

Department of Spatial Sciences

**Semi-Automated Registration with Direct Linear Transformation
and Quality Evaluation of Digital
Photogrammetry and Terrestrial Laser Scanning**

Kwanthar Lim

**This thesis is presented for the Degree of
Master of Philosophy (Surveying and Mapping)
of
Curtin University**

August 2012

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature:

Date:

ABSTRACT

Photogrammetry and Laser Scanning can be used to complement one another, during instances where digital images are taken of the object of interest with the intention to merge the 3D data and image in order to reconstruct photorealistic virtual models with photo quality and metric realism. Laser scanning acquires 3D data points and intensity information of objects but is unable to directly obtain photorealistic colour in most cases. To get photorealistic colour, some laser scanners come with an on-board camera, or alternatively a separate camera is used, and registration is required for both cases. One example uses a specially designed camera mounting for the laser scanner and another is to transfer colour information from 2D images captured from near the scanner to the 3D points using close-range photogrammetry. Currently limited methods exist for the registration of the data from multiple-sensors. This research outlines the evaluation and semi-automated registration of a single colour image to laser scanning point cloud data, using the canonical transformation and Direct Linear Transformation (DLT) methods for registration.

The method presented in this thesis is to directly reconstruct three dimensional data from a single image with the assistance of estimated depth information. Laser scanning point cloud information is used to supplement the recovery of the estimated depth information, which is then assigned to the image data. Two primary aspects for this research are (1) the Synthetic Camera Image, following on from previous work reported in the literature on utilising synthetic imagery created from point-clouds, and (2) the Direct Linear Transformation model, which is used to provide the transformation parameters between the 2D and 3D datasets.

Synthetic datasets were used to provide an indication of expected results in terms of range, incidence angle and image resolution. The image resolution is an important factor to consider. It should be as high as possible as it affects the outcome of precision. Testing with real data, the proposed method resulted in a precision of 2 mm for the data of a model T-Rex dinosaur dataset, and 19mm for a typical indoor scene. The variations in precision levels are due to different values for range, incidence angle and image resolution. Overall the results achieved the expectations producing a colour point cloud with metric assessment.

ACKNOWLEDGEMENTS

Foremost, I am grateful towards the many people, family, friends and lecturers whom have supported me throughout this project from the beginning to the finishing of this thesis.

I am grateful for my Mum and Dad, having helped me through the moments of being overwhelmed, with their kindness and concern.

Dr. Kwang-Ho Bae, my initial supervisor, deserves an honourable mention, whom I believe is central towards the success of this project, offering a first class level of guidance and support. He always went the extra mile for mentoring, teaching and guiding. He was always ready to help me gain a better understanding and directed me back on the right path when I lost focus.

Thanks to Professor Geoff West for supervising when Kwang-Ho's contract was discontinued.

David Belton is a good help in explaining some programming and mathematical concepts. Along with Kyle Chow and Johnny Lo, for the company and laughs.

Thanks to Alireza Shirkhani for proof reading the photogrammetry aspects, and providing your help as an external and independent set of eyes.

I thank my sisters, Yuemei for proof reading and encouragement, and Liimei for understanding.

I thank my girlfriend Fuyu Kwok for the encouragement she provided.

Finally, thanks go to the department of spatial sciences, especially the administrative staff, Caroline Rockliff, Pam Kapica, Lori Patterson and Meredith Mulcahy, whom deserve hearty thanks for supporting me during this project.

Overall, I thank God for the grace, the wisdom and the guidance throughout the thesis.

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF FIGURES	v
LIST OF TABLES	viii
1 INTRODUCTION	1
1.1 Aim.....	2
1.2 Significance and Approach	3
1.3 Dissertation Outline	4
2 BACKGROUND	6
2.1 Introduction.....	6
2.2 Photogrammetry	6
2.2.1 Camera, Images and scale	7
2.2.2 Calibration and Interior Orientation.....	8
2.2.3 Bundle adjustment and Image Orientation.....	12
2.3 Laser Scanning.....	13
2.3.1 Laser scanner data acquisition	15
2.3.2 Calibration of the laser scanner.....	17
2.4 Registration	19
2.5 Transformations	21
2.5.1 Homogeneous Coordinates and Helmert Transformation.....	21
2.5.2 DLT	22
2.6 Synthetic Camera Image (SCI) and Real Camera Image (RCI)	24
2.7 2D Iterative Closest Point	26
3 METHODOLOGY.....	29
3.1 Core Framework	30
3.1.1 Fundamental Expressions for Projection and Back-Projection.....	30
3.1.2 Correspondences and Collinearity	32
3.1.3 Understanding the DLT	33
3.1.4 Normalisation and 12 th Parameter Recovery	36
3.1.5 Perspective Centre Calculation	37
3.1.6 The Image Scale Component	37

3.2	Proposed Method	38
3.2.1	Rigorous estimation of the DLT parameters	38
3.2.2	RCI and SCI	39
3.2.3	Proposed Back-projection methods	44
3.3	Summary of Methodology	50
3.4	Chapter summary	52
4	RESULTS	53
4.1	Experiment with simulated datasets	53
4.1.1	Creation of simulated dataset	53
4.1.2	Camera Projection Concept used to create the SCI	57
4.1.3	Using the DLT method and the significance of control points	60
4.1.4	Comparison between the three methods of back-projection	63
4.1.5	Test with alternate forms of 3D data: Photogrammetry data	65
4.1.6	Simulated data with error information applied	67
4.1.7	Analysis and Evaluation with simulated dataset	69
4.2	Experiments with Real Datasets	76
4.2.1	Data Acquisition and Specifications	77
4.2.2	Results of Laboratory Dataset	77
4.2.3	Results of T-Rex Dataset	82
4.3	Chapter Summary	84
5	CONCLUSION	87
	REFERENCES	90
	APPENDIX A: Conversion calculation mm-to-pixel	96
	APPENDIX B: Derivative for DLT coefficients	98
	APPENDIX C: Rotation Matrix	100
	APPENDIX D: Conical and Corner Dataset Code	102
	APPENDIX E: Applying Error to Datasets	107
	APPENDIX F: Additional Interior Orientation Descriptions	110

LIST OF FIGURES

Figure 2.1: Geometry of a pinhole camera where the optical axis is defined as the centre line passes through the Perspective Centre (PC).....	7
Figure 2.2: Geometric locations of Interior Orientation parameters.....	9
Figure 2.3: Calibration field used for calibrating the camera (Australis, 2012).	12
Figure 2.4: Resected 3D points with position and orientation cameras with respect to the reconstructed object points (Australis, 2012).....	13
Figure 2.5: (a) Short range scanner (Konica Minolta, 2012) (b) Medium range scanner (Leica, 2012) (c) Long range scanner (Rieggl, 2012).....	16
Figure 2.6: Example of data from a laser scanner (a) Greyscale representation (b) Single-colour indicating visual issue (c) Intensity scale representation.....	17
Figure 2.7 Areas including corresponding points: (a) Digital photo (b) Point-cloud	20
Figure 2.8: Registration of (a) point-cloud to image (b) image to point-cloud	20
Figure 2.9: Visual representation of 3D points projected onto a plane.....	24
Figure 2.10: (a) A synthetic image (SCI) (b) A photo (or RCI)	26
Figure 2.11: (a) Large offset of points before ICP (b) Both datasets closer post ICP	27
Figure 2.12: Brief summary of the conventional ICP algorithm	28
Figure 3.1: A representation of the projection of control points to SCI points.....	40
Figure 3.2: The SCI and RCI control points to be closer after registration.	41
Figure 3.3: The RCI is transformed to match the SCI.	42
Figure 3.4: The effect of an offset in the image plane causing a misrepresentation of colour assignment in the point cloud.	42
Figure 3.5: RCI at pixel level: Each pixel may contain multiple SCI points.....	43
Figure 3.6: Geometric diagram to indicate multiple points per pixel as a function of range.....	43
Figure 3.7: Colour information from the RCI points is stored at the corresponding SCI points.....	44
Figure 3.8: Back-projection representing points to rays concept and to emphasise significance of the image depth value.....	45
Figure 4.1: Using similar triangles and ratios for the creation of the Roof Dataset. .	54
Figure 4.2: The Roof Dataset, a roof-like structure with views: (a) overall (b) top (c) side.	55

Figure 4.3: The Corner Dataset, a box corner structure with views: (a) overall (b) top (c) side.	55
Figure 4.4: The Conical Dataset, a cylindrical cone structure with views: (a) overall (b) top (c) side.	56
Figure 4.5: The Conical Dataset with control points highlighted in red and 3D map of the location with only control points.....	56
Figure 4.6: SCI image of the Roof Dataset and the Corner Dataset with numbers representing individual points of the dataset.....	58
Figure 4.7: SCI of the Conical and control points.	58
Figure 4.8: The back-projection of the SCI produces 3D points that form a plane (left). The red line in the left image indicates perspective direction which matches the control points (right).	59
Figure 4.9: Another example with Corner Dataset of the back-projected 3D points matching at a certain perspective.	59
Figure 4.10: Back-projected SCI points to the original location. Left is back- projected Roof Dataset and right is back-projected Corner Dataset.....	60
Figure 4.11: Image comparison of the SCI _{CPM} and SCI _{DLT} , and after ICP registration.	61
Figure 4.12 The RCI image of the Car Door dataset (Australis, 2012).	66
Figure 4.13 The ICP results of the RCI to SCI for the Car Door Dataset.....	66
Figure 4.14 The overlaid SCI onto the RCI including check points (indicated by the '+' symbol).....	67
Figure 4.15: The transformed RCI overlaid onto the SCI.....	68
Figure 4.16: Differences from the back-projection to the original, caused by the mismatch between the SCI and RCI, with a close-up of a selected region.....	69
Figure 4.17: RMS error of image and back-projection, in terms of range.....	72
Figure 4.18: The RCI of the Conical Dataset at range 60m, simulating reduced resolution.....	72
Figure 4.19: RMS error of image and back-projection, in terms of incident angle at range 25m.....	74
Figure 4.20: RMS error of image and back-projection, in terms of incidence angle at optimal range of 12m.	75
Figure 4.21: Image showing the cause of error at the limit of incident angle.....	75

Figure 4.22: The range should be chosen so that the pixel and coverage area will allow the desired precision.....	76
Figure 4.23: An image of the setup area indicating the Leica HDS targets (red), and printed Black and White targets (green), that are used as control points.....	78
Figure 4.24: A plan of the setup area.....	78
Figure 4.25: ICP results of the Lab Dataset in pixels.	79
Figure 4.26: SCI of the point cloud data in pixels.	80
Figure 4.27 The overlaid SCI onto the RCI with incorrectly stored data outlined within the red region.	81
Figure 4.28: The back projection points with colour assigned in metres.....	81
Figure 4.29: The RCI (left) and the point cloud for the T-Rex dataset (right).	83
Figure 4.30: The SCI of point cloud data overlaid onto RCI with red regions indicating data with incorrect colour assigned.....	83
Figure 4.31: The back projection of T-Rex data with colour assigned. The dimensions are in metres.....	84

LIST OF TABLES

Table 2.1: Summary of the interior orientation parameters	10
Table 2.2: Technical details of some commercial laser scanners (Minolta, 2012; Leica, 2012; Riegl, 2012).....	16
Table 3.1: Description of the sub-matrices related to the graphical transform.....	50
Table 3.2: Stages in the overall process of the proposed method.	51
Table 4.1: Parameters used for the three datasets.	57
Table 4.2: Difference of RCI coordinates of control points after ICP.....	62
Table 4.3: RMS error between RCI and SCI, and RMS error of back-projection results for control points having varying levels of introduced errors.	63
Table 4.4: Pseudo-inverse method: Difference between back projection and original control point coordinates.....	63
Table 4.5: Two-Step method: Difference between back projection and original control point coordinates.....	64
Table 4.6: Additional Fourth-Row method: Difference between back projection and original control point coordinates.	65
Table 4.7: Technical specifications of the camera used within the Australis software.	66
Table 4.8: Results and RMS for Back-projection using Pseudo-inverse, Two-Step or Additional Fourth-row methods.....	67
Table 4.9: RMS error results of Conical Data distance tests.	71
Table 4.10: RMS error results of Conical Data incident angle tests.....	73
Table 4.11: RMS error results of the Conical Data incident angle tests.....	75
Table 4.12: Technical specifications of the camera used (Nikon, 2012).....	77
Table 4.13: Technical specifications of the Terrestrial Laser Scanner used.....	77
Table 4.14: Results and RMS for Back-projection using Pseudo-inverse, Two-Step or Additional Fourth-row method.	82

1 INTRODUCTION

3D models of existing buildings, vehicles and other objects can be created through photogrammetric methods. These can have photo quality colour, texture, accurate details and other life-like features to depict their counterparts in reality. These are generally called photorealistic reconstructions or models. There is always a need to improve the representation of virtual models and generally the aspect of realism is sought after, which can be created by the addition of extra information such as colour or points. 3D modelling of existing objects is becoming increasingly popular, with the main applications in heritage preservation and city modelling. Multiple images along with photogrammetric techniques may be used to create metrically accurate while visually stimulating photorealistic models, though it requires a great amount of time, effort as well as experience.

One of the many advances in technology is the introduction of laser scanners, and there has been much research conducted into 3D laser scanning in terms of calibration, registration and segmentation (Vosselman and Maas, 2010). Laser scanning has recently gained popularity in the applications of cadastral and engineering surveying. Additionally they have been utilised with close-range photogrammetry for terrestrial, airborne and mobile applications. It offers a simpler approach to modelling, as it can directly acquire a 3D point cloud of objects in a relatively short time span. Photogrammetric 3D reconstruction can provide additional information, e.g. radiometric or reflective information of objects, to 3D point clouds from laser scanners. However, it has a limitation in that there is little flexibility (in terms of image resolution and perspective) in providing colour to its point cloud. A combination of laser scanner along with mounted cameras (whose location has been precisely obtained) as seen in Jansa et al. (2004), Sapkota (2008) and Kern (2001), can offer better flexibility and efficiency of photorealistic model creation.

This research presents an alternative method for photorealistic model reconstruction, focusing on flexibility, whereby the 3D model is obtained using a laser scanning device and the colour information is obtained using one image from any calibrated camera without the need of specialised mounting equipment. The major obstacle of 3D reconstruction of either single or multiple images is the recovery of depth

information of 3D points (Luhmann et al., 2006). In cases of multiple images, 3D photogrammetric resection (McGlone et al., 2004) with reflective targets has been successfully utilised in the last few decades in terrestrial and airborne applications. With close-range photogrammetry, relevant target detection, calibration and resection can be done with minimum user intervention using available commercial software (e.g. Australis, 2012; Adam Technology, 2012). However, a method for a single image has yet to be developed for 3D reconstruction or back-projection (Hartley and Zisserman, 2003), simply because it is not possible without some other external information such as that from laser scanner data.

This research was initiated from an idea that in the context of surveying applications, using terrestrial laser scanner data along with a single image, it may be possible to utilise 3D point clouds for estimating the depth information of 2D images. Two potential ideas were (1) to use Direct Linear Transformation (DLT) (Abdel-Aziz and Karara, 1971) and (2) to directly calculate the exterior orientation of a camera along with the camera calibration parameters with synthetic images (Forkuo and King, 2005), for the reconstruction of 3D points containing additional information. This allows us to obtain much more detailed or additional information to the 3D point cloud, not just the colour information. For example, in cases of mobile laser scanning systems that have been gaining in popularity, a procedure to estimate the depth information of 2D images is important since it provides a way of obtaining metric information of objects other than from 3D point clouds obtained from terrestrial laser scanners, which is usually installed next to the camera system, e.g. spherical multiple cameras. Although it may be possible to extract the depth information from multiple-images produced by this system, a more practical approach would be utilising 3D terrestrial laser scanners to assist this process instead of solely relying on the camera location from GNSS and INS.

1.1 Aim

The aim of this research is to develop methods that can utilise a single image and a 3D point cloud of the corresponding object/region to recreate a metrically precise photorealistic 3D object model. This research is in the field of photogrammetry and terrestrial laser scanning with the focus on registration of data from multiple sensors.

Terrestrial laser scanning, photogrammetry principles along with the Direct Linear Transformation (DLT) are the primary components used to assist the registration and back-projection between 3D and 2D spaces.

1.2 Significance and Approach

The fusion of photogrammetry and laser scanning can be seen in Forkuo and King (2005), Kang, Zlatanova and Gorte (2007), Al-Manasir and Fraser (2006), Briese, Pfeifer and Haring (2003) and Beraldin (2004). This research of combining and evaluating 2D and 3D data, intends to further the concept of photogrammetry and laser scanning fusion by opening possibilities for improvements in the future in terms of:

- (a) **Photorealistic 3D reconstructions:** This has the potential in aiding heritage preservation and city model reconstruction with known metric model quality. There has been much research in European Nations and Japan. For example see Briese, Pfeifer and Haring (2003), Liu et al. (2006), and Ikeuchi et al. (2003).
- (b) **Quality Control:** Having a metric measure presents an opportunity for registration results to be standardised to have a common index for comparison and an indicator of robustness.
- (c) **Feature detection:** Current feature detection relies on intensity, corners, edges, planes, points (Vosselman and Maas, 2010). This method will allow colour to be used to open new avenues of feature detection in 3D datasets, by way of transferring any 2D feature, lines or edges into 3D space as an initial estimation for 3D feature detection.
- (d) **Hardware:** To provide hardware flexibility, measurement quality and cost effectiveness in industrial applications such as surveying.
- (e) **Visualisations:** Improvements can be added to current software to automate this process to provide a quick photorealistic representation of the survey area.
- (f) **Automation:** Efficiency may potentially be increased with reduced times for manual selection and acquiring of commonly repeating tasks and geometric primitives.

In surveying and mapping, metric precision is valued over the visual appeal of image overlays and texture mapping. It is important to have the measurement data displayed accurately and precisely, so that it may reflect reality in terms of metrics.

Using a combination of photogrammetric techniques along with the DLT, colour information is registered from a single camera image onto a 3D point cloud. The distinction is that utilising a single image without the need of mounting brackets or fixtures, is registered to provide point clouds with colour information using the DLT approach.

1.3 Dissertation Outline

Chapter 1 provided the basis of this research, looking briefly at the current environment of registration fusion, which in this case is image and 3D point-cloud registration. The aim is to provide a solution that allows transformation from 3D to 2D, and vice-versa, with a single image and 3D point-cloud. Its significance carries into many areas such as heritage preservations and city modelling by creating metric photorealistic 3D reconstructions, for quality control, feature detection, hardware, visualisations and automation. The flexibility of the method is the distinguishing factor when compared to other approaches.

Chapter 2 introduces the main concepts and principles required in order to understand the research. A brief introduction into photogrammetry and laser scanning is presented, outlining even the fundamental workings of a camera and laser scanner. The key focus of the background is transformations with consideration to image capture and synthesis, as well as registration of datasets. Iterative Closest Point for image correspondence is described along with iterative least squares. Occlusion is discussed with brevity, as it does not pose any significance yet it is not to be overlooked.

Chapter 3 looks into the methodology explaining the core principle of 3D to 2D (projection) and 2D to 3D (back-projection) transformations. The DLT is explained in depth, covering its relationship to the collinearity equations, image scale component, perspective centre and normalisation of the 12th parameter. Each step is

thoroughly discussed from acquiring the data, to projection and registration, and back-projection. Data acquisition includes obtaining 2D and 3D data, from the camera and laser scanner respectively, along with corresponding control point information. Then in terms of projection, creation of a Synthetic Camera Image (SCI) is explained followed by the Real Camera Image (RCI), which contains the colour information, and then the SCI using the Iterative Closest Point (ICP) method. Finally, the back-projection concept is detailed with three methods, the Direct Pseudo-Inverse method, the Two-Step solution, and the Additional Fourth-row method.

Chapter 4 unveils the results of the proposed method using both synthesised data and real data. The initial testing of the simple back-projection shows that the three back-projection methods are needed in order to transform back to the original location. Random errors were added to the simulated datasets to mimic reality, which was then tested in terms of range and incidence angle. Tests with different ranges indicated that errors increased with distance. The incident angle tests showed that a maximum of 30° will maintain a good level of precision. Range and resolution were interdependent meaning that an increased distance required higher pixel density to obtain good precision. The two real datasets tested shows consistent results with the tested simulated datasets, with the caveat that some thresholds and/or deletion of data or a different image viewpoint would be required to minimise incorrectly transformed data.

2 BACKGROUND

2.1 Introduction

This chapter outlines the concepts and fundamental principles that are required for an extensive understanding for this research. The main concepts, which are Direct Linear Transformation (DLT), registration and Synthetic Camera Image (SCI), and the basics of Photogrammetry and Laser Scanning, are explained in detail. One can find further details of these topics in relevant text books and papers mentioned in this chapter, such as McGlone et al. (2004), Luhmann et al. (2006), Hartley and Zisserman (2003), Vosselman and Maas (2010), Mikhail, Bethel and McGlone (2001), Atkinson (1996), Wolf and Dewitt (2000), Kern (2001), Forkuo and King (2005), Abdel-Aziz and Karara (1971) and Karara (1989).

2.2 Photogrammetry

Photogrammetry is a measurement technology with a focus on terrestrial surveying, used in situations requiring remote observations and measurements of physical objects or phenomena. Therefore, its fundamental objective has always been the indirect and precise determination of the spatial position of objects using photographic images and relevant mathematics, e.g. least-squares (McGlone et al., 2004). Major advantages of using photogrammetry can be listed as follows: images being a common medium, a form of non-contact measurement, provision of high-precision of measurement and ability to achieve very small details (Kern, 2001). However its limitations are that corresponding points or features for 3D geometry are necessary and it is a two-phase process, which is taking images and then examining its geometry by merging the images (Kern, 2001). These processes can also be computationally intensive, and inhomogeneous (not constant) precision on non-target areas can persist even after rigorous calibration and reconstruction (Luhmann et al., 2006).

Photogrammetric techniques are often used to determine three dimensional (3D) digital or graphical representations, e.g. location, size, and shape, of objects by means of measuring and analysing two-dimensional (2D) photographs (McGlone et al., 2004; Luhmann et al., 2006; Jiang, Jáuregui, and White, 2008). In order to

produce a 3D representation of an object, which is usually called 3D Photogrammetric or Photorealistic reconstruction, a minimum of two images with sufficient corresponding points are necessary since the depth information of 3D points needs to be estimated. Alternative methods exist to obtain this 3D construction, such as instruments with a two camera setup, e.g. AICON (2012) or a single camera with a grid projector along with photogrammetric principles to construct 3D data, e.g. Luhmann et al. (2006), although reconstruction can be computationally intensive.

2.2.1 Camera, Images and scale

Understanding the basics of cameras, images and scale is essential as this allows the later concept of Synthetic Camera Images (SCI, Forkuo and King, 2005) to be well understood for this research. Briefly speaking, the SCI is an artificial 2D image created from 3D point clouds by Terrestrial Laser Scanners (TLS), which is to be explained in detail later on in this chapter (Section 2.5).

The basic pinhole model is commonly used to explain the process and effect of a camera capturing an image of an object as shown in Figure 2.1. A more in-depth look into the mathematics and geometry of the image capturing process has been presented in McGlone et al. (2004), Hartley and Zisserman (2003), Luhmann et al. (2006), Mikhail, Bethel and McGlone (2001), Atkinson(1996) and Karara (1989).

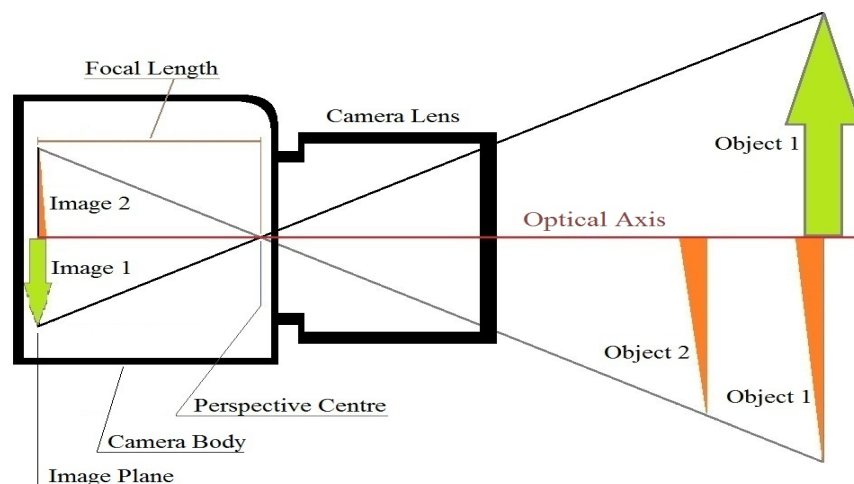


Figure 2.1: Geometry of a pinhole camera where the optical axis is defined as the centre line passes through the Perspective Centre (PC).

The pinhole camera model indicates that light rays of the objects of Figure 2.1 will travel through the Perspective Centre (PC) at a focal length (f) and captured as an image on the imaging device, e.g. film or CCD (Charge-Coupled Device) array. In short, an image is a scaled planar 2D representation of the 3D objects.

Depth information is in the form of the perpendicular distance of points on an object to the image plane, i.e. the surface of film or CCD and scale is the ratio between the focal length and the depth information (McGlone et al., 2004; Hartley and Zisserman, 2003). One can easily deduce that the depth information is lost during the photographic process presented in Figure 2.1. Furthermore, the objects with different depth, e.g. Objects 1 and 2 in Figure 2.1, can be represented as the same object in the 2D image space, i.e. Image 2 in Figure 2.1. In order to regain or estimate the depth information, multiple images with resection has been successfully utilised (McGlone et al., 2004; Hartley and Zisserman, 2003). With the assistance of 3D points from terrestrial laser scanners, this research presents an alternative solution using a single image.

2.2.2 Calibration and Interior Orientation

Camera calibration is performed in order to determine the geometric model of a camera, which is described by its interior orientation parameters (Luhmann et al., 2006). Evaluation of the performance and stability of a lens, determination of the optical and geometric parameters of a lens, lens-camera system, or of an imaging data acquisition system are several of the objectives met by the camera calibration (Atkinson, 1996).

Network design is considered a fundamental component of close range photogrammetry, as it impacts accuracy and precision of the final 3D point coordinates by photogrammetric reconstruction with the least-square exterior orientation. In order to maximise the resultant accuracy, multiple convergent images need to be used with a good geometric strength of the ray intersections, which is significantly dependent on the network design for both calibration and actual measurement as it is for control and tie point measurement in surveying. The strength of ray intersections is influenced by the number of rays determining each point and

the angles between them (Luhmann et al., 2006). However, this is primarily of importance to multiple image reconstructions of 3D points, whereby this research presents an idea of utilising a single image assisted by already determined 3D points from terrestrial laser scanners.

Calibration is also necessary in terms of reducing the effects of major distortions, e.g. radial lens distortion, for providing corresponding matches of the photo to the SCI. Notably, distortion will be present even with calibration, thus the automated matching method, e.g. Iterative Closet Point (ICP) (Besl and McKay, 1992), is implemented to assist the correspondence of closely-matched images, for this purpose of depth assignment. Metric cameras or other imaging systems with good camera stability are desirable to minimise the effects of distortion on the image, though it is not essential.

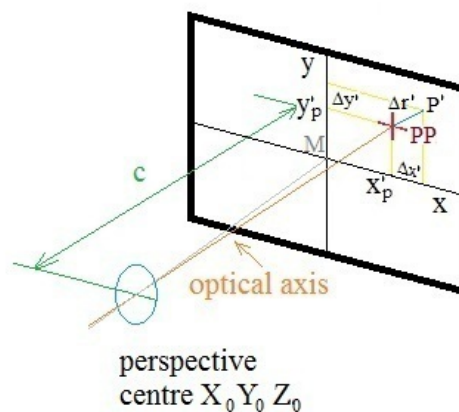


Figure 2.2: Geometric locations of Interior Orientation parameters

Interior orientation describes the parameters that model the passage of light rays through the lens onto the image plane, and it is essential to know, or to compensate for, these disparities with computations, before the use in photogrammetric projects (Karara, 1989; Atkinson, 1996). Elements of interior orientation are constants, which are needed so that accurate spatial information is to be determined from photographs (Wolf and Dewitt, 2000). The parameter set that the camera calibration process usually determines are: the principal distance (c) sometimes referred as the calibrated focal length (f), the Principal Point of Auto-collimation (PPA), affinity and shear, as well as the radial (symmetric) and decentring distortion of the lens (Karara, 1989; Wolf and Dewitt, 2000; Luhmann et al., 2006). These parameters will be

subsequently described in conjunction with Figure 2.2. The principal distance or the calibrated focal length (primarily referred to in the context of aerial photogrammetry when the focus is at infinity) is the perpendicular distance from the perspective centre of the lens system to the image plane (Karara, 1989; Atkinson, 1996). For close range photogrammetry, it is the principal distance that must be determined for every exposure if the camera is refocused, as the principal distance will vary (Atkinson 1996).

The following descriptions are linked to Table 2.1 and Figure 2.2. The Principal Point of Auto-collimation (PPA) is the ideal case where the optical axis of the lens perpendicularly intersects the image plane, tested with auto-collimation methods. However, the practical case favours the use of analytical methods. The analytical methods estimate the PPA based on the relative distortions present, shifted from the image centre, and uses alternative terminology, Principal Point (PP) or Principal Point Offset (Clarke, Wang and Fryer, 1998). For close range photogrammetry, the image centre (M) is approximately equal to the nadir of the perspective centre, i.e. $PPA \approx M$. The symmetric radial lens distortion (see Appendix F), which is the main constituent of image distortion for most camera systems, can be attributed to variations such as refraction, lens design, focusing distance and object distance at a constant focus. This distortion increases with the distance from the PPA. The tangential or decentring lens distortion (see Appendix F), formally known as radial-asymmetric distortion, is mainly caused by decentring and misalignment of the lens elements. The affinity and shear describe the deviations of the image coordinate system (see Appendix F), with respect to skew from orthogonal and uniform scale of the coordinate axes (Luhmann et al., 2006).

Title	Notation
Principal Point of Auto-collimation	x'_0 and y'_0
Image Coordinates	x'_p and y'_p
Radial Lens Distortion	$\Delta r'$ or $\Delta x'_{\text{radial}}$ and $\Delta y'_{\text{radial}}$
Tangential or Decentring Lens Distortion	$\Delta x'_{\text{tan}}$ and $\Delta y'_{\text{tan}}$
Affinity and Shear	$\Delta x'_{\text{aff}}$ and $\Delta y'_{\text{aff}}$

Table 2.1: Summary of the interior orientation parameters

The total correction combines the correction terms for imaging errors to apply to the image, to provide an ‘error-free’ image using equations (2.2) and (2.3):

$$\begin{bmatrix} \Delta x' \\ \Delta y' \end{bmatrix} = \begin{bmatrix} \Delta x'_{\text{radial}} + \Delta x'_{\text{tan}} + \Delta x'_{\text{aff}} \\ \Delta y'_{\text{radial}} + \Delta y'_{\text{tan}} + \Delta y'_{\text{aff}} \end{bmatrix} \quad (2.2)$$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x_p' - x_0' - \Delta x' \\ y_p' - y_0' - \Delta y' \\ -c \end{bmatrix}. \quad (2.3)$$

where equation (2.2) is the combined correction for the radial lens distortion, decentring lens distortion, affinity and shear. Equation (2.3) is then used to correct the image for the combined distortion previously mentioned along with the PPA and PPO. It is essential to have the corrections applied in order to obtain a highly accurate match during the correspondence process for registering the camera image to the SCI.

An experiment was set up in the photogrammetry laboratory of the Department of Spatial Sciences Studio at Curtin University and used for calibration of a typical camera. Figure 2.3 shows the calibration field that was used to calibrate the camera used for this project. Circular targets were used; with its centre representing the actual 3D point to be measured, while being ideal as it is rotation-invariant and scale-invariant over a wide range of image magnifications (Luhmann et al., 2006).

Retro-reflective targets (105 sets) were utilised for self-calibration of a consumer-grade camera (Nikon D80) with commercial software (Australis, 2012). Although there are not enough points in the depth direction to the camera axis, the resultant exterior orientation and values of camera calibration errors were acceptable.

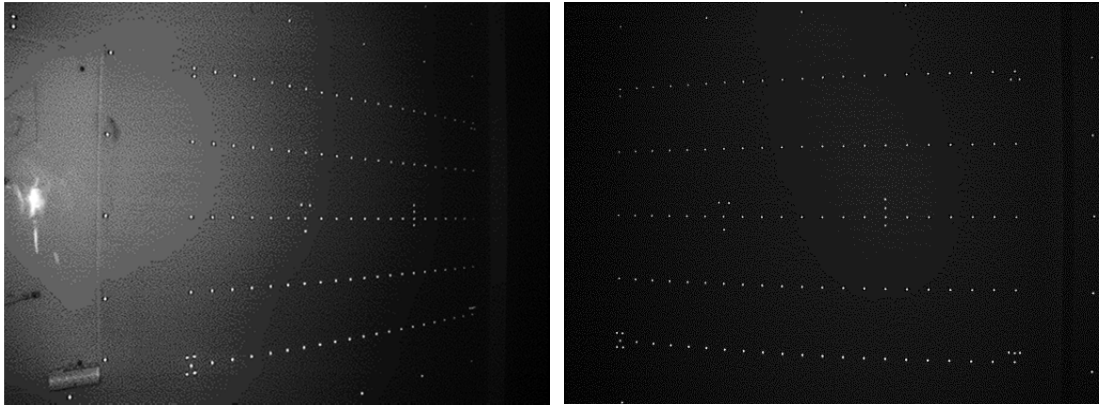


Figure 2.3: Calibration field used for calibrating the camera (Australis, 2012).

Circular retro reflective targets were chosen as it allows the target to be easily distinguished by providing a high contrast for edge detection of targets at a variety of angles and illumination conditions. Figure 2.4 shows an example of the output with the determined camera locations and orientations, in which the camera is represented by a small graphic consisting of a red body, green lens system and the orientation indicated by a small blue shutter button. The method of calibration used for this research is the ‘on-the-job’ calibration method used in conjunction with a test field, in which a field of object points as targets with known coordinates is imaged from several stations, as described by Luhmann et al. (2006).

Good ray intersections are ensured by capturing a series of images of the test field perpendicularly and obliquely with a relative rotation of 90° around the optical axis, as seen in Figure 2.4. The interior orientation parameters, exterior orientation parameters and adjusted test field coordinates are determined by processing the measured image coordinates and approximately known object data using bundle adjustment (see section 2.1.3). Text books such as Luhmann et al. (2006), Atkinson (1996), and Hartley and Zisserman (2003) provide greater detail for the other calibration methods in the context of Photogrammetry and Computer Vision.

2.2.3 Bundle adjustment and Image Orientation

Exterior orientation consists of six parameters describing the spatial positions and orientation of the camera as shown in Figure 2.4 with respect to the global object

coordinate system. Space resection is the procedure used to calculate exterior orientation parameters (Luhmann et al., 2006; McGlone et al., 2004).

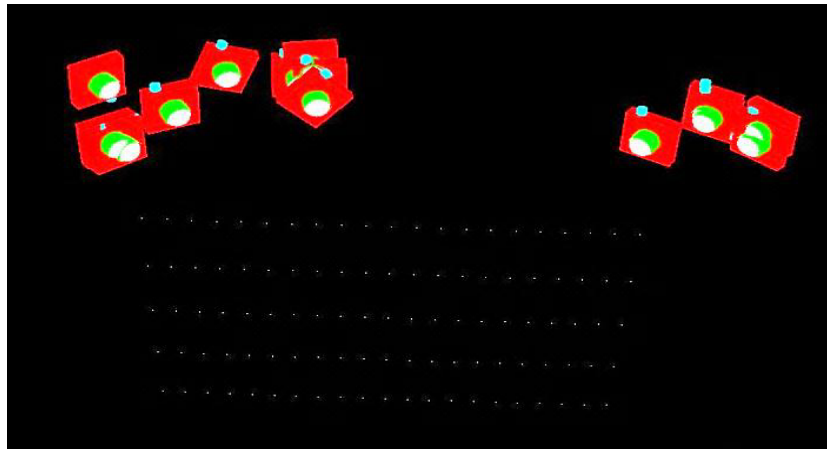


Figure 2.4: Resected 3D points with position and orientation cameras with respect to the reconstructed object points (Australis, 2012).

However, Wolf and Dewitt (2000) states that, generally the exterior orientation is not of interest to photogrammetrists, but is included for the purpose of consistency in the mathematical model. The context of that statement was in regards to bundle adjustment. Bundle adjustment (bundle block adjustment, multi-image triangulation, multi-image orientation) is a simultaneous numerical fitting method for an unlimited number of spatially distributed images that takes into account photogrammetric observations (measured image points), survey observations and an object coordinate system. The bundle adjustment is easily performed and is the most powerful and accurate method used in photogrammetry for determining image orientations and points. However, there are practical problems associated with bundle adjustment as follows: the solution is a large system of normal equations (up to few thousand unknowns), there is a need to generate approximate values of the unknowns and the need to detect and eliminate gross data errors (McGlone et al., 2004; Luhmann et al., 2006).

2.3 Laser Scanning

Laser scanners are non-contact measurement instruments used for both airborne and terrestrial cases (Vosselman and Maas 2010). Terrestrial laser scanning (TLS) is an innovative surveying technology allowing the user to capture large amounts of 3D

data directly, rapidly and with high accuracy (Reshetyuk, 2009). They capture and record the geometry and sometimes textural information of visible surfaces of objects and sites, producing a quantitative 3D digital representation (point-cloud) of the surface in a given field of view (Vosselman and Maas 2010). TLS can be used in many fields such as deformation monitoring, 3D modelling of buildings or other large scale scenes alongside photogrammetry and surveying methods (Liu et al., 2006). According to Jansa et al. (2004) and Kern (2001), laser scanners have the benefits of having high spatial resolution, very good spatial coverage, moderate reconstruction effort, high 3D point density and depth accuracy, while the limitations are its colour, texture reconstruction, and high instrument costs. In terms of cameras, the two main attributes that may assist in reducing the limitations of laser scanning are the ability to provide colour information and the means for texture reconstruction. Presently, laser scanning is being used in conjunction with photogrammetry in the following applications:

- (a) **Architecture and heritage preservation applications:** including city modelling, heritage preservation and restoration, and art and cultural analysis (Briese, Pfeifer and Haring, 2003; Liu et al., 2006; Ikeuchi et al., 2003).
- (b) **Engineering and surveying applications:** requiring measurement of deformations and change detection, tunnel profiles and concrete tanks, all of which require high measurement precision (Vosselman and Maas, 2010).
- (c) **Automotive applications:** requiring measurement of surface design models for parts analysis, for deformation and safety testing, and the inspection of parts (Vosselman and Maas, 2010).
- (d) **Industrial applications:** including pipe and machinery location for power stations; aircraft and aerospace requiring extremely high accuracy for measurement of corner fittings and mechanically and thermally stressed objects; forensics using photogrammetry for crime scene and accident measurements, and reconstructions that can potentially use laser scanning for a more efficient means of 3D reconstruction (Vosselman and Maas, 2010).

2.3.1 Laser scanner data acquisition

The laser scanner technology offers three basic methods of measurement, phase based, pulse based and triangulation (Vosselman and Maas 2010). Phase based measurements modulates the amplitude, power or wavelength of the light forming a continuous pattern, such as a sine wave, to measure the range based on the phase difference between the outgoing and the incoming wave (signal). The pulse based method repetitiously sends short bursts of light, and measurements are calculated based on the time difference between sent and received pulses. For shorter distances (usually under 5 metres) a triangulation approach is used, whereby measurement is based on the geometry of the laser and the detector at the instance of measurement. A detailed explanation of these methods and relevant history is provided in the textbook by Vosselman and Maas (2010). The laser scanner used for this research is a medium range scanner which uses the pulse based method. Table 2.2 gives the specifications for three different laser scanners shown in Figure 2.5.

Short range	Konica Minolta Specifications	
	Measuring Method	Triangulation, light block method
	Scan Range	0.6 to 1.0 m (In Standard mode) 0.5 to 2.5 m (In Extended mode)
	Laser Class	Class 2 (IEC60825-1)
	Accuracy	±0.05 mm
	Precision	±0.008 mm
	Image size	640x460 Pixels
Medium range	Leica HDS Specifications	
	Measuring Method	Pulsed, time-of-flight
	Scan Range	300 m
	Laser Class	Class 3R (IEC 60825-1)
	Accuracy	±6 mm
	Precision	±4 mm
	Image size	1024x1024 pixels

Long range	Riegl Specifications	
	Measuring Method	Pulsed, time-of-flight
	Scan Range	4000 m
	Laser Class	Class 3R (IEC 60825-1)
	Accuracy	±15 mm
	Precision	±10 mm
	Image size	2560 x 1920 Pixels

Table 2.2: Technical details of some commercial laser scanners (Minolta, 2012; Leica, 2012; Riegl, 2012).

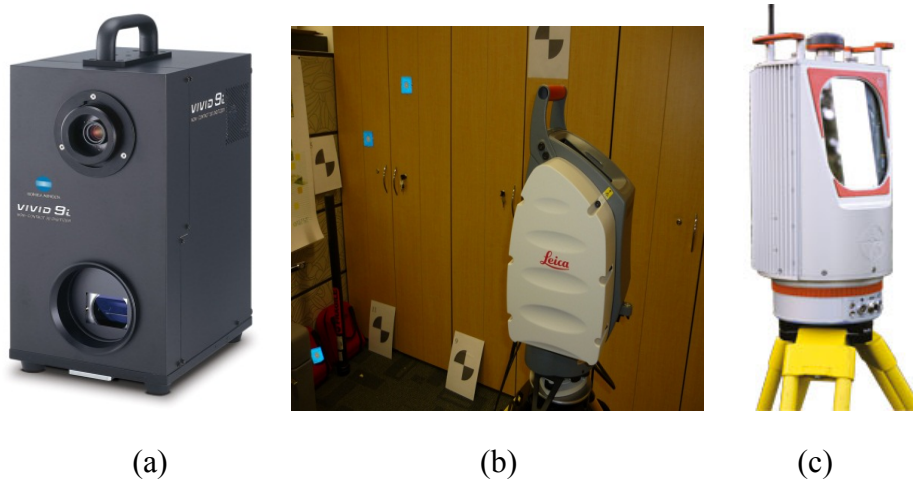


Figure 2.5: (a) Short range scanner (Minolta, 2012) (b) Medium range scanner (Leica, 2012) (c) Long range scanner (Riegl, 2012)

There are many laser scanning devices on the market that are used for the applications described previously. The three main classes of terrestrial laser scanner exist to provide the necessary accuracy required for the scanning range. Most short and medium range scanners also come equipped with a camera that is primarily used for inspection and visualisation purposes.

An example of the data acquired by a laser scanner, i.e. conventional 3D point-cloud data would be hard to visually distinguish as shown in Figure 2.6(b) without some form of animation or motion. Therefore, techniques have been developed to provide

a better visual representation for the point-cloud data, such as intensity mapping or scaling that may be applied as seen in Figure 2.6 (a) and (c), respectively.

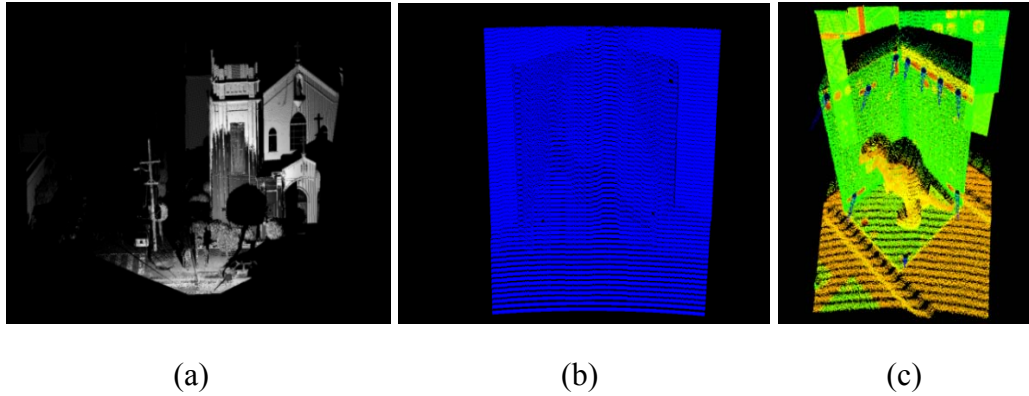


Figure 2.6: Example of data from a laser scanner (a) Greyscale representation (b) Single-colour indicating visual issue (c) Intensity scale representation

However, most current laser scanners contain a built-in camera to capture images to register as explained in Section 2.3. This provides a means of colour representation for the point-cloud as shown in Figure 2.8. The images acquired only offer a maximum image resolution of 1 megapixel and often with blurred images (i.e. poor sharpness), thus a need for this research offering the image resolution flexibility and clarity.

2.3.2 Calibration of the laser scanner

Although the research was conducted with un-calibrated point clouds, it is important to understand the possible systematic error sources of a terrestrial laser scanner since resultant 3D reconstruction errors in the proposed method will come from un-modelled systematic errors, positional and angular uncertainty of 3D point clouds and imperfectness of the camera calibration, to name a few. One can find details on rigorous TLS calibration methodology in Bae and Lichti (2008), Reshetyuk (2009), and Vosselman and Maas (2010).

The measurement principle of a terrestrial laser scanner is the same as an Electric Distance Measurement (EDM) or total station. In short, it measures the range, horizontal and vertical angles of a target. The spherical co-ordinate observations of

the i^{th} object space point in the j^{th} scanner space are range, ρ_{ij} , horizontal direction, θ_{ij} , and elevation (vertical) angle, α_{ij} , which are parameterised in terms of scanner space Cartesian co-ordinates (x_{ij} , y_{ij} , z_{ij}) as:

$$\rho_{ij} = \sqrt{x_{ij}^2 + y_{ij}^2 + z_{ij}^2} + \Delta\rho \quad (2.4)$$

$$\theta_{ij} = \tan^{-1} \left(\frac{y_{ij}}{x_{ij}} \right) + \Delta\theta \quad (2.5)$$

and

$$\alpha_{ij} = \tan^{-1} \left(\frac{z_{ij}}{\sqrt{x_{ij}^2 + y_{ij}^2}} \right) + \Delta\alpha \quad (2.6)$$

where $\Delta\rho_{ij}$, $\Delta\theta_{ij}$, $\Delta\alpha_{ij}$ are the radial, horizontal and vertical systematic error sources, respectively, including range offset, collimation and trunnion axis errors and so on (Reshetyuk, 2009; Vosselman and Maas, 2010).

There are two available methods for terrestrial laser scanner calibration, which are individual component calibration and complete system self-calibration, of which the latter is more ideal. Some limitations to the individual component calibration are that some scanners, due to their functions and design, cannot be centred over a known point, while the lack of a telescope does not allow manual determination of the systematic errors (Vosselman and Maas, 2010).

Terrestrial laser scanners sample a scene in uniform angular increments and the errors can be estimated through a highly redundant, network-based self-calibration procedure. Data capture of a large number of targets in different locations and orientations is required for this method, and then least squares is used to simultaneously estimate the object space coordinates of the targets, the exterior orientation of each scanner frame as well as all additional parameters (Vosselman and Maas, 2010).

2.4 Registration

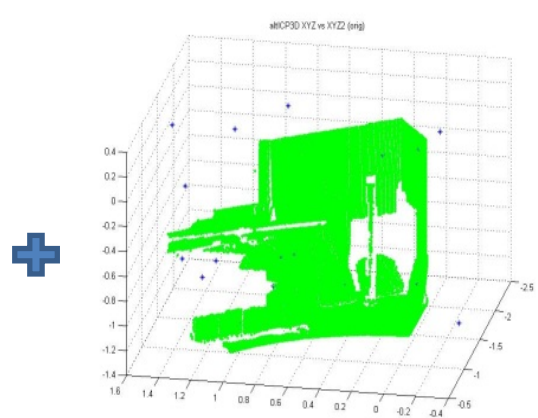
Registration is the process of transforming data from different sources and/or of different types into a reference system to allow for measurements and interpretations to take place. Generally registrations encompass components that relate to scale, skew, rotations and translations (Luhmann et al., 2006). Examples of the registration process can be found in Kang et al. (2007), Zitová and Flusser (2003), Al-Manasir and Fraser (2006) and Bae and Lichti (2008).

In the case of a large project involving multiple medium and long range laser scanning datasets, registration would be required to combine the data together into one dataset. Control points and other corresponding points or features may be used in this process to match and orientate the datasets into one main coordinate system. This research uses corresponding control points in the image and the 3D dataset for registration.

Figure 2.7 is an example of the registration process, where the image and point-cloud are registered to each other resulting in Figure 2.8 (a) or (b). Figure 2.8 (a) is an Real Camera Image (RCI) with targets in 2D manually identified (in blue), while Figure 2.7 (b) has the same targets (in 3D) that were obtained with the laser scanner's on-board target identifier. Targets are shown as blue dots and the 3D point cloud of the room is in green. Registration of the two using either of the two aforementioned transformation methods will result in either image data converted to the 3D point-cloud system, or 3D point-cloud information assigned to the RCI.



(a)

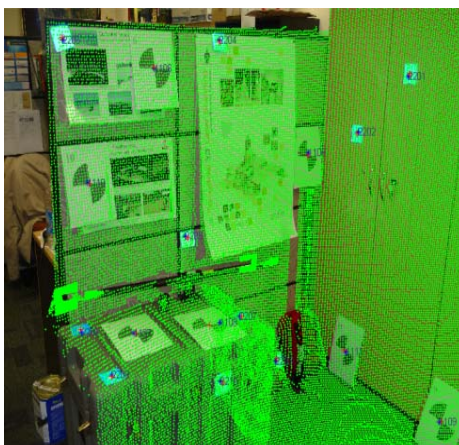


(b)

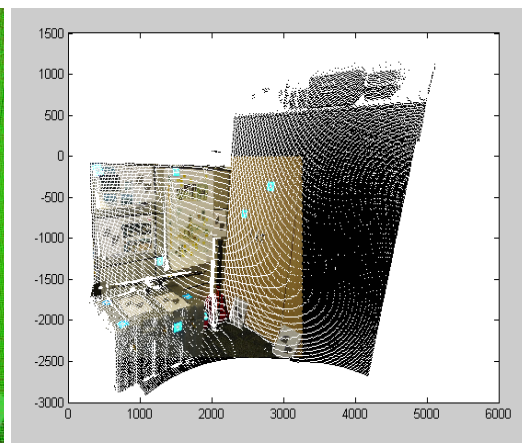
Figure 2.7 Areas including corresponding points: (a) Digital photo (b) Point-cloud

In most cases registration takes place during the post processing of data, but some systems have been proposed that process the registration on the fly (during data capture within the instrument) such as in Jansa et al. (2004), Sapkota (2008), and Kern (2001). The process of registration in terms of this research is summarised as follows:

- (a) Identification of points in image and point-cloud require to be co-located
- (b) Transformation parameters determined based on the co-located points
- (c) Transformation parameters applied to move one dataset into another
- (d) For reverse transformation, current transformation parameters are inverted
- (e) The inverse transformation parameters applied to bring the dataset back



(a)



(b)

Figure 2.8: Registration of (a) point-cloud to image (b) image to point-cloud

2.5 Transformations

Outlined in this section is the concept of Synthetic Camera Images (SCI) and its creation, along with the Helmert transformation and Direct Linear Transform (DLT), which are the two transformation methods used in this research.

2.5.1 Homogeneous Coordinates and Helmert Transformation

Homogeneous coordinates are commonly used in projective geometry, as they contain the relationship parameters that relate two differing sets of data together, especially as a set of ratios (Mikhail, Bethel and McGlone, 2001). It can be seen in the books by Mikhail, Bethel and McGlone (2001), Hartley and Zisserman (2003), as well as Luhmann et al. (2006) that the DLT is a form of the homogeneous coordinate model in that the 12 parameters arranged into a third row and fourth column matrix relates the image points to the 3D point-cloud data.

The Helmert transformation is a 3D to 3D transformation and is known by several different names, such as the similarity transformation, the 7-parameter transformation and the equiform transformation described by the equation:

$$\mathbf{X} = \mathbf{X}_0 + \lambda \mathbf{M} \mathbf{x} \quad (2.7)$$

or expanded into matrix form:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \lambda \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (2.8)$$

where $\mathbf{X} \in \mathbb{R}^{3 \times 1}$ represents the 3D points (X, Y, Z), $\mathbf{X}_0 \in \mathbb{R}^{3 \times 1}$ represents the translation components, λ is the scalar scale component, $\mathbf{M} \in \mathbb{R}^{3 \times 3}$ is the matrix containing 9 rotation parameters (calculation is in Appendix C) and $\mathbf{x} \in \mathbb{R}^{3 \times 1}$ is the image coordinates. The similarity transformation is defined by a set of 7 parameters, which are scale (λ), rotation (ω , ϕ , κ) and translation (t_x , t_y , t_z), which are used to transform 3D datasets into different 3D coordinate systems. In order to determine the 3D relationship between 3D datasets, a least squares approach may be implemented to obtain the 7 parameters, but this requires approximate initial values. Using the

homogeneous coordinate model, the 7-parameter transform method can be converted from Eq. 2.8 to 2.11 into the same form of representation as the DLT (Eq. 2.16):

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} m_{11} & m_{12} & m_{13} & 0 \\ m_{21} & m_{22} & m_{23} & 0 \\ m_{31} & m_{32} & m_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.9)$$

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda m_{11} & \lambda m_{12} & \lambda m_{13} & T_x \\ \lambda m_{21} & \lambda m_{22} & \lambda m_{23} & T_y \\ \lambda m_{31} & \lambda m_{32} & \lambda m_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.10)$$

and

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \lambda m_{11} & \lambda m_{12} & \lambda m_{13} & T_x \\ \lambda m_{21} & \lambda m_{22} & \lambda m_{23} & T_y \\ \lambda m_{31} & \lambda m_{32} & \lambda m_{33} & T_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (2.11)$$

Note that Eq. 2.8 is converted into Eq. 2.9, through application of matrix multiplication principles, which then multiplies into Eq. 2.10, thus having the homogeneous coordinates form. Eq. 2.11 is the condensed matrix version for the Helmert transformation having the redundant 4th row removed, which now offers a resemblance to the DLT matrix as outlined in the following Section 2.4.2.

2.5.2 DLT

The Direct Linear Transformation (DLT), an implementation of projective geometry, provides a linear function for modelling the transformation between the image pixel coordinate system and the object space coordinate system, and can be derived from collinearity equations (Atkinson, 1996; Mikhail, Bethel and McGlone, 2001; Luhmann et al., 2006). Originally proposed by Abdel-Aziz and Karara in 1971, the main advantage of the method is that it does not require approximate initial values, or a calibrated camera (Luhmann et al., 2006; Abdel-Aziz and Karara, 1971, Karara, 1989). In the book by Hartley and Zisserman (2000), DLT is viewed in terms of homography or homogenous equations. DLT is derived from collinearity equations, as shown below:

$$x'_i = x'_0 + \frac{m_{11}*(X-X_0)+m_{12}*(Y-Y_0)+m_{13}*(Z-Z_0)}{m_{31}*(X-X_0)+m_{32}*(Y-Y_0)+m_{33}*(Z-Z_0)} + \Delta x' \quad (2.12)$$

$$y'_i = y'_0 + \frac{m_{21}*(X-X_0)+m_{22}*(Y-Y_0)+m_{23}*(Z-Z_0)}{m_{31}*(X-X_0)+m_{32}*(Y-Y_0)+m_{33}*(Z-Z_0)} + \Delta y' \quad (2.13)$$

DLT has 12 unknown parameters (though the 12th is usually normalised as 1) used to transform without requiring any initial values, but requires a minimum of 6 reference points (McGlone et al. 2004; Luhmann et al., 2006). However, the DLT does not contain an easily interpretable form as used in the Helmert transform; that is scale, rotation and translation. However, approximated exterior orientation parameters may be derived, as described in Luhmann et al. (2006), Atkinson (1996), Mikhail, Bethel and McGlone (2001) and Karara (1989). Eqs. 2.14 and 2.15 are the functional equations for the DLT, one to describe the relationship of the 3D points to the image coordinate x , and another to describe the relationship of the 3D points to the image coordinate y . The coefficients of the equations can be put into a matrix format as seen in Eq. 2.16. Notice that the last row coefficients are common to the two functional equations and are assigned, w , in the resultant vector as it is a scaling parameter.

$$x_i = \frac{H_{11}*X_0+H_{12}*Y_0+H_{13}*Z_0+H_{14}}{H_{31}*X_0+H_{32}*Y_0+H_{33}*Z_0+H_{34}} \quad (2.14)$$

$$y_i = \frac{H_{21}*X_0+H_{22}*Y_0+H_{23}*Z_0+H_{24}}{H_{31}*X_0+H_{32}*Y_0+H_{33}*Z_0+H_{34}} \quad (2.15)$$

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.16)$$

The method of least-squares is used to estimate and calculate the parameters for the DLT. DLT primarily uses 11 parameters since the 12th parameter (H_{34}) is generally given the value of 1, indicating that there is a normalisation, but a 12 parameter solution is also possible when treated as a parameter (Hartley and Zisserman, 2003).

An in depth discussion of the DLT is given in Chapter 3, while the derivation for the DLT coefficients is contained in Appendix B.

2.6 Synthetic Camera Image (SCI) and Real Camera Image (RCI)

The Synthetic Camera Image (SCI) is an image created (with mathematical principles simulating a camera) from point-cloud data. Therefore it also contains scale information to transform it back into 3D. This synthetic image attempts to mimic the Real Camera Image (RCI). Forkuo and King (2005) provide the original explanation and concept. The RCI is a digital photo and the term is used to provide a distinction from simply using the word image, as both are images, but one is captured in reality (RCI) using a camera such as the Nikon D80 and the other is simulated or synthesised (SCI).

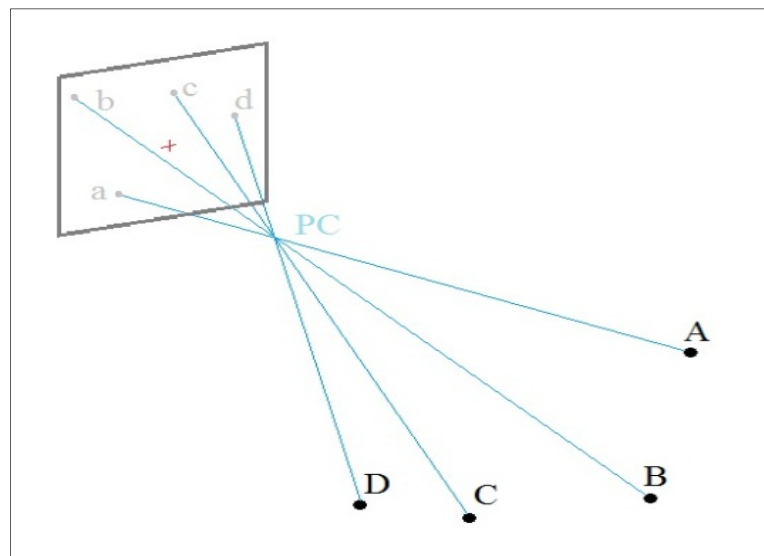


Figure 2.9: Visual representation of 3D points projected onto a plane

The creation of a synthetic image is based on a projection of an object onto a flat plane, which in principle is the same as capturing a photograph of an object. The light ray passes through a lens onto a flat surface (film or CCD), as indicated in Figure 2.9.

Creation of an SCI as seen in Figure 2.10 (a) is required to provide the platform whereby a relationship can be determined between the 3D point-cloud and RCI.

In order to mimic the RCI, the SCI created in this research uses the 2D camera projection model (sometimes referred as the canonical model; Hartley and Zisserman, 2003) to create the SCI of a point-cloud:

$$\mathbf{r}_i = \mathbf{k} [\mathbf{M}|\mathbf{Tr}] * \mathbf{r}_0 \quad (2.17)$$

where \mathbf{k} is the camera matrix, \mathbf{M} and \mathbf{Tr} are the rotation and translation matrix, respectively, and \mathbf{r}_i and \mathbf{r}_0 are the 2D image and 3D object points in the non-homogeneous notation, respectively. Similar to the similarity transform, the camera projection model can also be expressed using homogeneous coordinates so that it is visually similar to the DLT when changed into matrix format, as indicated in Section 2.4.2. Notably, the distinguishing difference between the camera projection model and the Helmert transformation is the scale matrix, whereby the camera projection model contains a matrix of interior orientation parameters (see Eq. 2.18) and the Helmert transformation has a singular scale (Eq. 2.8).

In order to simulate a specified camera, the camera matrix \mathbf{k} (Eq. 2.18) is required, which contains the calibration parameters to create a synthetic image with the properties of the specified camera:

$$\mathbf{k} = \begin{bmatrix} f & s & x_p \\ 0 & f & y_p \\ 0 & 0 & 1 \end{bmatrix} \quad (2.18)$$

where the focal length (f), skew (s) and principal point offset (x_p, y_p) are required in order to produce the SCI. In terms of the SCI, it is assumed that the calibration has been performed for the camera, thus requiring only the calibrated focal length to be entered and low principal point offset values. The RCI used would have the radial lens distortion applied allowing precise point to point correspondence solutions, thus reducing error propagation for 3D reconstruction. For the SCI x and y image coordinates, a millimetre to pixel ratio may be applied and then rounded to the closest integer to obtain a pixel based solution, or the pixel to millimetre ratio must be applied to the RCI for the two images to correspond.

The alternative approach for the creation of synthetic images is achieved using DLT. However, the synthetic images are created with a different approach whereby the DLT parameters are solved using least squares with corresponding points in the 2D image and 3D control points. Then an SCI is produced with the DLT equation, i.e. Eq. 2.16, in conjunction with the 3D point-cloud data, as seen in Figure 2.10 (a).

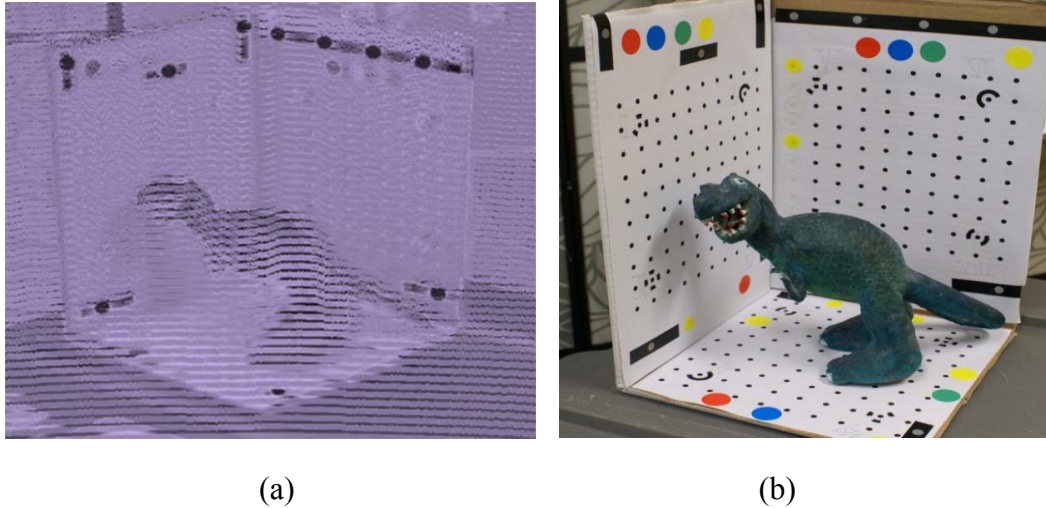


Figure 2.10: (a) A synthetic image (SCI) (b) A photo (or RCI)

2.7 2D Iterative Closest Point

The Iterative Closest Point (ICP) is fine registration method is used for correspondences of points, lines and surfaces (Besl and MacKay, 1992; Yang and Medioni, 1992; Salvi et al., 2007); a point based approach is implemented for 2D correspondences in this project. It is assumed that trials with the simulated data after transformations may obtain a singular solution. Other variants of ICP exist, e.g. Al-Manasir and Fraser (2006), Rusinkiewicz and Levoy (2001), and Vosselman and Maas (2010), but point-to-point distance is the method used for this research, which is acceptable given the density of the 3D point cloud and the ease with which good correspondences between the image and point cloud can be determined.

Outlined in Figure 2.11 is a brief overview of the ICP procedure represented by pseudo-code; beginning with point selection, then calculating and applying the transformation components, repeating the process until convergence is reached. Colour registration from 2D onto 3D or texture mapping refers to the application of

colour from images to point clouds, as seen in papers by Stamos and Allen (2002), Ikeuchi et al. (2003) and Liu et al. (2006). However, this research also allows the reverse of texture mapping (3D to 2D) for different properties, such as data from other sources like 3D point-clouds, to be registered to single images. A method is presented using DLT to attempt to preserve 2D image metric accuracy, while registering with 3D point-clouds.

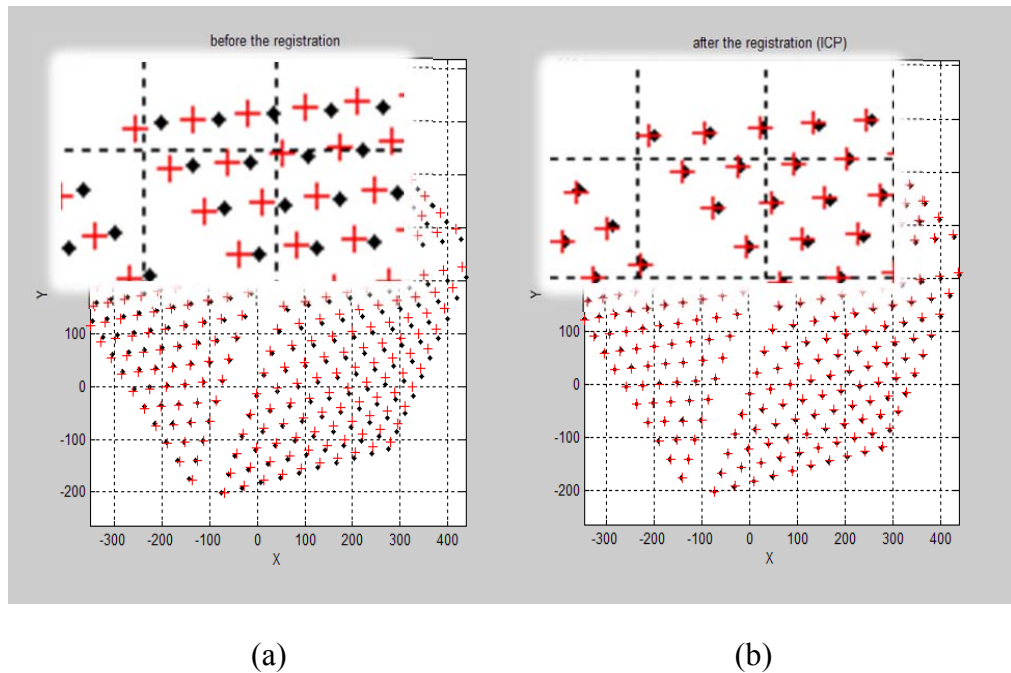


Figure 2.11: (a) Large offset of points before ICP (b) Both datasets closer post ICP

The ICP algorithm result is presented in Algorithm 1. Two images are compared and one is selected to be the base image, while the other image is selected to merge/transform onto the base image. ICP and its variants have been utilised for the automated registration of 3D point-clouds, medical images, and for pattern recognition. The most important and difficult part of the ICP algorithm is to choose the metric criteria, e.g. Euclidean and Mahalanobis distances with consideration of statistical inferences, for the corresponding points in either 3D or 2D datasets.

```
1: procedure ICP (Points  $X_{1ij}$ ,  $X_{2ij}$ )
2: for iteration = 1 to user specified do
3:   for j = 1 to length of X1 do
4:     find closest point (minimum distance) with the kdtree
5:     if distance < threshold distance
6:       then store corresponding points  $X_{1ij}$  and  $X_{2ij}$ 
7:       else end
8:   end
9:   calculate rotations and translations with least squares
10:  apply transformation
11:  reduce threshold distance
12:  if converged && small residuals
13:    then break
14:    else end
15: end procedure
```

Figure 2.12: Brief summary of the conventional ICP algorithm

3 METHODOLOGY

The generation of virtual representations of objects from sensor data requires 3D reconstruction or depth acquisition. Photogrammetry utilises multiple images and triangulation for calculation of its 3D points with colour assignment to each 3D point coming from one or more of the images. Alternative solutions incorporating photogrammetric principles exist such as AICON's (2012) two camera setup, a single camera with grid projection system, SwissRanger's 3D range camera (MesaImaging, 2012), or Terrestrial Laser Scanners (TLSs). The aforementioned methods of acquiring photorealistic models are faced with a combination of the following limitations: (i) range and resolution constraints, (ii) static arrangements of the sensors, (iii) sensitivity to lighting conditions, (iv) flexibility to incorporate different sensor types, and (v) time based restrictions (Liu and Stamos, 2005).

With these limitations in mind, this study provides a flexible solution for the 3D reconstruction and colour assignment using co-registered camera and TLS and solving for the registration. This registration approach focuses on the technique to provide the mapping of colour to point cloud, having flexibility by utilising any chosen camera and TLS setup. Firstly the camera is able to collect a myriad of colours in order to represent reality. Secondly the TLS, is able to acquire precise dense 3D point data of the object's surface with minimal reconstruction effort.

The conventional process of 3D reconstruction is to identify many corresponding points between multiple images and then triangulating with bundle adjustment to obtain 3D data points (McGlone et al., 2004; Luhmann et al., 2006; Atkinson, 1996). Although TLS has resolved the problem of obtaining 3D points, obtaining accurate colour information for the object still remains. Since cameras have the ability to provide much better colour information, camera images are used to supplement the limited colour acquisition of TLS. The method presented here focuses on using projective transforms as a means to achieve a set of parameters to relate the point cloud and image data together. As discussed in the previous chapters, the Direct Linear Transformation (DLT) is the main projective transform used, which is robust because of its linear solution of a non-linear problem and its ability to achieve a solution without requiring initial values. Discussed in this chapter is the relationship

between the various similar transforms: DLT, collinearity, camera projection equation, and Helmert transformation. This registration problem uses DLT. It is the most suitable technique, since it has the advantage of having a simple mathematical form and being relatively accurate, while offering camera setup flexibility (Chen, Armstrong and Raftopoulos, 1994; Dermanis, 1994a).

The different aspects of the DLT method are explored in this research, apart from its use in deriving initial parameters and in calibration adjustments (McGlone et al., 2004; Luhmann et al., 2006; Atkinson, 1996), it is used with the concept proposed by Forkuo and King (2005) to produce synthetic images (SCI). On a similar note, Ikeuchi et al. (2003) attempts using the DLT for a heritage modelling project, but discounts its potential in preference for development of the reflectance-edge based method because they perceived the DLT to be reliant on calibration, especially in terms of using calibration fixtures. However, this research takes an alternative path for the DLT, in which producing an SCI provides the transitioning medium for registration between 2D and 3D data, as correspondence and registration happen at a 2D-2D level. The transitioning medium only requires the DLT to obtain the relationship between camera images (RCI) and the 3D point cloud data. Therefore, combining point clouds from TLS and images from a camera can be achieved using the SCI method which offers the ability to register at a 2D to 2D level by using the DLT.

This chapter describes the entire process of the TLS and image registration methodology for creating photorealistic colour models, and depth information recovery. The core framework is presented first in which the key concepts are explained in detail to provide the foundational understanding. Then the method is proposed covering stages from data acquisition to the final back-projection. Lastly a concise summary of the entire methodology is presented.

3.1 Core Framework

3.1.1 Fundamental Expressions for Projection and Back-Projection

The DLT and the camera projection transformation are used to project 3D data into a 2D plane, thus creating an SCI. The fundamental concept is to be able to perform a

projection from 3D into 2D and then to back-project from 2D back into 3D. To be able to successfully register the two types of datasets, it is essential to understand the method by forming the mathematical relationship between the datasets. Initially simulated data is used to get an understanding and provide control of the variables, and a closed system used for the purposes of testing. The base idea of projection between 3D and 2D space can be expressed by Eqs. 3.1 and 3.2, which are the projection and back-projection respectively (Hartley and Zisserman, 2003):

$$\begin{bmatrix} x \\ y \\ w \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.1)$$

or more briefly, $\mathbf{x} = \mathbf{H}_{\text{Projective}} \mathbf{X}$, and:

$$\begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \\ w \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.2)$$

or more briefly, $\mathbf{H}_{\text{Projective}}^{-1} \mathbf{x} = \mathbf{X}$, where $\mathbf{H}_{\text{Projective}} \in \mathbb{R}^{4 \times 4}$ represents the matrix of projection parameters to be estimated, $\mathbf{x} \in \mathbb{R}^{4 \times 1}$ is the image coordinates and $\mathbf{X} \in \mathbb{R}^{4 \times 1}$ is the object coordinates in the homogeneous coordinate system (Hartley and Zisserman, 2003). This forms the fundamental concept of projection and back-projection, including an understanding of the mathematical relationship between 2D and 3D for the purpose of registration via projective means.

Eqs. 3.1 and 3.2 together are a simple representation of the process from projection to back-projection. Back-projection is not as simple as the forward projection from 3D onto 2D when it comes to the DLT. In theory, back-projection should be as simple as applying an inverse of the $\mathbf{H}_{\text{Projective}}$ matrix into Eq. 3.1, and the result should become Eq. 3.2 due to matrix identities $\mathbf{I} = \mathbf{H}_{\text{Projective}}^{-1} \mathbf{H}_{\text{Projective}}$, then the $\mathbf{H}_{\text{Projective}}^{-1}$ matrix multiplied by the $\mathbf{H}_{\text{Projective}}$ matrix results in an identity matrix. However, this cannot apply to this case since the $\mathbf{H}_{\text{Projective}}$ matrix in terms of the DLT comprises only 12 parameters, given as matrix \mathbf{H} . Since the DLT contains 12 parameters as expressed in Eq. 3.3, the \mathbf{H} matrix to be inverted is not a square matrix

instead but rather $\mathbf{H} \in \mathbb{R}^{3 \times 4}$, a standard matrix inverse cannot be performed. Each parameter may be represented either by their location within the matrix or as parameters H1 to H12 as in Eq. 3.3. Note these notations H_x and HX will be utilised in a mixed manner, whenever it is necessary, through this dissertation.

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \end{bmatrix} = \begin{bmatrix} H1 & H2 & H3 & H4 \\ H5 & H6 & H7 & H8 \\ H9 & H10 & H11 & H12 \end{bmatrix}. \quad (3.3)$$

To solve the issue regarding inverting the \mathbf{H} matrix, three methods, explained in Section 3.2.3, are devised which then allows for the back-projection. The main method is termed the pseudo-inverse method, as it applies a pseudo-inverse for the non-square DLT matrix; the second is a 2-Step solution, in which the projection matrix is split into two parts and the inverse performed to solve the rotation and translation separately; the third method is to apply an additional redundant row to treat the matrix as a type of 3D-3D transformation allowing a normal inverse to be performed. In essence, the simple representation provides only the conceptual understanding of the relationship matrix used for projection and back-projection, and further investigation is required in order for the back-projection to work.

3.1.2 Correspondences and Collinearity

A relationship between the two datasets needs to be established, using a set of corresponding points or features present in the image and the point-cloud. Then, the corresponding points will be used in the estimation of the \mathbf{H} matrix via a least squares solution. Upon obtaining the parameters relating the image and the point cloud, registration between the two datasets can take place to provide either supplementation of 3D information onto an image, or the colour information from the image to be stored onto 3D making a photorealistic model.

Collinearity can be represented by a set of equations that mathematically represents the ray path intersecting the object points, point of exposure and the image space, in three-dimensional space (Wolf and Dewitt, 2000). The general form of the collinearity can be expressed in two ways, one in a global coordinate system, e.g. Eq.

3.4 in Luhmann et al. (2006), while the other from a camera centric perspective, e.g. Eq. 3.5 in Karara (1971) or relevant expressions in McGlone et al. (2004) as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \lambda \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3.4)$$

or more briefly, $\mathbf{X}=\mathbf{X}_0+\lambda\mathbf{M}\mathbf{x}$, and:

$$\begin{bmatrix} x \\ y \\ -c \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}^{-1} \begin{bmatrix} X-X_0 \\ Y-Y_0 \\ Z-Z_0 \end{bmatrix} \quad (3.5)$$

or more briefly, $\mathbf{x}=\lambda^{-1}\mathbf{M}^{-1}(\mathbf{X}-\mathbf{X}_0)$. Several other derivative expressions of the collinearity exist whereby slight modifications are present and may be known as the camera projection equation, which includes the \mathbf{k} matrix of interior orientation parameters, Helmert or Similarity transformation that is primarily used as a 3D-3D shape-invariant transformation, and the DLT, which is a linear expression of the non-linear form.

3.1.3 Understanding the DLT

An understanding is required of the DLT with respect to collinearity in order to establish a firm understanding of the back-projection methods. The DLT equation is developed from the collinearity. The six elements of exterior orientation and five elements of interior orientation relate to the DLT parameters (Dermanis, 1994a). Therefore some similarities between the two concepts are expected to be present.

Similarities between collinearity and the DLT are evident when the equations are expressed in matrix form. An alternative expression exists (Dermanis, 1994a), for the conversion of the collinearity equation to the DLT. However, this method of expression is based on Karara (1971). The process of changing the collinearity solution to the DLT will be outlined, with explanation of the components that are merged into the DLT parameters. Therefore to begin, consider Eq. 3.5. This matrix form can be changed into the algebraic form (Eqs. 3.6 and 3.7), which is a more suitable representation to show the parameter changes:

$$x = c \frac{m_{11}(X-X_0)+m_{21}(Y-Y_0)+m_{31}(Z-Z_0)}{m_{13}(X-X_0)+m_{23}(Y-Y_0)+m_{33}(Z-Z_0)} \quad (3.6)$$

and:

$$y = c \frac{m_{12}(X-X_0)+m_{22}(Y-Y_0)+m_{32}(Z-Z_0)}{m_{13}(X-X_0)+m_{23}(Y-Y_0)+m_{33}(Z-Z_0)} \quad (3.7)$$

where x and y in the lower and upper cases represent the coordinates in the image and the object spaces, respectively, c is the principal distance of a camera, and m_{ij} is the individual component of the relative rotation matrix between the image and 3D point clouds spaces. Firstly, the equations are to be multiplied out using:

$$x = c \frac{m_{11}X+m_{21}Y+m_{31}Z + (-m_{11}X_0-m_{21}Y_0-m_{31}Z_0)}{m_{13}X+m_{23}Y+m_{33}Z + (-m_{13}X_0-m_{23}Y_0-m_{33}Z_0)} \quad (3.8)$$

and:

$$y = c \frac{m_{12}X+m_{22}Y+m_{32}Z + (-m_{12}X_0-m_{22}Y_0-m_{32}Z_0)}{m_{13}X+m_{23}Y+m_{33}Z + (-m_{13}X_0-m_{23}Y_0-m_{33}Z_0)} \quad (3.9)$$

where translation parts, i.e. $-m_{ij}X_0$, $-m_{ij}Y_0$ and $-m_{ij}Z_0$, can be grouped together (indicated in brackets in Eqs. 3.8 and 3.9) as a constant to establish the DLT coefficients H_4 , H_8 and H_{12} . Next, the rotation and translation elements can be multiplied together and treated as coefficients. Eqs. 3.10 to 3.12 and Eqs. 3.11 to 3.13, relate to the remaining DLT parameters. Consider:

$$x = \frac{c m_{11}X+c m_{21}Y+c m_{31}Z+c H_4}{m_{13}X+m_{23}Y+m_{33}Z+H_{12}} \quad (3.10)$$

and

$$y = \frac{c m_{12}X+c m_{22}Y+c m_{32}Z+c H_8}{m_{13}X+m_{23}Y+m_{33}Z+H_{12}} \quad (3.11)$$

in which subsequently the principal distance can be incorporated into the equation. It can be also seen that parameters H_1 , H_2 , H_3 , H_5 , H_6 and H_7 contain the principal

distance and orientation elements whereas the parameters H9, H10 and H11 only contain rotation elements. This is notable when orientation angles are required to be determined as:

$$x = \frac{H1 X + H2 Y + H3 Z + H4}{H9 X + H10 Y + H11 Z + H12} \quad (3.12)$$

and:

$$y = \frac{H5 X + H6 Y + H7 Z + H8}{H9 X + H10 Y + H11 Z + H12} \quad (3.13)$$

which are the equivalent expressions to Eqs. (3.10) and (3.11). The DLT is commonly regarded as having 11 parameters, because the parameter H12 undergoes normalisation. The parameter H12 is divided through the two DLT equations, thus being expressed as unity as:

$$x = \left(\frac{1}{H12} \right) \frac{H1 X + H2 Y + H3 Z + H4}{H9 X + H10 Y + H11 Z + H12} = \frac{H1 X + H2 Y + H3 Z + H4}{H9 X + H10 Y + H11 Z + 1} \quad (3.14)$$

and

$$y = \left(\frac{1}{H12} \right) \frac{H5 X + H6 Y + H7 Z + H8}{H9 X + H10 Y + H11 Z + H12} = \frac{H5 X + H6 Y + H7 Z + H8}{H9 X + H10 Y + H11 Z + 1} \quad (3.15)$$

which are the equivalent DLT expressions to Eqs. (3.12) and (3.13).

The elements that contribute to the DLT components can be summarised as:

$$\begin{bmatrix} H1 & H2 & H3 & H4 \\ H5 & H6 & H7 & H8 \\ H9 & H10 & H11 & H12 \end{bmatrix} = \begin{bmatrix} c m_{11} & c m_{21} & c m_{31} & c m T_{xyz} \\ c m_{12} & c m_{22} & c m_{32} & c m T_{xyz} \\ m_{13} & m_{23} & m_{33} & m T_{xyz} \end{bmatrix}, \quad (3.16)$$

$$\mathbf{k M}^{-1} (\mathbf{X} - \mathbf{X}_0) = \begin{bmatrix} \mathbf{k M}^{-1} & \\ & \mathbf{k M}^{-1} \mathbf{Tr} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (3.17)$$

and

$$\begin{bmatrix} H1 & H2 & H3 & H4 \\ H5 & H6 & H7 & H8 \\ H9 & H10 & H11 & H12 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{k M}^{-1} & \mathbf{k M}^{-1} \mathbf{Tr} \\ 3,3 & 3,1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \quad (3.18)$$

Representing the components in this perspective aids in understanding the calculation of the perspective centre used for the back projection. Eq. 3.17 is a matrix representation of the collinearity equation to demonstrate that similarity exists, represented in Eq. 3.18, where $\mathbf{k M}^{-1} \in \mathbb{R}^{3 \times 3}$ is representative of H1, H2, H3, H5, H6, H7, H9, H10 and H11, and $\mathbf{k M}^{-1} \mathbf{Tr} \in \mathbb{R}^{3 \times 1}$ is representative of H4, H8 and H12.

3.1.3.1 Limitations of the DLT

The DLT offers a flexible method for the projection and back-projection, and furthermore, the resection in the case of multiple images leading to better opportunities in developing automated procedures for image analysis (Hartley and Zisserman, 2003). However, for the purpose of the back-projection from 2D to 3D space, the reverse solution through applying the inverse to the DLT matrix has a limitation due to its weakness of lacking the additional depth information for each image pixel (Hartley and Zisserman, 2003, Chen, Armstrong and Raftopoulos, 1994). Through an understanding of the fundamental relationship between the photogrammetric interpretation of a camera and the DLT, three methods can be used for providing the necessary additional information to the DLT for the back-projection of 2D images to 3D point clouds.

3.1.4 Normalisation and 12th Parameter Recovery

The 12th DLT parameter, i.e. H12, is commonly normalised as unity, but it may need to be extracted or recovered for instances such as determining orientation angles and also for successful back-projection. The best estimation for the 12th parameter uses the third component of the calculated perspective centre, termed Z_c , as the perspective centre models the camera location or camera translation. This is based upon comparing the DLT and the collinearity equation, finding that similarity exists (see Eq. 3.18) for the translation side. The estimated 12th parameter needs to be applied to every component as represented in Eq 3.19. This is due to the

normalisation step, i.e. Eq. 3.14 and 3.15, thus needing to apply the inverse of the H12 factor for the entire function:

$$(H_{34}) \begin{bmatrix} x \\ y \\ w \end{bmatrix} = (H_{34}) \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.19)$$

3.1.5 Perspective Centre Calculation

The perspective centre calculation of the DLT (Karara, 1989; Luhmann et al., 2006; Atkinson, 1996) is important as it provides the necessary translation parameters required for the proposed three back-projection methods in this dissertation. It is the base concept for the 2-step back-projection method. The formulation of the perspective centre requires the link represented in Eq. 3.18, which shows that the DLT matrix is split into two parts similar to the collinearity matrix. Based on the knowledge that the perspective centre is the location of the camera with respect to the origin of the object coordinate system, the values for X, Y and Z are given the value of zero (Eq. 3.18). Following though the calculation gives the solution in Eq. 3.20, which then is rearranged to give the solution for the perspective centre in Eq. 3.21:

$$\begin{bmatrix} H4 \\ H8 \\ 1 \end{bmatrix} = \mathbf{k} \mathbf{M}^{-1} \mathbf{Tr} = \mathbf{k} \mathbf{M}^T \mathbf{X}_0 = \begin{bmatrix} H1 & H2 & H3 \\ H5 & H6 & H7 \\ H9 & H10 & H11 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (3.20)$$

and

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} H1 & H2 & H3 \\ H5 & H6 & H7 \\ H9 & H10 & H11 \end{bmatrix}^{-1} \begin{bmatrix} H4 \\ H8 \\ 1 \end{bmatrix}. \quad (3.21)$$

3.1.6 The Image Scale Component

The image scale component, termed w , needs to be understood as it is an integral component to the back-projection process. The image scale component is another normalised parameter just like the H12 coefficient of the DLT, as normalisation is required to transform via scaling the calculated image data or SCI back to its original image points (see Eq. 3.22):

$$\left(\frac{1}{w}\right) \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \quad (3.22)$$

in the non-homogeneous representation. The image scale component differs for each image point, since it acts as a 3D component, often seen as the collinearity parameter z in books by McGlone et al. (2004) and Luhmann et al. (2006). Before normalisation of the SCI or calculated image data, the coordinates are in terms of 3D space or the object space. Thus when normalisation occurs, the calculated image data is changed to the image coordinate system, being independent from the other data. To bring the image data into a 3D form, a method using the DLT is used to create SCIs, which allows registration between 2D and 3D to occur.

The inner constraints in conjunction with the bundle adjustment has been widely utilised in close-range photogrammetry, e.g. Fraser (1997) and Dermanis (1994b) or commercial software like Australis (2012), in which observations including control points and parameters are dealt with in the context of a free network (Sillard and Boucher, 2001). In this study, the accuracy of the control points can be assumed to be much higher than the observations for both the collinearity and DLT approaches. The control points are obtained by using the laser scanner software, which will perform a precise scan of the target to determine the centre; also, a total station may be used to precisely collect the control point data. Therefore, much emphasis was placed on the recovery of the depth information of each pixel with the assistance of 3D point clouds as presented in this chapter. Thus the free network with the inner constraint was not investigated.

3.2 Proposed Method

3.2.1 Rigorous estimation of the DLT parameters

The DLT is a form of projective transform, derived from the collinearity equations, which directly relates the Euclidean coordinates of the object and image points without the need of approximate initial values (McGlone et al., 2004; Abdel-Aziz and Karara, 1971). The camera projection model cannot be used in most cases because it is difficult to obtain the camera orientation parameters for a single image.

Thus the DLT is implemented to provide the relationship parameters because of its robust linear form (Luhmann et al., 2006). The camera projection matrix is only used initially in order to understand the mechanics of the DLT, as both are similar in presentation, being derived from the collinearity equation. It is inferred that the DLT comprises an alignment part, which deals with rotations, and a depth or z-axial projection part, which acts like translation. As mentioned previously in Section 3.1.3, using Eq. 3.16 as a reference, the scale and error components are intermixed within the 12 DLT matrix parameters.

Determining the DLT projection matrix simply requires the 3D points and their corresponding 2D image points for a solution (Luhmann et al., 2006). The DLT coefficients can be determined with iterative least squares and is used for the purpose of obtaining the relationship parameters between the 2D and 3D control points, for generating an SCI and back-projection. Notably in this method, the solution obtained will show a good match for the 3D points to the image under normal circumstances and a near perfect match for simulated cases. However, as explained earlier, the DLT transformation coefficients will not allow the solution to easily revert back to the original 3D, unless the scaling component (w) is retained. Even with the scaling component, some form of manipulation is still required for back-projection, which is presented in the later sections.

3.2.2 RCI and SCI

3.2.2.1 Create SCI with control points only

Initially, only the control points are required to create the SCI, which is done by multiplying the DLT projection matrix with the point cloud control points according to Eq. 3.28. This step is basically in preparation for registration of the RCI to the SCI using ICP. There is no reason to create the SCI for the whole point cloud because the ICP registration only requires a few corresponding points in order to match the two datasets and provide the transformation between the two. Figure 3.1 shows that points are projected using the camera projection model to create an SCI, which functions similarly to the DLT as explained in Chapter 2. The SCI is a mapping of the control points to the image system, which will then allow data manipulations to

take place in the same system and eventually to be utilised for the estimation of all the DLT parameters.

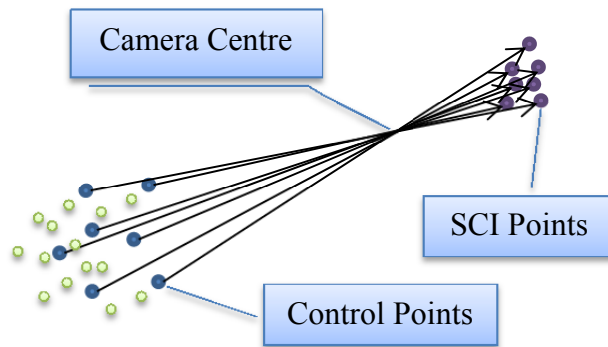


Figure 3.1: A representation of the projection of control points to SCI points.

3.2.2.2 Coarse and Fine Registration of RCI to SCI using ICP

Registration at this stage requires an image matching method, e.g. Besl and McKay's (1992) ICP, to be implemented on the RCI and SCI. The registration algorithm using ICP will produce a 2D transformation that is to be applied to the entire RCI making it ready for colour assignment in a later step. It is essential to obtain a very close match at this stage for high precision and accurate back-projection. A deep understanding of some of the requirements is required to ensure a good registration.

The control points in the RCI and the SCI are factors that will determine the outcome of the transformation. Having sufficient and evenly distributed control points is needed to provide the best outcome (Chen, Armstrong and Raftopoulos, 1994). Control points within the RCI as well as the corresponding TLS control points are required to be located as accurately as possible as mentioned in the previous step. This is to allow a good DLT solution in order to create the control point based SCI for registration.

Removal of image related errors in the RCI such as misalignments, radial lens distortion, skew, and so on, is needed in order to obtain a very close point-to-point correspondence with the SCI. As displayed in Figure 3.2, ICP can then be used to further match the RCI and SCI. In an ideal case, the RCI and the SCI would result in a perfect match when registering the two image datasets together. However, in reality

many factors, especially errors, known and unknown would influence the disparity between the two datasets and should be removed when possible.

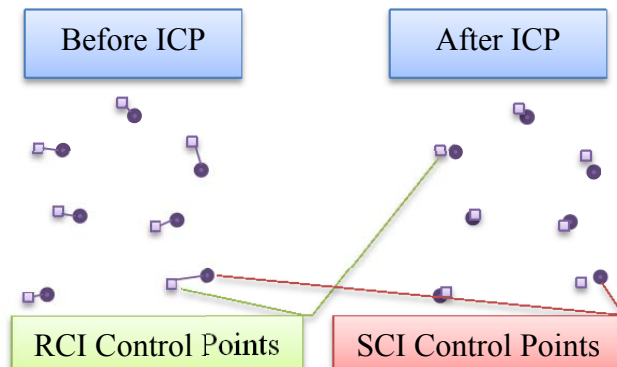


Figure 3.2: The SCI and RCI control points to be closer after registration.

Registration between the control points of the RCI and SCI will obtain the transformation parameters required, which is a by-product of the ICP algorithm. Figure 3.3 shows conceptually that the entire RCI would be transformed based on the parameters acquired in this step, thus rotated, scaled and skewed for colour assignment to the SCI at a later stage of this process. The RCI provides the colour information while the SCI contains the depth information, thus a good match between the RCI and SCI will allow an accurate colour model representation after back-projection.

A very close match between the RCI and SCI is required, otherwise the location of the colour desired for the object may be offset or shifted, thus forming a misrepresentation or causing other misalignment effects. Figure 3.4 shows this problem visually as the original points that are expressed as \mathbf{X} representing the coordinate set, i.e. X, Y, Z, have translated after the back-projection caused by a mismatch or low level precision of corresponding points in the image. This representation is modified based on the version of the points-to-rays back-projection concept (Hartley and Zisserman, 2003) for providing some understanding of the problems in the back-projection caused by registration at this stage.

Based on Figure 3.4, an assumption can be made that range and incidence angle from the PC to the object will be influencing factors causing back-projection errors. In

terms of range, as the distance from the object increases, the effect caused by the offset also increases. In a similar manner, as the angle of incidence increases, the likelihood of mismatches occurring also increases. In order to achieve a solution with minimal misrepresentations after back-projection, it is important to understand the effect on the solution due to the registration at this stage.

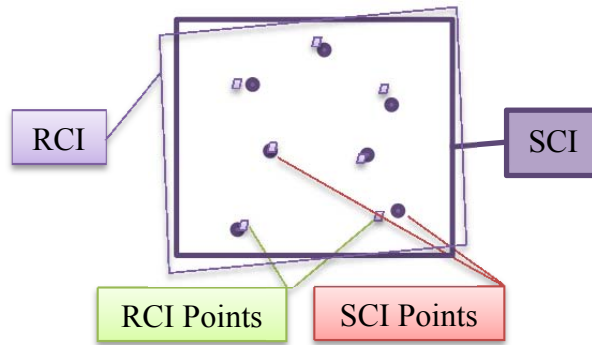


Figure 3.3: The RCI is transformed to match the SCI.

Before proceeding further, an additional refinement step may be desired to achieve closer matches between the SCI and the RCI, since having a closer match will enable the back-projection to have better accuracy. Therefore, iteration from step 2 to step 5 may be considered (refer to Table 3.5). If iteration is implemented, the RCI control points must be updated and the DLT parameters recalculated. However, in performing the iteration the RCI may become overly skewed and distorted from the original due to multiple transformations being applied, thus compromising data quality.

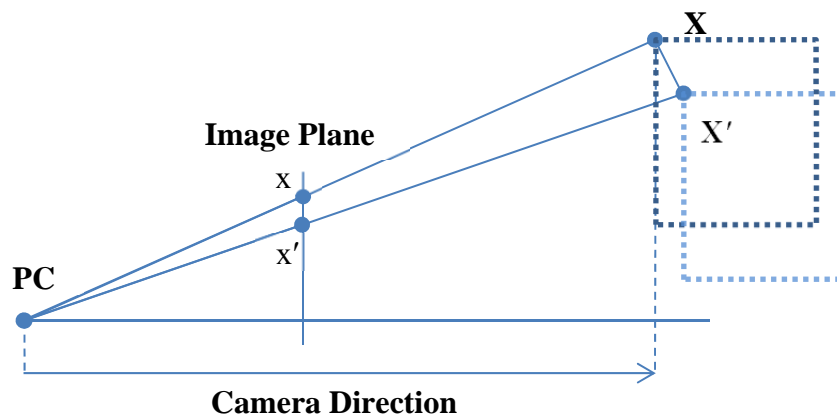


Figure 3.4: The effect of an offset in the image plane causing a misrepresentation of colour assignment in the point cloud.

3.2.2.3 Create an SCI of the entire point cloud

This step requires the entire point cloud to be converted into an SCI, which is done by multiplying the point cloud coordinates with the DLT as shown in Eq. 2.16 (Chapter 2, pp. 23). There may be multiple points per pixel, which is dependent on three factors: (1) point cloud density, (2) image resolution and (3) distance of image to object; as pixel size enlarges or reduces based on range. Figure 3.6 illustrates the multiple points per pixel phenomenon.

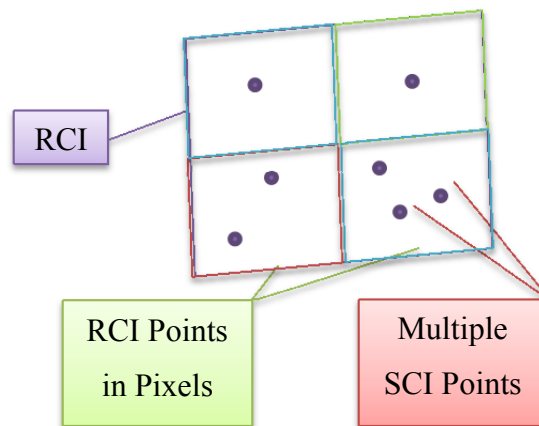


Figure 3.5: RCI at pixel level: Each pixel may contain multiple SCI points.

For dense point clouds or low image resolution, it can be seen that multiple points can be projected onto an individual pixel and the outcome should be similar to Figure 3.5, and an increase in point cloud density should result in greater amount of points per pixel. To decrease the amount of multiple SCI points per pixel value, logically the density of the pixels has to increase or a higher resolution image obtained. Then there would be more pixels to dilute the SCI density.

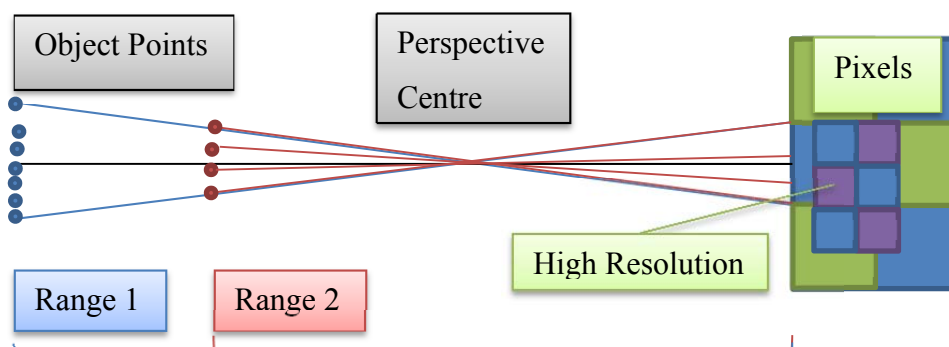


Figure 3.6: Geometric diagram to indicate multiple points per pixel as a function of range.

Range is an important factor as both the previous situations may be manipulated simply by changing the distance between image and object. Referring to Figure 3.6 at Range 1 there is an opportunity to fit more points onto a single pixel than at Range 2. The RCI are integer values if using pixels, while SCI points may be in decimal values if no rounding is applied to conform to the RCI values.

3.2.2.4 Assigning of Colour Information (Additional Data Registration)

The colour information from the (closely) matching RCI may now be stored onto the SCI, ready for back-projection. Colour information for each coordinate value of the RCI can be stored at the same coordinate value of the SCI. In other words, colour information can be merged to the SCI, based on the matching pixels in the SCI and RCI. Some pixels may be assigned with several SCI points and some pixels may not have any points. This is dependent on point density, RCI perspective, RCI resolution and range.

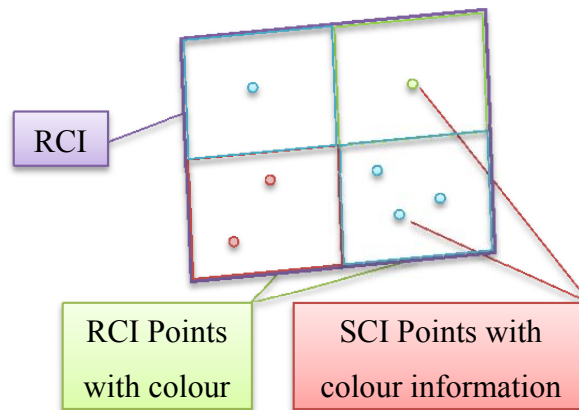


Figure 3.7: Colour information from the RCI points is stored at the corresponding SCI points.

3.2.3 Proposed Back-projection methods

Back-projection is a term used when explaining the reverse concept of forward projection, which is to project from image to object point (Eq. 3.2). The back-projection concept in this method is based on that proposed by Hartley and Zisserman (2003). Referring to the representation given in Figure 3.8, the acquisition of an object point X from its corresponding image point x , requires a collinear ray

path determined from the perspective centre (PC) passing through the image point and a depth parameter to define a limit to the depth of the object point.

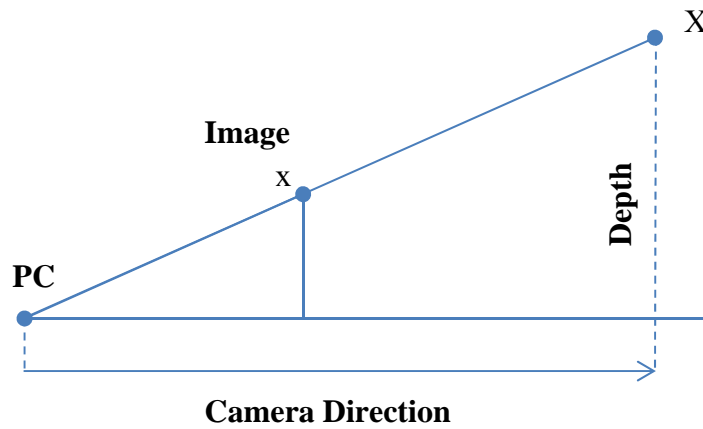


Figure 3.8: Back-projection representing points to rays concept and to emphasise significance of the image depth value.

The camera direction is defined within the DLT parameters, enabling the acquisition of the perspective centre. The ray path is defined by performing an inverse on the DLT parameters. Depth is defined partly through the acquisition of parameter w , in creation of an SCI, from the DLT projection ($\mathbf{x} = \mathbf{HX}$), i.e. Eq. 2.16, and the other part a translation/shift based on the method of back-projection.

For further clarification, 3D points are projected into the 2D digital photo (RCI) and matched by points falling into the respective pixel region. This is not to merely project and back-project, but to define an associative relationship between the 2D and 3D, if it is solely projecting back and forward, one can also use standard photogrammetry functions and principles to achieve desired results.

There is a method using OpenGL, which is a simple case of tracing the ray path, assigning the closest pixel colour and storing to the 3D point. However, this primary aim is to find the relationship between 2D and 3D for colour assignment with DLT.

3.2.3.1 Direct Pseudo-Inverse of H

This method uses a pseudo inverse to calculate a back-projection matrix for the DLT (Hartley and Zisserman 2003). A pseudo inverse is a method used for inverting non-

square matrices. Calculating the pseudo inverse for the DLT matrix is given by Eq. 3.29, taken from Hartley and Zisserman (2003), where a pseudo inverse of the matrix is denoted with a superscript plus (+) symbol as follows:

$$\mathbf{H}^+ = \mathbf{H}^T (\mathbf{H} \mathbf{H}^T)^{-1} \quad (3.29)$$

The DLT pseudo inverse matrix, $\mathbf{H}^+ \in \mathbb{R}^{4 \times 3}$, is to be used with the coordinates of the SCI for back-projection (Eq. 3.30). Applying the pseudo inverse transformation will result in a set of four coordinate values, of which the first three pertain to the X, Y and Z image location coordinates while the last value (W) is the depth scaling parameter used to project or translate the points into object locations, which can be expressed as follows:

$$\begin{bmatrix} X^+ \\ Y^+ \\ Z^+ \\ W \end{bmatrix} = \mathbf{H}^+ \begin{bmatrix} x \\ y \\ w \\ 1 \end{bmatrix} \quad (3.30)$$

The back-projected points without the application of the depth scaling parameter will only align and rescale the SCI image plane into the 3D point cloud coordinate system. As mentioned in section 3.1, the SCI is an image representation in the global coordinate system and that an RCI in the image system is independent. In order to complete the back-projection process, first the depth scaling parameter will need to be subtracted by 1 because of normalisation, see Eq. 3.31, as follows:

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = (\mathbf{W} - 1) \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (3.31)$$

and

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X^+ \\ Y^+ \\ Z^+ \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}. \quad (3.32)$$

The recalculated depth scales are to be multiplied to the perspective centre to give the additional shift/translation, which are to be added to the non-depth back-projected points, as in Eq. 3.32. Thus a set of coloured 3D back-projected points has

been created. A possible visual explanation based on Figure 3.8, is that the normalised unity defines a set depth, in which all points will relate to when applying only the pseudo-inverse. When plotting the pseudo-inverse object points (X^+ , Y^+ , Z^+), the result will appear to be an enlarged image plane. Therefore a further step is required, based on the acquisition of the depth scaling parameter to shift individual object points.

3.2.3.2 Two-Step solution with the DLT

An alternative technique for back projection is the two step solution that resembles the calculation of the perspective centre with the DLT as seen in Atkinson (1996), Karara (1989) and Luhmann et al. (2006). This method of back-projection is unique as it is based on the calculation of the perspective centre (Eq. 3.34), whereby the DLT matrix is split into two parts (Eq. 3.33). The first part deals with the alignment or rotation element and the second with the translation. Comparing Eq. 3.34 with the collinearity Eq. 3.5, the resemblance can be seen when at the origin. A further explanation can be found by referring to Figure 3.8, where the PC is found by finding the camera direction. The camera direction is defined by the ray path from the PC to the origin and implying that H_4 , H_8 and H_{12} , is the location of the principal point. Understanding that H_{12} is 1 due to normalisation, the parameter set resembles image coordinates $(H_4, H_8, 1) = (x, y, 1)$, thus this equation may be modified whereby Part 2 (Eq.3.33) can be substituted with SCI values as expressed as:

$$\text{Part 1: } \begin{bmatrix} H_1 & H_2 & H_3 \\ H_5 & H_6 & H_7 \\ H_9 & H_{10} & H_{11} \end{bmatrix} \text{ and Part 2: } \begin{bmatrix} H_4 \\ H_8 \\ H_{12} \end{bmatrix} \quad (3.33)$$

which leads to:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} H_1 & H_2 & H_3 \\ H_5 & H_6 & H_7 \\ H_9 & H_{10} & H_{11} \end{bmatrix}^{-1} \begin{bmatrix} H_4 \\ H_8 \\ H_{12} \end{bmatrix} \quad (3.34)$$

In order to calculate the perspective centre, which in essence is the camera location, an inverse of Part 1 in Eq. 3.33 must be obtained and then multiplied by Part 2 of Eq. 3.33, which gives Eq. 3.34. Next is to perform the back projection, whereby the

initial step is to multiply each set of SCI points with the inverse of Part 1 as shown in Eq. 3.35:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} H_1 & H_2 & H_3 \\ H_5 & H_6 & H_7 \\ H_9 & H_{10} & H_{11} \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \\ w \end{bmatrix}. \quad (3.35)$$

This will produce a full scale 3D point cloud at the origin. The next step is to use the results of the first step and translate them by the perspective centre, as shown in Eq. 3.36:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}. \quad (3.36)$$

3.2.3.3 Addition of Fourth-Row to the Original DLT Matrix

This concept can be understood from a combination of two perspectives, first being the Helmert transformation, which is explained in Chapter 2, and second through graphical transformations (Luhmann et al., 2006). Although the above two are mentioned in that order, it should be noted that the order is irrelevant and that a fuller understanding is obtained through understanding both perspectives mentioned. A simple generalisation of this additional row concept is that the key to the development of this method was from a deeper look into the component nature that is inherent in projective transforms and applied to this DLT case.

Firstly, it is explained in terms of the Helmert transformation that is defined by 7 parameters: a scale, 3 rotations and 3 translations. When applying the Helmert transformation, the 3 rotation angles are required to be converted and used in the rotation matrix. Once the parameters are in a suitable form, they can be presented in the matrix form as shown in Eq. 2.11 (Chapter 2, pp. 22), which has the same matrix size as the DLT. In Eq. 2.10 (Chapter 2, pp.22), the same equation is represented, but as a square matrix of 4 rows and 4 columns, in which the entire bottom row is an addition only to modify it to become a homogeneous matrix. One of the benefits of this is that the matrix is square and a standard inverse can be performed to allow a back-projection, which is something desired for the DLT. The similarities between

the Helmert transformation and the DLT presents the possibility of having the DLT matrix inverted by applying the additional row.

Secondly, the graphical transformation function relates to the DLT in two ways. Firstly the function is in homogeneous form and secondly, decomposition shows they consist of the same elements. The graphical transformation function is given below, which is taken from Luhmann et al. (2006). On a side note, comparing the decomposed formula with the Helmert transform shows near identical components:

$$\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & :t_{14} \\ t_{21} & t_{22} & t_{23} & :t_{24} \\ t_{31} & t_{32} & t_{33} & :t_{34} \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ t_{41} & t_{42} & t_{43} & :t_{44} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{11} & : \mathbf{T}_{12} \\ 3,3 & : 1,3 \\ \dots\dots\dots \\ \mathbf{T}_{21} & : \mathbf{T}_{22} \\ 3,1 & : 1,1 \end{bmatrix} \quad (3.37)$$

and:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_X & 0 & 0 & 0 \\ 0 & s_Y & 0 & 0 \\ 0 & 0 & s_Z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} & 0 \\ m_{21} & m_{22} & m_{23} & 0 \\ m_{31} & m_{32} & m_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.38)$$

$$= \begin{bmatrix} s_X m_{11} & s_X m_{12} & s_X m_{13} & T_x \\ s_Y m_{21} & s_Y m_{22} & s_Y m_{23} & T_y \\ s_Z m_{31} & s_Z m_{32} & s_Z m_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

or more briefly, $\mathbf{X} = \mathbf{T}$ and $\mathbf{x} = \mathbf{T}_T \mathbf{T}_S \mathbf{T}_R \mathbf{x}$.

Luhmann et al. (2006) states that the reverse transformation for the graphical transformation may not be changed as the projections are not necessarily commutative and the reverse transformation is given by $\mathbf{x} = \mathbf{T}^{-1} \mathbf{X} = \mathbf{T}_R^{-1} \mathbf{T}_S^{-1} \mathbf{T}_T^{-1} \mathbf{X}$. Although the commutative properties of this transformation are of little importance to this method, the main part is that this transformation matrix can be reversed. Therefore, after investigating that the DLT matrix is similar in nature to this transform and the Helmert transform, the possibility of having an additional row to the DLT matrix arises to allow an inverse to be performed. The graphical transformation provides the meaning of the additional row. Referring to Eq. 3.37, the content of the sub-matrices can be identified as shown in Table 3.1.

Sub-matrix	Content
\mathbf{T}_{11}	Scaling, reflection in a line, rotation
\mathbf{T}_{12}	Translation
\mathbf{T}_{21}	Perspective
\mathbf{T}_{22}	Homogeneous scaling

Table 3.1: Description of the sub-matrices related to the graphical transform.

The sub-matrix \mathbf{T}_{21} deals with the perspective, in which the values can remain as zeroes as it is not required. Therefore it may be implemented for the DLT case. In understanding the similarities between the DLT and the two transformations mentioned previously, the additional row appears to be a viable solution that will allow the DLT function to be reversed.

Following the registration process, the additional row inverse method requires the DLT matrix to be reconstructed into a 4 by 4 square matrix. This is done by adding to the DLT matrix a fourth row, as presented in Eq. 3.39. After that, it is multiplied by the SCI coordinate set. When examining the DLT matrix after the inverse has been applied, this method seems to closely resemble the 2-step method seeing that there is an alignment and perspective translation part as:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \quad (3.39)$$

3.3 Summary of Methodology

The overview of the proposed methodology for the back-projection from 2D image to 3D space is outlined in Table 3.2, which shows the method as a sequence of eight steps. First, in terms of data acquisition, a selected area or object that requires modelling is photographed and scanned, thus providing an RCI and a point cloud. Next is to use the calibration parameters to create a rectified (error-removed) image of the RCI, as well as to locate the control points within the image and the corresponding control points from the point cloud. Having these corresponding

points allows the DLT parameters to be determined using iterative Least-Squares. Next, an SCI of the control points must be created in order to perform the 2D-2D registration using ICP. Multiplying the DLT matrix with the control points will result in creating the SCI of the control points. The SCI coordinates need to be normalised by the division of the coordinate values by the individual scale value (w). Once the SCI coordinates are changed into image space coordinates, the ICP registration can be performed on the RCI control points to match the SCI. The ICP registration algorithm contains the transformation required to be applied to the whole RCI, thus completion of the registration will allow the transformation to be applied on the RCI. Next is to create the SCI of the entire point cloud, which is performed using the DLT matrix and the entire point cloud, allowing the assigning of colour information onto the whole point cloud SCI from the corresponding RCI coordinates. At this stage, an SCI with colour information is obtained. Next, the back-projection can take place, using any of the three methods: the Pseudo-inverse method, the Two-Step method or the Additional row method.

Data acquisition and administration			
Step 1	RCI (Corrected 2D Image)		3D point cloud from a TLS
	RCI (Located 2D control points in pixels or metric units)		Point cloud (Located 3D control points)
Registration process from 3D to 2D			
Step 2	Obtain DLT parameters with the iterative Least-Squares		
Step 3	Create SCI (control points only)		
Step 4	Register RCI to SCI (control points) using ICP		
Step 5	Apply transformation to RCI		
Step 6	Create SCI of the entire point cloud		
Step 7	Assign RCI information onto SCI		
Back-projection from 2D to 3D			
Step 8	Primary	Alternatives	
	<u>Method 1</u>	<u>Method 2</u>	<u>Method 3</u>
	Pseudo-inverse method with DLT	Two-Step method with DLT and six-parameter transformation	Inverse with the additional row of DLT matrix

Table 3.2: Stages in the overall process of the proposed method.

3.4 Chapter summary

For this research, the method for performing single image and point cloud depth information recovery is outlined in this chapter. The core idea is based on transforming 3D to 2D then from 2D to 3D. The projection and back-projection equations are explained first before exploring the DLT. The DLT is explained in terms of the collinearity equation, and covers normalisation of its 12th parameter. The perspective centre and the image scale component of the back-projection are explained, as they form the main elements of the method.

Each process of the proposed method is outlined in detail, before proceeding to the summary outlining the nine steps required to register 2D and 3D datasets together, from data acquisition, to projection and registration, and finally back-projection. The process begins with obtaining the data and control point information from devices such as TLS for 3D and cameras for 2D. Then an SCI of the control points is to be created for registration, with ICP used to register the RCI to the SCI. The result of the ICP algorithm provides the transformation parameters to be applied to the RCI. Next, the colour information from the RCI is stored on the matching SCI points. Lastly, the SCI is back-projected using one of the presented methods, which are, the Direct Pseudo-Inverse method, the Two-Step solution, or the Additional Fourth-row method.

4 RESULTS

4.1 Experiment with simulated datasets

This section presents the results of the camera projection and back-projection process using simulated data. Simulated data was used to provide adequate control over the test 3D data, in terms of errors and orientation. The basic concept is also shown to be applied in practical applications using the DLT method along with the three proposed back-projection solutions.

4.1.1 Creation of simulated dataset

To begin the process, the creation of a preliminary set of 3D points (e.g. box or conical shape) to represent a point cloud is conducted both manually with an arbitrary coordinate system, as well as using software such as Matlab 7.12. The simulated point clouds created have a small sample size in comparison to a real TLS point cloud at this preliminary stage. The first dataset, to be referred as the Roof Dataset, is a roof-like structure that is manually created using similar triangles and the camera scale and ratio principles as presented in Eq. 4.1:

$$\frac{\text{Range}}{\text{Focal Length}} = \frac{\text{Field of View}}{\text{Sensor Length}} \quad (4.1)$$

Firstly, consideration was focussed on the scale to be achieved in terms of the sensor/image size, which meant that the object points are required to be at a certain distance or within the field of view. Image size, focal length, camera location (at origin or offset) and location of data points had to be taken into account. This was done by obtaining the imaging camera (in this case a Nikon D80) parameters for focal length (18mm) and sensor length (23.6mm), and choosing an approximate camera to object distance (10m) to calculate the viewable length. Figure 4.1 provides a visual explanation of this process.

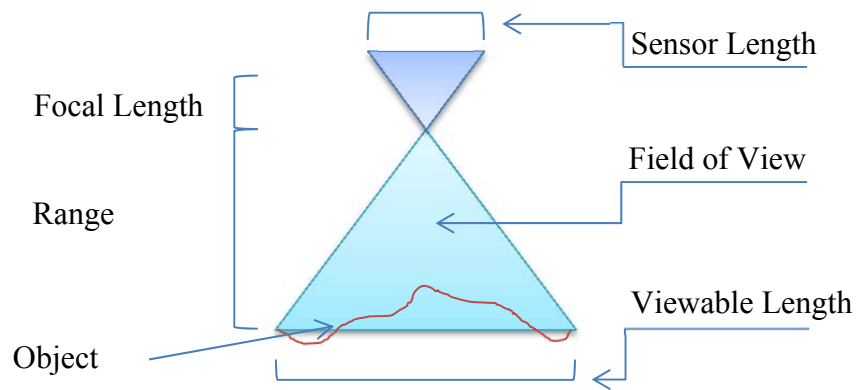
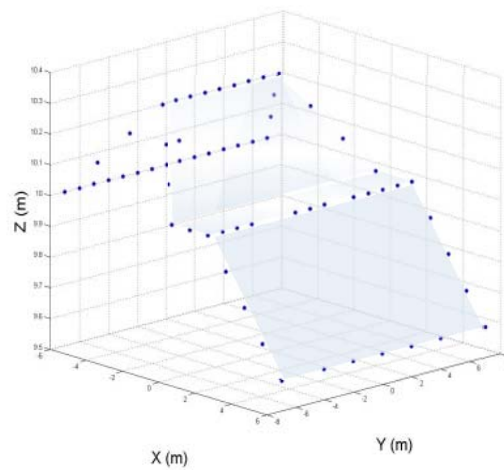


Figure 4.1: Using similar triangles and ratios for the creation of the Roof Dataset.

The other two datasets are created using Matlab 7.12 (see Appendix D), in which the Corner Dataset simulates a corner and the Conical Dataset represents a cylindrical cone shape. The Corner Dataset was intended to be used for the DLT, thus it contains several control points, which were manually selected. Figures 4.2, 4.3, 4.4 and 4.5 show the three simulated sets of data: Roof Dataset, Corner Dataset, and Conical Dataset; and the Conical Dataset with control points, respectively.



(a)

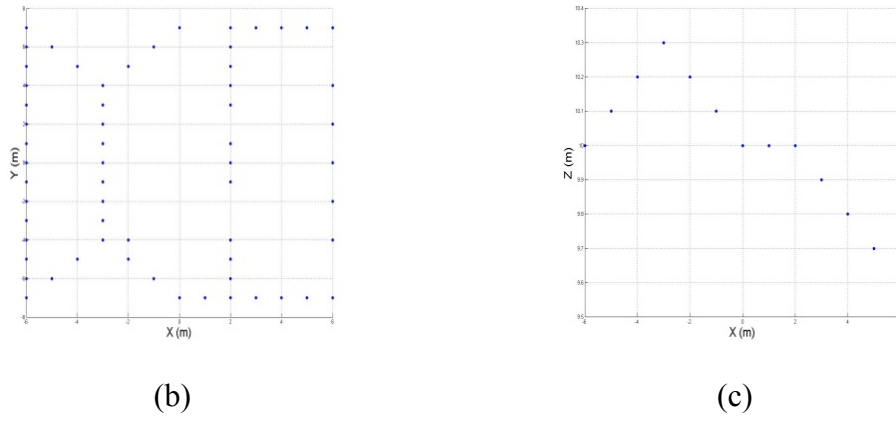
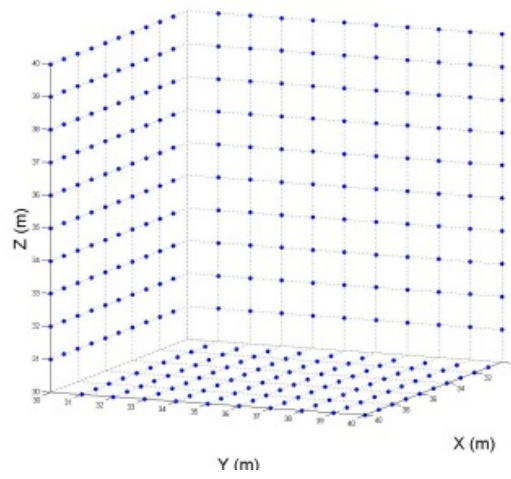


Figure 4.2: The Roof Dataset, a roof-like structure with views:
(a) overall (b) top (c) side.



(a)

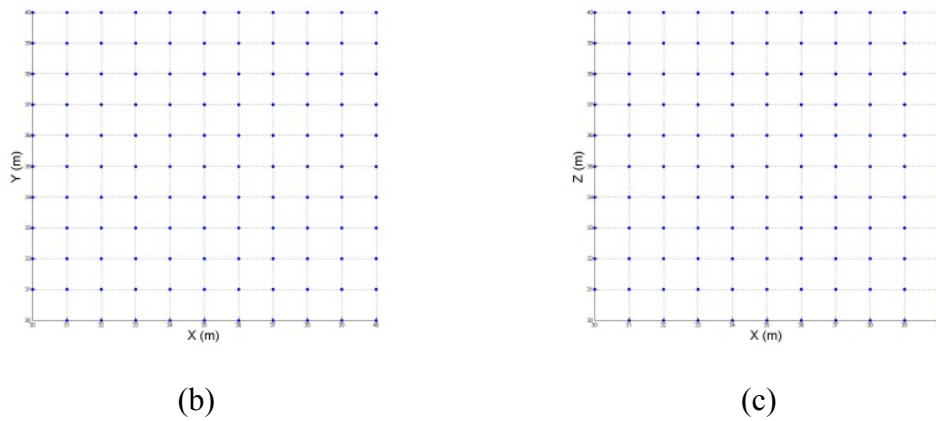
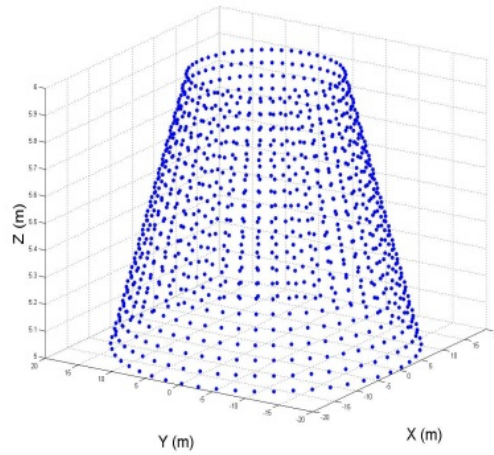
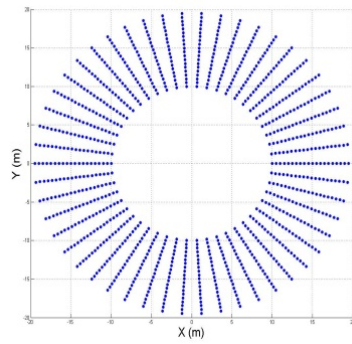


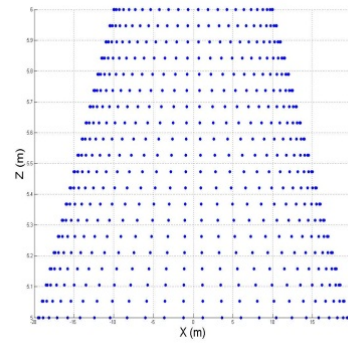
Figure 4.3: The Corner Dataset, a box corner structure with views:
(a) overall (b) top (c) side.



(a)



(b)



(c)

Figure 4.4: The Conical Dataset, a cylindrical cone structure with views: (a) overall (b) top (c) side.

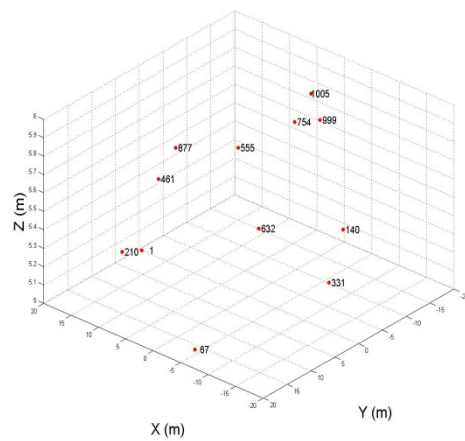
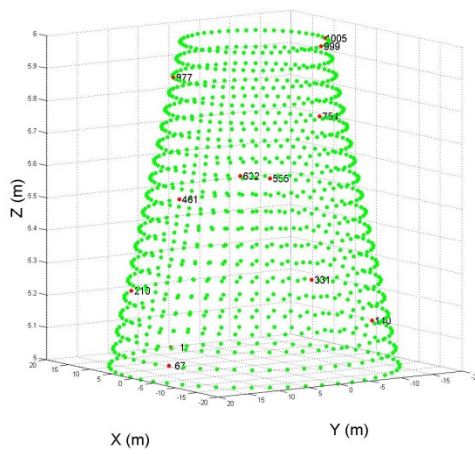


Figure 4.5: The Conical Dataset with control points highlighted in red and 3D map of the location with only control points.

4.1.2 Camera Projection Concept used to create the SCI

The experimentation commenced with creating an SCI of all three datasets. For this preliminary stage, the known camera projection matrix is used for creating the SCI. However, note that with the Conical Dataset and the final process, the DLT is used instead of the camera projection matrix for the SCI. However, the SCI of the Conical Dataset including its control points is created to act as an RCI for the DLT method (see Figure 4.7). Error components were not essential as the main objective was to create a suitable 2D and 3D pair, and it is assumed that calibration should remove the major error sources. Table 4.1 shows the parameters used in creation of the SCI of the three datasets.

Dataset	Focal length (m)	Orientation			Translation (m)			Total number of points
		ω°	ϕ°	κ°	T_x	T_y	T_z	
Roof	0.018	0°	0°	0°	0	0	2	61
Corner	0.018	311°	34°	11°	0	0	5	300
Conical	0.500	5°	8°	10°	1	1	5	1020 (Control = 12)

Table 4.1: Parameters used for the three datasets.

For this initial procedure, an RCI was not obtained as the primary focus was basically to test the simple case of projection and back-projection. It may be assumed that in this case the RCI matches the SCI perfectly, because it is in an ideal environment. The projection produced the following SCIs (Figure 4.6) for the two datasets, based on the specifications mentioned earlier. Therefore there is also no need to perform the ICP algorithm for registering the RCI and SCI together.

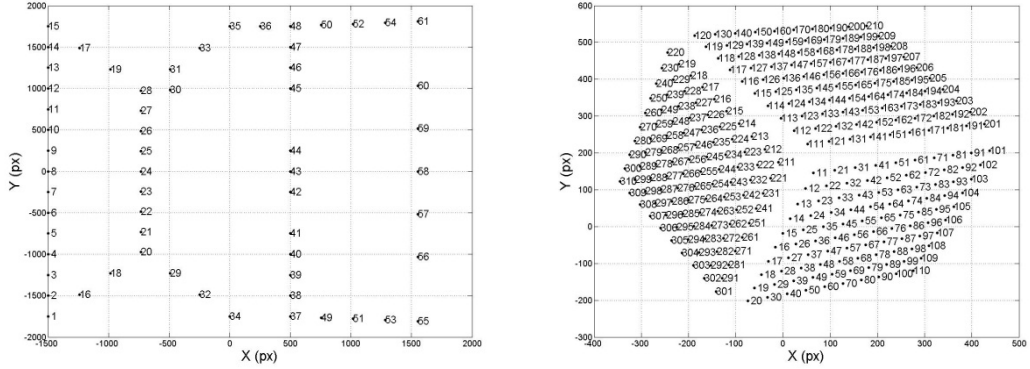


Figure 4.6: SCI image of the Roof Dataset and the Corner Dataset with numbers representing individual points of the dataset.

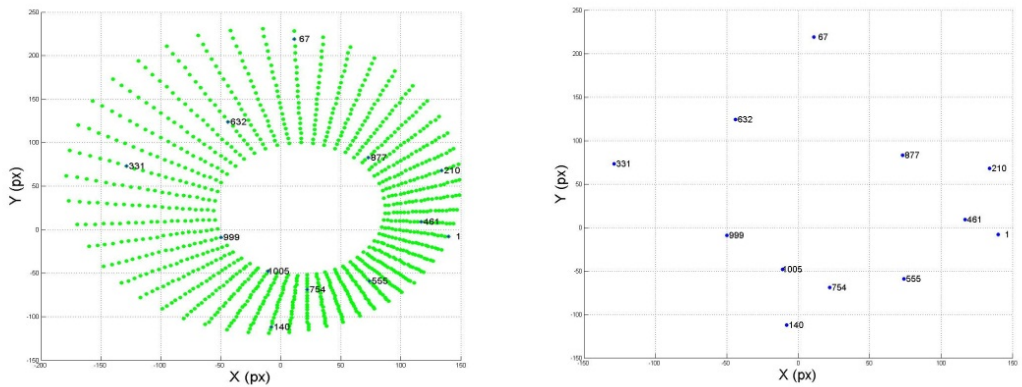


Figure 4.7: SCI of the Conical and control points.

The simple case of back-projection by using the pseudo inverse onto the projection matrix produced results that are in the incorrect location, but notice that the data lies on a plane as shown in Figures 4.8 and 4.9. If the results are viewed from a specific angle, the results appear to match. The viewpoint is from roughly the centre of the incorrectly located results to the origin of the original 3D data points. Thus, it is required to find the viewpoint direction to the back-projected points and obtain the distance translation for the points by finding the appropriate scale value for the direction, which is explained in Section 3.2.3.

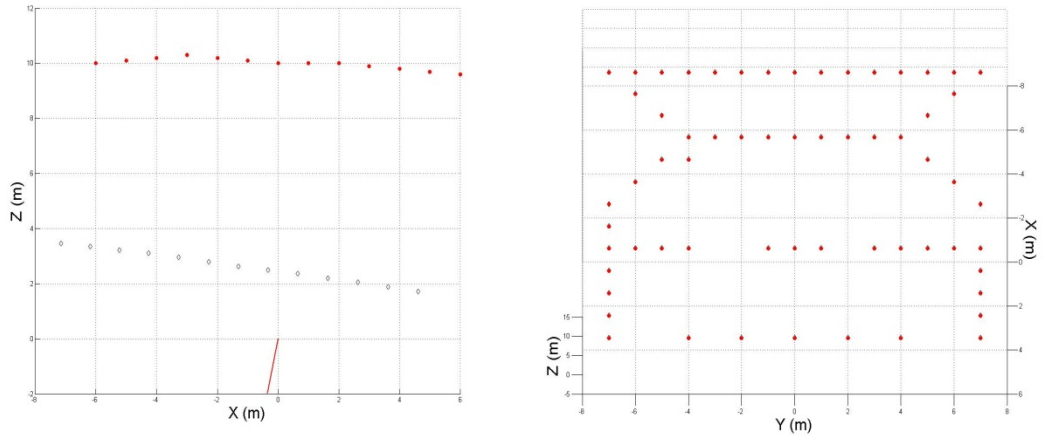


Figure 4.8: The back-projection of the SCI produces 3D points that form a plane (left). The red line in the left image indicates perspective direction which matches the control points (right).

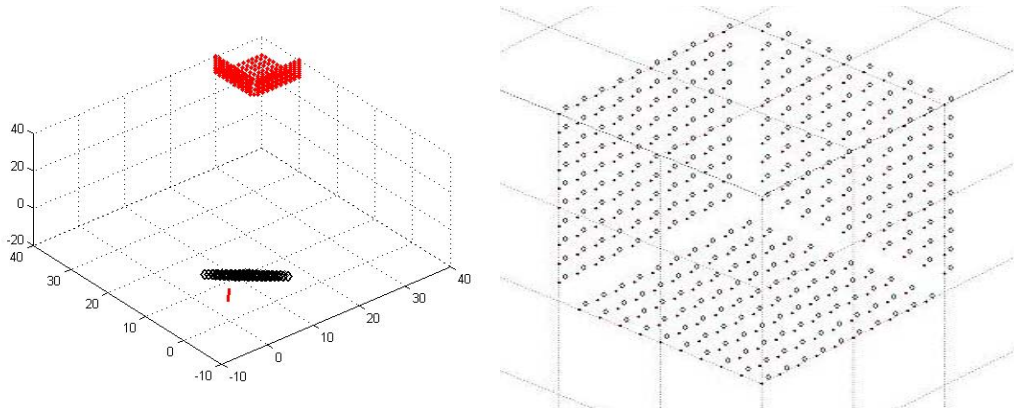


Figure 4.9: Another example with Corner Dataset of the back-projected 3D points matching at a certain perspective.

The simple back-projection does form a trend as seen in Figures 4.8 and 4.9, which shows that despite having an alternate 3D location, the points lie in a plane. When the points are viewed from a certain perspective, a matching result seems to be displayed. This trend may be regarded as a transformation of the image into the object coordinate system. From this point, it can be seen that a transformation into its correct object points requires additional data along with the back-projection concept mentioned in Section 3.2.3. Thus, implementing the Pseudo-inverse method, the Two-Step method or the Additional Fourth-Row method will achieve a successful back-projection as indicated in Figure 4.10.

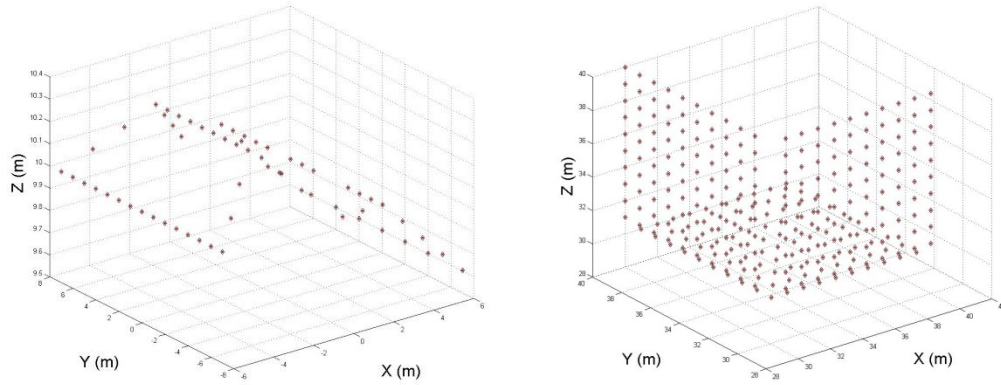


Figure 4.10: Back-projected SCI points to the original location. Left is back-projected Roof Dataset and right is back-projected Corner Dataset.

Figure 4.10 shows a perfect example of a successful back-projection, in which the RMS errors are zero. This case is as stated earlier: a back-projection of the SCI coordinates. To achieve this case when using the RCI, it is required that the RCI matches the SCI perfectly. However, this case does not occur in a practical scenario, and to achieve this requires the minimisation of the error of the RCI. This in theory will give the best case for back-projection with minimal errors.

4.1.3 Using the DLT method and the significance of control points

This section will use the Conical Dataset for explanation, which includes 12 selected control points. Thus far, the Conical Dataset contains a point cloud with a set of control points and a corresponding image (SCI) with control points. The conic is 1 metre in length with a radius of 10 metres at the top and 20 metres at the base. As explained earlier, several points were randomly selected from the point cloud to be control points. A practical case involving errors for the simulated point cloud and the control points will be introduced later in Section 4.1.6. For this section, the focus will be on obtaining a general understanding of the effect of mismatches between the RCI and SCI, which incorporates the use of ICP.

Since this is a simulated environment, an RCI from a real camera cannot be obtained. The method used to overcome this issue is to use the synthesised SCI created with the camera projection model to represent an RCI for this purpose, which also provides some control over the desired RCI results. In this case the distinction

between the two SCI will be SCI_{CPM} , which is created with the Camera Projection Model (CPM), as described in Chapter 2 and the SCI_{DLT} , which is created using the DLT as part of the registration method. The DLT parameters are obtained using the control points of the SCI_{CPM} and 3D points via Least Squares, and used to create the SCI_{DLT} . A comparison image is shown in Figure 4.11, along with the post ICP registration in order to match the two images together.

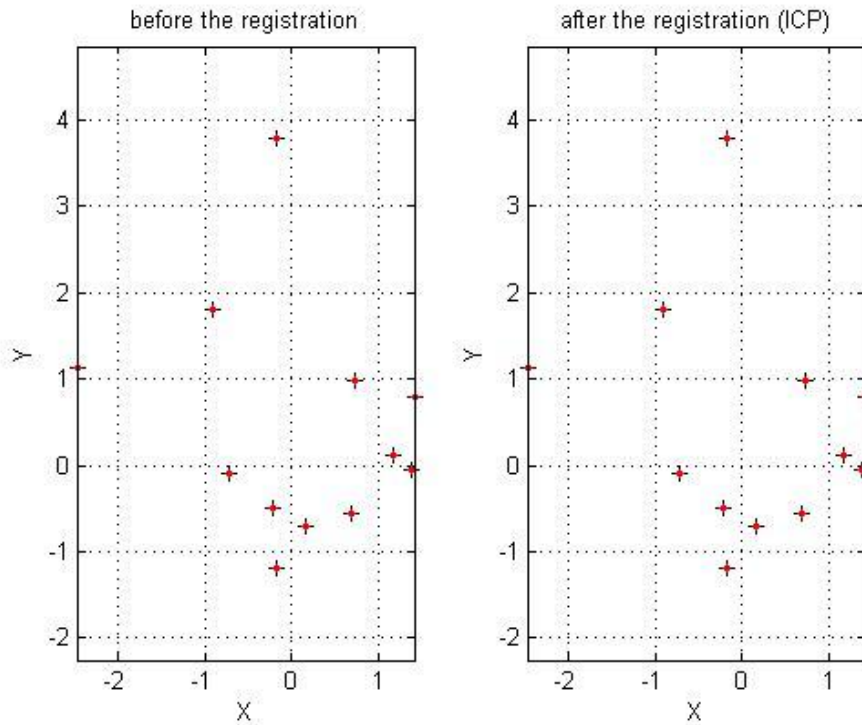


Figure 4.11: Image comparison of the SCI_{CPM} and SCI_{DLT} , and after ICP registration.

There is not much difference when viewed at this scale, but metric differences can be seen in the results shown in Table 4.2. This initial set of results does not have any errors added to the data, yet there are still minor differences. These differences are primarily produced by the use of two different projection models, the DLT and the camera projection equation. The calculated RMS of the difference between the old SCI_{CPM} control points and the new SCI_{CPM} using ICP are 0.004px and 0.004px, for X and Y coordinates respectively and 0.004px for the total RMS, where px is one pixel in the image. These sub-pixel results indicate that there is minor variation for the image.

Control Points	Original X (px)	ICP: δX (px)	Original Y (px)	ICP: δY (px)
1	145	-0.003	-9	0.004
2	12	-0.005	217	-0.008
3	-7	0.005	-112	0.003
4	138	-0.005	70	0.001
5	-124	0.004	71	-0.007
6	121	-0.003	9	0.002
7	76	0.001	-60	0.003
8	-43	-0.001	123	-0.007
9	75	-0.004	84	-0.002
10	23	0.003	-70	0.002
11	-48	0.004	-9	-0.002
12	-10	0.003	-48	0.001

Table 4.2: Difference of RCI coordinates of control points after ICP.

Another test (Table 4.3) shows that the DLT maintains a consistent error despite having variations in the precision of the control points. The precision error has been set at varying levels ranging from 0.001m up to 0.050m. The comparison test is calculated between the SCI_{CPM} and SCI_{DLT} and their respective back-projection results. Having this consistency of error is expected, as using the DLT method of matching provides the relationship between the datasets irrespective of the errors involved as well as the image resolution of the RCI being a factor influencing the precision, which will be discussed in Section 4.1.7.1.

Control point introduced precision error (m)	RMSe of ICP	RMSe of back-projection
0.001	0.004	0.027
0.002	0.004	0.026
0.003	0.004	0.026
0.004	0.004	0.026
0.005	0.004	0.027
0.006	0.004	0.026

0.007	0.004	0.027
0.008	0.004	0.026
0.009	0.004	0.026
0.010	0.004	0.027
0.050	0.004	0.027

Table 4.3: RMS error between RCI and SCI, and RMS error of back-projection results for control points having varying levels of introduced errors.

4.1.4 Comparison between the three methods of back-projection

As explained in the background, the SCI contains the parameters that relate back to 3D. After the ICP registration of the SCI_{CPM} to the SCI_{DLT} , the entire SCI_{CPM} may be transformed in order for the colour information to be assigned to the SCI_{DLT} . Then the SCI_{DLT} may be back-projected with any of the three back-projection methods to obtain the 3D solution. The resulting back-projection is given in Tables 4.4, 4.5 and 4.6, which includes the difference between the original and back-projected results.

	X (m)	δX (m)	Y (m)	δY (m)	Z (m)	δZ (m)
1	22.500	0.049	1.000	-0.056	5.000	-0.012
2	-2.871	0.018	19.070	-0.005	5.053	-0.003
3	1.838	0.010	-17.463	-0.003	5.105	-0.002
4	17.158	-0.061	11.286	0.023	5.211	0.011
5	-13.370	-0.001	3.068	0.005	5.316	0.001
6	17.882	0.006	2.880	0.031	5.474	0.002
7	13.570	-0.016	-8.926	0.059	5.526	0.008
8	-6.841	-0.043	10.241	0.017	5.632	0.008
9	7.684	0.034	10.953	-0.030	5.895	-0.008
10	5.342	-0.041	-11.279	-0.025	5.737	0.004
11	-5.763	0.056	-3.818	-0.015	6.000	-0.010
12	-0.090	-0.012	-8.511	-0.002	6.000	0.002

Table 4.4: Pseudo-inverse method: Difference between back projection and original control point coordinates.

The calculated RMS of the difference between the Pseudo-inverse back-projected RCI control points and the original control points are 0.035m, 0.029m and 0.007m, for X, Y and Z coordinates respectively and the total RMS is 0.027m.

	X (m)	δX (m)	Y (m)	δY (m)	Z (m)	δZ (m)
1	22.500	0.049	1.000	-0.056	5.000	-0.012
2	-2.871	0.018	19.070	-0.005	5.053	-0.003
3	1.838	0.010	-17.463	-0.003	5.105	-0.002
4	17.158	-0.061	11.286	0.023	5.211	0.011
5	-13.370	-0.001	3.068	0.005	5.316	0.001
6	17.882	0.006	2.880	0.031	5.474	0.002
7	13.570	-0.016	-8.926	0.059	5.526	0.008
8	-6.841	-0.043	10.241	0.017	5.632	0.008
9	7.684	0.034	10.953	-0.030	5.895	-0.008
10	5.342	-0.041	-11.279	-0.025	5.737	0.004
11	-5.763	0.056	-3.818	-0.015	6.000	-0.010
12	-0.090	-0.012	-8.511	-0.002	6.000	0.002

Table 4.5: Two-Step method: Difference between back projection and original control point coordinates.

The calculated RMS of the difference between the Two-Step method back-projected RCI control points to the original control points are 0.035m, 0.029m and 0.007m, for X, Y and Z coordinates respectively and the total RMS is 0.027m. These results do not differ from the previous results for the Pseudo-inverse method, which is expected as each of the back projection methods only invert the projection process. The differences are based on the ICP result at the RCI and SCI correspondence stage.

	X (m)	δX (m)	Y (m)	δY (m)	Z (m)	δZ (m)
1	22.500	0.049	1.000	-0.056	5.000	-0.012
2	-2.871	0.018	19.070	-0.005	5.053	-0.003
3	1.838	0.010	-17.463	-0.003	5.105	-0.002
4	17.158	-0.061	11.286	0.023	5.211	0.011
5	-13.370	-0.001	3.068	0.005	5.316	0.001

6	17.882	0.006	2.880	0.031	5.474	0.002
7	13.570	-0.016	-8.926	0.059	5.526	0.008
8	-6.841	-0.043	10.241	0.017	5.632	0.008
9	7.684	0.034	10.953	-0.030	5.895	-0.008
10	5.342	-0.041	-11.279	-0.025	5.737	0.004
11	-5.763	0.056	-3.818	-0.015	6.000	-0.010
12	-0.090	-0.012	-8.511	-0.002	6.000	0.002

Table 4.6: Additional Fourth-Row method: Difference between back projection and original control point coordinates.

The calculated RMS of the difference between the Additional Fourth-Row back-projected RCI control points and the original control points are 0.035m, 0.029m and 0.007m, for X, Y and Z coordinates respectively and the total RMS is 0.027m. As previously mentioned these results are the same as the back-projection methods that were developed to transform the SCI back to its original location. The back-projection results will differ if the ICP results change, because of the differences due to the misalignment between the SCI_{CPM} and SCI_{DLT} .

4.1.5 Test with alternate forms of 3D data: Photogrammetry data

To provide a practical dataset similar to the above test case, 3D points were acquired from Australis software (2012). This also demonstrates the versatility of this registration method as it may be used with alternate forms of 3D data, and not only with TLS point cloud information. The dataset was for a car door, as pictured in Figure 4.12. Bundle adjustment was used resulting in a total of 34 points, of which 23 were randomly selected to be control points and 11 as normal points for checks. The image was processed using millimetres instead of pixels as a variation. The camera used was a BMP-DemoCam (see Table 4.7) as specified below, where the details were obtained from the Australis software.

The DLT parameters were obtained using the selected control points for the image and 3D data. Then an SCI was created of the control points before the ICP was applied. After applying the ICP (Figure 4.13), the entire set of transformed RCI points was then back-projected (see Figure 4.14). The RMS error of the difference

between the original 3D points and the back-projected points are calculated with the results shown in Table 4.8.

BMP-DemoCam	
Sensor size (CCD)	13.7×9.1 mm
Image sizes	1524×1012 (10.0 MP)
Focal Length	0.020m
Pixel size (horizontal; vertical)	0.009 (mm per pixel)

Table 4.7: Technical specifications of the camera used within the Australis software.



Figure 4.12 The RCI image of the Car Door dataset (Australis, 2012).

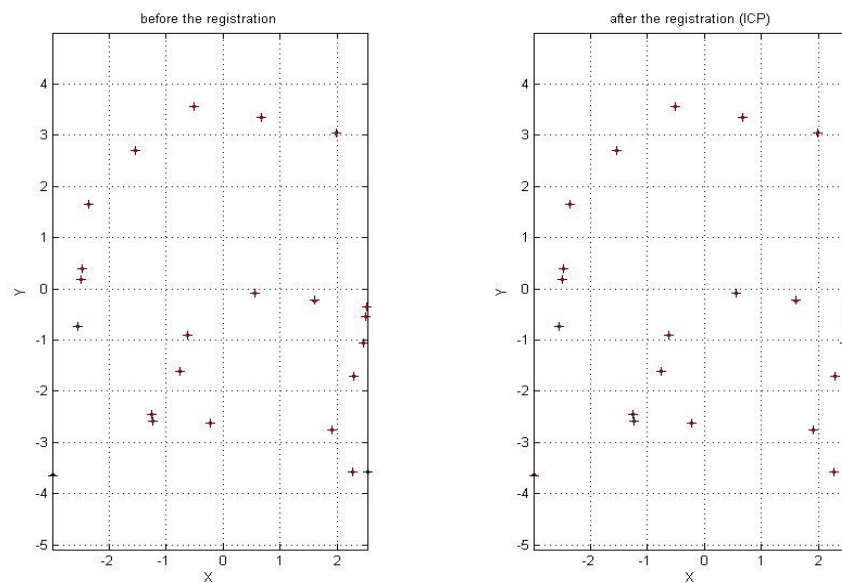


Figure 4.13 The ICP results of the RCI to SCI for the Car Door Dataset.

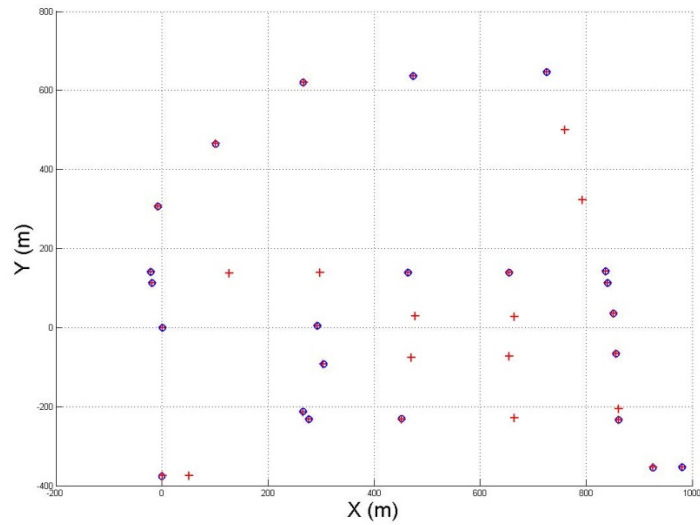


Figure 4.14 The overlaid SCI onto the RCI including check points (indicated by the '+' symbol).

The overall results in Table 4.8 show that a small scale error in the image of 0.002mm can produce up to 0.246m of error, dependent on image resolution and distance between image and object. Thus in order to achieve a desirable 3D colour transformation, one has to take into account the image resolution and distances.

RMS	Image (mm)			3D point cloud (m)			
	δX	δY	Total	δX	δY	δZ	Total
Control	0.002	0.002	0.002	0.255	0.243	0.185	0.230
Checks	0.003	0.003	0.003	0.288	0.330	0.194	0.277
Overall	0.002	0.002	0.002	0.266	0.274	0.188	0.246

Table 4.8: Results and RMS for Back-projection using Pseudo-inverse, Two-Step or Additional Fourth-row methods.

4.1.6 Simulated data with error information applied

In order to apply errors, the Conical Dataset and its control points were converted from the Cartesian coordinates (XYZ) to spherical coordinates, which is given by Eqs. 2.4, 2.5 and 2.6, for range, horizontal direction and elevation angle respectively. When the data is converted to spherical coordinates, an error value is generated

which is based on the Leica HDS specifications, as outlined in Table 4.13. The precision of the control points were $\pm 0.002\text{m}$ for the X, Y and Z coordinates, which is based upon the TLS performance specifications (Leica, 2012). After the error values were applied, the data was converted back into Cartesian form. The conversion and error synthesis process is given as code in Appendix E. An SCI of the new dataset was created and overlaid with the SCI of the dataset without errors, as shown in Figure 4.15. After the ICP process, the SCI_{CPM} was back-projected and compared to the dataset with errors, as shown in Figure 4.16. For this dataset, the total RMS error was 0.039m, consisting of RMS error in X, Y and Z, which are 0.050m, 0.045m and 0.008m respectively. Noticeably the RMS values have increased, which is expected as additional errors have been added. The Z component RMS error is small in comparison to the X and Y components, because the simulated camera is aligned with the Z axis. If the camera was in a different location then the Z component RMS error will change accordingly.

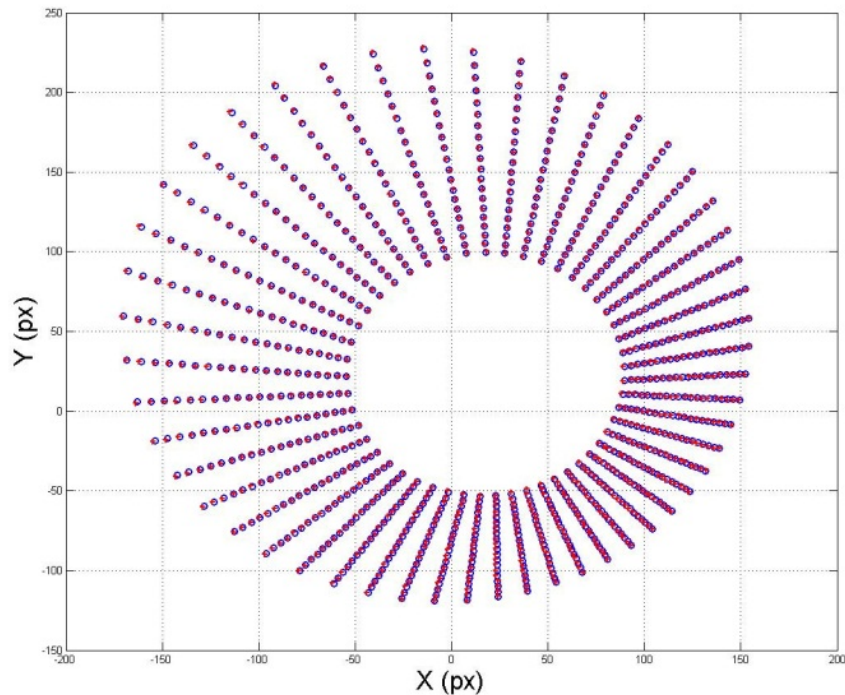


Figure 4.15: The transformed RCI overlaid onto the SCI.

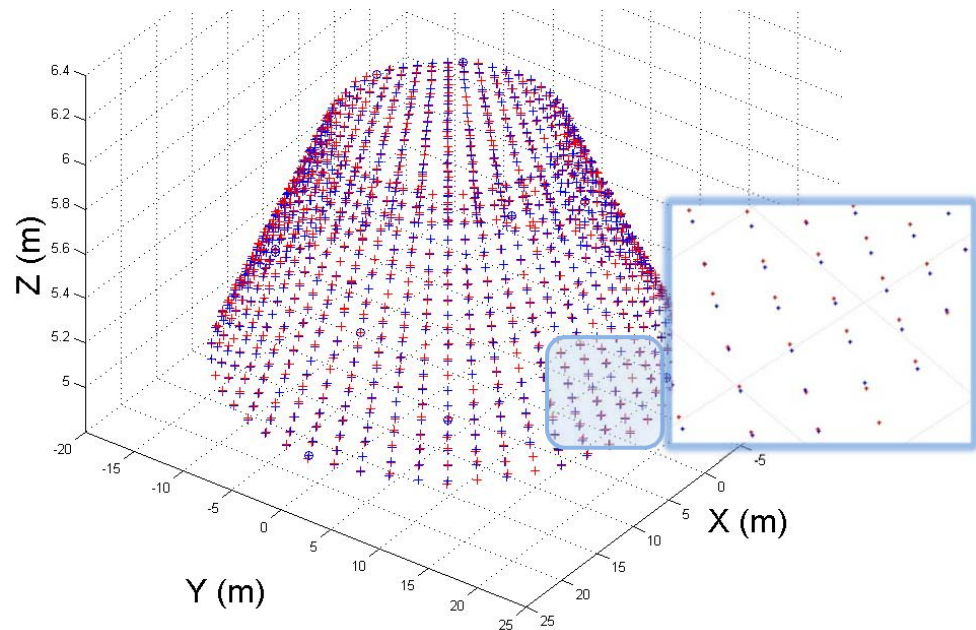


Figure 4.16: Differences from the back-projection to the original, caused by the mismatch between the SCI and RCI, with a close-up of a selected region.

4.1.7 Analysis and Evaluation with simulated dataset

Evaluation of the back-projected points is necessary to see the quality of the results obtained. The analysis is performed for the control points as these are the distinguishable references that can be assessed after the back-projection. This means that from the point of merging the RCI information to the SCI until the back-projection, the control points are the only points that remain identified. The testing of the method would be a function of range and incidence angle. These sets of results would serve as indicators, which can be used as a guide of the influences of error for the image captured compared to the point cloud in terms of range and incidence angle. This is based on the Conical Dataset and the conditions are outlined under each of the sections.

4.1.7.1 Distance

The range component would begin at a distance of 6 metres from object to the camera. Having the camera close to the laser scanner, at this distance the projection of the object onto the image is maximized in size. Calculations will be performed in increments of 1 metre initially until 15 metres where the image dimensions change to smaller values. Thereafter the distance increases in increments of 5 metres to 60

metres, where there is little variation of the image resolution. The variations in the image RMS errors are most likely due to the ICP matching process. Referring to Table 4.9, as the distance increased, the RMS of the image did not appear to have a noticeable change, but the RMS error of the back-projected coordinates shows an increase. The cause for this is the image resolution, because as the distance increases, the image resolution is required to also increase to account for the detail required to be covered per pixel. In other words, the image resolution is a limiting factor as the distance increases, as detail will be lost when each pixel begins to cover a greater area of the object. Thus as distance increases it would be ideal to have adequate pixel resolution to achieve the desired results.

Figure 4.17 is derived from Table 4.9, and indicates some possible trends for this method based on distance. According to the graph, we can see that the increase in range from the object causes the back-projected points to increase in RMS error. This is most likely due to the decreasing size of the object in the image, which may also be referred to as a decrease in image resolution, causing a pixelated effect as pictured in Figure 4.18. Thus, having a lower number of pixels to define an image would mean lesser pixels to define the precision within the image, leading to a misrepresentation when back-projected. There are also less pixels defining the object and therefore it is expected that the RMS error for the image should be less. The optimal distance for this dataset, which is about 12m, appears to show that the image dimensions up to approximately 262px produce a balanced back-projected RMS error for that image resolution. In other words, based on these results, the threshold for image dimensions should be about 262px before the RMS error becomes too high. It also can be inferred that higher image resolutions will result in lower back-projection RMS error values. Therefore, the outcome of the error in back-projection has a dependence upon image resolution, which is influenced by range.

Distance (Z m)	Average image dimensions (px)	RMS Total	
		Image (px)	Back-Projection (m)
1	3017	0.0037	0.0027
2	1544	0.0054	0.0063
3	1037	0.0012	0.0069
4	781	0.0087	0.0073
5	626	0.0036	0.0110
6	523	0.0082	0.0118
7	449	0.0062	0.0121
8	393	0.0041	0.0149
9	350	0.0022	0.0145
10	315	0.0019	0.0211
11	287	0.0042	0.0286
12	262	0.0013	0.0227
13	242	0.0043	0.0276
14	225	0.0021	0.0291
15	210	0.0058	0.0413
20	158	0.0024	0.0486
25	126	0.0014	0.0442
30	105	0.0016	0.0580
35	90	0.0024	0.0740
40	79	0.0025	0.0811
45	70	0.0028	0.0978
50	64	0.0043	0.1209
55	58	0.0033	0.1189
60	53	0.0039	0.1485

Table 4.9: RMS error results of Conical Data distance tests.

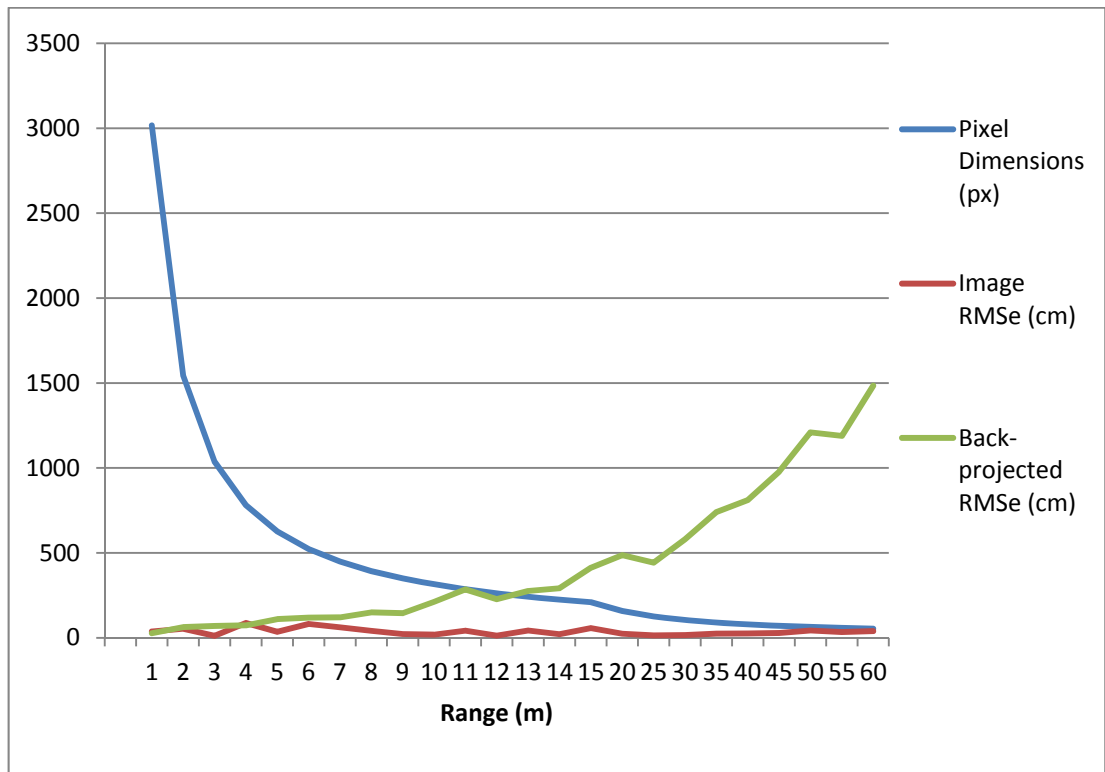


Figure 4.17: RMS error of image and back-projection, in terms of range.

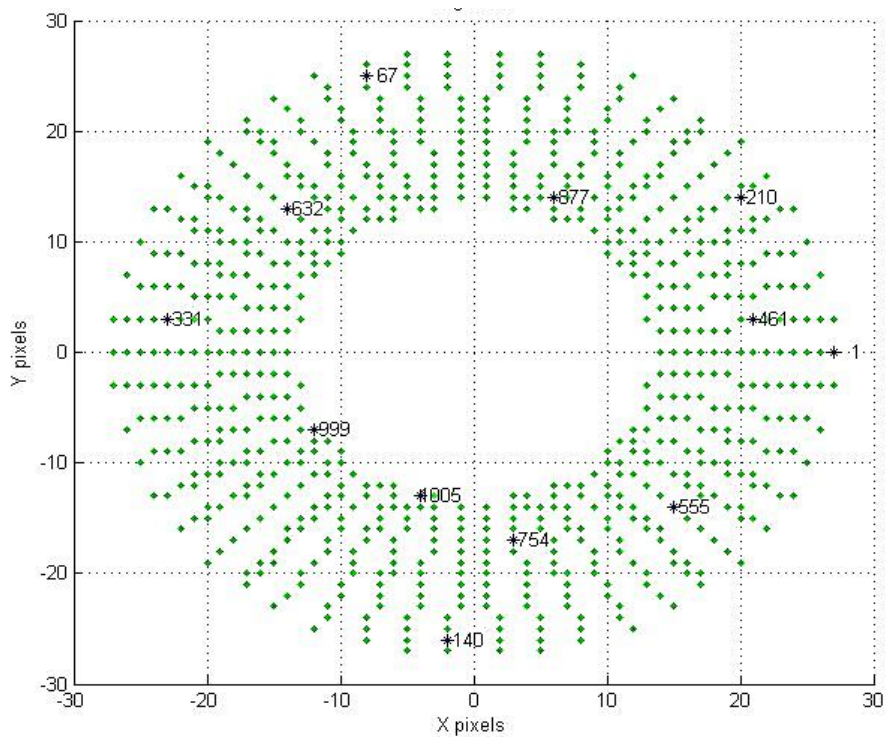


Figure 4.18: The RCI of the Conical Dataset at range 60m, simulating reduced resolution causing a pixelated effect compared to the original (Figure 4.4 (b)).

For variations in incidence angle, tests may be done from a perpendicular perspective and moving right or left until up to a maximum of 80° incident angle with changes by 10° at a distance of 25m. This test method determines the effect when the image is taken from angles away from nadir. Table 4.10, shows the opposite effect to the range tests, in which for this case, the image results seem to show an increase in error as the incident angle increases, with the back-projection showing no apparent trends. Figure 4.19 visually shows this trend and indicates that at 40° to 50° the back-projection has a high RMS error. It is likely that from 40° onwards, the angle causes too many misalignments in the SCI and RCI. The optimal angle for image capture seems to be from nadir to 30°. Based on these results, it is best to limit the angle of capture within a 30° incident angle from the TLS data capture to obtain desirable results.

Angle (ω°)	RMS Total	
	Image (px)	Back-Projection (m)
0	0.0031	0.0745
10	0.0016	0.0484
20	0.0037	0.0645
30	0.0049	0.0503
40	0.0073	0.0911
50	0.0280	0.0837
60	0.0380	0.0622
70	0.0276	0.0551
80	0.0775	0.0671

Table 4.10: RMS error results of Conical Data incident angle tests.

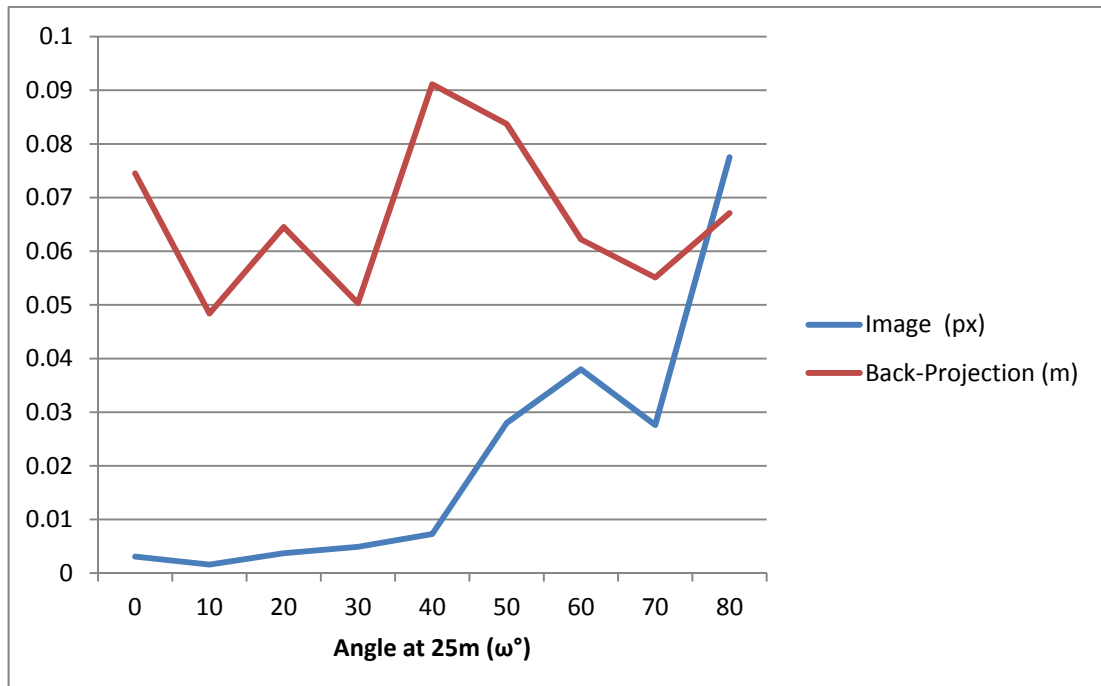


Figure 4.19: RMS error of image and back-projection, in terms of incident angle at range 25m.

The following results were produced for the more optimal range of 12m. Similar to the previous distance of 25m, 30° was the limit for this dataset as Figure 4.21 visually shows that this angle causes many of the points to be clustered into one area. Figure 4.20 provides a smoother indication than Figure 4.19 of the behaviour of the results up to the point where the image causes too much error. For this case, Table 4.11 shows that within an incident angle of 10° , millimetre accuracy is achievable with an optimal range and image resolution ($<0.0024\text{m}$).

Angle (ω°)	RMS Total	
	Image (px)	Back-Projection (m)
0	0	0.0013
5	5	0.0024
10	10	0.0021
15	15	0.0118
20	20	0.0373
25	25	0.0258

30	30	0.0651
33	33	0.5794

Table 4.11: RMS error results of the Conical Data incident angle tests.

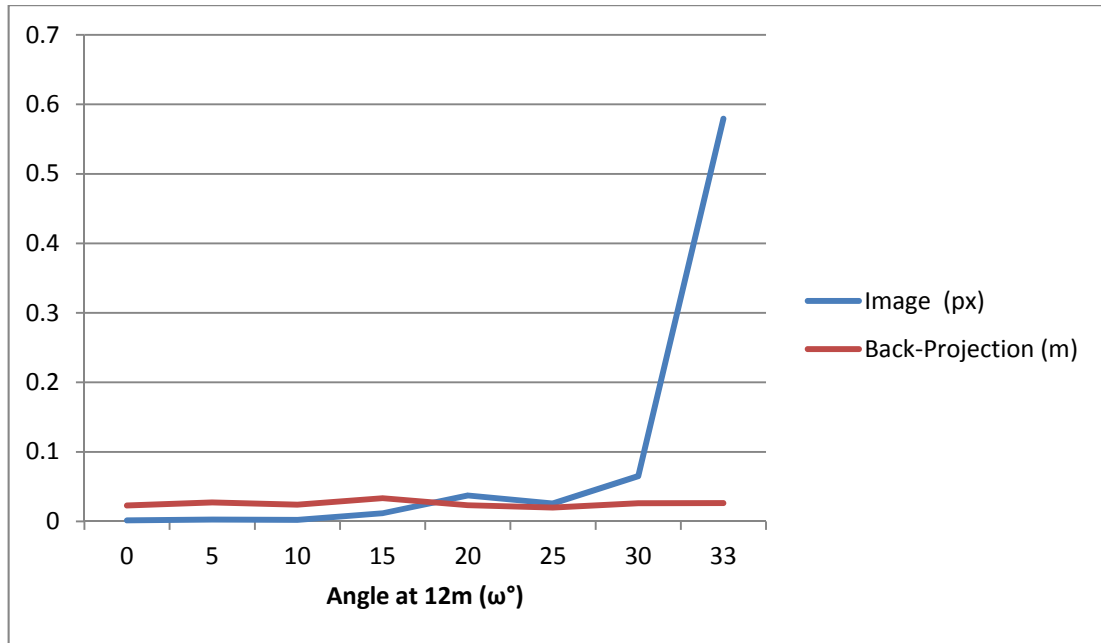


Figure 4.20: RMS error of image and back-projection, in terms of incidence angle at optimal range of 12m.

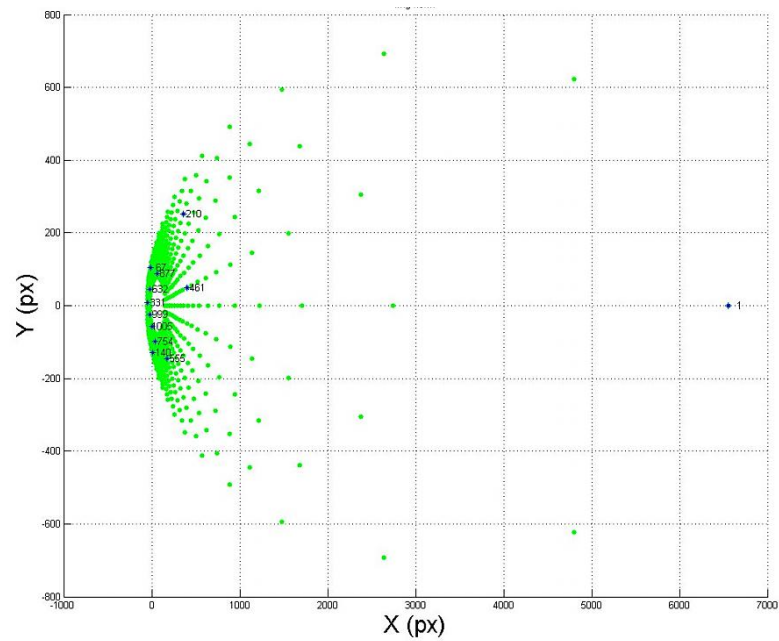


Figure 4.21: Image showing the cause of error at the limit of incident angle.

4.1.7.2 Pixel Coverage

It is important to note that the object area will require sufficient pixel coverage to enable better resulting precision. This is because the back-projection is dependent on good correspondence between the RCI and SCI. From nadir up to 30° should be the best angles to use, but it is not always possible as factors such as obstructions, lighting conditions and limited accessibility exist that will influence the location of the captured image. For this discussion, it is assumed that the angle of incidence is taken within the limits to preserve the shape of the coverage area (object) for the pixel, so that the focus can be on scale, range and image resolution.

At a certain range, each pixel will cover a scaled region of the object; this will be referred to as the coverage area. This means that a certain range may cover a 1cm square or even a 50cm square for a pixel. This can be understood by the example in Figure 4.22, which shows the range alters the scale and therefore the density of pixels in the object area. Thus depending on range and focal length each pixel covers a certain region of the object. To reduce the coverage area, the range may be reduced, or a greater pixel density (better camera) used. Therefore when the coverage area is reduced, the result is to allow a greater level of precision because the RCI and SCI have a greater level of detail for the correspondence process.

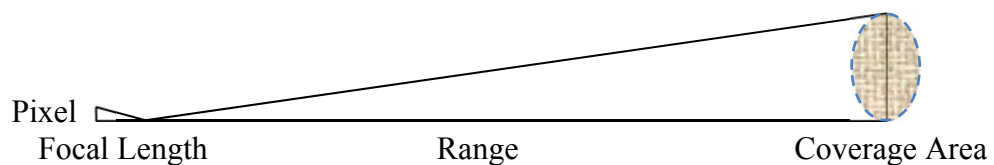


Figure 4.22: The range should be chosen so that the pixel and coverage area will allow the desired precision.

4.2 Experiments with Real Datasets

This section considers real datasets, comprising a TLS point cloud with control points and a digital photo of the area/object taken with a consumer grade camera. The specifications of the equipment used for this test are outlined in Section 4.3.1.

4.2.1 Data Acquisition and Specifications

The data acquired from the camera and the TLS has to be processed and formatted into an organised form, whereby the corresponding control point data for the image and point cloud must be labelled and arranged into a suitable format for processing to take place. The RCI obtained from a consumer-grade SLR camera (Nikon D80), needs to have the distortion parameters applied to obtain a corrected image with minimal errors. The table below outlines the technical details of the camera and TLS used for this research.

Nikon D80	
Sensor size (CCD)	23.6×15.8 mm
Image sizes	3872×2592 (10.0 MP)
Focal Length	0.018m
Pixel size (horizontal; vertical)	0.006 (mm per pixel)

Table 4.12: Technical specifications of the camera used (Nikon, 2012).

Leica HDS Specifications	
Measuring Method	Pulsed and time-of-flight
Scan Range	300 m
Laser Class	Class 3R (IEC 60825-1)
Accuracy	±6 mm
Precision	±4 mm
Horizontal Angle Precision	60 µrad
Vertical Angle Precision	60 µrad
Image size	1024×1024 pixels

Table 4.13: Technical specifications of the Terrestrial Laser Scanner used.

4.2.2 Results of Laboratory Dataset

The test setup area as shown in Figure 4.23 consists of 17 targets which were a mixture of Leica HDS targets (10 targets) and printed black and white targets (7

targets). The approximate distance from the scanner to the objects was relatively close for this medium range scanner. The image was taken to cover the set of targets, while taking into account capturing within 30° and as close as possible attempting to effectively use the maximum pixel resolution for this area. As mentioned before in Section 4.1.7.1 it is important to have as much pixel density as possible for the given range to produce better back-projection results. This is a preliminary field test and Figure 4.24 is a plan of the setup area.

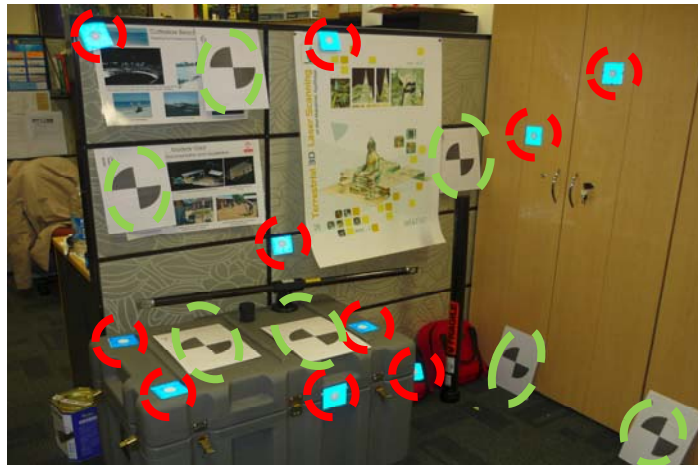


Figure 4.23: An image of the setup area indicating the Leica HDS targets (red), and printed Black and White targets (green), that are used as control points.

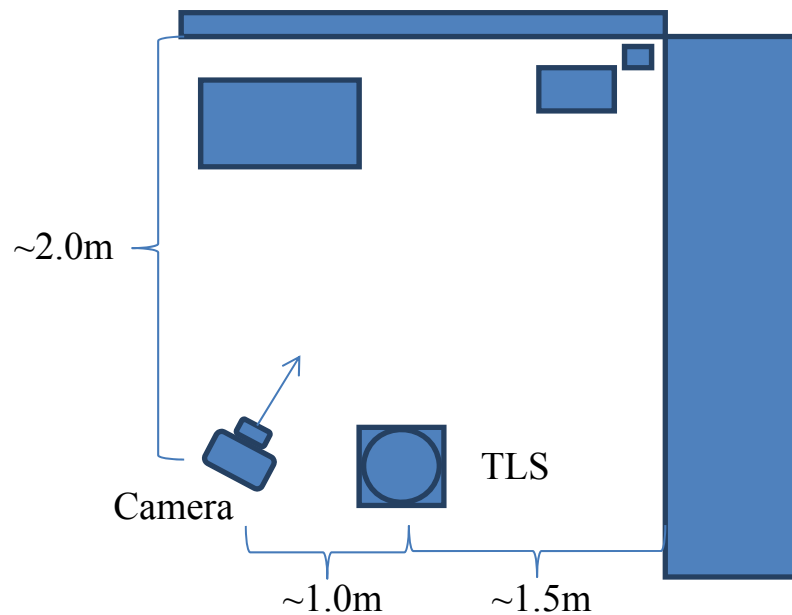


Figure 4.24: A plan of the setup area.

This dataset acquired using the Leica HDS, contains 88,216 points at the distance of 1.5m, covering the available testing area. Figure 4.27 shows the image of the area with the corresponding point cloud data superimposed using the DLT. The corresponding control points for the image need to be identified during the data capture using the target detection feature of the TLS instrument. This is to ensure that the control points obtained are determined as accurately and precisely as possible. The RCI control points can be obtained using Australis (2012) or similar software with a centroid function to detect target centres. In order to obtain millimetre or pixel coordinate values, a conversion value may be calculated to convert the control point values of the SCI (Appendix A). In this project, pixel values are primarily used for the RCI control points.

The registration of the RCI control points to the SCI is performed using ICP, which in turn provides the transformation parameters required to transform the entire RCI. The results of the ICP between the RCI and SCI control points are shown in Figure 4.25.

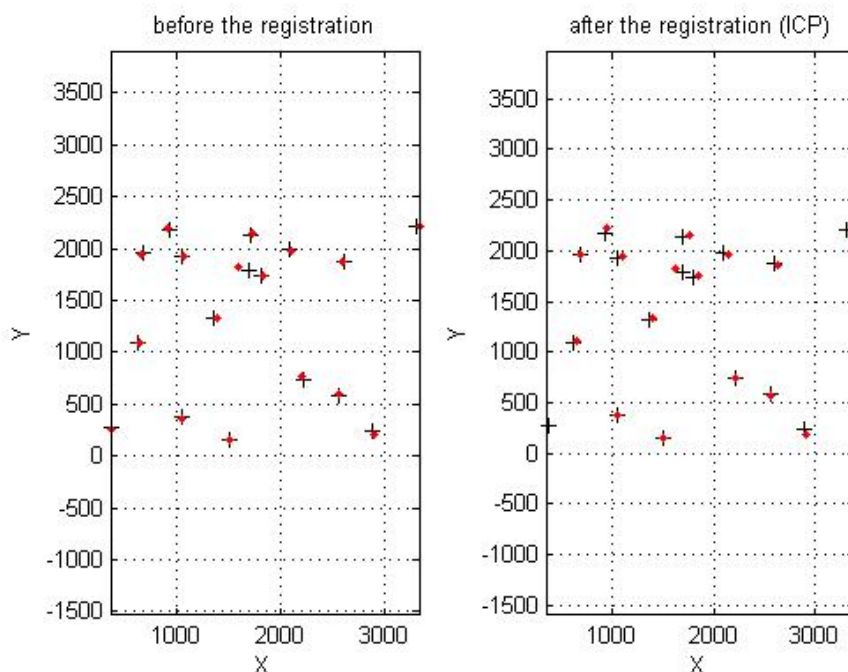


Figure 4.25: ICP results of the Lab Dataset in pixels.

The transformation parameters acquired from the ICP are then applied to the entire RCI, before proceeding to creating the SCI of the entire point cloud. Using the DLT parameters, the entire point cloud is converted into an SCI, as shown in Figure 4.26, including the control points. Next the transformed RCI colour information is assigned to the corresponding SCI points. As a visual example of this process, Figure 4.27 shows the SCI points overlaid onto the RCI, whereby the green dots indicate the corresponding SCI points to be assigned the colour information.

Some apparent limitations are erroneously stored colour data, outlined in red in Figure 4.27 where the region of the point cloud is not visible in the image, but present in the point cloud because of the slightly different viewpoints of the camera and TLS. The hidden region present in the point cloud causes the SCI to use the RCI colour information for that area. A future method suggested to avoid this occurrence is to apply threshold parameters, temporarily manually or automatically delete the data points that do not associate with the RCI, or obtain the image (or a second image) from a different perspective.

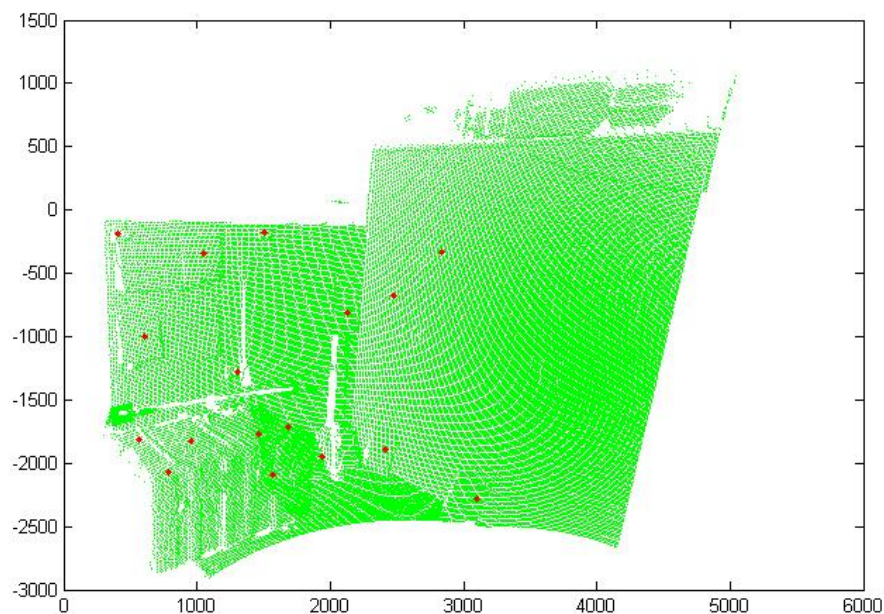


Figure 4.26: SCI of the point cloud data in pixels.

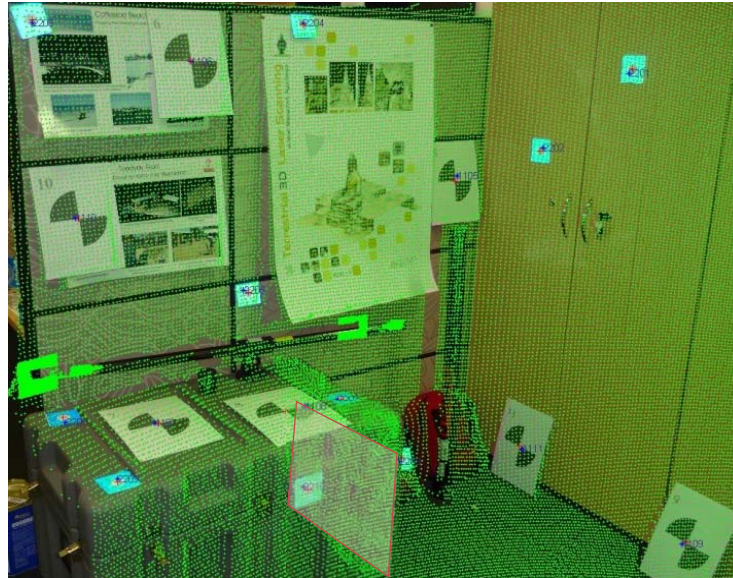


Figure 4.27 The overlaid SCI onto the RCI with incorrectly stored data outlined within the red region.

Once the colour information from the RCI has been stored onto the SCI, it is ready for back-projection. Figure 4.28, shows the back-projected coloured point-cloud which any of the three methods will produce. However it is important to view the differences metrically as well as visually. Table 4.14, shows the back-projection of the control points to 3D data space using the Pseudo-inverse method. Results for the Two-Step method and the Additional Fourth-row method are not shown as they produce the same results. The errors or mismatches after back-projection are a result of the mismatch of RCI to SCI control points, as ICP is the measure implemented to minimise the effect of the mismatch.

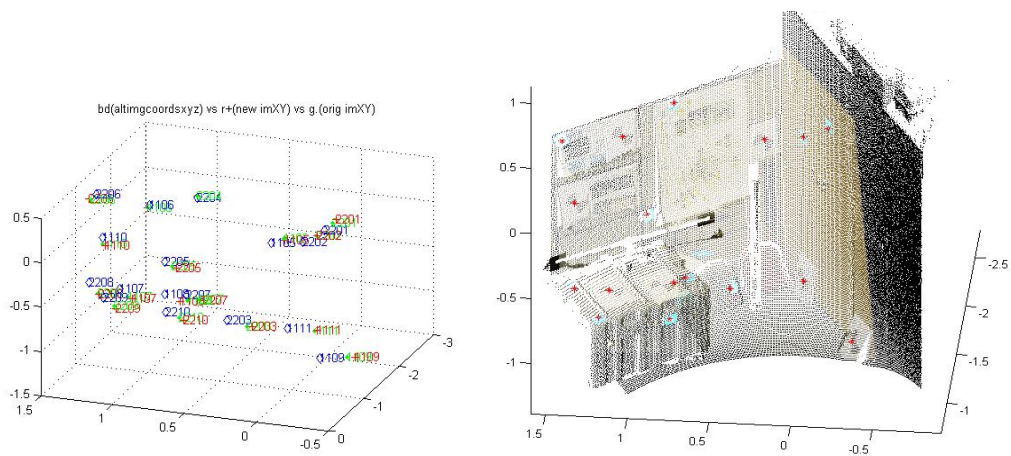


Figure 4.28: The back projection points with colour assigned in metres.

	X (m)	δX (m)	Y (m)	δY (m)	Z (m)	δZ (m)
1	0.321	-0.008	-2.094	0.004	-0.250	-0.011
2	1.125	0.000	-1.522	0.000	0.150	0.000
3	1.156	-0.029	-1.084	0.006	-0.708	-0.013
4	0.773	0.046	-1.191	0.001	-0.708	-0.022
5	-0.310	-0.057	-1.307	0.002	-1.183	0.013
6	1.400	-0.020	-1.351	0.006	-0.252	-0.014
7	0.058	-0.024	-1.890	-0.002	-1.174	0.017
8	-0.116	-0.021	-1.703	-0.008	0.153	0.039
9	0.064	-0.006	-1.965	-0.001	-0.122	0.006
10	0.492	-0.041	-1.743	0.001	-1.151	0.010
11	0.839	0.000	-1.757	0.000	0.254	0.000
12	0.993	-0.033	-1.658	0.006	-0.540	-0.012
13	1.473	0.011	-1.318	-0.001	0.246	0.001
14	0.707	-0.033	-1.261	0.005	-0.715	-0.005
15	1.369	-0.006	-1.092	0.002	-0.717	-0.004
16	1.187	-0.014	-0.809	0.008	-0.714	-0.024
17	0.784	-0.037	-1.002	0.008	-0.840	-0.015
RMS		0.028		0.005		0.015
Total RMS						0.019

Table 4.14: Results and RMS for Back-projection using Pseudo-inverse, Two-Step or Additional Fourth-row method.

The RMS errors of the results were within expectations for the given range (2m – 3m) and image resolution. The RMS error of 0.019m for the control points would indicate the approximate error for the entire dataset.

4.2.3 Results of T-Rex Dataset

This dataset was acquired using the same instruments as mentioned in Section 4.2.1. The control points were manually obtained by picking the central points of the targets in both the image and the point cloud data. The image control point coordinates were obtained by using image viewing software showing pixel

coordinate values, while Cyclone v.6.0 was used to select the 3D point cloud coordinates. Figure 4.29 gives an indication of the 3D viewing space used to select the control points and its corresponding RCI.

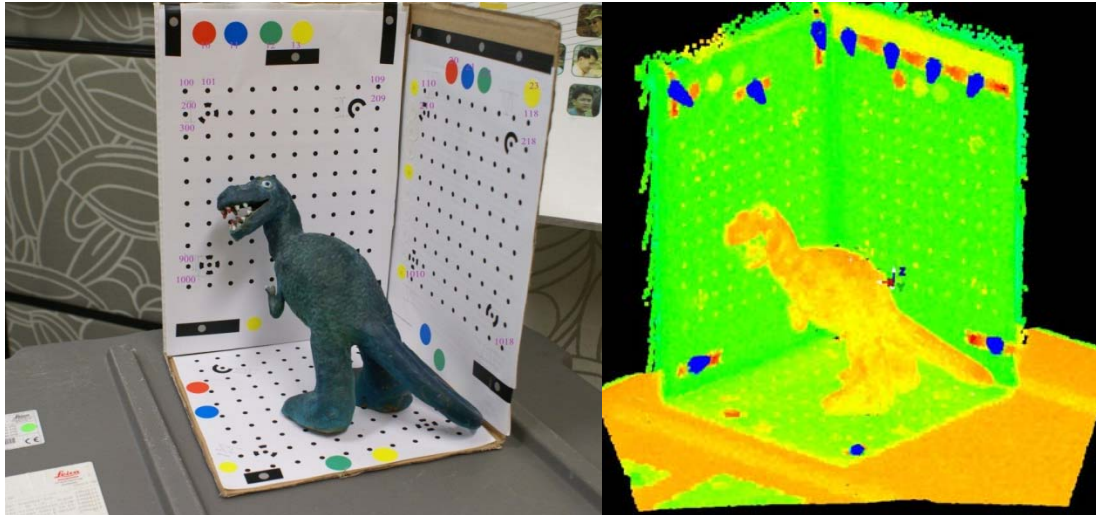


Figure 4.29: The RCI (left) and the point cloud for the T-Rex dataset (right).

Next, the datasets were processed to produce the SCI of the point cloud (129923 points), which produced the result shown in Figure 4.30. Some points of the SCI have erroneously assigned data, which is expected and highlighted in red. One method may be to manually delete the points that are misrepresented on the image. Another method that may limit the misrepresentation is to capture the image with an incident angle closer to nadir.

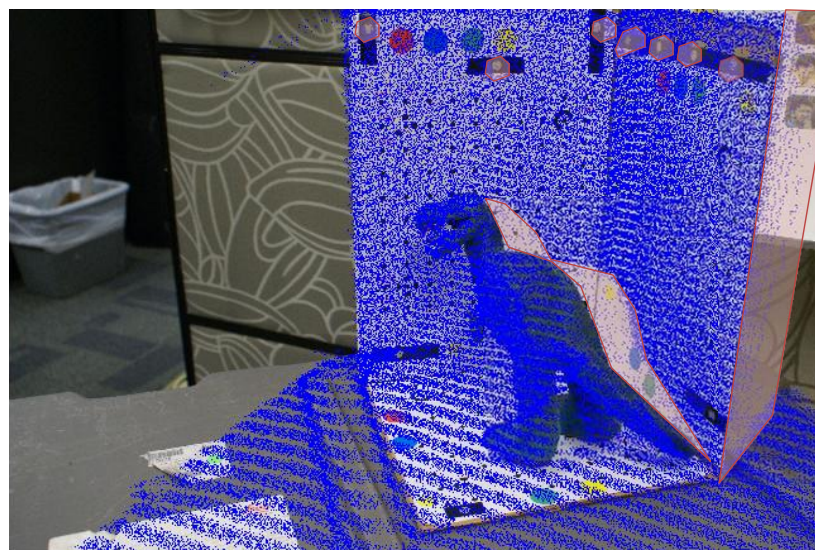


Figure 4.30: The SCI of point cloud data overlaid onto RCI with red regions indicating data with incorrect colour assigned.

The T-Rex Dataset was acquired at a distance of approximately 1.3m from the object, yielding image RMS error of 1.5px after ICP (15.3px before ICP) and back-projection RMS error of 2mm. These results are consistent with the findings in Section 4.1.7.1, whereby having a high resolution for the given range allows achievement of millimetre precision. With regard to incidence angle, having taken an image within the 30° limitations has also contributed to this solution obtaining millimetre precision. The results of the successful back-projection of this dataset are displayed in Figure 4.31 from a different angle:

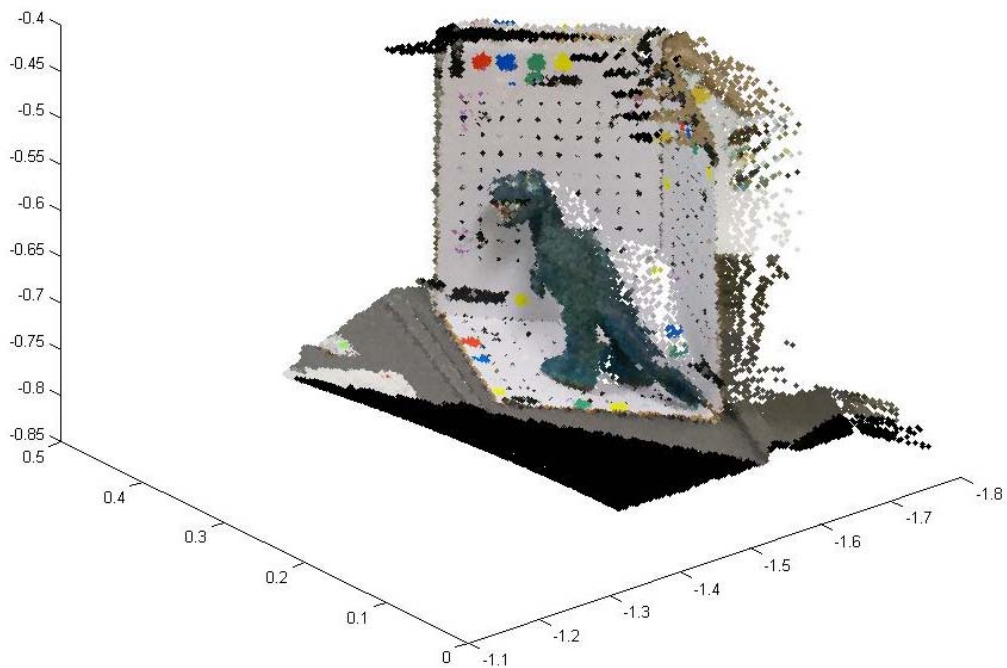


Figure 4.31: The back projection of T-Rex data with colour assigned. The dimensions are in metres.

4.3 Chapter Summary

Several datasets were created in order to evaluate the back-projection techniques. One was manually calculated and two were generated in Matlab 7.12. The first test was to trial a simple back-projection by inverting the projection matrix, but the results were not transformed to their original location as expected. However a trend was identified, which was a pre-cursor to the development of the three back-projection methods.

The three back-projection methods were tested to determine their performance, resulting in obtaining the same results for all three methods. Thus any of the three methods would be suitable in order to obtain a successful back-projection. In order to confirm the method works in a practical application, a dataset taken from Australis was used.

The next step was to use a simulated dataset with some random errors incorporated based on the equipment specifications to simulate a more realistic dataset. The RCI was still simulated for this case. Then the DLT components were derived using least squares, followed by matching with ICP, and finally back projecting. This dataset was used to analyse the effect of range and incidence angle on the back-projection results. Showing that as a function of range, the errors in the back-projection increased when the distance increased. For the incidence angle, the effect was mainly on the error between the RCI and SCI, i.e. the image RMS error increased as the angle deviated away from the object. Based on the image resolution an optimal range may be determined, which is dependent on the precision requirements. Also, the incident angle of the captured RCI should be kept within 30° in order to maintain a good level of precision.

The significance of the image resolution is discussed, since the results of the back-projection relies on good ICP matching between the RCI and the SCI, thus it is important that consideration is taken for the pixel coverage of the object space. When millimetre precision is required for the back-projection, sufficient image resolution for the distance of image capture is required for consideration as each pixel contributes to the level of precision that can be obtained.

Lastly, the proposed methodology was carried out in full using a realistic dataset comprising of a consumer grade digital SLR camera and TLS. Firstly the image data is acquired from a Nikon D80 and the point cloud data from the Leica laser scanner. Then the control points are identified for the image and point cloud data before using least squares to determine the DLT parameters. Next the SCI is created and ICP applied to produce the transformation required for the RCI to be matched. After ICP registration, the colour was registered onto the SCI and back projected to produce a

coloured point cloud. The RMS error of the control points was calculated to provide an indication of the mismatching of the colour data on the point cloud.

The Laboratory dataset and the T-Rex dataset were the realistic datasets. Both datasets were able to meet the expectations based on the range and image resolutions that were used to capture these datasets. However, both had shown that some thresholds and/or deletion of data or an image from an alternative viewpoint would be required as some sections of the point-cloud had colour stored when no image data was available for those points.

5 CONCLUSION

This dissertation proposed a method to utilise a single image and 3D point cloud for the provision of additional depth information using Direct Linear Transformation (DLT) in order to recreate a photorealistic 3D object model.

The registration and back-projection applies both terrestrial laser scanning (TLS), photogrammetry principles along with DLT in order to create the 3D colour model. This DLT method also presents the opportunity to not only stay within the bounds of TLS and cameras for acquisition of data, but also utilise other sensors that are able to provide the same types of information, i.e. providing 2D and 3D information.

Depth information recovery and flexibility are the main attributes of this method, which allows the use of a single image without the need of mounting brackets or fixtures, in registering point clouds with colour information using the DLT approach. The core idea is based on transforming 3D to 2D then from 2D to 3D, which is termed as projection and back-projection respectively. Since DLT is the basis of the method, an in-depth understanding is provided, outlining its relationship with the collinearity equation, its limitations, and normalisation. These main aspects of DLT are required to understand the relationship that is present using this method, between the 2D and 3D datasets.

In terms of projection, it is quite a straightforward process, which is determining the DLT parameters using least squares to convert the point cloud data into image space. The image scale component is an important value arising from the use of DLT to calculate the Synthetic Camera Image (SCI), which is normally lost during the projection process. In order to achieve the transformation from 2D to 3D, the image scale component is the necessary depth parameter recovered enabling back-projection.

There are two main aspects that are common for back-projection, the image scale component and the perspective centre. The depth parameter is given by the image scale component, while the perspective centre provides the translation component to shift the coordinates into the appropriate locations. These two components form the

basis of the three methods of back-projection, which are the pseudo-inverse method, Two-Step method and the Additional Fourth-Row method.

The process of the proposed method begins with acquiring the data and control point information from devices such as TLS for 3D and cameras for 2D. Then the DLT parameters are obtained by means of a least squares solution, which is used to create an SCI of the control points for registration. Registration of the SCI to the Real Camera Image (RCI) uses the corresponding control points in the SCI and the RCI with ICP. The ICP process will transform the RCI to match the SCI as close as possible, in which the colour information will be assigned to the SCI. After assigning the colour information, the SCI is then back-projected using one of the methods, e.g. pseudo-inverse method, which results in a photo-realistic colour model.

In order to examine the proposed method, synthetic and real datasets were used and analysed in terms of range and incidence angle. The synthetic datasets are simulated using the manual calculation and automated methods, while the real datasets were obtained using photogrammetric software as well as the TLS and camera combination. The preliminary test showed that the simple back-projection via a matrix inverse did not return the projected 3D points to their original location, which is expected, hence the development of the three proposed back-projection methods. These three methods of back-projection were designed to obtain the original 3D coordinates from the SCI solution, thus using any of these methods can achieve the desired back-projection results.

The synthetic datasets provided initial indications of factors that attributed to the back-projection and colour assignment, which were ICP correspondence of the RCI and SCI control points, range, incidence angle and image resolution. The registration stage involving ICP indicated that it was necessary to achieve a close match between the RCI and the SCI to ensure back-projection results with minimal error. The ideal case is to have a perfect match of the RCI to the SCI, however for practical scenarios, minimizing the components of error will provide the best results.

Testing of the range component shows that there is an optimal maximum range versus error performance, and that reducing the range from the optimal will produce better results. More importantly, the image resolution presents itself as the major component and that range is an influencing factor. This test accentuates that the image resolution has to be considered in order to achieve precise results.

The incidence angle tests the effect of capturing the image at a variety of angles. Results indicate that having angles of image capture above 30° from the face of the object produced poor/high values for precision, which were caused by clustering of data. Thus, the photographed object should be taken within the 30° incidence angle to achieve good levels of precision.

An important aspect to consider is pixel coverage, which will aid in producing results with good levels of precision. This aspect is to make an awareness that the influencing factors of range and incidence angle are to be taken into account for the image resolution. The main aim is to reduce the pixel coverage for the intended project, which can be achieved by either having a very high image resolution (determined by the camera used) or a shorter range.

After testing with synthetic datasets, two real datasets were acquired to test the proposed method. Both the datasets were acquired with the Nikon D80 (image) and the Leica Scanstation laser scanner (point cloud). The DLT parameters were obtained using the corresponding control points, followed by the creation of the SCI of the control points. Then colour information from the RCI is registered onto the SCI with ICP. Next, the SCI can be back-projected with any of the three methods to produce a coloured point cloud. Lastly, the RMS error was calculated from the control points to provide a metric indication of the amount of mismatching. Both datasets met the expectations based on range and image resolutions, and a successful coloured point-cloud with metric indication of quality was achieved.

REFERENCES

- Abdel-Aziz, Y. I. and H. M. Karara (1971). "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry", Proceedings of the Symposium on Close-Range Photogrammetry. Falls Church, VA, American Society of Photogrammetry: 1-18.
- Adam Technology (2012). <http://www.adamtech.com.au/>, accessed on 01 June 2012.
- AICON (2012). <http://www.aicon3d.com/>, accessed on 10 March 2012.
- Al-Manasir, K. and C. S. Fraser (2006). "Registration of terrestrial laser scanner data using imagery", Photogrammetric Record 21(115): 255-268.
- Atkinson K. B. (1996). "Close Range Photogrammetry and Machine Vision", Whittles Publishing, 371 pages.
- Australis (2012). <http://www.photometrix.com.au/>, accessed on 10 March 2012.
- Bae, K.-H. and D. D. Lichti (2008). "A method for automated registration of unorganised point clouds", ISPRS Journal of Photogrammetry and Remote Sensing 63(1): 36-54.
- Beraldin, J.-A. (2004). "Integration of laser scanning and close-range photogrammetry - The last decade and beyond", Proceedings of ISPRS (The Photogrammetry, Remote Sensing and Spatial Information Science) Congress, Commission VII, Istanbul, Turkey: 972-983.
- Besl, P. J. and N. D. McKay (1992). "A method for registration of 3-D shapes", Transactions on Pattern Analysis and Machine Intelligence. 14(2): 239-256.
- Briese, C., N. Pfeifer, and A. Haring (2003). " Laser scanning and photogrammetry for the modelling of the statue of Marc Anton", International Archive of Photogrammetry and Remote Sensing and Spatial Information Sciences, Antalya, Turkey. <http://cipa.icomos.org/fileadmin/template/doc/antalya/139.pdf>, accessed at 10 March, 2012.

Chen, L., C. W. Armstrong, and D. D. Raftopoulos (1994). "An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique", *Journal of Biomechanics* 27(4): 493-500.

Dermanis A. (1994b). "The Photogrammetric Inner Constraints", *ISPRS Journal of Photogrammetry and Remote Sensing*, 49 (1), 25-39.

Dermanis A. (1994a) "Free network solutions with the DLT method", *ISPRS Journal of Photogrammetry and Remote Sensing*, 49 (2), 1-12.

Forkuo, E. K. and B. King (2005). "Automatic fusion of photogrammetric imagery and laser scanner point clouds", *Proceedings of the XXth International Society for Photogrammetry and Remote Sensing Congress, Istanbul, Turkey*: 921-926.

Fraser, C. S. (1997). "Digital camera self-calibration", *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(4): 149-159.

Hartley, R. and A. Zisserman (2003). "Multiple view geometry in computer vision", *Cambridge University Press*, 672 pages.

Ikeuchi, K., A. Nakazawa, K. Hasegawa and T. Ohishi (2003). "The Great Buddha Project: Modeling cultural heritage for VR Systems through observation", *Proceedings of the Second IEEE/ACM International Symposium on Mixed and Augmented Reality*: 7-16.

Jansa, J., N. Studnicka, G. Forkert, A. Haring and H. Kager (2004). "Terrestrial laser scanning and photogrammetry – acquisition techniques complementing one another", *Proceedings of the XX ISPRS Congress, Istanbul, Turkey*. <http://www.isprs.org/congresses/istanbul2004/comm5/papers/184.pdf>, accessed on 10 March 2012.

Jiang, R., D. V. Jáuregui, and K. R. White (2008). "Close-Range Photogrammetry Applications in Bridge Measurement: Literature Review", *Journal of the International Measurement Confederation* 41, no. 8: 823-34.

Kang, Z., S. Zlatanova, and B. Gorte (2007). "Automatic Registration of Terrestrial Scanning Data Based on Registered Imagery", *Proceedings of FIG Working Week*:

Strategic Integration of Surveying services, Hong Kong SAR, China. http://www.fig.net/pub/fig2007/papers/ts_8c/ts08c_02_kang_etal_1317.pdf, accessed on 10 March 2012.

Karara H. M. (1989). "Non-Topographic Photogrammetry", Second edn, ASPRS (American Society of Photogrammetry and Remote Sensing), 445 pages.

Kern, F. (2001). "Supplementing Laser Scanner geometric data with photogrammetric images for modeling", Processing of the International CIPA Symposium, Potsdam, Germany: 454-461.

Konica Minolta (2012). <http://sensing.konicaminolta.us/>, accessed on 10 March 2012.

Leica_Geosystems (2012). <http://hds.leica-geosystems.com/>, accessed on 10 March 2012.

Liu, L. and I. Stamos (2005). "Automatic 3D to 2D registration for the photorealistic rendering of urban scenes", Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2: 137-143.

Liu, L., I. Stamos, G. Yu, G. Wolberg and S. Zokai (2006). "Multiview geometry for texture mapping 2D images onto 3D range data", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2: 2293-2300.

Luhmann, T., S. Robson, S. Kyle and I. Harley (2006). "Close Range Photogrammetry: Principles, Methods and Applications", Whittles Publishing, 510 pages.

Matlab (2012). <http://www.mathworks.com.au/>, accessed on 10 March 2012.

MesaImaging (2012). <http://www.mesa-imaging.ch/>, accessed 01 June 2012.

Mikhail, E. M., J. S. Bethel and J. C. McGlone (2001). "Introduction to Modern Photogrammetry", John Wiley & Sons Inc., 479 pages.

McGlone, J. C., E. M. Mikhail, J. S. Bethel and R. Mullen (2004). "Manual of Photogrammetry", Fifth Ed., ASPRS (American Society of Photogrammetry and Remote Sensing), 1151 pages.

Nikon (2012). <http://www.nikon.com.au/>, accessed on 10 March 2012.

Reshetyuk, Y. (2009). "Self-calibration and direct georeferencing in terrestrial laser scanning" PhD dissertation, Kungliga Tekniska högskolan (KTH), 162 pages.

Riegl (2012). <http://www.riegl.com/>, accessed on 10 March 2012.

Rusinkiewicz, S. and M. Levoy (2001). "Efficient Variants of the ICP Algorithm", Proceedings of the Third International Conference on 3-D Digital Imaging And Modeling, Quebec City, Quebec: 145 - 152.

Salvi, J., C. Matabosch, et al. (2007). "A review of recent range image registration methods with accuracy evaluation", *Image and Vision Computing* 25(5): 578-596.

Sapkota, P. P. (2008). "Segmentation of Coloured Point Cloud Data" M. Sc dissertation, International Institute for Geo-Information Science and Earth Observatin (ITC), Enschede, The Netherlands, 68 pages.

Sillard and Boucher (2001). "A review of algebraic constraints in terrestrial reference frame datum definition", *Journal of Geodesy* 75: 63-73.

Stamos and Allen (2002). "Geometry and Texture Recovery of Scenes of Large Scale", *Computer Vision and Image Understanding (CVIU)* 88 (2): 94-118.

Vosselman, G. and H. G. Maas (2010). "Airborne and terrestrial laser scanning", Whittles Publishing, 318 pages.

Wolf, P. R., and B. A. Dewitt (2000). "Elements of Photogrammetry with Applications in GIS", McGraw-Hill, 607 pages.

Yang, C. and G. Medioni (1992). "Object modelling by registration of multiple range images", *Image and Vision Computing* 10(3): 145-155.

Zitova, B. and J. Flusser (2003). "Image registration methods: a survey", *Image and Vision Computing* 21: 977-1000.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

APPENDIX A CONVERSION CALCULATION MM-TO-PX

APPENDIX A: Conversion calculation mm-to-px

This shows the conceptual working through for the calculation of the pixel to millimetre ratio conversion, when inversed will provide the converse, millimetre to pixel ratio:

$$\frac{\text{mm}_v}{\text{px}_v} = \frac{(\text{vertical sensor length in mm})}{(\text{vertical image length in px})} \quad (\text{A.1})$$

$$\frac{\text{mm}_h}{\text{px}_h} = \frac{(\text{horizontal sensor length in mm})}{(\text{horizontal image length in px})} \quad (\text{A.2})$$

$$x_i (\text{mm}) = \frac{\text{mm}_h}{\text{px}_h} x_i (\text{px}) \quad (\text{A.3})$$

$$y_i (\text{mm}) = \frac{\text{mm}_v}{\text{px}_v} y_i (\text{px}) \quad (\text{A.4})$$

The sensor length can be found from the manufacturer's website or in the device manual, while the image extents can be found via software or as mentioned previously. In most cases, the sensors are square, which means that the calculated ratios are equal for the horizontal image and sensor length, and the vertical image and sensor length.

$$\frac{\text{mm}_v}{\text{px}_v} = \frac{\text{mm}_h}{\text{px}_h} \quad (\text{A.5})$$

APPENDIX B
DERIVATIVE FOR DLT COEFFICIENTS

APPENDIX B: Derivative for DLT coefficients

The following are the derivations for the DLT parameters used in the least squares design matrix, where F_1 and F_2 are the two DLT equations:

$$M = H_1 X + H_2 Y + H_3 Z + H_4 \quad (\text{B.1})$$

$$N = H_5 X + H_6 Y + H_7 Z + H_8 \quad (\text{B.2})$$

$$U = H_9 X + H_{10} Y + H_{11} Z + 1 \quad (\text{B.3})$$

$$F_1 = \frac{M}{U} \quad (\text{B.4})$$

$$F_2 = \frac{N}{U} \quad (\text{B.5})$$

The following are derivatives with respect to the DLT coefficients of Eq. B.4., the function relating to the image coordinates x :

$$\frac{dF_1}{dH_1} = \frac{X}{U} \quad \frac{dF_1}{dH_2} = \frac{Y}{U} \quad \frac{dF_1}{dH_3} = \frac{Z}{U} \quad \frac{dF_1}{dH_4} = \frac{1}{U} \quad (\text{B.6})$$

$$\frac{dF_1}{dH_9} = -x \frac{X}{U} \quad \frac{dF_1}{dH_{10}} = -x \frac{Y}{U} \quad \frac{dF_1}{dH_{11}} = -x \frac{Z}{U} \quad (\text{B.7})$$

The following are derivatives with respect to the DLT coefficients of Eq. B.5., the function relating to the image coordinates y :

$$\frac{dF_2}{dH_5} = \frac{X}{U} \quad \frac{dF_2}{dH_6} = \frac{Y}{U} \quad \frac{dF_2}{dH_7} = \frac{Z}{U} \quad \frac{dF_2}{dH_8} = \frac{1}{U} \quad (\text{B.8})$$

$$\frac{dF_2}{dH_9} = -y \frac{X}{U} \quad \frac{dF_2}{dH_{10}} = -y \frac{Y}{U} \quad \frac{dF_2}{dH_{11}} = -y \frac{Z}{U} \quad (\text{B.9})$$

APPENDIX C
ROTATION MATRIX

APPENDIX C: Rotation Matrix

The orientation of the image space relative to the object space can be determined using the rotation matrix as described in books by Luhmann et al. (2006) and McGlone et al. (2004). The rotation matrix, M , consists of nine direction cosines as a function of orientation three angles (ω , ϕ , κ) as the Cardan angle sequence (Eq. A.4). The calculation is applied as follows:

$$M(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{bmatrix} \quad (C.1)$$

$$M(\phi) = \begin{bmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{bmatrix} \quad (C.2)$$

$$M(\kappa) = \begin{bmatrix} \cos \kappa & \sin \kappa & 0 \\ -\sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (C.3)$$

$$M(\omega, \phi, \kappa) = M(\kappa) * M(\phi) * M(\omega) \quad (C.4)$$

APPENDIX D
CONICAL AND CORNER DATASET CODE


```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear X2 Y2 Z2
control_ID=[1 67 140 210 331 461 555 632 877 754 999 1005];
for i=1:size(control_ID,2)
    Ctrl_X2(i)=all_X2(control_ID(i));
    Ctrl_Y2(i)=all_Y2(control_ID(i));
    Ctrl_Z2(i)=all_Z2(control_ID(i));
    counter(i,1:2)=[i control_ID(i)];
end

XYZ2=[Ctrl_X2;Ctrl_Y2;Ctrl_Z2];
all_XYZ2=[all_X2;all_Y2;all_Z2];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
omega=5;%degrees
phi=8; %degrees
kappa=10;%degrees
mat_R = rotation_matrix(omega,phi,kappa);
% with calibration done there should be 0,0 for Xp,Yp
mat_K=[0.50/0.006 0 0.0
        0 0.50/0.006 0.0
        0 0 1];
mat_Tr=[tr_x tr_y tr_z];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%CREATE IMAGE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Ctrl_img=mat_K*[mat_R mat_Tr']
*[Ctrl_X2;Ctrl_Y2;Ctrl_Z2;ones(1,size(Ctrl_X2,2))];

img02=mat_K*[mat_R mat_Tr']*[all_XYZ2;ones(1,size(all_XYZ2,2))];

for i=1:size(Ctrl_img,2)
    Ctrl_img_norm(1:3,i)=round(Ctrl_img(1:3,i)/Ctrl_img(3,i));
end
for i=1:size(img02,2)
    img_norm02(1:3,i)=round(img02(1:3,i)/img02(3,i));
end

```

CornerData.m

```

clear all

sides=0;
count=0;
box_size=10;

%%% create the box bounded by vertices

    for i=1:box_size;
        for j=1:box_size;
            count=count+1;
            box(count,1)=count;
            box(count,2)=i;
            box(count,3)=j;
            box(count,4)=0;

        end
    end

    for i=1:box_size;
        for j=1:box_size;
            count=count+1;
            box(count,1)=count;
            box(count,2)=i;
            box(count,3)=0;
            box(count,4)=j;

        end
    end

    for i=1:box_size;
        for j=1:box_size;
            count=count+1;
            box(count,1)=count;
            box(count,2)=0;
            box(count,3)=i;
            box(count,4)=j;

        end
    end

%%% translation
box=[box(:,1) box(:,2)+20 box(:,3)+20 box(:,4)+20];

box=box+10;

save data('file');

%%% Creating Synthetic Camera Images from DLT

%%% Load Data
data = load('file');

control_points_ID = data(:,1);
control_points = [data(:,2) data(:,3) data(:,4)];%(m)

```

```
% apply_calibration parameters
% k_matrix
focal_length=0.018;
shear=0;
mx=1/0.000006;%pixels per unit in x (m)
my=1/0.000006;%pixels per unit in y (m)
offset_x=0.00;%principal point offset x (m)
offset_y=0.00;%principal point offset y (m)
K_matrix = k_matrix(focal_length, shear, mx, my, offset_x, offset_y)

% rotation_matrix
omega=311;
phi=34;
kappa=11;
mat_R = rotation_matrix(omega,phi,kappa)

% translation
translation = [0.0 0.0 5.0]';

% k_projection.m
[projection, temp_points, image_points] = k_projection(K_matrix,
mat_R, translation, control_points);
projection
```

APPENDIX E
APPLYING ERROR TO DATASETS

APPENDIX E: Applying Error to Datasets

Matlab code for applying randomised error to the datasets.

XYZ2TLS_w_error.m

```

%%%Initial items
clear all

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
data = load('file');
data2 = load('file');
data3 = load('file');
data4 = load('file');
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

ctID = data(:,1);
ct_X = data(:,2);
ct_Y = data(:,3);
ct_Z = data(:,4);

allID = data3(:,1);
all_X = data3(:,2);
all_Y = data3(:,3);
all_Z = data3(:,4);

%convert to TLS system
for i=1:size(ctID,1)
    %just control points
    X=ct_X(i);
    Y=ct_Y(i);
    Z=ct_Z(i);

    ct_range(i,1)=sqrt(X^2 + Y^2 + Z^2);
    ct_horizontal_direction(i,1)=atan(Y/X);
    ct_elevation_angle(i,1)=atan(Z/sqrt(X^2+Y^2));
end

for i=1:size(allID,1)
    %all points
    X=all_X(i);
    Y=all_Y(i);
    Z=all_Z(i);

    all_range(i,1)=sqrt(X^2+Y^2+Z^2);
    all_horizontal_direction(i,1)=atan(Y/X);
    all_elevation_angle(i,1)=atan(Z/sqrt(X^2+Y^2));
end

%add error
for i=1:size(allID,1)
    %for the range error
    % millimeter error within 0.004 range equal to Leica HDS
    Rerror=random('bino',4,0.5)/1000;

```

```

random_n=random('bino',1,0.5);
    if random_n==1
        Rerror=Rerror*-1;
    end
Eall_range(i,1)=all_range(i,1)+Rerror;
%for the horizontal direction error at 60 urad (micro-radians)
Herror=(random('bino',60,0.5))/1e6; %only for urad
random_n=random('bino',1,0.5);
    if random_n==1
        Herror=Herror*-1;
    end

Eall_horizontal_direction(i,1)=all_horizontal_direction(i,1)+Herror;
%for the elevation angle error at 60 urad
Error=(random('bino',60,0.5))/1e6; %only for urad
random_n=random('bino',1,0.5);
    if random_n==1
        Error=Error*-1;
    end
Eall_elevation_angle(i,1)=all_elevation_angle(i,1)+Error;

end

% Thanks to Reshetyuk, Y. @ 3rd IAG / 12th FIG Symposium, Baden, May
22-24, 2006
for i=1:size(allID,1)
    sign=all_X(i,1)/abs(all_X(i,1));

    all_X_new(i,1)=sign*abs(Eall_range(i,1)*cos(Eall_horizontal_directio
n(i,1))*cos(Eall_elevation_angle(i,1)));
    sign=all_Y(i,1)/abs(all_Y(i,1));

    all_Y_new(i,1)=sign*abs(Eall_range(i,1)*sin(Eall_horizontal_directio
n(i,1))*cos(Eall_elevation_angle(i,1)));
    sign=all_Z(i,1)/abs(all_Z(i,1));

    all_Z_new(i,1)=sign*abs(Eall_range(i,1)*sin(Eall_elevation_angle(i,1
))));
end

```

APPENDIX F
ADDITIONAL INTERIOR ORIENTATION DESCRIPTIONS

APPENDIX F: Additional Interior Orientation Descriptions

Symmetric radial lens distortion:

This is the main constituent of image distortion for most camera systems, can be attributed to variations such as refraction, lens design, focusing distance and object distance at a constant focus. This distortion increases radially with the distance from the principal point. The distortion curve is described with the Seidel polynomial series, having the distortion parameters K_1 to K_n . The equation below includes an additional constant K_0 term, which is applied to remove the correlations with image scale or principal distance. (Luhmann et al., 2006)

$$\Delta r'_{\text{rad}} = K_0 + K_1 r'^3 + K_1 r'^5 + K_1 r'^7 + \dots \quad (\text{F.1})$$

The tangential or decentring lens distortion:

The tangential or decentring lens distortion, based on Luhmann et al. (2006), formally known as radial-asymmetric distortion, is mainly caused by decentring and misalignment of the lens elements. Generally this distortion is minimal for high quality lenses and is much lesser than the radial lens distortion. However, for low-cost lenses that appear mostly for video and surveillance systems, this error may be significant. The following equations are used to compensate for this distortion:

$$\Delta x'_{\text{tan}} = B_1 (r'^2 + 2x'^2) + 2B_2 x' y' \quad (\text{F.2})$$

$$\Delta y'_{\text{tan}} = B_2 (r'^2 + 2y'^2) + 2B_1 x' y' \quad (\text{F.3})$$

The affinity and shear:

The affinity and shear describe the deviations of the image coordinate system, with respect to skew from orthogonal and uniform scale of the coordinate axes (Luhmann et al., 2006). Digital imaging systems are usually affected by this when its sensors are rectangular rather than square. The following functions are used to describe affinity and shear:

$$\Delta x'_{\text{aff}} = C_1 x' + C_2 y' \quad (\text{F.4})$$

$$\Delta y'_{\text{aff}} = 0 \quad (\text{F.5})$$