.

These were transformed through a rotation of 55 degrees then translated through $t = [56; -107]$. This produced an ideal set of data. Noise with a standard deviation of 0.2 and RMS of 0.3 was added to make it more realistic. The standard deviation value and the RMS distance was varied in order to check up to what value ICP would continue functioning. The original set of data was considered as set X . The transferred data set was considered as set X' , where number of points $N_X = N_{X'} = 50$. The ICP algorithm was implemented on the two sets of data to align the data. The initial rotation matrix was set to an identity matrix. The midpoint of the centers of mass of the two sets of points was used as the initial translation vector.

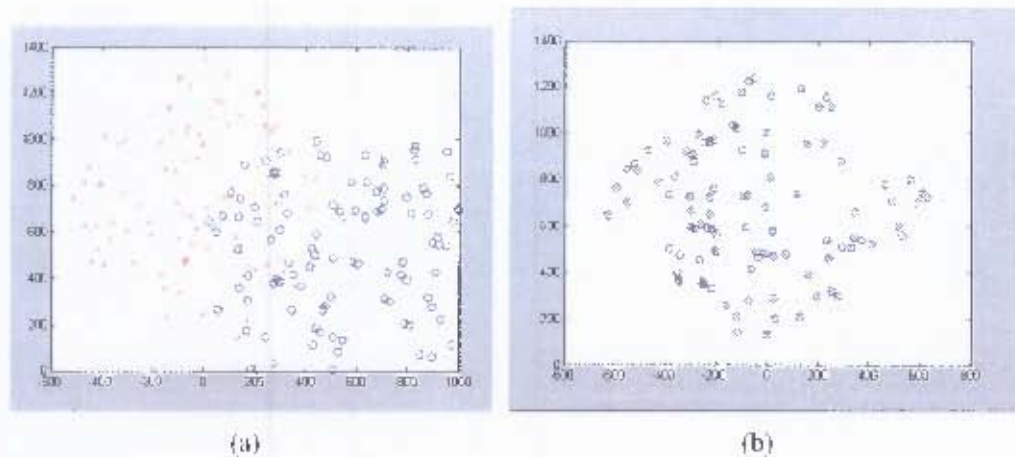


Figure 42: Figure(a) shows two sets of randomly generated data in 2-D, before alignment, (b) shows how the two data sets align after being registered using ICP algorithm.

Table 5: Table showing output results from aligning two data sets of different amounts of points using the ICP algorithm in 2-D.

No. of Random Points	No. of iterations	Average <i>distance</i> error	Duration of ICP algorithm (sec)
50	7	0.0207	1.592
100	8	0.0189	3.766
150	20	0.0187	13.049
250	36	0.0191	22.582

Figure 42(a) shows the two sets of data X and X' prior to registration, where set X is the red data and set X' is the blue data. Figure 42(b) shows the two data sets after registration. The average error between corresponding points is 0.0207 and the average time taken to carry out the ICP algorithm is 1.592sec.

The amount of point cloud data was increased from 50 to 250 in steps. This was to observe the effect an increase in data has on the average distance error and time taken for implementing the ICP algorithm. Since randomly generated data was used, dimensions (i.e. millimeters/meters) on the error results was not introduced. The results of the ICP algorithm for the different numbers of points are given in Table 5. The table presents the time taken to carry out an ICP algorithm in seconds. The number of iterations the algorithm loops through before reaching a local minima and the average root mean square distance are also tabulated.

After implementing the ICP algorithm on the generated 2-D data, 3-D data was then considered. The same procedure as before was used on the 3-D data and the results are shown in Figure 43(b) and Table 6 .

10.1.2 Point set matching on 3-D

The same procedure in generating the 2-D data was also used to generate the 3-D data points (taking standard deviation and RMS distance into account). The aligned 3-D points are shown in Figure 43(b). The calculated results obtained during the algorithm are shown in Table 6.

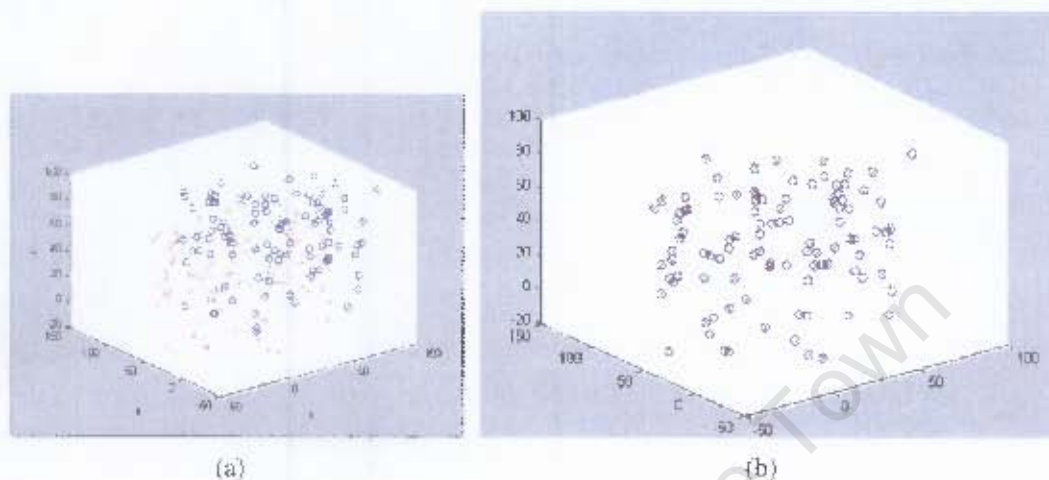


Figure 43: (a) shows two sets of randomly generated data in 3-D before alignment, (b) shows how the two data sets align after being registered using ICP algorithm

Table 6: Output results from ICP of different amounts of input points in 3-D.

No. of random points	No. of iterations	Average <i>dist_{rms}</i> error (mm)	Duration of ICP algorithm (sec)
50	6	0.0226	3.9360
100	11	0.0248	3.9260
150	12	0.0248	5.9990
250	21	0.0237	8.3420
450	28	0.0247	14.5010

10.1.3 Analysis of the randomly simulated data

The algorithm was tested on the synthetic data and showed good convergence between points as discussed below. From Figures 42(b) and 43(b) the corresponding data sets were able to align with an average Euclidean distance error of 0.019 and 0.021 in 2-D and 3-D respectively. The error computed was the average RMS distance computed between the closest points obtained. The optimal situation is that all the distances are zero. From Tables 5 and 6, the distance error results obtained show that for each randomly generated data a correct local minimum was reached. From Tables 5 and 6 it is observed that, the number of iterations and

time taken increases with an increase in the number of points in ICP algorithm. The increase in iteration is not linearly related to the increase in number of points. Tables 5 and 6 show that for a set of 50 points, locating the closest points takes longer for two data sets when the dimension is increased from 2-D to 3-D i.e 2-D 50 points = 1.59sec, 3-D 50 points = 2.93sec. This is because another dimension is introduced increasing on computation time.

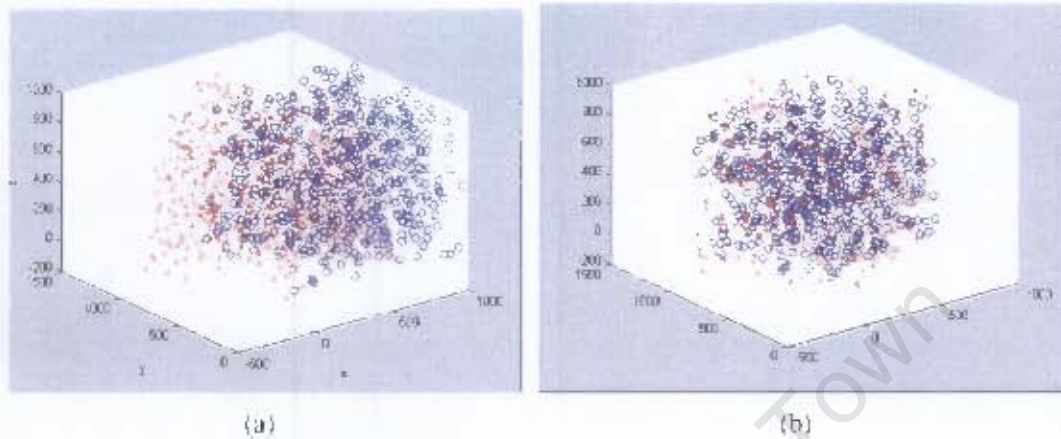


Figure 44: (a) shows a larger set of randomly generated data in 3-D before alignment, (b) shows how the two data sets do not register after being registered using ICP algorithm

Figure 44(a) shows the results of trying to align a large data set of 1000 randomly generated points. Figure 44(b) shows that the correct local minimum was not reached in this example. It was also observed that the data sets did not align when the standard deviation was increased to 0.3.

10.1.4 Conclusion on the synthetically generated data

From Table 5 and 6, it can be concluded that with an increase in points the number of iterations increases, hence increasing the time taken to perform the ICP iterations. The algorithm reaches a local minima which is non-zero. This is because of the noise present in the data.

It is also stated in [1] that with an ambiguous shaped data set, it is hard to obtain registration of a large data set if the data set is large. This is the case and is shown in Figure 44(b). When the data set of a synthetic 3-D data set was increased to 1000 points, it became difficult to find alignment. The average RMS error in the above case was 69.41. However with data with structure this is not the case, as will be shown with an example from the blast plate point data in the coming section.

10.2 Applying ICP on measured data set

In this section the ability of ICP algorithm to register the real measured data of the metal plate is demonstrated. Different numbers of points are obtained for each scan of the blast plate.

The ICP algorithm was written such that it takes the same number of points for the two sets of data. In order to use the same number of inputs the same number of points were chosen randomly from the two data sets. Figure 45 shows the first two sets of measured point cloud data prior to registration. The green crosses ('x') denote set X and the red dots (·) set X'. The number of outliers in the input data was minimized as this leads to bigger distance errors or causes convergence to the incorrect local minimum. A different number of points were taken from the two sets and increased to show how this affects the time taken. An initial set 500 of

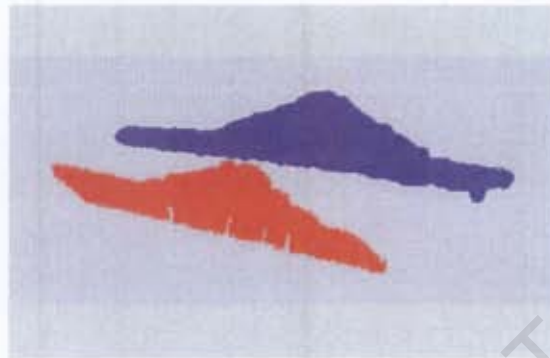


Figure 45: Real complete points cloud data of set X and X' before ICP.

points was randomly chosen from both set X and X' as shown in Figure 46. In order to be sure that at least fairly even number of points are selected from all the areas of the surface, the data set is divided into ten equal parts. This sectioning is possible as the data is stored in an orderly manner i.e. from the negative x-coordinates to the positive x-coordinates. The same random number is chosen from each of these sections to make put the 500 points. The ICP

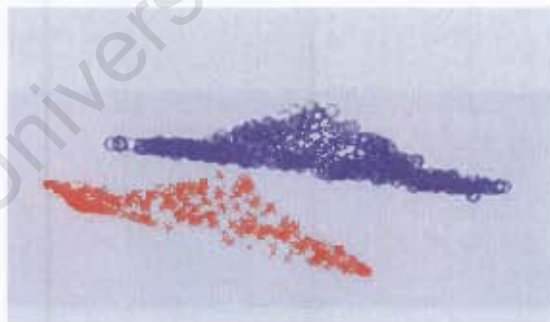


Figure 46: 500 randomly chosen points from both set X and X'.

algorithm was applied to this new set of 500 points with the following termination criteria:

- Iterate until the point distance error is less than $\frac{1}{10000}$.
- For six consecutive iterations the point distance error does not change in the hundred thousandth decimal place. $\frac{1}{1000000}$.

After running the ICP algorithm with an initial transformation $[\mathbf{R}|\mathbf{t}] = [\mathbf{I}|\mathbf{t}]$, the results shown in Table 7 were obtained. The rotation matrix and translation vector that was used to align

Table 7: Results of ICP for different numbers of randomly selected 3-D point cloud data from a complete measured data.

No. of random points	No. of iterations	Average RMS error (mm)	Duration of ICP algorithm (sec)
500	38	1.7748	19.76
1000	112	1.0751	100.05
2000	91	1.1131	139.69
4000	149	1.0926	319.98

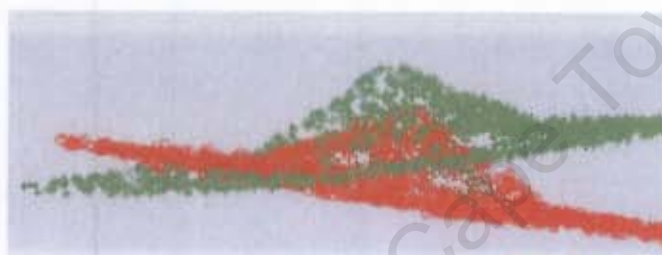


Figure 47: Figure showing process of the aligning subsets of X and X' .

subset X on subset X' was applied on the whole of set X to see the complete alignment of set X on set X' . Figure 48 shows set X aligned on set X' using the transformation obtained from aligning the 500 randomly chosen points. This is to investigate if, with a wider coverage of points, the root mean square error (RMSE) will be different. The result of this analysis is shown in Table 7.

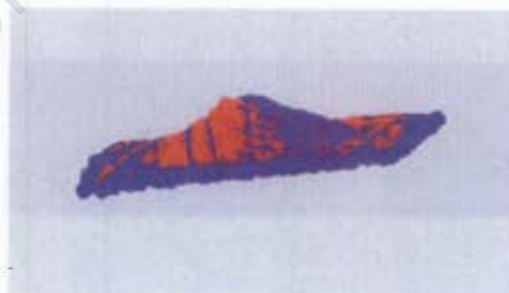


Figure 48: Alignment of set X and X' after transformation obtained from using 500 points.

10.3 Analyzing results of registered point cloud data of the blast plate

10.3.1 Analyzing choice of local minima

From the results of the registration the average error was greater than 1mm. By examining the results in Figure 48, it can be seen that the right half of set X is not exactly aligned onto set X' . The flat section of the plate on the left is however properly aligned. In this case the point cloud data found its local minimum on the left half in the flat section of the plate. The top section is not well aligned; this is caused by outliers. If the randomly selected sets have more distribution of points in a particular section the algorithm weights that section more heavily. To minimize this effect a number of random selections were made, each time with a different number of points. The results of the alignments are presented in Table 7.

The algorithm reaches a local minimum with an average distance error of 1.26mm as calculated from Table 7. As was shown with the synthetically generated data, an increase in the number of points results in an increase in the time taken to perform ICP.

10.3.2 Analyzing computational approaches and time taken

The algorithm does not take into account the extra points or the ordering of the points because it tries to pair a given point to its closest point from the other set. The initial estimation of the rotation and translation is very important, as this determines whether the data (registration) will converge or diverge. In order to try different initial rotation matrices and translation vectors, CPU time for local matching and point order insensitivity is compromised. If every possible correspondence is tested in order to find the best local match, this registration would require

$$\frac{N_{x'}!}{(N_{x'} - N_x)!} \quad (39)$$

operations of the basic least square match [1]. An example to explain this is shown below.

From one of the previous experiments on 2-D registration, given $N_{x'} = 50$ and $N_{px} = 50$, this task would take $50! = 3.0414e^{+064}$ registration operations, if every correspondence for every point is considered. Using ICP the above registration required only 7 iterations and took a time of 1.592 seconds as shown in Table 5, since the initial state was already a member of the equivalence class of global minimum. However, if other options of initial rotation and translation are considered, as suggested in [1] Besl, these registrations would each have a certain number of iterations depending on whether the initial transformations are in the global minimum.

This would be time consuming, so an initial rotation of zero degrees and a translation of the center of mass of the two sets of data is instead used because from Figure 46 the positioning of the two data sets in the initial position is not far apart. Hence a translation is used instead as part of the initial transformation. Only one initial rotation and translation are used in this task since the chosen transformation arrived at its local minimum at the first try.

10.4 ICP registration on surface reconstruction of other objects

After looking at how the ICP algorithm works on the point cloud data of the blast plate it was tested on the point cloud data of other objects to see how it adapts. A tennis ball was chosen because of its simple form, i.e. round and smooth.

10.4.1 Registration of cloud point data of a tennis ball



Figure 49: Image of the tennis ball that is introduced for comparison.

A tennis ball was clamped on a stand as shown in Figure 49 and was scanned from two different angles using the scanner. 3-D reconstructions of the tennis ball from the different views were obtained. This was done in order to obtain different orientations of the tennis ball's surface reconstruction. The surface reconstructions obtained from the scanning are shown in Figure 50.

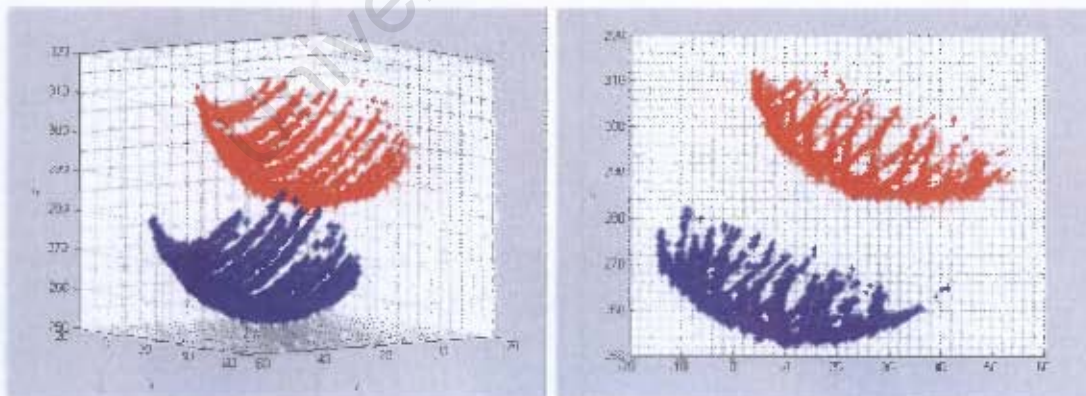


Figure 50: 3-D point cloud data obtained from scanning a tennis ball at different orientations. (The red points show the first scan in one view and the blue scan show the point cloud data of another view).

The ICP algorithm was then used to observe whether the two scans of the same object could be aligned. As explained before, the time taken by the ICP algorithm depends on the amount

of input data. There were approximately 18,000 3-D points for one data set and 16,000 for the other. Since the algorithm takes in the same number of points, the same number of points were randomly selected from the two sets of data to estimate a transformation that aligns set one with set two. This is equally done for different numbers of points to investigate how it affects the time taken and the point distance error. These results are recorded in Table 8. An image of the two sets of aligned 3-D points of the tennis ball aligned are shown in Figure 51.

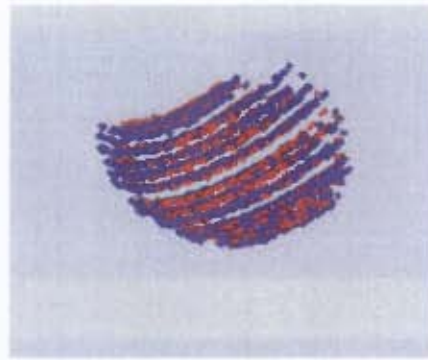


Figure 51: Image showing the results obtained from aligning point set-one (red points) with point set-two (blue points) of different surface reconstructions.

Table 8: Results of ICP for different number of randomly selected 3-D point cloud data from the data points of the tennis ball

No. of random points	No. of iterations	Average RMS error (mm)	Duration of ICP algorithm (sec)
500	26	0.9169	16.50
1000	30	0.9013	29.96
3000	106	0.9088	260.13
4000	100	0.8931	290.01
15000	60	0.8912	446.79

10.4.2 Analyzing the registration

The number of iterations is not directly related to the time taken as shown in Table 8. Iterations depend on the number of transformations the measured points go through to be aligned. The time taken depends on the total time it takes to calculate the closest distance and then estimate a transformation for every iteration.

10.5 Conclusions made on the ICP algorithm alignment

From the deduction made from the ICP experiments the following conclusion were drawn:

1. The average mean square error of 0.9169mm obtained from aligning the tennis ball point cloud data is smaller than that of the metal plate. This is because of the simplicity of the reconstruction of the ball.
2. The reason why the error is not well minimized is because of outliers. Even after removing background noise when preprocessing the captured images as explained in Section 7.5, there are still some outliers which are too close to the rest of the data set.

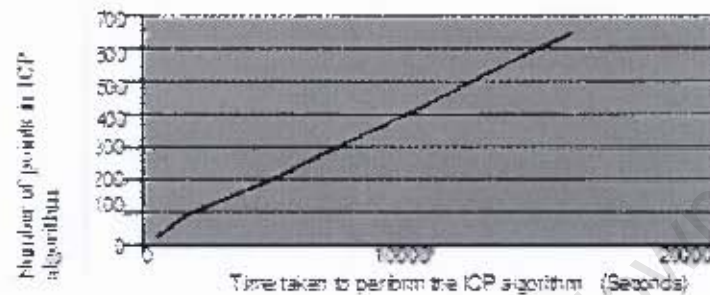


Figure 52: Time taken to perform ICP algorithm using different number of points chosen as subsets

3. The error on average does not reduce with the number of random points selected for the initial transformation. This is because the ICP algorithm continuously finds the closest points it can align the data sets. Hence whenever there is a new randomly selected data the possibility of having an exact alignment will not increase with the number of points. This is because the increased random points chosen do not necessarily have the correct corresponding points in the other set. This uncertainty makes the error calculated unpredictable.
4. When choosing a large set of random points (i.e. 600 and above) in the reconstructed point cloud data, the point distance error estimated is less than the point distance error estimated from the synthetic data. This is because, even if the points in the reconstructed data are randomly chosen, they define a shape. The synthetically generated data have no model shape and hence do not obey the surface equation that would model the object. Since the synthetic points are featureless the error distance between points will be greater than those of the measured data.

11 Conclusions

The conclusions presented below are made on the findings obtained in this thesis.

The overall conclusion that can be made from the investigations in this thesis is that a surface reconstruction of a blast metal plate can be made using a low cost stereo-rig. The accuracy of the results obtained depends on the conditions under which the reconstruction is made. The conditions include controlled lighting of the experiment room, surface properties of the metal surface reconstructed and careful calibration of the stereo-rig. The stereo-vision method using two cameras is ideal as it reduces the influence of mechanical moving parts on the accuracy and the stability of the system.

Different reconstruction results are acquired under different conditions. As the blast plate has specular properties the point cloud data obtained from reconstructing an unpainted plate was incomplete. This is because information is lost during scanning as the plate surface has poor reflection properties. These properties lead to:

- false image reflection hence false reconstructed data and measurements.
- poor laser image due to mirror-like reflection on some parts of the metal surface and hence incomplete reconstruction.

For the scanning experiments results using a laser spot, it is concluded that specular reflection is view dependent and can cause large intensity differences at corresponding points in stereo images. This is why there were not always corresponding points in both images for every laser spot directed on the metal surface when the angle of incidence is changed. However, different concavity of the surface can lead to false reflections which lead to mismatching, and eventually cause significant depth errors.

Apart from outliers, a correct and complete reconstruction was obtained from a painted blast plate. This is because the optical surfaces of the object determines the performance of the laser scanner, as shown by the reconstructions obtained for the painted and unpainted metal plate. A total Lambertian surface with a highly reflective index is ideal for laser scanning.

The ICP algorithm used to align two sets of data worked better for the randomly generated data compared to the measured data. This is because of the presence of outliers in the measured data. After using ICP algorithm to align the different reconstructions it can be concluded that the different surface reconstructions obtained at the different instances are consistent. This is because of the small average root mean square distance between points. From Figure 27 the two sets of data are seen to align. The error obtained after the algorithm reached the local minimum was introduced by the fact that every point in one reconstruction did not necessarily have a corresponding point in the other reconstruction, hence introducing a distance error from its closest point.

The accuracy of the system is highly dependent on the accuracy of the calibration. The calibration procedure involved corner extraction from a checkerboard. For images tilted from

the camera's view, it was observed that a re-projection error of up to 3 pixels was obtained depending on the angle of tilt. Because of lens distortion it is difficult to estimate tilted corner points when doing corner extraction. If these images are used in the calibration process they add the pixel error. From this it can be concluded that the type of checkerboard images used in the calibration process should be considered for accuracy.

Accurate laser line extraction is one of the key processes of reconstruction. In this project, the laser was mostly extracted using the maximum intensity-value method. This means that the reconstructed results depend on the camera resolution. For a row on the laser image that had more than one column of a threshold intensity, an averaging of the pixel was done. This contributed to error in the measurements.

University of Cape Town

References

- [1] P.J Besl and N.D McKay. A method for registration of 3D shapes. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.
- [2] Dinkar N. Bhat and Sheer K. Nayar. Binocular stereo in the presence of specular reflection. In *Proceedings of ARPA Image Understanding Workshop*, Monterey, November 1994.
- [3] Jean-Yves Bouguet. *Description of the calibration parameters*. PhD thesis, www.vision.caltech.edu/bouguetj/calib/doc/htmls/parameters., last visited 20 November 2006. Last updated August 10th, 2005.
- [4] Jean-Yves Bouguet. *Camera Calibration Toolbox for Matlab*. PhD thesis, www.vision.caltech.edu/bouguetj/calib/doc. Last visited 20 November 2006.
- [5] Forsyth D.A. and Ponce J. *Computer Vision - A modern approach*,. Prentice-Hall, 2003.
- [6] Farlex. The free dictionary. [http://encyclopedia.thefreedictionary.com/computer vision](http://encyclopedia.thefreedictionary.com/computer%20vision). Last visited on 28 November 2006.
- [7] F.E.Nicodemus, J.C.Richmond, J.J.Hsia, I.W.Ginsberg, and T. Limperis. Geometrical considerations and nomenclature for reflectance. Technical report, NBS Monograph 160 National Bureau of Standards, October 1977.
- [8] Natasha Gelfand, Leslie Ikemoto, Marc Levoy, and Szymon Rusinkiewicz. Geometrical Stable Sampling for the ICP Algorithm. Stanford University, Princeton University.
- [9] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004.
- [10] Katsushi Ikeuchi. Determining surface orientations of specular surfaces by using the photometric stereo method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(6):661–669, 1981.
- [11] Davis James and Chen Xing. A laser range scanner designed for minimum calibration complexity. In *Proceedings of the 3rd International Conference on 3-D Imaging and Modeling*, 2001.
- [12] Forest J.and Salvi J., Cabruja E., and Pous C. Laser stripe peak detector for 3D scanners: A FIR filter approach. *IEEE*, 2004.
- [13] Park J. and Kak A. C. Multi-peak range imaging for accurate 3D reconstruction of Specular Objects. In *In 6th Asian Conference on Computer Vision*, 2004.
- [14] Park J. and Kak A. C. *Specularity Elimination in Range Sensing for Accurate 3D Modeling of Specular Objects*. Robot Vision Lab, Purdue University, West Lafayette, USA.
- [15] Stanford University Computer Graphics Laboratory. The Stanford 3D Scanning Repository. www.graphics.stanford.edu/data/3Dscanrep. Last visited on 23 June 2006.

- [16] Ryan Charles Mathews. Calibration and implementation of a laser strip profiler. Undergraduate thesis, Department of Electrical Engineering, University of Cape Town, 2004.
- [17] Bruno Merven. Person Tracking in 3D using Kalman Filtering in Single and Multiple Camera Environments. Master's thesis, University of Cape Town, August 2004.
- [18] Mukani Bikani Moyo. Three dimensional object reconstruction using structured lighting. Undergraduate thesis, Department of Electrical Engineering, University of Cape Town, 2004.
- [19] Andrews Myers. Introductory literature review of surface reconstruction from three dimensional range data. Mics. notes on surface reconstruction, April 1999.
- [20] Levoy M., Pulli K., Curless B., Rusinkiewicz S., Koller D., Pereira L., and Ginzton M. The digital Michelangelo Project: 3D scanning of large statues. *Computer Graphic, SIGGRAPH'00*, 2000.
- [21] Oren M. and Nayar S. K. A theory of specular surface geometry. *International Journal of Computer Science*, 24(2):105–124, September 1996.
- [22] Shree K. Nayar, Katshushi Ikeuchi, and Takeo Kanade. Surface reflection: Physical and geometrical perspectives. Technical Report 15213, The Robotics Institute, Carnegie Mellon University, March 1989.
- [23] G.N. Nurick. The measurement of the deformation response of a structure subjected to an explosive load using a light interference technique. In *Proceedings of the SEM Spring*, 1986.
- [24] G. N. Nurick and J.B. Martin. Deformation of thin plates subjected to impulsive loading. a review part 2: Experimental studies. *Int. J. Impact Engng*, 8(2):171–186, December 1989.
- [25] M.D. Oslon, G. N. Nurick, and J.R. Fagnan. Deformation and ruptures of blast loaded square plates Prediction and experiments. *Int. J Impact Engng*, 13(2):279–291, 1993.
- [26] Jarvis R.A. A perspective on range finding techniques for computer vision. *IEEE Transaction on Pattern Analysis and Machine Intelligence*., PAMI-5, no. 2., March 1983.
- [27] S. Rusinhkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proceedings 3 Dimensional*, 2001.
- [28] Gergeley Vass and Tamas Perlaki. Applying and removing lens distortion in post production. In *The Second Hungarian Conference on Computer Graphics and Geometry*, 2003. Budapest.
- [29] R.J. Woodham. Photmetric stereo: A reflectance map technique for determining surface orientation from image intensity. In *SPIE vol 155*, 1978.
- [30] <http://psych.hanover.edu/krantz/art/texture.html>, 24 November 2006.