THE MENSURATION OF TEXTURED SURFACES USING CLOSE-RANGE DIGITAL PHOTOGRAMMETRY FOR APPLICATIONS IN INDUSTRY

Submitted to the University of Cape Town in fulfilment of the requirements for the Degree of Master of Science in Engineering

By
Pierre Geoffrey John Hoffa

Department of Surveying and Geodetic Engineering
April 1996

DECLARATION

I hereby declare that this thesis is my original work and has not been submitted in any form to another university.

Pierre Hoffa

ABSTRACT

The mapping of rock surfaces in the open cast mining industry is of importance in assessing the fragmentation distribution of blasted rock, as well as for the profiling of vertical rock faces. This thesis reports on the development of a close range digital photogrammetric surface measurement system for the semi-automatic measurement of rock faces in open cast mining environments. The system, designed for use by non-photogrammetrists, allows for a rapid and simple field component of the photogrammetric process. This ensures that the system, although relatively costly to set up in terms of camera and computer equipment, is relatively inexpensive to run.

Various photogrammetric and image processing algorithms were investigated and compared. Of particular note is a comparison of a number of techniques for determining approximate exterior orientation parameters of the images of the rock surface. An orientation method based on the collinearity equations, was found to be suitably robust for use in the measurement system. An investigation into the suitability of various interest operators, for extracting points of interest on textured surfaces, was carried out. The Canny edge operator proved to be the most suitable in terms of selecting a large number of well distributed points, that are representative of the surface. Image restitution was carried out by means of a free network bundle adjustment. Multi-photo geometrically constrained matching with multi-image correlation was used to determine the conjugate positions of the extracted interest points. Many of the algorithms described in this thesis can generally be applied to the measurement of any surface containing sufficient texture.

Three test surfaces were successfully mapped. The results indicate that the system is capable of relatively high accuracies. Using a Kodak DCS 420 still video camera to capture the images, the average relative precision in the XY plane was of the order of 1:15000, and for the depth 1:7000. When a low cost, 'off the shelf' ITC CCD video camera was used for image acquisition, the relative precision in the XY plane was approximately 1:4500, and 1:2600 in the depth. The still video camera was found to have significant practical advantages over the CCD video camera.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor Heinz Rüther for providing me with the opportunity to undertake this research project and for all the advice and guidance that he gave me throughout my studies.

I would also like to thank Val Atkinson, Sue Binedell, Michael Haywood, Kari Laatikainen, Dr. Scott Mason and Sidney Smith of the Department of Surveying and Geodetic Engineering for their help with the numerous problems and tasks related to this project.

To my fellow post-graduate students Ulrike Brüssler, Mike Calitz, Mark Cammidge, Dirk Craigie, Mark Matthews, Julian Smit and Graeme van der Vlugt, many thanks for all the discussions and advice. A special acknowledgement is due to Malcolm Dingle for all the assistance that he has given me over the course of this project. A further special thanks to Henty Waker for his assistance in helping to solve computer related problems that frequently occurred.

To my family and friends, thanks for all the support and encouragement over the past few years.

Finally, I would like to acknowledge the Foundation for Research Development for providing the funding that made this project possible.

TABLE OF CONTENTS

ABSTRACT i

ACKNOWLEDGEMENTS ii

LIST OF FIGURES v

LIST OF TABLES vii

1. INTRODUCTION 1

2. LITERATURE REVIEW 3

3. THE FUNDAMENTALS OF PHOTOGRAMMETRY 5
- 3.1 IMAGE GEOMETRY 5
- 3.2 CAMERA CALIBRATION AND ORIENTATION 7
- 3.3 THE DIRECT LINEAR TRANSFORMATION 9
- 3.4 THE COLLINEARITY EQUATIONS 11
- 3.5 THE BUNDLE ADJUSTMENT 11
- 3.6 THE COPLANARITY EQUATION 12
- 3.7 EPIPOLAR GEOMETRY 14
  - 3.7.1 Epipolar Lines Using the Coplanarity Equation 14
  - 3.7.2 Epipolar Lines Using the Collinearity Equations 15
- 3.8 ADDITIONAL PARAMETERS 17

4. DIGITAL IMAGE PROCESSING TECHNIQUES 19
- 4.1 MEASURING OF TARGETS 20
  - 4.1.1 Thresholding 21
  - 4.1.2 Edge Following 22
  - 4.1.3 Weighted Centre of Mass 23
- 4.2 EDGE DETECTION 24
  - 4.2.1 Sobel Edge Operator 25
  - 4.2.2 Canny Edge Operator 26
  - 4.2.3 Moment Preserving Method 27
- 4.3 INTEREST OPERATORS 30
  - 4.3.1 Förstner Interest Operator 30
  - 4.3.2 Moravec Interest Operator 32
  - 4.3.3 Angle Based Interest Operator 33
- 4.4 IMAGE MATCHING 33
- 4.5 MULTI-PHOTO GEOMETRICALLY CONSTRAINED MATCHING 36

5. APPROXIMATE IMAGE ORIENTATION TECHNIQUES 40
- 5.1 THE DIRECT LINEAR TRANSFORMATION 43

5.2 RELATIVE ORIENTATION WITH LINEAR TRANSFORMATION 44
5.3 SMITH'S EXPLICIT SPACE RESECTION SOLUTION 50
5.4 THE COLLINEARITY EQUATIONS 55
5.5 FURTHER METHODS OF CAMERA ORIENTATION 58

6. THE SELECTION OF FEATURES OF INTEREST 59
6.1 CANNY EDGE OPERATOR 60
6.2 FÖRSTNER INTEREST OPERATOR 61
6.3 MORAVEC INTEREST OPERATOR 65
6.4 ANGLE BASED INTEREST OPERATOR 68
6.5 CONCLUSIONS OF CHAPTER 69

7. SYSTEM DESCRIPTION 72
7.1 CHARACTERISTICS OF THE MEASURING ENVIRONMENT 72
7.2 EQUIPMENT 72
7.2.1 Computer Hardware 73
7.2.2 Kodak Still Video CCD Camera 73
7.2.3 ITC CCD Video Camera 73
7.2.4 Additional Equipment 74
7.3 NETWORK DESIGN 74
7.3.1 Camera Configuration 75
7.3.2 Number of Images per Camera Station 75
7.3.3 Control Point Configuration 76
7.3.4 Size and Nature of Targets 76
7.4 SOFTWARE 77
7.4.1 Programming Language 78
7.4.2 Software Structure and Measurement Procedure 78
7.4.3 Downloading of Images 80
7.4.4 Image Reduction 80
7.4.5 Target Measurement and DLT Software 82
7.4.6 Bundle Adjustment Software 84
7.4.7 MPGC Matching Software 84
7.5 CAMERA CALIBRATION PROCEDURE 86

8. SURFACE MEASUREMENT TEST RESULTS 90
8.1 CAPE TOWN HARBOUR EXCAVATION 90
8.2 ROCK WALL 94
8.3 TYGERBERG HILLS QUARRY SITE 97

9. CONCLUSIONS 102

REFERENCES 104

BIBLIOGRAPHY 108

LIST OF FIGURES

3.1	The projection of an object point onto an image plane.	5
3.2	Typical photogrammetric system for (a) calibration (b) measurement.	8
3.3	The geometric configuration of a 2-camera system indicating coplanarity.	13
3.4	The intersection of epipolar lines in a 4-image system.	16
4.1	The metric and pixel image coordinate systems.	19
4.2	Graph of the binary thresholding function.	21
4.3	The chain coding search sequence used for edge following.	23
4.4	Masks for the Sobel edge operator (a) x gradient (b) y gradient.	25
4.5	The Gaussian and the first derivative of the Gaussian.	26
4.6	Parameters associated with the Gaussian convolution mask.	27
4.7	An ideal step edge matched to sample data.	28
4.8	Resampling by means of bilinear interpolation.	29
4.9	Flowchart of the matching procedure developed by Van der Vlugt and Rüther.	37
4.10	The MIC process.	38
5.1	An example of the distribution of targets in an open cast mining environment.	41
5.2	The control (calibration) frame.	43
5.3	The relationship between the object space and image space coordinate systems.	43
5.4	Configuration of control points used in Smith's explicit space resection.	50
6.1	The sequence of the Canny edge detection process. (a) The original test image. (b) The detected binary edges. (c) The detected edges shown as a grey scale, proportional to their strength. (d) The thinned points using a 7 x 7 pixel window.	62
6.2	(a) The original test image. (b) The detected Förstner interest points.	63
6.3	Graphs indicating the effect on the number of pixels selected, n, by varying (a) qlim, (b) c, (c) the window size used to evaluate the centre pixel, (d) σ.	64
6.4	The distribution of the detected Förstner interest points for (a) c=1 and (b) c=2, applied to the test image in Figure 5.2(a).	65
6.5	(a) The original test image. (b) Interest clumps detected using the Moravec interest operator.	66
6.6	The distribution of the detected Moravec interest clumps with the size of the window used to evaluate the interest value at the centre pixel, equal to (a) 3 x 3 pixels and (b) 5 x 5 pixels, applied to the test image.	67
6.7	(a) The original test image. (b) Interest clumps detected using the angle based interest operator.	68
6.8	An aerial image of an urban environment with Förstner interest points overlaid, to illustrate the effect of the detection of non interest points and the non detection of obvious interest points.	70

6.9 A comparison of the distribution of selected pixels using the four feature extractors investigated, based on a portion of the test image. (a) Canny edge operator. (b) Förstner interest operator. (c) Moravec interest operator. (d) Angle based interest operator. These examples were extracted from Figures 5.1(d), 5.2(b), 5.6(a) and 6.7(b) respectively. In (a) and (b) each point represents a single interest pixel. In (c) and (d) interest clumps are shown. 70

7.1 A flowchart of the measurement procedure indicating the relationship between the various software components of the measurement system. The first two vertical branches of the flowchart relate to the process of using either the still video camera or the CCD video camera. The bold print indicates the specific software used in the system. 81

7.2 Two potential targets, a ring target and a solid target, located by means of thresholding, are subject to the target parameter criteria independently. 83

8.1 The network geometry of the Cape Town harbour test surface (dimensions are approximate). 91

8.2 Digital image of the test surface at the harbour excavation site as seen from the central camera station. 92

8.3 An orthographic plot of the harbour excavation test surface generated by the Surfer surface display package. 93

8.4 The network geometry of the rock wall test surface (dimensions are approximate). 94

8.5 Digital image of the rock wall test surface as seen from the central camera station. (The dark area towards the top right hand side of the image is a shadow of a light pole.) 95

8.6 An orthographic plot of the rock wall test surface generated by the Surfer surface display package. 97

8.7 The network geometry of the Tygerberg Hills quarry test surface (dimensions are approximate). 98

8.8 Digital image of the test surface at the Tygerberg Hills quarry site. 99

8.9 An orthographic plot of the Tygerberg Hills quarry test surface generated by the Surfer surface display package. 101

LIST OF TABLES

5.1 Summary of the base vector errors resulting from the RLT with nine simulated images using 18 well distributed points on the control frame. 48

5.2 Summary of the relative rotation angle errors resulting from the RLT with nine simulated images using 18 well distributed points on the control frame. 48

5.3 Summary of the relative rotation angle errors resulting from the RLT with five simulated images of the Cape Town harbour excavation site. 49

5.4 Summary of the perspective centre errors resulting from the Smith method using nine images with four non-coplanar points on the control frame. 53

5.5 Summary of the rotation angle errors resulting from the Smith method using nine images with four non-coplanar points on the control frame. 53

5.6 Summary of the perspective centre errors resulting from the Smith method using nine images with four coplanar points on the control frame. 54

5.7 Summary of the rotation angle errors resulting from the Smith method using nine images with four coplanar points on the control frame. 54

5.8 Summary of the perspective centre errors resulting from the Smith method using nine images with four non-coplanar points on the control frame, where the coordinates have been randomly varied by up to 2,5% of the control frame size. 55

5.9 Summary of the rotation angle errors resulting from the Smith method using nine images with four non-coplanar points on the control frame, where the coordinates have been randomly varied by up to 2,5% of the control frame size. 55

5.10 Summary of the rotation angle errors resulting from the collinearity equation method of image orientation using nine images with three control points on the control frame, where the coordinates have been randomly varied by up to 10% of the control frame size. 57

7.1 Summary results of typical camera calibration adjustments using images acquired with the Kodak still video camera and the ITC CCD video camera. The first improvement factor serves as a comparison between the two calibration sets using the Kodak images. The second improvement factor indicates the improvement of the second calibration set over the third. 88

8.1 Summary results of the free network bundle adjustment used to orient the 5 images of the harbour test surface. 92

8.2 Summary results of the MPGC matching on the harbour test surface. 92

8.3 The approximate relative accuracies in the X, Y and Z directions for the harbour test surface. 93

8.4	Summary results of the free network bundle adjustment used to orient the 5 images of the rock wall test surface.	95
8.5	Summary results of the MPGC matching on the rock wall test surface.	96
8.6	The approximate relative accuracies in the X, Y and Z directions for the rock wall test surface.	96
8.7	Summary results of the free network bundle adjustment used to orient the 5 images of the quarry test surface.	100
8.8	Summary results of the MPGC matching on the quarry test surface.	100
8.9	The approximate relative accuracies in the X, Y and Z directions for the quarry test surface.	100

1. INTRODUCTION

In the last decade photogrammetry has made a transition from the use of analogue images using analytical photogrammetric techniques to the digital domain. This has been made possible by the rapid advancement of digital technology, in particular the development of powerful computers, digital algorithms and imaging devices capable of producing digital images, such as Charge Coupled Device (CCD) video cameras and, more recently, still video CCD cameras. Haralick and Shapiro (1992) define the process of digital photogrammetry as 'the computer processing of perspective projection digital images with analytical photogrammetry techniques as well as other computer techniques for the automatic interpretation of scene or image content'. In this definition the two main components of digital photogrammetry are identified, namely analytical photogrammetry and image processing. Digital photogrammetry has realised the potential for high accuracies within highly automated, and in some cases fully automated, near real-time environments.

This thesis reports on the development of a digital photogrammetric measurement system for the mensuration of textured surfaces for industrial applications. The specific industrial application for which the system was designed is the measurement of rock surfaces in open cast mining environments. Of particular interest to the mining industry is the development of a system that can be used for assessing the fragmentation distribution of blasted rock as well as for other applications such as profiling of vertical rock faces and to a lesser extent, volume determination. Although the system developed for this thesis was designed within the constraints of a working excavation environment, many of the algorithms used here can generally be applied to the measurement of any other surface that contains sufficient surface texture.

The design of the system was based on a number of objectives and constraints:

(i) The amount of time taken for the field component of the photogrammetric process had to be minimised in order to reduce downtime at the excavation site.
(ii) The system should be suitable for use by non-photogrammetrists.
(iii) The system should require as little human interaction as possible (i.e. it must be highly automated).
(iv) Software available within the Department of Surveying and Geodetic Engineering at UCT was to be used where possible. If software did not exist then it was to be developed. All the software to be developed was to be user-friendly.
(v) The system was to be implemented and tested on a number of surfaces.

In this thesis both a still video camera and a CCD video camera were used for image acquisition. (The terms 'still video CCD camera' and 'still video camera' will be used interchangeably in this thesis.) The still video camera belonged to the Chamber of Mines and was on loan to the Department of Surveying and Geodetic Engineering at UCT for a limited time. For the measurements of two of the test surfaces this camera

was used. After this camera had been returned a CCD video camera was used for image capture of the third test surface.

The thesis starts by briefly reviewing the relevant literature. Chapter 3 discusses the theory of analytical photogrammetric techniques that relates to the research undertaken for this thesis. Chapter 4 describes the theory of the image processing techniques that were investigated. In chapter 5 a comparison is made of the methods of determining provisional values for exterior orientation parameters of the images of the surface to be measured. Chapter 6 compares the performance of various feature extractors on a test image. The measurement system that was developed, is described in chapter 7. Aspects such as a description of the measuring environment, the equipment, network design and the software that was used is discussed. Included in this chapter is a section on the method of camera calibration that was used. Chapter 8 presents the results of the measurements of three test surfaces. Finally, in chapter 9, conclusions are drawn.

2. LITERATURE REVIEW

Various close range digital photogrammetric surface measurement systems have been reported in the academic literature over the past few years. Most of these systems have industrial or medical applications where the surfaces to be measured are smooth and featureless. In such cases, it is common for surface definition to be created by placing targets directly onto the surface. Fraser (1995) reports on the successful measurement of deformations on a large coal dredger using 200 circular retroreflective targets to define the surface. The area of deformation under investigation was 10 x 13 m with the centre of this region 25 m above the ground. Cameras were placed on an embankment, 20 m high and at an average distance of 50 m from the dredger. Approximate object space coordinates of the targets were determined by inspecting one of the images taken from the central camera station. Three images from each of seven camera stations were acquired with a Kodak DCS 200 still video camera. A free network bundle adjustment was used for image restitution. Relative precisions of the targets of 1:50000 were obtained.

Fraser and Shortis (1995) describe the measurement of a final assembly jig for the rudder tab of a Boeing 777 aircraft. The jig measured 4.5 x 5 m. 60 circular retroreflective targets were placed on the surface. A Kodak DCS 200 still video camera was used to acquire eight images from each of six camera stations. Image restitution was carried out by means of a free network, self calibrating bundle adjustment. An independent check was made using theodolite observations. Discrepancies between the object space coordinates calculated using the two methods were in the order of 1:80000.

Beyer (1992a) describes the successful implementation of a system to measure the deformation of car bodies in crash tests. The system was tested under factory floor conditions at a motor manufacturing plant. 73 circular targets made of retroreflective material were placed on the surface of a car. Low cost CCD video cameras were used to capture the images. The identification of targets was semi-automatic in that four to six targets on each image had to be manually identified after which the remaining targets were automatically located and measured. The precise determination of the image coordinates of the targets was performed using least squares matching. Coordinate accuracies of less than 1 mm were attained within a measurement volume of 5 x 2 x 2 m^3. The positions of the targets were checked with a theodolite.

An alternative method of defining featureless surfaces is to create surface texture by projecting a pattern or grid onto the surface. Van der Vlugt and Rüther (1994) report on the development of an automated PC-based surface measurement system that uses multi-photo geometrically constrained (MPGC) matching to determine conjugate image points and three dimensional coordinates. Image matching requires sufficient image detail. The system uses either the natural surface texture or if this is insufficient, creates artificial texture by projecting a pattern onto the surface using either a slide or overhead projector. The images are captured using an ITC CCD video camera. Some operator interaction is required for the exterior orientation of the

images. The relative precisions of the surface points were in the order of 1:2000 in the plane of the surface and 1:700 in the depth. A refinement of the system described above is reported in Van der Vlugt (1995) in which significantly better precisions of the surface points were obtained. The relative precisions were approximately 1:5000 in the plane of the surface and 1:1700 in the depth.

Maas (1991) reports on the use of the intersection of epipolar lines to establish multi-image correspondences between tracer particles suspended in water for the purpose of turbulent flow research. Three CCD video cameras were synchronised to capture the images simultaneously with 25 images triples acquired per second. 700 particles were successfully tracked by means of this method. Dold and Maas (1994) employed the same technique of using the intersection of epipolar lines to establish correspondences between discrete points on a masonry wall. The points consisted of 380 randomly placed circular retroreflective targets. 18 images were taken of the 2 x 4 metre test section with a large format Rollei LFC camera. The image coordinates were measured using a digital automatic monocomparator. All of the targets were correctly matched and relative precisions of 1:200000 were obtained.

A highly automated measurement system for the all round mapping of a human body is presented in Wong et al. (1992). Surface texture was created by projecting dot patterns onto the body using three slide projectors. Nine CCD cameras arranged in three stereo triples were used to capture images covering the entire body. The cameras were calibrated and oriented automatically with the aid of circular targets each of which was surrounded by a bar code for automatic identification. The targets were placed on a portable three dimensional control field. After calibration the person to be measured was photographed in six different poses. For each pose a DEM was created using epipolar geometry to establish correspondences between the extracted dots. The measurement of retroreflective targets placed on the person was used to transform the DEM's into a single coordinate system. Accuracies of better than 5% of the measuring volume was achieved. A major limitation of the system, caused by hardware constraints, was that the time required to capture the images for each pose was one minute. Even small movements of the person during this time period would certainly lead to a deterioration of the measurement accuracy.

No reference could be found to a surface measurement system specifically designed for the measurement of rock faces in a working mine environment.

3. THE FUNDAMENTALS OF PHOTOGRAMMETRY

3.1 IMAGE GEOMETRY

The perspective projection is the mathematical model that projects an object in three-dimensional space (object space) onto a two-dimensional plane (image space). Figure 3.1 shows the relationship between point P(X,Y,Z) in object space and its corresponding point p(x,y) on a positive image taken by a single camera.

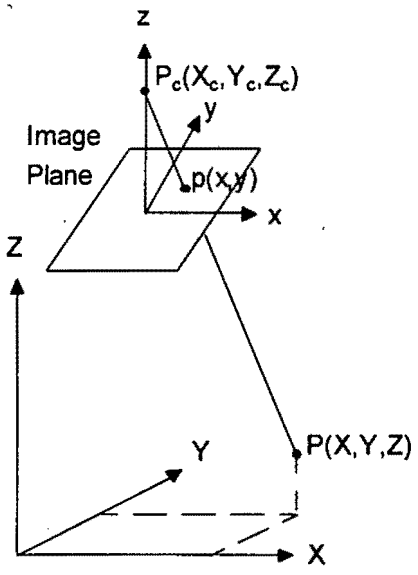


Figure 3.1: The projection of an object space point onto an image plane.

The perspective centre, indicated as $P_c(X_c,Y_c,Z_c)$ and measured in object space, can be defined as the centre of the lens through which light from the object passes in order to be projected onto the image. The image coordinate system x,y,z has its origin at the principal point (x_p,y_p), where this point is defined as the intersection of the perpendicular from the perspective centre to the image plane. The principal point is the projection of the perspective centre onto the image, and is the point relative to which all image coordinates are measured. The perpendicular distance between the perspective centre and the image plane is referred to as the principal distance, and is denoted by the letter c. Often the term 'focal length' is used interchangeably with the principal distance, but this refers to a principal distance of infinity which is often not the case in close-range photogrammetry.

The two fundamental principles of image geometry that need to be satisfied are:

(i) the image points all lie on one plane (image plane);
(ii) any object point, the perspective centre and its image point lie on a straight line.

The latter condition is that of collinearity upon which much of photogrammetry is based. In theory these conditions hold, but in practice factors such as lens distortion, deviation of the image carrier from the plane, refraction distortions and others occur. Effects such as lens distortion can be reduced by introducing appropriate parameters to model corrections to the image coordinates. This will be discussed in more detail in section 3.8.

The transformation from point (x_i,y_i,z_i) in image space, with z_i on the image plane, to its projected position (X_i,Y_i,Z_i) in object space can be modelled by three translations, three rotations and a scale factor. Mathematically this can be represented by:

$$\begin{pmatrix} x_i - x_p \\ y_i - y_p \\ c \end{pmatrix} = sR \begin{pmatrix} X_i - X_c \\ Y_i - Y_c \\ Z_i - Z_c \end{pmatrix} \qquad 3.1$$

where s is a scale factor that relates the relative distances between the two systems and R is an orthogonal rotation matrix that is a function of the rotations of the image plane about the three-dimensional coordinate system in object space. The nine elements of the rotation matrix are functions of the three rotations about the X, Y and Z axes. These rotation angles are necessary to bring the image coordinate system parallel to the object space coordinate system.

For the purpose of this thesis the following rotation angles are used: a rotation ω measured about the X axis, κ measured about the Y axis and ϕ measured about the Z axis. A positive rotation is measured clockwise when seen from the origin of the coordinate system. The orthogonal transformation matrices for each positive rotation about fixed axes in a right handed coordinate system are:

$$R_X(\omega) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{pmatrix} \qquad 3.2$$

$$R_Y(\kappa) = \begin{pmatrix} \cos\kappa & 0 & \sin\kappa \\ 0 & 1 & 0 \\ -\sin\kappa & 0 & \cos\kappa \end{pmatrix} \qquad 3.3$$

$$R_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad 3.4$$

The full rotation matrix R is the product of the three individual rotation matrices R_X, R_Y and R_Z (i.e. $R_{(\omega,\kappa,\phi)} = R_Z(\phi)R_Y(\kappa)R_X(\omega)$). The order in which the matrices are multiplied affects the resultant rotation matrix. For the rotation sequence $R_{(\omega,\kappa,\phi)}$, the elements of the rotation matrix are:

$$\begin{aligned} r_{11} &= \cos\kappa\cos\phi \\ r_{12} &= -\cos\kappa\sin\phi \\ r_{13} &= \sin\kappa \\ r_{21} &= \cos\omega\sin\phi + \sin\omega\sin\kappa\cos\phi \\ r_{22} &= \cos\omega\cos\phi - \sin\omega\sin\kappa\sin\phi \\ r_{23} &= -\sin\omega\cos\kappa \\ r_{31} &= \sin\omega\sin\phi - \cos\omega\sin\kappa\cos\phi \\ r_{32} &= \sin\omega\cos\phi + \cos\omega\sin\kappa\sin\phi \\ r_{33} &= \cos\omega\cos\kappa \end{aligned} \qquad 3.5$$

and

$$R_{(\omega,\kappa,\phi)} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \qquad 3.6$$

3.2 CAMERA CALIBRATION AND ORIENTATION

A fundamental component of the photogrammetric process is that of camera calibration and orientation. This can be performed either prior to the measurement of unknown object space coordinates or as part of a self-calibration process in which the camera calibration and orientation parameters, as well as the object space coordinates of the object to be measured, are determined simultaneously. In order for measurements to take place the relative positions of two or more cameras need to be known. A typical two-camera configuration for the separate photogrammetric calibration and photogrammetric measurement is shown in Figure 3.2.

In Figure 3.2(a) the orientation of the cameras are unknown. The calibration frame contains targets of which the three dimensional positions in object space are known. From this the unknown orientation parameters can be determined by means of a space resection. The reverse, space intersection or measurement process, shown in Figure

3.2(b), uses the known camera orientation parameters to determine the object space coordinates of the object to be measured.

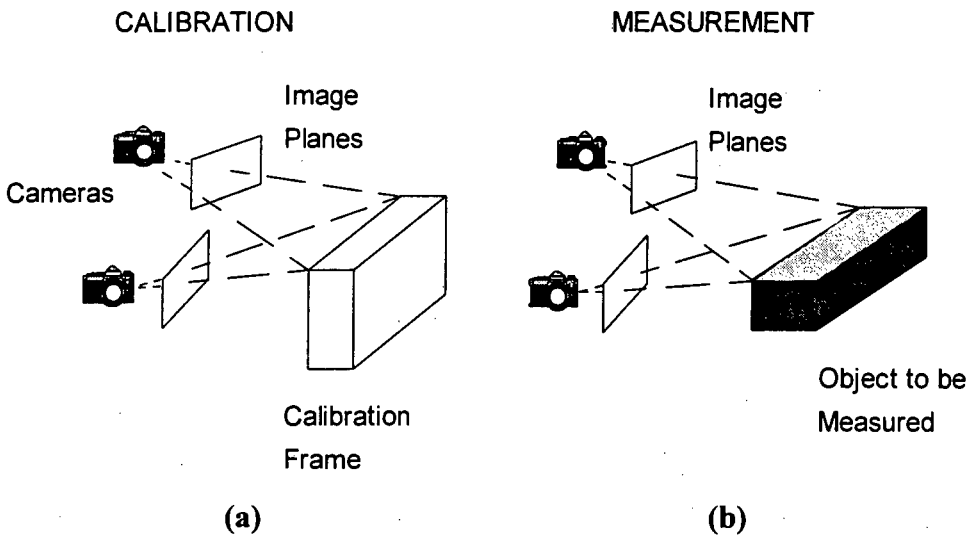


Figure 3.2: Typical photogrammetric system for (a) calibration (b) measurement.

The calibration of a camera system involves three different types of orientation, namely interior, relative and exterior (also referred to as absolute) orientation.

Interior orientation of a camera refers to the relationship between the perspective centre of a single camera and the image coordinate system and is modelled using the following parameters:

(i) the principal distance of the camera;
(ii) the position of the principal point on the image;
(iii)additional parameters (for example lens distortion parameters).

Additional parameters will be discussed in section 3.8.

Relative orientation determines the relationship between two or more images relative to each other. The relative orientation parameters are:

$$
\begin{aligned}
B_X &= X_2 - X_1 \\
B_Y &= Y_2 - Y_1 \\
B_Z &= Z_2 - Z_1 \\
\Delta\omega &= \omega_2 - \omega_1 \\
\Delta\phi &= \phi_2 - \phi_1 \\
\Delta\kappa &= \kappa_2 - \kappa_1
\end{aligned}
\qquad 3.7
$$

where subscripts $_1$ and $_2$ refer to the first and second cameras respectively. The vector B defines the change in position between the two perspective centres. Relative orientation is known if any five elements of relative orientation are known.

Exterior orientation is defined by the position of the perspective centre in three-dimensional object space and the rotations of the image plane of the camera about the X, Y and Z axes. If the exterior orientation of two or more cameras is known, then the relative orientation is also known.

3.3 THE DIRECT LINEAR TRANSFORMATION

The direct linear transformation (DLT) can be used as a direct method of calculating transformation parameters between image space and object space coordinates for a single image. The equations are:

$$x_i = \frac{b_{11}X_i + b_{12}Y_i + b_{13}Z_i + b_{14}}{b_{31}X_i + b_{32}Y_i + b_{33}Z_i + 1} \qquad 3.8$$

$$y_i = \frac{b_{21}X_i + b_{22}Y_i + b_{23}Z_i + b_{24}}{b_{31}X_i + b_{32}Y_i + b_{33}Z_i + 1} \qquad 3.9$$

where x_i, y_i are the image coordinates of point i, X_i, Y_i, Z_i are the corresponding object space coordinates, and the b_{xx} terms are the unknown transformation parameters. This linearly dependent set of equations contains nine linearly independent parameters and two linearly dependent parameters. Ignoring the linear dependence of the parameters allows for an approximation. Less accurate results are thereby attained compared to other traditional models, such as the collinearity equations.

Lens distortion parameters can be incorporated by adding Δx_i and Δy_i terms to the image coordinates in equations 3.8 and 3.9 respectively (see section 3.8).

The primary advantage of using the DLT is that neither a calibrated camera nor any knowledge of the exterior orientation parameters of the camera is required, since no initial approximations for the unknown parameters need be known. A further advantage of the DLT lies in its fast convergence even with poor initial coordinates. The DLT is often used for determining provisional values for the bundle adjustment (see section 3.5).

For the calibration of a camera a minimum of five and a half measured image coordinates relating to six coordinated control points are required to solve for the unknown parameters. A parametric case of the least squares adjustment can be used if redundant information is available. If two or more images are used a space intersection can be performed to measure points in object space. Again, if there are more observations than

unknowns, a parametric least squares adjustment can be used to calculate the three-dimensional coordinates of the points. The minimum solution of the intersection problem requires only three ordinates measured in image space.

A disadvantage of the DLT model is that the control points must be well distributed in three dimensions, preferably in the form of a cube (Fraser, 1992).

The 11 DLT parameters can be converted into the nine linearly dependent physical parameters of interior and exterior orientation. These conversion equations for the rotation sequence given in section 3.1 are (McGlone, 1989):

$$\lambda^2 = 1/(b_{31}^2 + b_{32}^2 + b_{33}^2)$$

$$x_p = \lambda^2(b_{11}b_{31} + b_{12}b_{32} + b_{13}b_{33})$$

$$y_p = \lambda^2(b_{21}b_{31} + b_{22}b_{32} + b_{23}b_{33})$$

$$c_x^2 = \lambda^2(b_{11}^2 + b_{12}^2 + b_{13}^2) - x_p^2$$

$$c_y^2 = \lambda^2(b_{21}^2 + b_{22}^2 + b_{23}^2) - y_p^2$$

$$c = (c_x + c_y)/2$$

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}^{-1} \begin{pmatrix} -b_{14} \\ -b_{24} \\ -1 \end{pmatrix}$$

$$r_{31} = \pm\lambda b_{31}$$

$$r_{32} = \pm\lambda b_{32}$$

$$r_{33} = \pm\lambda b_{33}$$

$$r_{21} = \pm\lambda(b_{13}r_{32} - b_{12}r_{33})/c_x$$

$$r_{22} = \pm\lambda(b_{11}r_{33} - b_{13}r_{31})/c_x$$

$$r_{23} = \pm\lambda(b_{12}r_{31} - b_{11}r_{32})/c_x$$

$$r_{11} = \pm\lambda(b_{22}r_{33} - b_{23}r_{32})/c_y$$

$$r_{12} = \pm\lambda(b_{23}r_{31} - b_{21}r_{33}) / c_y$$

$$r_{13} = \pm\lambda(b_{21}r_{32} - b_{22}r_{31}) / c_y$$

$$\kappa = \sin^{-1}(r_{13})$$

$$\omega = \tan^{-1}(-r_{23} / r_{33})$$

$$\phi = \tan^{-1}(-r_{12} / r_{11}) \qquad 3.10$$

3.4 THE COLLINEARITY EQUATIONS

The collinearity equations make use of the most fundamental principle of photogrammetry i.e. that any object point, the perspective centre and the corresponding image point all lie on the same straight line. They form the basis for most restitution solutions in analytical photogrammetry. From equation 3.1 the collinearity conditions can easily be derived. Mathematically they can be represented by:

$$x_i - x_p = c\frac{r_{11}(X_i - X_c) + r_{12}(Y_i - Y_c) + r_{13}(Z_i - Z_c)}{r_{31}(X_i - X_c) + r_{32}(Y_i - Y_c) + r_{33}(Z_i - Z_c)} \qquad 3.11$$

$$y_i - y_c = c\frac{r_{21}(X_i - X_c) + r_{22}(Y_i - Y_c) + r_{23}(Z_i - Z_c)}{r_{31}(X_i - X_c) + r_{32}(Y_i - Y_c) + r_{33}(Z_i - Z_c)} \qquad 3.12$$

where the r_{xx} terms are the elements of the rotation matrix as indicated in equation 3.5. The nine linearly independent elements of the collinearity equations require that a minimum of four and a half measured image space coordinates, relating to five coordinated non-coplanar control points, are available when solving for the parameters of these equations. When redundant observations are available, as is almost always the case, a parametric least squares adjustment can be carried out to solve for the unknowns.

3.5 THE BUNDLE ADJUSTMENT

The collinearity equations can be used to solve for the camera orientation parameters provided good approximations to the unknowns are available and that sufficient control points are used. Additional parameters, such as those of lens distortion, that cause departures from collinearity can be included. The equations can also be used to solve for unknown object space coordinates provided the interior and exterior orientation parameters are known and that good approximations to the unknown coordinates are available. It is common to combine the two photogrammetric processes by solving for both the orientation parameters and object space coordinates simultaneously in what is

referred to as a bundle adjustment. The bundle adjustment approach has the additional advantage of not requiring a good control point distribution in three dimensions (Fraser, 1992).

When the interior orientation parameters are solved simultaneously with the exterior orientation parameters and the object space coordinates, such an adjustment is referred to as self-calibration. One of the advantages of this method is that a control point field is not required.

For the simultaneous solution of the bundle adjustment the following conditions must be satisfied (Van der Vlugt, 1991):

(i) every point must be visible in at least two images;
(ii) a minimum of five well distributed points in image space must be visible on each image;
(iii) a minimum of two full control points plus an ordinate from a third point must be known in order to provide a datum for the network;
(iv) the number of observation equations must be equal to or exceed the number of unknowns.

In order to compensate for random observation errors redundant observations must be included in the bundle adjustment. This can be done by increasing the number of images used or by increasing the number of control points in the network.

The bundle adjustment is a very powerful and flexible technique of solving the restitution problem and is referred to extensively in this project.

3.6 THE COPLANARITY EQUATION

The coplanarity equation uses the condition that for two images taken of the same point, the object point, the two perspective centres and the two image points all lie on the same plane. This plane is referred to as an epipolar plane. The coplanar principle is demonstrated in Figure 3.3.

The coplanarity equation is useful when the relative orientation of two images is of interest. This has applications for camera calibration and epipolar geometry as described in sections 3.2 and 3.7 respectively.

The equation is described in terms of the base vector B between the two perspective centres, and the image point vectors, a_1 and a_2, of each image. Mathematically these can be represented by:

$$B = \begin{pmatrix} B_X \\ B_Y \\ B_Z \end{pmatrix} = \begin{pmatrix} X_{C2} - X_{C1} \\ Y_{C2} - Y_{C1} \\ Z_{C2} - Z_{C1} \end{pmatrix} \quad 3.13$$

$$a_1 = \begin{pmatrix} u_1 \\ v_1 \\ w_1 \end{pmatrix} = R_1' \begin{pmatrix} x - x_{p1} \\ y - y_{p1} \\ -c_1 \end{pmatrix} \quad 3.14$$

$$a_2 = \begin{pmatrix} u_2 \\ v_2 \\ w_2 \end{pmatrix} = R_2' \begin{pmatrix} x - x_{p2} \\ y - y_{p2} \\ -c_2 \end{pmatrix} \quad 3.15$$

where R_1' and R_2' are transposed rotation matrices for images 1 and 2 respectively.

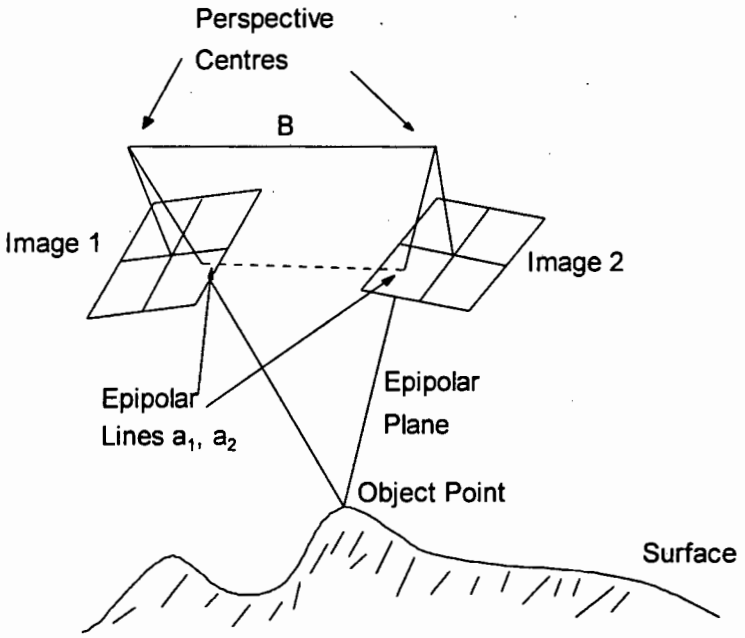


Figure 3.3: The geometric configuration of a 2-camera system indicating coplanarity.

Since the three vectors lie on a plane the volume of the parallelepiped that is formed is equal to zero. This can be written in the form of a determinant:

$$\begin{vmatrix} B_X & B_Y & B_Z \\ u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \end{vmatrix} = 0 \quad 3.16$$

or in an expanded form as:

$$B_X(v_1w_2 - v_2w_1) - B_Y(u_1w_2 - u_2w_1) + B_Z(u_1v_2 - u_2v_1) = 0 \qquad 3.17$$

Assuming that the interior orientation elements are known, there are 12 unknowns in each equation (three translations and three rotations for each image). No approximations for the object points are needed as these coordinates do not appear in the coplanarity equation.

3.7 EPIPOLAR GEOMETRY

If the interior and relative orientation parameters of two images are known, then for a point in image space on the first image, its conjugate point can be calculated to fall along a unique line in the second image. Such a line is referred to as an epipolar line. Epipolar lines are formed by the intersection of the image planes and the plane passing through an object space point and the perspective centres of the cameras as illustrated previously in Figure 3.3. Due to lens distortion the epipolar line will not be a straight line, but will be curved (Dold and Maas, 1994).

Two methods of determining conjugate points with epipolar lines were investigated, namely that using the coplanarity equation and that using the collinearity equations. While the first approach uses the relative orientation parameters, the second approach requires that the exterior orientation of the images be known.

3.7.1 Epipolar Lines Using the Coplanarity Equation

Considering two images whose relative orientation is known, and given the image space coordinates of a point in the first image, a x coordinate in the second image can be arbitrarily chosen. The y coordinate relating to this second x coordinate can be calculated so as to fall on the epipolar line corresponding to the point on the first image. By choosing a second x value on the second image and calculating the resultant y value, an epipolar line can be formed. If stereo pairs are used, the length of the epipolar line can be restricted since the location of the epipolar line in the second image can be estimated. A correlation function can then be used to find the position of the conjugate point in the second image along the epipolar line. Such a method is described by Wong and Ho (1986).

Using the coplanarity equation as given in equations 3.13, 3.14 and 3.15 and with the aid of the expanded form of the coplanarity equation given in equation 3.17, the formulae for calculating the y coordinate in the second image can be readily derived. The result is

given in equation 3.18. For simplicity the elements of the equations for the second image will be denoted by a bar.

$$\bar{y} = \frac{-Qq - Ss - Tt}{\bar{r}_{21}T + \bar{r}_{22}S + \bar{r}_{23}Q} \qquad 3.18$$

where

$$\begin{aligned} Q &= B_X v_1 - B_Y u_1 \\ S &= B_Z u_1 - B_X w_1 \\ T &= B_Y w_1 - B_Z v_1 \end{aligned} \qquad 3.19$$

and

$$\begin{aligned} q &= \bar{r}_{13}(\bar{x} - \bar{x}_p) - \bar{r}_{23}\bar{y}_p - \bar{r}_{33}\bar{c} \\ s &= \bar{r}_{12}(\bar{x} - \bar{x}_p) - \bar{r}_{22}\bar{y}_p - \bar{r}_{32}\bar{c} \\ t &= \bar{r}_{11}(\bar{x} - \bar{x}_p) - \bar{r}_{21}\bar{y}_p - \bar{r}_{31}\bar{c} \end{aligned} \qquad 3.20$$

The r_{xx} elements denote the rotation matrix elements of R_2 as given in equation 3.15.

The correlation function that Wong and Ho used to find the location of the conjugate point along the epipolar line was the normalised cross-correlation function given by:

$$R_{xy} = \frac{mn\sum g_t g_s - \sum g_t \sum g_s}{\sqrt{mn\sum g_t^2 - \left(\sum g_t\right)^2}\sqrt{mn\sum g_s^2 - \left(\sum g_s\right)^2}} \qquad 3.21$$

where g_t and g_s are the grey scale values in the target and search arrays respectively, and m and n are the respective number of rows and columns in the two arrays. The value of R_{xy} ranges from -1 to +1, where +1 represents exact similarity.

3.7.2 Epipolar Lines Using the Collinearity Equations

If the image space coordinates of a point on the first image is known, as well as the camera orientation parameters of both images, the X and Y object space coordinates of the point can be calculated for a given Z, or depth, by means of the collinearity equations. Using these X, Y, Z coordinates, the coordinates of the point in image space on the second image can be calculated. By varying the Z coordinate an epipolar line on the second image can be formed. If the approximate depth range of the object space point is known the length of the epipolar line can be restricted. A certain tolerance can be added to the epipolar line to form a narrow two dimensional window in image space within which a search for the conjugate point can take place. A correlation function, such as that

given by equation 3.21, can be used to evaluate the position of the conjugate point along the epipolar line.

A second method of establishing correspondences using either the coplanarity equation or the collinearity equations is discussed, where the points to be matched comprise many discrete points, as opposed to a continuous surface. This method uses the intersection of epipolar lines to identify conjugate points on multiple images. In such a case the possibility is great that more than one of the discrete points will fall within the narrow search window in image space. Dold and Maas (1994) describe a robust method using four images to reduce ambiguity within the search area. This method is demonstrated in Figure 3.4.

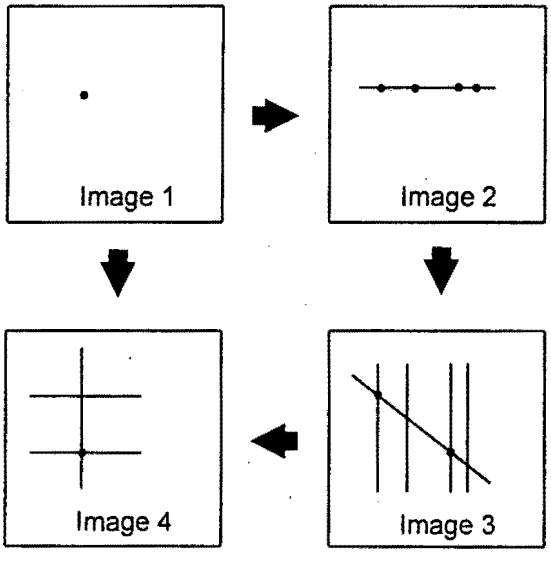


Figure 3.4: The intersection of epipolar lines in a 4-image system.

Referring to Figure 3.4, an epipolar line in image 2 is formed based on a point in image 1. In this example four candidates are found within the search window in image 2. The epipolar lines for these candidates are then projected into image 3. The intersection of these lines with the epipolar line formed using the point from image 1 reduces the number of candidates, for example, to two. Continuing in image 4, the intersection of the appropriate epipolar lines will lead to an unambiguous solution.

The use of epipolar geometry can provide good initial approximations for the positions of conjugate points for subsequent least squares matching as described in sections 4.4 and 4.5.

3.8 ADDITIONAL PARAMETERS

The assumption of precise collinearity that has so far been made is only correct in theory. In practise there are departures from collinearity that have an effect on measured image coordinates. These deviations can be modelled to a large extent by introducing additional parameters into the various adjustment equations such as the collinearity or DLT equations. These additional parameters include lens distortions, atmospheric refraction, and differences in the scaling of the x and y axes on the image. For analytical photogrammetry using analogue film, focal plane distortion can be included as an additional parameter. However, some uncertainties exist about the effects of this distortion when the imaging sensor is a large area CCD array (Fraser and Shortis, 1995). This distortion can occur due to unflatness of the physical chip surface or variations in depth of the light sensitive surface and is likely to limit photogrammetric accuracies significantly as the CCD chip arrays get larger and camera fields of view gets wider. Shortis and Hall (1989) recommend that this effect is best reduced by the careful selection of an appropriate camera.

Lens distortion is a major source of deviation from collinearity and is generally categorised into two main types, namely radial lens distortion and decentring distortion. A commonly used lens distortion model is the Brown model which can be mathematically represented as follows (Beyer, 1992b):

$$\Delta x = \bar{x}(k_1 r^2 + k_2 r^4 + k_3 r^6) + P_1(r^2 + 2\bar{x}^2) + 2P_2 \bar{x}\bar{y}$$

$$\Delta y = \bar{y}(k_1 r^2 + k_2 r^4 + k_3 r^6) + 2P_1 \bar{x}\bar{y} + P_2(r^2 + 2\bar{y}^2) \qquad 3.22$$

where

$$\bar{x} = x - x_p$$

$$\bar{y} = y - y_p$$

$$r^2 = \bar{x}^2 + \bar{y}^2 \qquad 3.23$$

Δx and Δy are the correction terms to image coordinates x and y. r is the radial distance from the principal point. The k_i terms represent the symmetrical radial lens distortion and the P_i terms represent the coefficients of the decentring distortion.

In terms of the radial lens distortion, for low to medium accuracy CCD camera applications, it is usually sufficient to only solve for the k_1 term (Fraser, 1992). For higher accuracy applications it may be necessary to solve for the k_2 and k_3 terms.

Additional parameters relating to the non-perpendicularity and scaling differences of the x and y image axes can be included in the distortion model by introducing the following additional correction terms:

$$\Delta x = \bar{x}s_x + \bar{y}a$$

$$\Delta y = \bar{x}a \qquad 3.24$$

where s_x is the correction parameter to the scale factor in the x direction and a is a shearing parameter. For the purposes of this thesis equations 3.22 and 3.24 were combined to model the deviations from collinearity. The correction terms are corrections to the image coordinates and thus the appropriate terms must be added to the image coordinates of the adjustment equations.

4. DIGITAL IMAGE PROCESSING TECHNIQUES

Digital images are made up of pixels (PICture ELementS). Using eight bits per pixel, colour can be represented as a grey scale value between 0 (black) and 255 (white). The location of each pixel in an image is given by its row and column numbers. Low resolution, low cost, 'off the shelf' CCD cameras that are available in the Department of Surveying and Geodetic Engineering have 512 rows by 512 columns. Digital still video cameras, that are becoming increasingly popular with digital photogrammetrists, make use of high resolution CCD arrays. The Kodak DCS 460, for example, has 2036 rows by 3060 columns.

The origin of the coordinate system in digital images is different to that of conventional metric images used in analytical photogrammetry. These differences are demonstrated in Figure 4.1. One method of transformation between the two coordinate systems has the effect of a translation of the origin, a reflection of the positive vertical axis and a scale change between pixels and millimetres. Separate horizontal and vertical scaling parameters compensate for the fact that the pixels in CCD chips often have different horizontal and vertical scales.

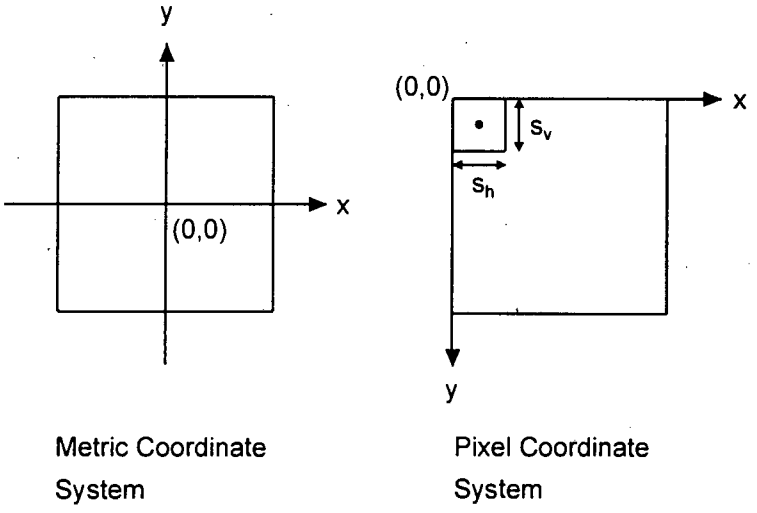


Figure 4.1: The metric and pixel image coordinate systems.

The equations for this transformation are:

$$x_{mm} = s_h\left(x_{pix} - \frac{h}{2}\right) \qquad 4.1$$

$$y_{mm} = s_v\left(\frac{v}{2} - y_{pix}\right) \qquad 4.2$$

where x_{mm} and y_{mm} are in the conventional coordinate system, x_{pix} and y_{pix} are in the pixel coordinate system, h and v are the number of columns and rows in the digital image, and s_h and s_v are the horizontal and vertical scales respectively. Although the manufacturers of CCD cameras provide values for s_h and s_v, these are often unreliable for photogrammetric purposes. Additional parameters can be used to correct for this scaling effect during camera calibration.

A second method of transforming between pixel and metric coordinate systems is that of modelling the transformations by means of an affine transformation, that allows for a rotation and a small shear angle. This can be represented by:

$$\begin{pmatrix} x_{mm} \\ y_{mm} \end{pmatrix} = \begin{pmatrix} -s_h \dfrac{h}{2} \\ s_v \dfrac{v}{2} \end{pmatrix} + \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix} \begin{pmatrix} s_h x_{pix} \\ -s_v y_{pix} \end{pmatrix} \qquad 4.3$$

where the elements x_{mm}, y_{mm}, x_{pix}, y_{pix}, h, v, s_h and s_v are the same as in equations 4.1 and 4.2. β is the small shearing factor, relating to the non-perpendicularity of the x and y axes, and α is the rotation angle. The relative rotation of the coordinate systems is assumed to be zero, with the effect that the rotation matrix becomes an identity matrix. It is possible to account for the shearing factor by solving for it as an additional parameter, as described in section 3.8. If the shearing factor is excluded from equation 4.3, a similar formulation to equations 4.1 and 4.2 is achieved.

For the purpose of this thesis the first method of coordinate transformation was used. The shear factor was applied by including it as an additional parameter.

4.1 MEASURING OF TARGETS

The measurement of image coordinates of targets to sub-pixel accuracy is required in close-range digital photogrammetry. This is necessary for the calibration of the cameras where previously coordinated targets in a control point field need to be measured on the image plane. In many applications the unknown points to be measured on an object are also defined by targets on the surface of the object, as opposed to natural features. These targets are usually circular and are often made of retroreflective material. Although it is generally accepted that target diameters should range between five and ten pixels (Beyer, 1992b), target diameters of between three and 15 pixels are not uncommon. Due to the perspective of the image, the targets are nearly always distorted and appear elliptical on the image. The target centring algorithms used must allow for this distortion.

An efficient method for the measurement of circular targets on the image plane is described. This method takes place in three stages:

(i) target detection, for example, by using a thresholding technique;
(ii) determining the extent of the target, for example, by using an edge following algorithm;
(iii) determining the sub-pixel centre of the target, for example, by using the weighted centre of mass algorithm.

4.1.1 Thresholding

Target detection can easily be performed by means of thresholding the digital image. Targets that are made of retroreflective material and carefully illuminated will typically have grey scale values close to 255 (i.e. white). By assigning all pixels with a grey scale value of below a pre-determined threshold value to zero, the image can be thresholded and the targets can be identified as any area with grey scale values exceeding the threshold. A variation of this method is to assign all pixels with values above the threshold to 255, thereby creating a binary image. This concept is illustrated in Figure 4.2. The success of thresholding in the target detection process is influenced greatly by image content. If the background contains areas close to 255, these areas often cannot be thresholded out without degrading the quality of the real targets. Some methods of removing such non-targets are discussed in section 7.4.5.

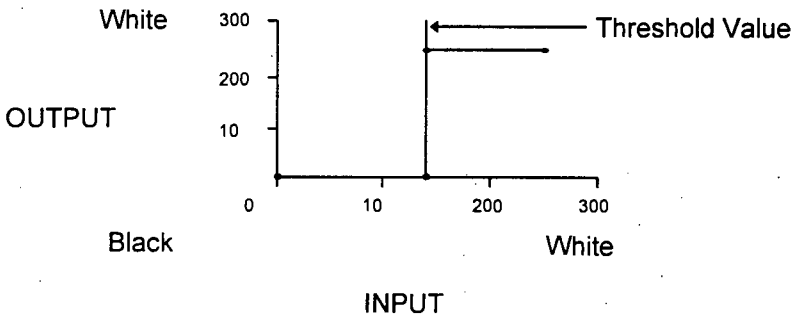


Figure 4.2: Graph of the binary thresholding function.

The selection of the threshold value is easily done interactively by the user while viewing the thresholded image on a computer screen. An alternative method is the automatic selection of a threshold value. Wong and Ho (1986) use the equation:

$$t = \frac{g_{\min} + g_{mean}}{2} \qquad 4.4$$

where t is the threshold value, g_{min} is the minimum grey scale value, and g_{mean} is the mean grey scale value in a selected window. This equation is best applied to a small window surrounding the target. Its use over the entire image does not generally give good results due to the large variations in the minimum and maximum grey scale values. With a small window the background pixel values are typically significantly higher than the minimum grey scale value taken over the whole image. Another method for the automatic selection of a threshold value is given by:

$$t = g_{min} + \frac{g_{max} - g_{min}}{4} \qquad 4.5$$

where t and g_{min} are the same as for equation 4.4, and g_{max} is the maximum grey scale value. This method is also best applied to a small window around the target, although it is less dependent on the window size. Another threshold method, developed by Otsu (1979), determines a threshold value for the full image.

4.1.2 Edge Following

In order to create an optimal rectangular window around the detected targets required for the centre of mass algorithm, the target extent in both vertical and horizontal directions need to be found. One method of calculating the target extent is that of edge following. Using the thresholded (or binary) image, the edge of every detected target is tracked.

One commonly used method of edge following is that of chain coding as described by Rubinstein (1990). This algorithm uses a defined sequence indicated by the next rotation ring, shown in Figure 4.3, to search for neighbouring edge pixels (edgels). Following thresholding the first edgel located on a detected target becomes the current edgel. A sequential clockwise search is carried out starting in the direction where the first pixel on the target is found, typically in direction 3, until the next edgel is located. This edgel now becomes the current edgel. The search for the next edgel starts in the direction indicated by the next rotation ring in Figure 4.3. For example, if the last edgel was found in direction 3, the search for the next edgel will start in direction 2, moving sequentially clockwise. For every new edgel located the minimum and maximum horizontal and vertical extents of the target can be re-evaluated. The algorithm terminates when the search returns to the starting edgel.

The chain coding algorithm searches each pixel outside of the thresholded area only once and the perimeter is always found in a clockwise direction. The search routine is robust and complex edges can be tracked. With the Freeman chain coding method, the search always starts in the direction in which the last edgel was found. This method can fail if the target has a complex shape.

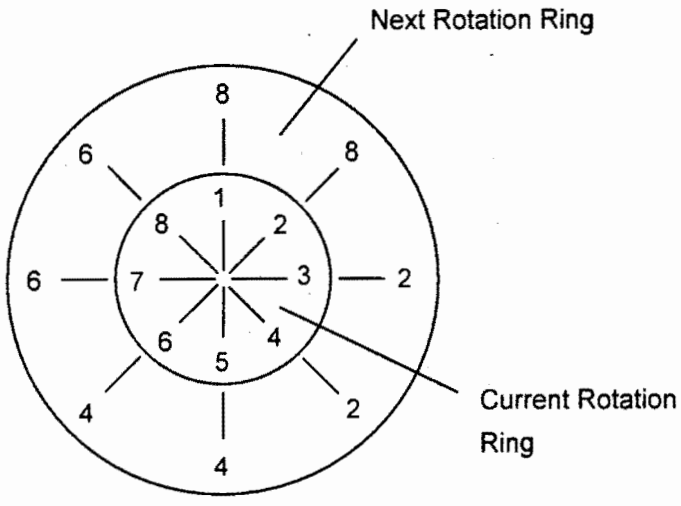


Figure 4.3: The chain coding search sequence used for edge following.

The extent of the targets can also be used to eliminate detected non-target areas whose geometry does not match a template target. For example, detected targets whose extents in x and y are not within a certain tolerance can be rejected. Further discussion on the elimination of targets can be found in section 7.4.5.

4.1.3 Weighted Centre of Mass

The target centring algorithms can be classified in terms of area and edge based algorithms. Area based algorithms are based on the weighted centre of mass. Edge based algorithms are evaluated in a two step process. Firstly, by finding the edge around the target, and secondly by finding the centre of the target based on the edge information. The centring algorithm can take the form of intersecting diameters, best fitting ellipses and circles, and the use of the modified Hough transform to detect partially obscured targets.

Rubinstein and Rüther (1991) identified a modified weighted centre of mass algorithm as a good method of target measurement. The equations for this algorithm are:

$$x = \frac{\sum_i \sum_j g_{ij} . i}{\sum_i \sum_j g_{ij}} \qquad 4.6$$

$$y = \frac{\sum_i \sum_j g_{ij} \cdot j}{\sum_i \sum_j g_{ij}} \qquad 4.7$$

where $g_{ij} = e_{ij}$ for e_{ij} >= Threshold
$g_{ij} = 0$ for e_{ij} < Threshold
e_{ij} is the grey scale value at pixel location (i,j)

This algorithm is applied to a window, a few pixels larger than the target size, placed around each target. For targets in real images, the background pixels with non-zero grey scale values were found to have large effects on the accuracy of the centre of mass algorithm unless the modification of assigning all background pixels to zero was applied. This modification provides an effective solution to the problem. It allows for a translation of the threshold level to zero while retaining the information about the target. Of all the methods evaluated by Rubinstein and Rüther, this method proved to have the most precise target centring capability (based on synthetic targets) as well as being easy to code and having the fastest execution time.

Rubinstein and Rüther go on to note that the choice of threshold is critical for the accurate performance of the algorithm. The threshold value should be chosen 'just above' the background level in the target window. For an optimal threshold, on synthetic images, the accuracy of the centre of between 1/200 and 1/1000 th of a pixel can be attained.

Another high accuracy method of obtaining the centres of circular targets is that of least squares template matching (Gruen, 1985). See section 4.4 for a discussion about area based least squares matching.

4.2 EDGE DETECTION

Edge detection serves as a process 'to simplify the analysis of images by drastically reducing the amount of data to be processed, while at the same time preserving useful structural information about object boundaries' (Canny, 1986). Edges can be seen as distinguishing features of an object and consequently can be used to define the shape of textured surfaces. On this basis, the number of points to be mapped can be significantly reduced while retaining the general shape of the surface. Edge detectors are good low level feature extractors.

Abrupt grey scale value changes indicate the location of an edge in an image. Two classes of edge detectors exist, namely gradient operators and second derivative operators (Huertas and Medioni, 1986). Gradient operators detect a broad edge, a number of pixels wide, which requires the non-maxima suppression of the edge in order to increase its resolution. Second derivative operators use the zero-crossing at the edge location. Edge detectors can also be categorised in terms of pixel and sub-pixel accuracy. In this thesis

only gradient operators were investigated due to their relative simplicity and effectiveness. The Sobel and Canny edge detectors to single pixel accuracy, as well as the moment preserving method of sub-pixel edge detection were investigated.

4.2.1 Sobel Edge Operator

The Sobel operator is a low level edge detector that convolves two 3 x 3 pixel masks (also referred to as kernels) for the x and y gradients. These masks are shown in Figure 4.4.

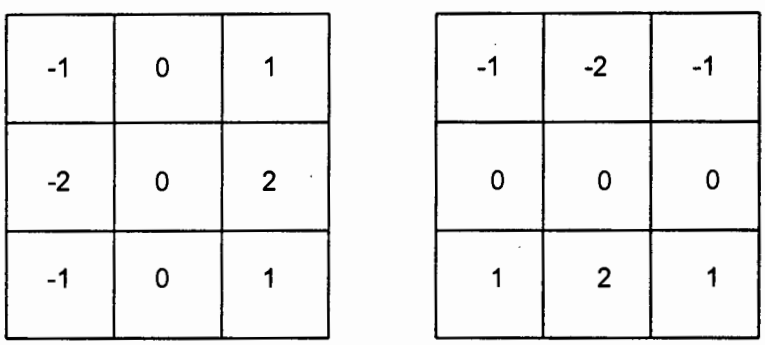


Figure 4.4: Masks for the Sobel edge operator (a) x gradient (b) y gradient.

Mathematically the convolution in one direction can be expressed in the form:

$$G(x_0, y_0) = \sum_{j=-1}^{1}\sum_{i=-1}^{1} M(i,j)g(x_0 + i, y_0 + j) \qquad 4.8$$

where $G(x_0,y_0)$ is the gradient in either the x or y directions, M(i,j) is the Sobel value at point (i,j), and g(x,y) is the grey scale value at (i,j).

Working separately with the edges found in the x and y directions, broad bands of edges, a number of pixels wide, are located using a threshold, whereby all pixels with convolution values exceeding the threshold are classified as edges. Non-maxima suppression of edges is now needed to locate the edges to single pixel accuracy. This can be achieved using a search routine to locate the pixel with the maximum convolution value within the edge band. The gradient magnitude and direction can be calculated as a function of the two Sobel convolution values, calculated in the two directions.

The Sobel operator is easy to implement and computationally inexpensive. However, it is a low pass filter with which noise is reduced at the expense of some reduction of the signal. Low pass filters that incorporate a larger convolution region tend to have smaller

effects on signal reduction but are computation intensive as in the case of the Canny operator.

4.2.2 Canny Edge Operator

When computation time is not a significant factor, such as in non near-real time photogrammetry, the Canny edge detector performs well as a low pass edge operator to single pixel accuracy. Canny (1986) describes the creation and performance of an optimal operator. He shows that an efficient approximation to this optimal filter is that of the first derivative of the Gaussian since this function can be computed with much less computational effort than his optimal filter. For real images the difference in performance of these two operators is hard to detect.

The Canny edge operator is based on a linear gradient of the signal (pixels on the image) with Gaussian smoothing to give the low pass filtering characteristic. The Gaussian function in one dimension is given by:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-x^2}{2\sigma^2}\right) \qquad 4.9$$

The first derivative of the Gaussian is:

$$\frac{\partial G}{\partial x}(x) = \frac{-x}{\sigma^2} G(x) \qquad 4.10$$

Figure 4.5 indicates the relationship between the Gaussian function and its first derivative.

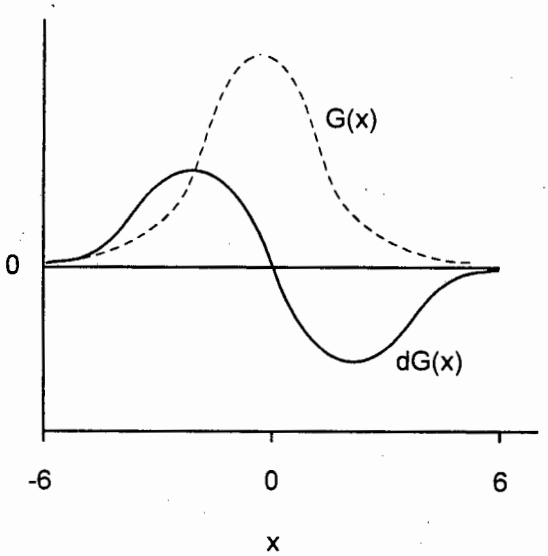


Figure 4.5: The Gaussian and the first derivative of the Gaussian.

The level of smoothing and size of the linear convolution mask is determined by the space constant σ of the Gaussian function. Pixels further away from the pixel of interest have less effect on the convolution value. The width of the central excitatory region of the operator is given in equation 4.11 and is shown in Figure 4.6. Huertas and Medioni (1986) suggest a mask size of 3w relating to 99.7% of the area under the one dimensional Gaussian.

$$w = 2\sqrt{2}\sigma \qquad 4.11$$

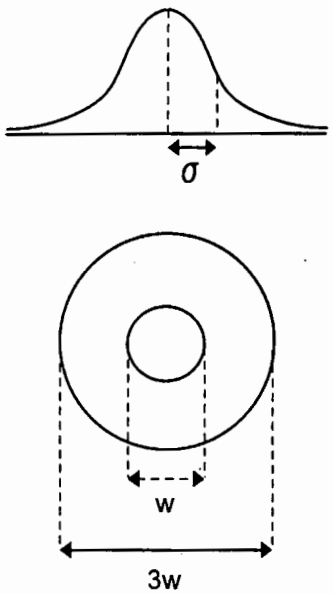


Figure 4.6: Parameters associated with the Gaussian convolution mask.

The convolution of the first derivative of the Gaussian is carried out separately in the x and y directions. A procedure similar to that described in section 4.2.1 for the location of the edges to single pixel accuracy, using a threshold value to detect broad bands of edges and a subsequent non-maxima suppression technique, can be implemented. This provides edges based on the convolution values in the x and y directions. A convolution magnitude value and the gradient direction for each edgel is a function of the values in the x and y directions.

4.2.3 Moment Preserving Method

Tabatabai and Mitchell (1984) describe a moment preserving method of edge detection to sub-pixel accuracy. For a one dimensional sequence of pixels making up a step edge an

operator is defined such that upon application of this step edge the first three sample moments are preserved. The sample moments are defined as:

$$m_i = \frac{1}{n}\sum_{j=1}^{n} x_j^i = \sum_{j=1}^{2} p_j h_j^i \qquad \text{i} = 1,2,3 \tag{4.12}$$

where

$$p_1 = \frac{k}{n} \tag{4.13}$$

and

$$p_1 + p_2 = 1 \tag{4.14}$$

h_1 and h_2 denote the two grey scale levels of the step edge and k denotes the number of h_1 values in the step edge as depicted in Figure 4.7. n is the number of pixels in the sample. The unknown value k defines the location of the edge.

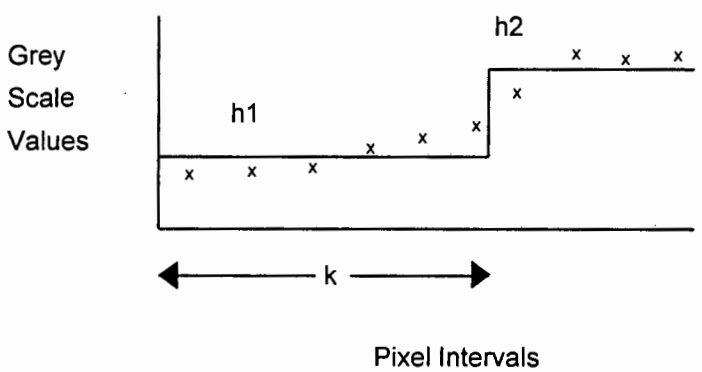


Figure 4.7: An ideal step edge matched to sample data.

Preserving the three moments is equivalent to solving for equation 4.12. The solution to these equations is given by:

$$h_1 = m_1 - \sigma\sqrt{\frac{p_2}{p_1}} \tag{4.15}$$

$$h_2 = m_1 + \sigma\sqrt{\frac{p_1}{p_2}} \tag{4.16}$$

$$p_1 = \frac{1}{2}\left(1 + s\sqrt{\frac{1}{4+s^2}}\right) \qquad 4.17$$

where

$$s = \frac{m_3 + 2m_1^3 - 3m_1 m_2}{\sigma^3} \qquad 4.18$$

$$\sigma^2 = m_2 - m_1^2 \qquad 4.19$$

It can be seen that the location of the edge, $k = np_1$ may be a non-integer (i.e. sub-pixel).

Any single pixel accuracy edge detector can be used to determine the approximate position of the edge. The method described above detects sub-pixel edges using a one dimensional sequence of pixels. On a two dimensional image a one dimensional sequence of pixels must be resampled at one pixel intervals in the direction of the maximum grey scale gradient (i.e. orthogonal to the edge). The maximum grey scale gradient can be determined from convolution processes in the x and y directions at the single pixel accuracy edge. A common method of resampling is that of bilinear interpolation as shown in Figure 4.8.

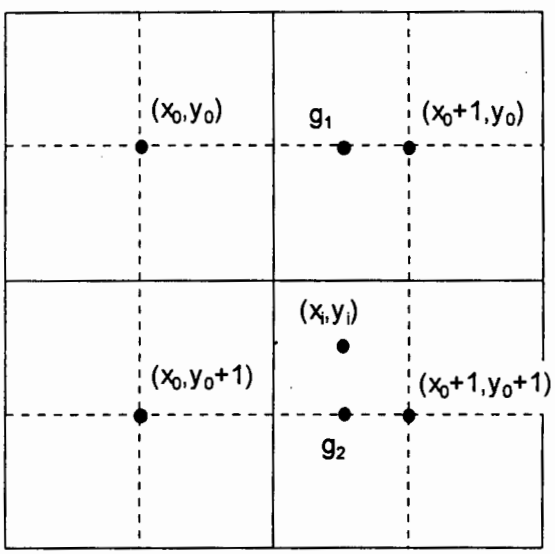


Figure 4.8: Resampling by means of bilinear interpolation.

Referring to Figure 4.8, let $g(x_i,y_i)$ represent the grey scale value at point (i,j). $g(x_i,y_i)$, at non-integer x,y pixel coordinates is resampled based on the four surrounding pixels at x,y integer values. The bilinear interpolation equations are given by:

$$g_1 = g(x_0, y_0) + \big(g(x_0 + 1, y_0) - g(x_0, y_0)\big)(x_i - x_0) \quad 4.20$$

$$g_2 = g(x_0, y_0 + 1) + \big(g(x_0 + 1, y_0 + 1) - g(x_0, y_0 + 1)\big)(x_i - x_0) \quad 4.21$$

$$g(x_i, y_i) = g_1 + (g_2 - g_1)(y_i - y_0) \quad 4.22$$

Alternatively these bilinear interpolation equations can be reformulated such that g_1 and g_2 are calculated along the y axis between (x_0,y_0) and (x_0,y_0+1), and (x_0+1,y_0) and (x_0+1,y_0+1) respectively. $g(x_i,y_i)$ can then be calculated based on these new values for g_1 and g_2.

4.3 INTEREST OPERATORS

A second class of feature extractors is that of interest operators. As with edge operators, interest operators serve to identify points of interest thereby reducing the amount of data to be processed. An objective of any feature extractor should be to identify points that allow for the good reconstruction of the object surface. Three different types of interest operators were investigated. Firstly, the intersection of straight lines, by means of the Förstner operator. Secondly, regions with large scale variations in their grey scale values, such as with the Moravec operator. A third, angle based interest operator was also considered. A comparison of the performance of these operators can be found in chapter 6.

4.3.1 Förstner Interest Operator

The operator as described by Förstner and Gülch (1987) is designed to detect distinct points, corners and centres of circular features. The detection of corners (the intersection of straight lines) is considered here.

Two main steps are required for feature extraction with the Förstner operator:

(i) determining optimal windows that contain a point or feature of interest;
(ii) solving for the position of the interest pixel within the selected windows.

When selecting optimal windows two conditions must be satisfied:

(i) the error ellipse should be small;
(ii) the error ellipse should be near-isotropic.

Selecting the optimal windows requires the formation of a N matrix which is dependent on the feature type to be located. For corner location this matrix is given by:

$$N = \begin{pmatrix} \sum g_r^2 & \sum g_r g_c \\ \sum g_r g_c & \sum g_c^2 \end{pmatrix} \qquad 4.23$$

where g_r and g_c are row and column convolution values calculated using a Gaussian filter as described in section 4.2.2. Other types of filters, such as approximations to the Gaussian or box filters, can also be used. The convolutions are applied to the whole image. For every pixel on the image, N is formed. Förstner and Gülch recommend using a N matrix window size of 5 x 5 or 7 x 7 pixels. The first condition above relating to the size of the error ellipse can be described in terms of the weight of the point which is defined as:

$$w = \frac{1}{trace(N^{-1})} = \frac{\det N}{traceN} \qquad 4.24$$

The second condition relating to the form or roundness of the error ellipse can be measured by the value:

$$q = \frac{4 \det N}{trace^2 N} \qquad 4.25$$

The determination of the interest value, being the preliminary weight for each window position, is given by:

$$\begin{aligned} &\text{if } q(r,c) > q_{\text{lim}} \text{ and } w(r,c) > w_{\text{lim}} \text{ then } w^*(r,c) = w(r,c) \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{else } w^*(r,c) = 0 \end{aligned} \qquad 4.26$$

Förstner and Gülch suggest a value for q_{lim} of between 0.5 and 0.75 corresponding to the ratios 2 and $\sqrt{3}$ of the semi-minor and semi-major axes of the error ellipse. The threshold value w_{lim} suppresses windows containing areas with low grey scale variations. A recommend value is based on:

$$w_{\text{lim}} = c.w_{med} \qquad 4.27$$

where the value of the critical value c, is 5 and w_{med} is the median of the weights taken over the whole image. The calculation of w_{med} is a limitation of this method due to slow median location algorithms. A less rigorous approach is to take the mean of the weight. This threshold, however, has the disadvantage of unpredictability due to its dependence on edge content, sharpness of the edges and noise level in the image.

Once the interest values have been thresholded and the interest pixels located, local non-maxima pixels are suppressed by setting the function w*(r,c) to zero at local non-maxima's. Regions of interest points prior to non-maxima suppression are commonly

referred to as interest clumps. A window size for the suppression must be selected. A suppression algorithm that can be implemented is to sequentially inspect pixels in the suppression window around the pixel of interest. If the interest value of the pixel of interest is less than any other interest value within the window, w*(r,c) is assigned to zero. Otherwise the next interest pixel is inspected. To achieve a stronger separation of interest points, the window size can be increased. All pixels for which w*(r,c) is not zero are extracted.

If sub-pixel interest points are required these can be calculated by solving for r_0 and c_0 using the normal equation:

$$\begin{pmatrix} \sum g_r^2 & \sum g_r g_c \\ \sum g_r g_c & \sum g_c^2 \end{pmatrix} \begin{pmatrix} r_0 \\ c_0 \end{pmatrix} = \begin{pmatrix} \sum g_r^2 r + \sum g_r g_c c \\ \sum g_r g_c r + \sum g_c^2 c \end{pmatrix} \qquad 4.28$$

Similar equations can be formulated for the sub-pixel detection of points and centres of circular features.

4.3.2 Moravec Interest Operator

The Moravec interest operator identifies regions on the image with large grey scale variations. Fuller and Ehlers[1] describe it as being suitable for digital photogrammetric image processing tasks in terms of its speed and accuracy. The Moravec operator identifies high contrast interest clumps by calculating the sum of the squares of grey scale value differences in the four principal directions within a given window. The interest value, IV, at the centre of the window, is the minimum value of the four squared sums and is given by:

$$IV = \min \begin{pmatrix} \sum \left(g_{i,j} - g_{i,j+1} \right)^2 \\ \sum \left(g_{i,j} - g_{i+1,j} \right)^2 \\ \sum \left(g_{i,j} - g_{i+1,j+1} \right)^2 \\ \sum \left(g_{i,j} - g_{i+1,j-1} \right)^2 \end{pmatrix} \qquad 4.29$$

where $g_{i,j}$ is the grey scale value at point (i,j).

As with the Förstner operator, local non-maxima suppression of the interest points needs to follow, as described in section 4.3.1.

---

[1] The reference source for this paper is unknown to the author of this thesis. An attempt to locate the name and date of the publication in which this paper appeared was unsuccessful.

4.3.3 Angle Based Interest Operator

Calitz (1995) describes an angle based interest operator originally designed to extract corners of buildings in aerial images. The operator is based on the comparison of gradients of neighbouring edgels. An edge operator is passed over the image and a binary edge image is formed. At every edgel location a 3 x 3 pixel window is placed on the original image and the gradients in the x and y directions are calculated using the Sobel masks as described in section 4.2.1.

The angle of the gradient with respect to the x axis is given by:

$$\alpha(i,j) = \arctan\left(G_y(i,j) / G_x(i,j)\right) \qquad 4.30$$

where $G_x(i,j)$ and $G_y(i,j)$ are the Sobel gradients in the x and y directions respectively at point (i,j). The angle $\alpha(i,j)$ is calculated for every pixel in the window and compared to the angle calculated at the central pixel. A counter c, which is initialised to zero, is incremented according to the condition:

$$\begin{array}{lll} \text{if } |\alpha(i,j) - \alpha(1,1)| < \theta & \text{then} & c = c + 1 \\ \text{or} \qquad\qquad\qquad \geq \theta & \text{then} & c = c \end{array} \qquad 4.31$$

where θ is a preselected threshold. $\alpha(1,1)$ is the angle at the central pixel in the window and $\alpha(i,j)$ is the angle at pixel location (i,j) within the 3 x 3 pixel window. Interest points are selected as those where the counter is less than or equal to four.

The principle of this operator is that edgels with significant changes in gradient direction are selected as interest points. The threshold θ determines the number of points selected. As θ tends to zero the greater the number of edgels selected. If $\theta = \pi/2$ only a few interest points are identified. According to Calitz, for aerial images of buildings, this operator selects a better distribution of points than the Förstner or Moravec operators. A disadvantage of the angle based operator is that it produces interest clumps as a final result. The suppression of local non-maxima interest pixels is not possible since no interest values are determined.

4.4 IMAGE MATCHING

For photogrammetric object restitution, conjugate image space coordinates for every object space point to be measured must be found in at least two images. This requires that an image matching technique be used to identify conjugate points in the images. A powerful and well recognised sub-pixel method commonly employed is that of area-based least squares matching (Gruen, 1985). This technique establishes correspondences between points by minimising grey scale differences between a reference patch and a corresponding region of the same size in a second, search image. A patch is a small

region of an image, normally square in shape. The reference patch is matched to its corresponding search patch with the aid of an affine transformation. Alternatively a template could replace the reference patch. In this section, matching on only two digital images will be considered. This method can, however, be extended to multi-image matching.

The reference and search patches are given as discrete two dimensional functions f(x,y) and g(x,y). Correlation is established if :

$$f(x,y) = g(x,y) \tag{4.32}$$

However, due to the effects of random noise, radiometric differences, perspective distortions and other factors, a true error function e(x,y) is added:

$$f(x,y) - e(x,y) = g(x,y) \tag{4.33}$$

The objective is to find the location of g(x,y) such that equation 4.33 is satisfied. This equation can be considered a non-linear observation equation which models the observation vector f(x,y) with a function g(x,y) whose location needs to be estimated. This can be determined using the parametric case of the least squares adjustment. The location is described by shift parameters Δx and Δy, which are taken with respect to an initial position at $g^0(x,y)$. To allow for systematic image deformations, caused by different perspectives of the images, image shaping parameters are introduced. If the grey scale values are considered to be on a grid, image shaping can be performed by resampling $g^0(x,y)$ over the transformed grid. The affine transformation can be used for image shaping. The equations for the affine transformation are given by:

$$x = a_{11} + a_{12}x_0 + a_{21}y_0$$

$$y = b_{11} + b_{12}x_0 + b_{21}y_0 \tag{4.34}$$

The shift parameters Δx and Δy, are denoted by a_{11} and b_{11}. Also included are two scale factors, one rotation and a small shearing angle.

A radiometric shift correction can also be included to account for different lighting conditions under which the two images were taken. The shift parameter, r_s, can be represented by:

$$r_s = \left(\sum (g(x,y) - f(x,y))\right) / n \tag{4.35}$$

where n is the total number of pixels in the patches.

Substituting equation 4.34 into equation 4.33 and linearising using a Taylor series expansion, yields:

$$f(x,y) - e(x,y) = g^0(x,y) + \frac{\partial g^0(x,y)}{\partial x} dx + \frac{\partial g^0(x,y)}{\partial y} dy \qquad 4.36$$

where

$$dx = da_{11} + x_0 da_{12} + y_0 da_{21}$$

$$dy = db_{11} + x_0 db_{12} + y_0 db_{21} \qquad 4.37$$

Simplifying and adding the radiometric shift parameter results in:

$$f(x,y) - e(x,y) = g^0(x,y) + G_x da_{11} + G_x x_0 da_{12} + G_x y_0 da_{21} + \\ + G_y db_{11} + G_y x_0 db_{12} + G_y y_0 db_{21} + r_s \qquad 4.38$$

where

$$G_x = \frac{\partial g^0(x,y)}{\partial x} = \frac{g(x+1,y) - g(x-1,y)}{2}$$

$$G_y = \frac{\partial g^0(x,y)}{\partial y} = \frac{g(x,y+1) - g(x,y-1)}{2} \qquad 4.39$$

G_x and G_y represent the across pixel grey scale gradients at each point in the patch in the x and y directions respectively.

Equation 4.38 can be solved by means of the parametric case of the least squares adjustment of the form:

$$x = (A^T PA)^{-1} A^T P\ell \qquad 4.40$$

where the solution vector x is:

$$x^T = \left(da_{11}, da_{12}, da_{21}, db_{11}, db_{12}, db_{21}, r_s\right) \qquad 4.41$$

and the vector difference ℓ is:

$$\ell = f(x,y) - g^0(x,y) \qquad 4.42$$

Equation 4.33 is non-linear and thus requires an iterative solution. The unknown parameters for the first approximations can be set to:

$$a_{11}^0 = b_{11}^0 = a_{21}^0 = b_{12}^0 = 0$$
$$a_{12}^0 = b_{21}^0 = 1$$

Since the radiometric shift is linear a priori, r_s can be calculated prior to the adjustment by including equation 4.35 in equation 4.42.

4.5 MULTI-PHOTO GEOMETRICALLY CONSTRAINED MATCHING

Image matching can be divided into two stages. The first stage is that of determining the approximate position of the search patch in the search image. In the second stage this approximate position is used as a provisional value in a fine matching process in which the accurate position of the search patch is calculated. Given good interior and relative orientation parameters, fine matching can be accomplished using multi-photo geometrically constrained (MPGC) matching proposed by Gruen (1985) and thoroughly investigated by Baltsavias (1991). This powerful technique is a refinement of the method of image matching described in section 4.4. MPGC matching exploits the known camera orientation parameters to limit the extent of the search in multiple search images. If two images are used the search area can be geometrically constrained to fall along the epipolar lines. When three or more images are used (i.e. a multi-photo approach) the search area can be further reduced by forcing the search patches to move along the epipolar lines at certain intervals.

In order for MPGC matching to take place good provisional values for the search image coordinates and the object space coordinates of the reference image patch need to be determined. Many methods of determining approximate positions of the conjugate points exist. Examples of such techniques include using epipolar matching (Wong and Ho, 1986), or using of a combination of area and feature based matching (Van der Merwe and Rüther, 1994). Another method, described below, uses surface extrapolation from a DEM image. However, none of these techniques are robust enough to work under any circumstances (i.e. they often fail when measuring complex or poorly defined objects). Van der Vlugt and Rüther (1994) developed a technique of solving the problem of determining the approximate positions of the conjugate points. The method, known as multi-image correlation (MIC), is more reliable than other correlation methods due to the simultaneous use of more than one search image. The search patches are not just constrained along the epipolar lines, but using geometric constraints, to certain positions within each line. The image matching procedure used by Van der Vlugt and Rüther is described below. A flowchart indicating the matching procedure is shown in Figure 4.9. It differs slightly from the earlier approach developed by Baltsavias in the way that approximate positions of the conjugate points are determined.

The matching process is driven by patches of interest in the reference image. These patches can be, for example, patches of pixels centred on points of interest that were selected using a feature extractor. The approach used by Baltsavias (1991) requires a

good provisional Z coordinate (depth) for the first patch. Using this Z value, provisional X and Y object space coordinates of a reference patch are calculated using the collinearity equations. Provisional positions of the search patches are then calculated by projecting the object space point onto the plane of the search images. Additional parameters are included in this step to account for the curvature of the epipolar lines caused by lens distortion. MPGC matching now takes place. If the matching is successful the Z coordinate is stored in a 'DEM image' at the same pixel location as the pixel at the centre of the patch under consideration in the reference image. The object space coordinates of the successfully matched point is stored in a surface DEM. The Z coordinate of the next patch that is to be matched is approximated by using the height of the closest point in the DEM image. This method works well with smooth continuous surfaces and a dense interest point distribution. However, problems exist in determining the first Z coordinate and when there are surface discontinuities.

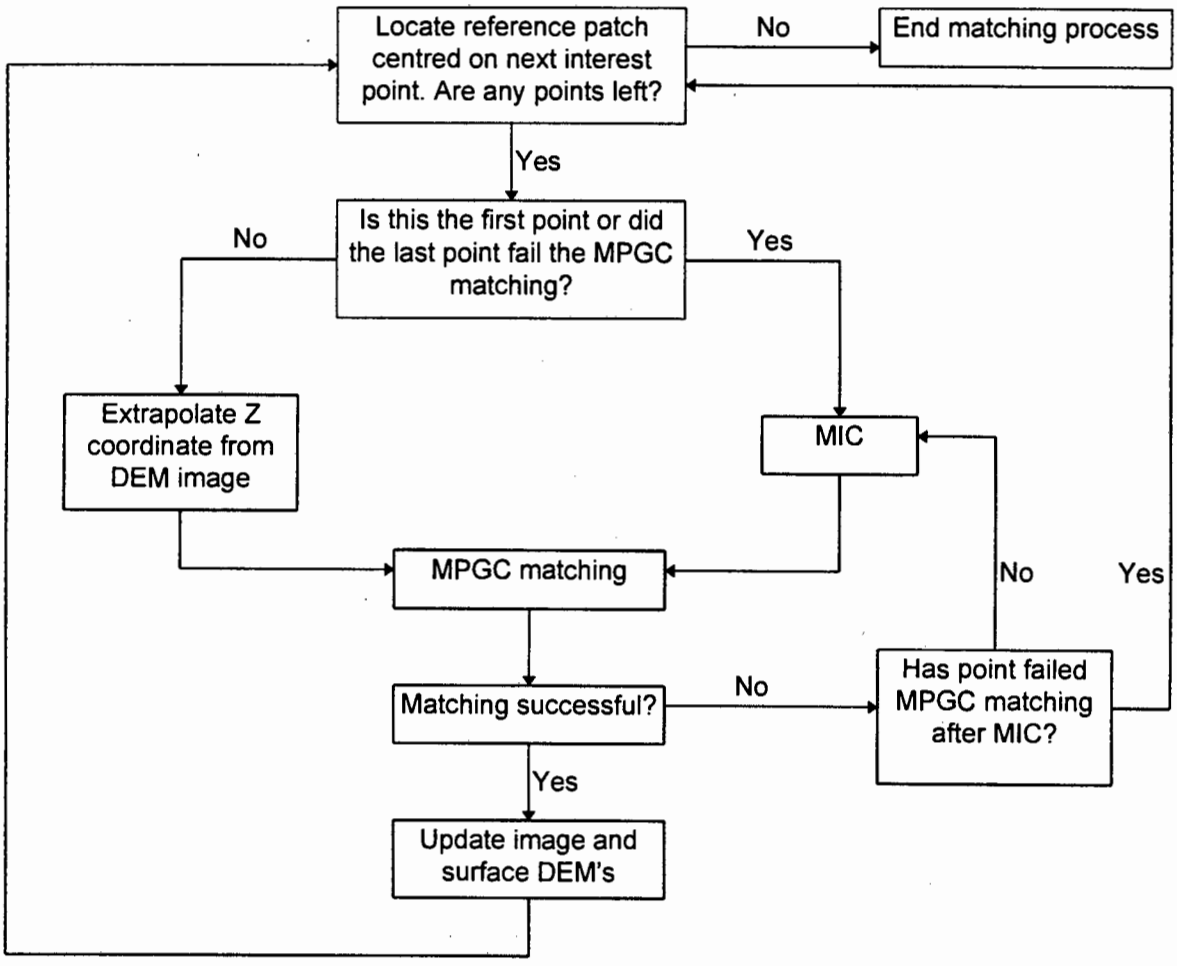


Figure 4.9: Flowchart of the matching procedure developed by Van der Vlugt and Rüther.

The method of MIC was developed to overcome these two problems. Firstly, MIC is used to estimate the Z coordinate of the first patch. And secondly, it is used when the MPGC matching of a point fails to converge. This occurs in cases of poor provisional values, as a result of surface discontinuities. MIC attempts to calculate a better estimate of the provisional values. If MPGC matching fails after MIC, the DEM image is not updated. For the determination of the provisional values of the next point, MIC is used as opposed to the extrapolation of the Z coordinate from the DEM image. Surface extrapolation becomes increasingly unreliable as the distance from the neighbouring point on the DEM increases. The process of MIC is shown in Figure 4.10. The position of the reference patch in the reference image as well as a Z coordinate is used to determine the X and Y coordinates of a fictitious object space point along a reference ray. The reference ray can be defined as the imaging ray passing through the centre of the reference patch. The fictitious object space point is back projected into all of the search images. Each search patch is correlated to the reference patch and for the particular Z coordinate used, a correlation value is determined as the average of the correlation values for each search patch. For this the normalised cross-correlation function, given in equation 3.21, is used. The Z coordinate is now incremented and the process of determining the new correlation value is repeated. Additional parameters are included in the MIC process described above. The best provisional positions of the search patches are indicated by the search interval with the highest correlation value.

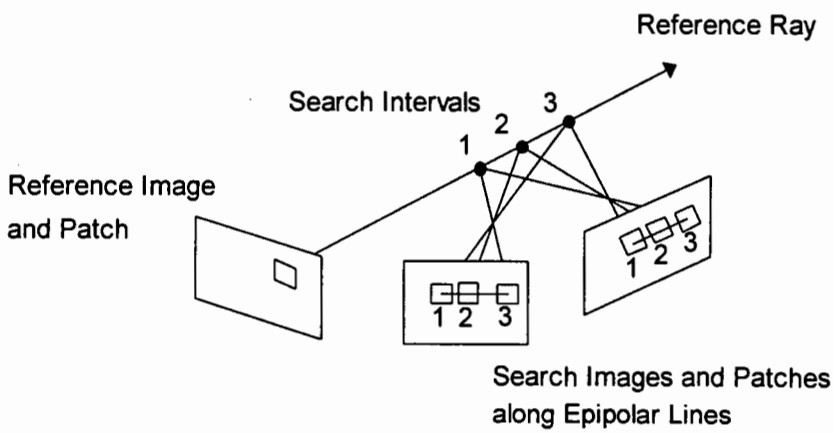


Figure 4.10: The MIC process.

The difference in the minimum and maximum Z coordinates of the control points can be used to determine the depth range over which the Z coordinate is incremented. A square reference patch will be distorted in the search images for two reasons, that of camera perspective and due to the difference in heights on the surface within the reference patch. Search patch shaping can be applied to model the first effect. The second effect cannot be modelled since a detailed knowledge of the surface is not available.

Van der Vlugt and Rüther (1994) report successful implementation of their method of MPGC matching. MIC was found to be a significant improvement in the ability to measure discontinuous surfaces, that contain rapid changes in surface height.

5. APPROXIMATE IMAGE ORIENTATION TECHNIQUES

The photogrammetric restitution of an object requires a knowledge of the image orientation parameters. The relationship between system calibration (involving interior, relative and exterior orientation) and the measurement of the object is discussed in section 3.2. Interior orientation parameters of a camera can be precisely determined by a process of camera calibration using a control frame in laboratory conditions. (This technique is discussed fully in section 7.5.) In this thesis the exterior orientation parameters were solved by means of a free network bundle adjustment in which all the images were adjusted simultaneously. For the bundle adjustment good approximations for the unknown parameters are required. In this chapter various methods for the determination of approximate values for the unknown orientation parameters are investigated. It will be assumed that the interior orientation parameters of the camera are known.

All of the methods investigated require the image coordinates of a number points on the images of the object to be measured. The minimum number of these points that are required, as well as their distribution, is dependent on the method used. For exterior orientation the object space coordinates of the points are also required. These points are then referred to as control points. Relative orientation methods do not require the object space coordinates of the points.

The points referred to above can take any form and can include prominent, natural points of detail on the surface of the object, or special targets placed on or around the object for the purpose of image orientation. For this thesis circular targets, placed on the surface of the object, were used. Circular targets have significant advantages over 'natural' points. Firstly, circular targets allow for rapid and semi-automatic identification and measurement of these points (see section 7.4.5). Such a high level of automation is not possible for natural conjugate points. Secondly, high accuracy target centring is easily achieved using circular targets (refer to section 4.1.3). The high accuracy measurement of the image coordinates of natural features is considerably more difficult.

The operational constraints in a typical industrial environment, such as an open cast mine, require that the amount of time necessary in setting up the targets and acquiring the images must be minimised. The setting up and measurement of the object space coordinates of the targets (for exterior orientation) constitutes the most time consuming stage of the field component of the photogrammetric process. The time required for the survey of the control points increases significantly with the increase of the required accuracy of the control points.

A further constraint that must be considered is the positioning of targets in the measuring environment. The rock surface to be measured is generally vertical with a certain amount of loose rock or rubble at the base. Although the targets can be well distributed in the plane of the vertical section of the rock surface, the positioning of the targets to create a good distribution in terms of depth, is limited, unless some sort of structure is used to create a better depth field. Such a structure, for example scaffolding or poles projected upwards, could obscure parts of the surface, would involve time in setting it up, and depending on certain conditions, such as wind, may be unstable. Furthermore, it would involve additional costs for the measurement system. These consequences are undesirable. Instead, a method of determining the approximate values of the image orientation parameters valid for a poor distribution of targets, that can be placed directly onto the surface, was sought. A typical example of the configuration of the targets that could be expected in an open cast mining environment is shown in Figure 5.1. It is possible to locate the bottom row of targets at ground level, thus allowing for the targets to be positioned further towards the camera stations. The middle and top rows of targets, in Figure 5.1, have been located on the vertical section of the rock surface.

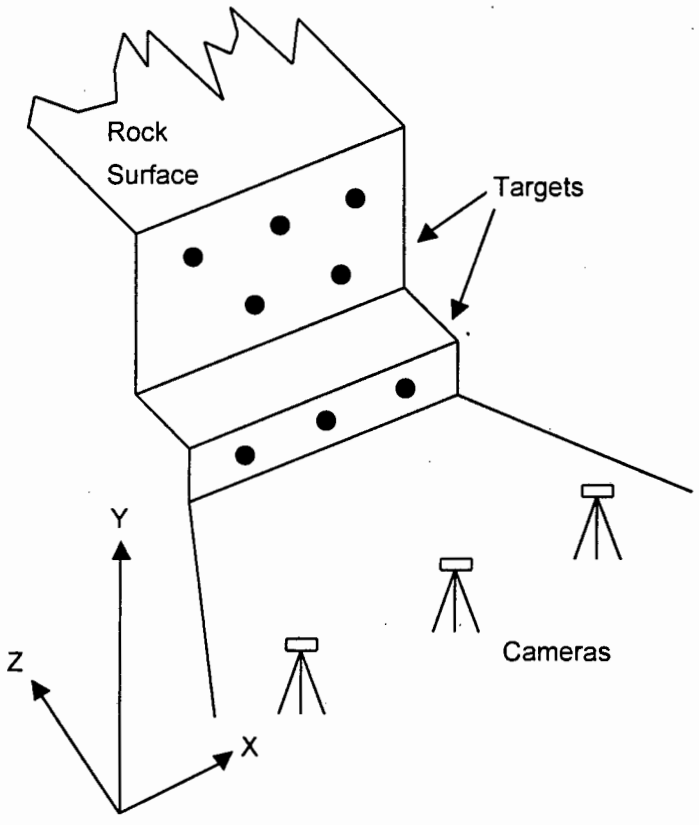


Figure 5.1: An example of the distribution of targets in an open cast mining environment.

Based on the constraints described above, it was necessary to identify a robust method of determining the approximate values of the image orientation parameters. An ideal method should have the following features:

(i) As few conjugate points (for relative orientation) or control points (for exterior orientation) as possible, should be used;
(ii) It should have the ability to reach a solution using a poor configuration of conjugate points or control points;
(iii)It should be capable of a solution using low accuracy control point coordinates (for exterior orientation methods).

The techniques used in this chapter for testing the various orientation methods are briefly described. One of these techniques used for the rigorous testing of the method of relative orientation with linear transformation (see section 5.2), was that of creating simulated image coordinates. The advantages of using simulated images are:

(i) The user can easily choose which targets are to be used, without limiting factors, such as occlusions or bad illumination, effecting the results;
(ii) All the observations have very low residuals, relating only to rounding errors. Lens distortion errors are eliminated;
(iii)A large degree of control and flexibility can be provided for testing of the techniques under different configurations and conditions.

The simulated image coordinates were created by adopting values for the interior and exterior orientation parameters, as well as appropriate object space coordinates of targets and camera stations, and substituting these values into the collinearity equations.

Other techniques used in the testing of the orientation methods included the use of real images of a control (calibration) frame, as well as pseudo-randomly altering the numerical values of the control point coordinates, so as to simulate poor control. Tests were carried out on both a well distributed set of control points as well as on a control point configuration similar to that shown in Figure 5.1.

The images of the control frame constituted a set of nine convergent images taken in three rows with three images in each row. The camera stations were positioned so as to fill the field of view with the frame. In order to compare the results of the methods investigated in this chapter with a well recognised image orientation method, the exterior orientation parameters of all nine images were calculated using the method of camera calibration described in section 7.5. This camera calibration method was based on a bundle adjustment solution. The parameters from the calibration procedure, as well as the positions of selected control points on the frame, were also used as the input parameters required for the creation of simulated image coordinates. The control frame, of dimensions 700 x 700 x 250 mm, is shown in Figure 5.2.

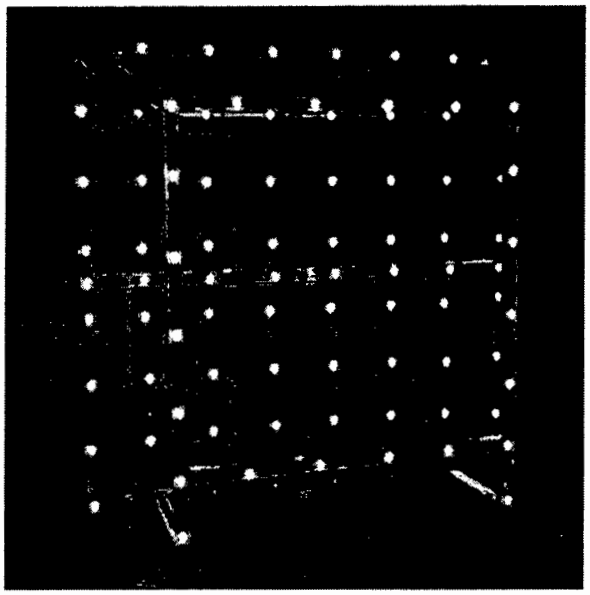


Figure 5.2: The control (calibration) frame.

For all of the tests in this chapter, and for the entire thesis, a coordinate system, as described in section 3.1, was used. This coordinate system ensures that the rotation angles of the images are small. The relationship between the object space and image space coordinate systems is shown in Figure 5.3.

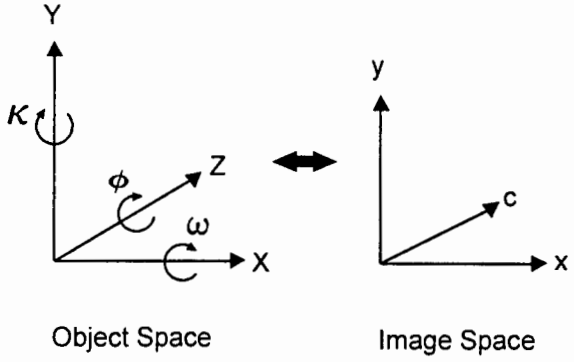


Figure 5.3: The relationship between the object space and image space coordinate systems.

5.1 THE DIRECT LINEAR TRANSFORMATION

The DLT, previously described in section 3.3, can be used as a direct method of camera orientation, which generally provides a good means of obtaining exterior orientation parameters for a single image. The DLT has the advantage of not requiring initial approximations for the unknown parameters. However, it requires a good distribution, in three dimensions, of at least six control points. The DLT provided very poor results

when applied to points on the control frame in a configuration similar to that shown in Figure 5.1. In many cases the DLT iterated to an incorrect solution or failed to solve. This method of orientation was considered impractical for poor control point configurations.

5.2 RELATIVE ORIENTATION WITH LINEAR TRANSFORMATION

Methods of relative orientation have an advantage over methods of exterior orientation in that object space coordinates of the conjugate points are not required. The transformation of the coordinates of an object measured using relative orientation parameters (given in equation 3.7), into an external coordinate system (i.e. as if measured using exterior orientation parameters), can be accomplished by means of a three dimensional transformation comprising three translations, three rotations and a scale factor. This transformation is of the form:

$$x = x_0 + \lambda R\bar{x} \qquad 5.1$$

where x is a vector of points in the target system, $\bar{x}$ is a vector of points in the object system (not equivalent to object space), with translation vector x_0, rotation matrix R and scale factor λ.

A closed form, or direct, solution of relative orientation as described by Shih (1994), was investigated. The method, named relative orientation with linear transformation (RLT), is outlined below.

Closed form solutions have an advantage over solutions based on equations that are non-linear with respect to the unknown parameters, since they are independent of initial approximations to the unknowns. A model based on non-linear equations can be solved in a least squares adjustment where the solution vector is made up of the corrections to the unknowns. An iterative approach is needed in this case, and initial values for the unknowns must be approximated. (In the case of the DLT, where the equations are non-linear, corrections to the unknowns are also calculated. However, the same initial approximations are sufficient in all circumstances, thereby allowing the DLT to be regarded as not requiring initial approximations to the unknowns.) Where the model is already linear the full unknown elements replace the corrections to the unknowns.

The minimum requirements for the RLT are eight non-coplanar conjugate points, as well as a knowledge of the interior orientation parameters of the cameras.

The coplanarity equation as described previously in section 3.6 forms the condition:

$$L_1yx' + L_2yy' - L_3yc' + L_4cx' + L_5cy' - L_6cc' + L_7xx' + L_8xy' - L_9xc' = 0 \qquad 5.2$$

where

$$
\begin{aligned}
L_1 &= B_X r_{31} - B_Z r_{11} \\
L_2 &= B_X r_{32} - B_Z r_{12} \\
L_3 &= B_X r_{33} - B_Z r_{13} \\
L_4 &= B_X r_{21} - B_Y r_{11} \\
L_5 &= B_X r_{22} - B_Y r_{12} \\
L_6 &= B_X r_{23} - B_Y r_{13} \\
L_7 &= B_Z r_{21} - B_Y r_{31} \\
L_8 &= B_Z r_{22} - B_Y r_{32} \\
L_9 &= B_Z r_{23} - B_Y r_{33}
\end{aligned}
\qquad 5.3
$$

r_{ij} are the elements of an orthogonal rotation matrix. x,y and x′,y′ refer to image coordinates on the first and second images respectively, already adjusted for their respective principal point displacements. c and c′ denote the principal distances of the first and second images respectively.

Since equation 5.3 is homogeneous, one parameter can be set constant. Choosing L_5, equation 5.2 now becomes:

$$L_1'yx' + L_2'yy' - L_3'yc' + L_4'cx' - L_6'cc' + L_7'xx' + L_8'xy' - L_9'xc' + cy' = 0 \qquad 5.4$$

where

$$L_i' = L_i / L_5 \qquad 5.5$$

The model can be solved directly without any knowledge of approximate values for the unknown parameters since equation 5.4 is already linear with respect to the L_i' parameters. The solution vector, consisting of the full unknown elements, can be solved in a least squares adjustment of the form:

$$x = \left(A^T P A\right)^{-1} A^T P \ell \qquad 5.6$$

The weight matrix P, is an identity matrix.

An alternative approach is to solve the model using the combined case of the least squares adjustment. From equation 5.4 it can be noted that there is more than one observation in each condition equation. Approximations to the unknowns set to 0.1, tend to be a suitable value from which to calculate corrections to the unknowns. Thus, even with the combined least squares adjustment, the solution can be viewed as effectively

being independent of initial approximations to the unknowns. The solution vector has the well known form:

$$x = -\left(A^T\left(BP^{-1}B^T\right)^{-1}A\right)^{-1}A^T\left(BP^{-1}B^T\right)^{-1}w \qquad 5.7$$

The model can be identified as the quasi-parametric adjustment, significantly reducing the computational effort and computer memory required in reaching a solution. The solution vector can thus be reduced to:

$$x = -\left(A^T\bar{P}A\right)^{-1}A^T\bar{P}w \qquad 5.8$$

where the quasi-weight matrix is:

$$\bar{P} = \left(BP^{-1}B^T\right)^{-1} \qquad 5.9$$

The sign of L_5 is ambiguous and needs to be determined. Shih describes a mathematically rigorous method of evaluating L_5. A less elegant technique is that of substituting the calculated L_i values back into equation 5.2 for both positive and negative L_5 and evaluating the residuals. Summing the absolute values of the residuals provides a method of determining the correct sign of L_5. The ratio of the sums with the correct to incorrect L_5 values is commonly of the order 1 to 100 000 (for simulated images). These residuals also act as a check of the adjustment.

The decomposition of the unknown parameters into the physical relative orientation parameters, using the re-scaled L_i parameters, can be achieved by setting B_X to unit length and computing B_Y and B_Z from equations 5.10 and 5.11. (For purposes of completeness, in equations 5.10 to 5.12, the B_X parameter is shown instead of its numerical value of one.)

$$B_Y = -\frac{L_1L_7 + L_2L_8 + L_3L_9}{B_X} \qquad 5.10$$

$$B_Z = \frac{L_4L_7 + L_5L_8 + L_6L_9}{B_X} \qquad 5.11$$

The rotation matrix elements can be determined from:

$$r_{11} = \frac{L_3 L_5 - L_2 L_6 - B_Z L_1 - B_Y L_4}{B_X^2 + B_Y^2 + B_Z^2}$$

$$r_{12} = \frac{L_1 L_6 - L_3 L_4 - B_Z L_2 - B_Y L_5}{B_X^2 + B_Y^2 + B_Z^2}$$

$$r_{13} = \frac{L_2 L_4 - L_1 L_5 - B_Z L_3 - B_Y L_6}{B_X^2 + B_Y^2 + B_Z^2}$$

$$r_{21} = \frac{B_Y r_{11} + L_4}{B_X}$$

$$r_{22} = \frac{B_Y r_{12} + L_5}{B_X}$$

$$r_{23} = \frac{B_Y r_{13} + L_6}{B_X}$$

$$r_{31} = \frac{B_Z r_{11} + L_1}{B_X}$$

$$r_{32} = \frac{B_Z r_{12} + L_2}{B_X}$$

$$r_{33} = \frac{B_Z r_{13} + L_3}{B_X} \qquad 5.12$$

The decomposition of the rotation matrix results in the relative rotation angles $\Delta\omega$, $\Delta\kappa$ and $\Delta\phi$.

To test this method, software was written based on the quasi-parametric case of the least squares adjustment. Nine simulated images were created using 18 points, well distributed in three dimensions. As mentioned earlier in this chapter, the parameters required for the creation of the simulated image coordinates were based on the calculated parameters from a camera calibration procedure (carried out using the control frame), as well as on the positions of selected targets on the frame. The RLT method described above determines the relative orientation of two images. In order to orient multiple images relative to each other, one of the images was used consistently in every set of RLT adjustments. Using nine images, eight image pairs were adjusted.

The base vector components of each image pair was based on B_X of unit length. Since the correct value for B_X was known, the B_Y and B_Z values, calculated by means of the RLT, could be scaled to the original object space scale. (i.e. After scaling, B_Y and B_Z is based on the true value of B_X as opposed to the unit length value of B_X.) A comparison was made of the values for B_Y and B_Z, determined using the RLT (and scaled to the original object space scale), to the true values of B_Y and B_Z. This comparison indicates the error of the RLT adjustment. Using the known angles, the errors in $\Delta\omega$, $\Delta\kappa$ and $\Delta\phi$ were also calculated. A summary of the errors of the RLT parameters for all eight image pairs is shown in Tables 5.1 and 5.2. In Table 5.1 the percentage of the control frame size

refers to the comparison of the root mean square (RMS) error, based on all eight adjustments, to the dimensions of the control frame in the XY plane. These dimensions, as mentioned earlier, are 700 x 700 mm. (Note again that coordinates of targets on the control frame were used to create the simulated image coordinates.) The error range referred to in Table 5.1 gives the range between the most accurate and least accurate calculations of the base vector components for all eight image pairs. The angle range referred to in Table 5.2 shows the difference in rotation angles for the most convergent images. This indicates the degree to which the orientation of the nine images differed. (Throughout this thesis the RMS error is based on the following equation:

$$\text{RMS error} = \sqrt{\frac{\sum_{i=1}^{n} \varepsilon_i^2}{n}} \qquad 5.13$$

where ε is the error associated with a single value in a data set, containing n values.)

	RMS Error (mm)	Error Range (mm)	% of Control Frame Size
B_Y	15	6 - 23	2%
B_Z	21	2 - 28	3%

Table 5.1: Summary of the base vector errors resulting from the RLT with nine simulated images using 18 well distributed points on the control frame.

	RMS Error (degrees)	Error Range (degrees)	Angle Range (degrees)
$\Delta\omega$	0.33	0.03 - 0.57	17.82
$\Delta\kappa$	0.17	0.02 - 0.29	43.25
$\Delta\phi$	1.58	0.04 - 2.61	3.50

Table 5.2: Summary of the relative rotation angle errors resulting from the RLT with nine simulated images using 18 well distributed points on the control frame.

Taking an average of all eight adjustments, the back substitution of the re-scaled L_i parameters into equation 5.2 yielded a mean absolute value residual of 4e-6 mm, relating both to rounding errors and to errors in the transformation model. The values of the base vector components and relative rotation angles that were obtained, were certainly acceptable as approximate values for a subsequent bundle adjustment.

The RLT algorithm was also tested using simulated images with the minimum of eight conjugate points. The configuration of the points, in object space, was similar to what

would be expected in a typical open cast mining environment as indicated in Figure 5.1. The results were of the same order as those in Tables 5.1 and 5.2.

The approximation in this method is a result of solving for the dependent exterior orientation parameters using a linear transformation by assuming linear independence. The formation of the RLT model assumes the orthogonality of the rotation matrix. However, the increase in magnitudes of $\Delta\omega$, $\Delta\kappa$ and $\Delta\phi$ cause the rotation matrix to deviate from orthogonality. This can be seen by forming R^TR. (For orthogonal matrices $R^TR = I$.) To test this orthogonality condition, the algorithm was tested using simulated stereo images where $\Delta\omega$, $\Delta\kappa$ and $\Delta\phi$ were zero. As predicted, the RLT produced negligible errors for all of the calculated parameters.

A set of images taken of a rock surface at an excavation site at the Cape Town harbour was used to further test the RLT. A full description of the configuration of the targets and camera stations can be found in section 8.1. The image orientation parameters and the object space coordinates of the targets were determined using a technique described later in this chapter (see section 5.4), followed by a free network bundle adjustment. These calculated values were used to simulate image coordinates of the eight targets in five images. Simulated image coordinates were used in preference to the real image coordinates so as to exclude the effects of lens distortion and inaccuracies in target centring. Using the configuration in this test case, the RLT produced base vector components that did not in any way correspond to the known camera positions. However, the relative rotation angle results corresponded relatively well to the known values. A summary of the relative rotation angle results, based on all four adjustments, is shown in Table 5.3.

	RMS Error (degrees)	Error Range (degrees)	Angle Range (degrees)
$\Delta\omega$	3.92	0.1 - 6.8	7.6
$\Delta\kappa$	0.07	0.0 - 0.1	48.9
$\Delta\phi$	2.38	1.2 - 3.9	10.8

Table 5.3: Summary of the relative rotation angle errors resulting from the RLT with five simulated images of the Cape Town harbour excavation site.

The error of the relative orientation angles deteriorated progressively as the angle, κ about the Y axis increased. Tests of the RLT with measured image coordinates of the harbour excavations were also carried out. Although lens distortion parameters and principal point displacements were accounted for, the RLT failed to produce valid results for any of the relative orientation parameters in any combination of images. The fact that this camera orientation technique failed in this configuration indicates that the RLT is not robust enough for practical use in this thesis.

5.3 SMITH'S EXPLICIT SPACE RESECTION SOLUTION

An explicit solution for the space resection of a single image is given by Smith (1965). This method requires no initial knowledge of the six unknown parameters of exterior orientation. In order to provide a solution three non-collinear control points, as well as the interior orientation parameters of the camera, are required. However, in the general case, four solutions are obtained. In order to identify the unique solution either the camera orientation must be approximately known or a fourth control point must be introduced. Whereas iterative solutions may iterate to an incorrect solution, explicit methods will examine all of the possible solutions. Smith's explicit method is outlined below.

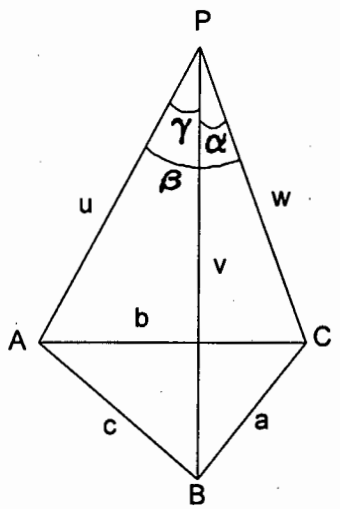


Figure 5.4: Configuration of control points used in Smith's explicit space resection.

In Figure 5.4 A, B and C denote control points with object space coordinates (X_1,Y_1,Z_1), (X_2,Y_2,Z_2) and (X_3,Y_3,Z_3) respectively. P is the perspective centre of the camera with a principal distance of f (used in this section instead of the symbol c, for purposes of clarity). The sides of ΔABC are a, b and c. The image coordinates of A, B and C are (x_1,y_1), (x_2,y_2) and (x_3,y_3) respectively.

The angles α, β, γ are subtended at P by ΔABC. α can be given by:

$$\cos\alpha = \frac{(x_2x_3 + y_2y_3 + f^2)}{(x_2^2 + y_2^2 + f^2)^{1/2}(x_3^2 + y_3^2 + f^2)^{1/2}} \qquad 5.14$$

Similarly, equations for $\cos\beta$ and $\cos\gamma$ can be formulated.

Letting the lengths PA = u, PB = v and PC = w, it can be shown that:

$$u = c(1 - 2m_1 \cos\gamma + m_1^2)^{-1/2} \quad 5.15$$

$$v = m_1 u \quad 5.16$$

$$w = m_2 u \quad 5.17$$

where m_1 can be obtained by solving:

$$Am_1^4 + Bm_1^3 + Cm_1^2 + Dm_1 + E = 0 \quad 5.18$$

and

$$m_2 = \frac{(a^2 + c^2 - b^2) + 2(b^2 - a^2)m_1 \cos\gamma - (b^2 + c^2 - a^2)m_1^2}{2c^2(\cos\beta - m_1 \cos\alpha)} \quad 5.19$$

The terms A, B, C, D, E are given by:

$$A = (b^2 + c^2 - a^2)^2 - 4b^2c^2 \cos^2\alpha$$

$$B = K - 2A\cos\gamma$$

$$C = A + E - 2K\cos\gamma + 4c^4(\cos^2\alpha + \cos^2\beta + \cos^2\gamma - 2\cos\alpha\cos\beta\cos\gamma - 1)$$

$$D = K - 2E\cos\gamma$$

$$E = (a^2 + c^2 - b^2)^2 - 4c^2a^2\cos^2\beta \quad 5.20$$

where

$$K = 2(b^2 + c^2 - a^2)(a^2 + c^2 - b^2)\cos\gamma + 4c^2(a^2 + b^2 - c^2)\cos\alpha\cos\beta \quad 5.21$$

The roots of equation 5.18, a fourth order polynomial in m_1, can be solved by any one of a number of methods. Bairstow's eigenvalue method is a numerical method suggested by Smith. For each value of m_1, the corresponding value of m_2 is calculated. Four sets of values for (u,v,w) can now be determined. If any of the terms of (u,v,w) are either negative or imaginary, the set can be rejected. A fourth control point, together with two of the three original control points, is now used to resolve the ambiguity of (u,v,w) by solving for the roots of equation 5.18 using this second set of control points and comparing them to the roots that were previously calculated.

If Δ denotes the area of ΔABC, the direction cosines of the normal to the plane ABC can be represented by:

$$\lambda_1 = \frac{\Delta_1}{\Delta}, \quad \lambda_2 = \frac{\Delta_2}{\Delta}, \quad \lambda_3 = \frac{\Delta_3}{\Delta} \quad 5.22$$

where

$$\Delta_1 = \pm\frac{1}{2}\left(Y_1(Z_2 - Z_3) + Y_2(Z_3 - Z_1) + Y_3(Z_1 - Z_2)\right)$$
$$\Delta_2 = \pm\frac{1}{2}\left(Z_1(X_2 - X_3) + Z_2(X_3 - X_1) + Z_3(X_1 - X_2)\right)$$
$$\Delta_3 = \pm\frac{1}{2}\left(X_1(Y_2 - Y_3) + X_2(Y_3 - Y_1) + X_3(Y_1 - Y_2)\right) \qquad 5.23$$

The sign of the direction cosines are either all positive or all negative depending on which side of plane ABC point P falls.

It can be shown that the length of the perpendicular from P onto the plane is given by:

$$h = \frac{uvw}{2\Delta}\left(1 + 2\cos\alpha\cos\beta\cos\gamma - \cos^2\alpha - \cos^2\beta - \cos^2\gamma\right)^{\frac{1}{2}} \qquad 5.24$$

The coordinates of $P(X_P, Y_P, Z_P)$ can be determined by solving the following set of linear equations:

$$h = \lambda_1(X_P - X_1) + \lambda_2(Y_P - Y_1) + \lambda_3(Z_P - Z_1) \qquad 5.25$$

$$v^2 - u^2 = 2X_P(X_1 - X_2) + 2Y_P(Y_1 - Y_2) + 2Z_P(Z_1 - Z_2) + \\ + X_2^2 + Y_2^2 + Z_2^2 - X_1^2 - Y_1^2 - Z_1^2 \qquad 5.26$$

$$w^2 - u^2 = 2X_P(X_1 - X_3) + 2Y_P(Y_1 - Y_3) + 2Z_P(Z_1 - Z_3) + \\ + X_3^2 + Y_3^2 + Z_3^2 - X_1^2 - Y_1^2 - Z_1^2 \qquad 5.27$$

The elements of the rotation matrix can be determined by solving the following three linear equations:

$$r_{11}(X_1 - X_P) + r_{12}(Y_1 - Y_P) + r_{13}(Z_1 - Z_P) = \frac{ux_1}{(x_1^2 + y_1^2 + f^2)^{1/2}} \qquad 5.28$$

$$r_{11}(X_2 - X_P) + r_{12}(Y_2 - Y_P) + r_{13}(Z_2 - Z_P) = \frac{vx_2}{(x_2^2 + y_2^2 + f^2)^{1/2}} \qquad 5.29$$

$$r_{11}(X_3 - X_P) + r_{12}(Y_3 - Y_P) + r_{13}(Z_3 - Z_P) = \frac{wx_3}{(x_3^2 + y_3^2 + f^2)^{1/2}} \qquad 5.30$$

The remaining rotation matrix elements can be obtained from similarly formulated sets of equations. The extraction of the rotation angles can now follow. An extension to this method involves including redundant control in a least squares adjustment (see Smith (1965)).

The Smith algorithm, based on the minimum solution of four control points, was tested using software written by Mark Cammidge of the Department of Surveying and Geodetic Engineering at UCT. Nine real images of the control frame, shown in Figure 5.2, were used with the same camera configuration that was described earlier in this chapter. A bundle adjustment solution adjusting all nine images simultaneously provided a comparison for the Smith method. Three well distributed, non-coplanar control points were used in an initial test of the method. A fourth control point, in the centre of the control frame, provided a check for determining the unique solution. A summary of the exterior orientation parameter results taken from all nine images is given in Tables 5.4 and 5.5.

	RMS Error (mm)	Error Range (mm)	% of Control Frame Size
X	4.8	0.3 - 8.6	1%
Y	11.9	1.2 - 19.4	2%
Z	11.8	7.9 - 16.0	2%

Table 5.4: Summary of the perspective centre errors resulting from the Smith method using nine images with four non-coplanar points on the control frame.

	RMS Error (degrees)	Error Range (degrees)	Angle Range (degrees)
ω	0.65	0.11 - 1.33	17.82
κ	0.40	0.23 - 0.77	43.25
ϕ	2.32	0.02 - 5.12	3.50

Table 5.5: Summary of the rotation angle errors resulting from the Smith method using nine images with four non-coplanar points on the control frame.

Relatively high accuracy results were obtained by this method, considering that the minimum solution was employed and a wide range of camera positions and rotation angles were used. The results obtained were more than adequate for use as approximate parameters for a subsequent bundle adjustment. It was found that when the fourth control point (used to determine the unique solution) and the perspective centre of the camera had approximately the same position in the XY plane, an incorrect solution was chosen. By selecting another control point in a different position in the XY plane, the correct solution was selected.

The Smith method has an advantage over other methods of space resection, such as the DLT, in that it allows for the control points to be coplanar. Four well distributed coplanar control points were selected on the control frame and the algorithm was tested

using the same nine images as in the previous test. A summary of these results is displayed in Tables 5.6 and 5.7.

	RMS Error (mm)	Error Range (mm)	% of Control Frame Size
X	4.4	0.5 - 7.5	1%
Y	4.7	1.2 - 6.5	1%
Z	7.9	5.6 - 11.5	1%

Table 5.6: Summary of the perspective centre errors resulting from the Smith method using nine images with four coplanar points on the control frame.

	RMS Error (degrees)	Error Range (degrees)	Angle Range (degrees)
ω	0.34	0.02 - 0.72	17.82
κ	0.13	0.01 - 0.28	43.25
ϕ	1.92	0.05 - 4.97	3.50

Table 5.7: Summary of the rotation angle errors resulting from the Smith method using nine images with four coplanar points on the control frame.

Many other tests were carried out with the same nine images using varying positions of the control points for both coplanar and non-coplanar configurations. From these tests (including the results of the examples given in Tables 5.4 to 5.7) it appears that a coplanar configuration of points provides a better solution to that of a non-coplanar configuration. Both the RMS errors and the range of errors, for both perspective centre positions and rotation angles, tended to be smaller when coplanar points were used.

In order to investigate the required accuracy of the control point coordinates in object space, a number of tests were carried out by randomly varying the coordinates of the points on the control frame. One example is given below in which the control point coordinates of four non-coplanar points were varied by up to 2.5% of the control frame size. The same nine images as for the previous tests in this section, were used. A summary of these results is shown in Tables 5.8 and 5.9.

From Table 5.8 it can be seen that the RMS error of the perspective centre coordinates (calculated from all nine images) is very large when taken as a percentage of the control frame size. The higher end of the error range indicates that the error of some of the perspective centre coordinates are up to 35% of the control frame size. Although the RMS errors of the rotation angles are generally within acceptable limits for use as approximate values for the subsequent bundle adjustment, the error range indicated in Table 5.9, shows that a number of the rotation angles errors are very large. The errors

produced from the above test are generally too large for these parameters to be used as initial values in a subsequent bundle adjustment, especially when considering the upper end of the error ranges. The errors for all six exterior orientation parameters deteriorated rapidly as the control point coordinates were varied beyond 2.5% of the control frame size. A least squares adjustment making use of redundant control would very likely reduce the level of accuracy required for the control point coordinates.

	RMS Error (mm)	Error Range (mm)	% of Control Frame Size
X	134.3	36.8 - 216.0	19%
Y	99.4	3.4 - 243.3	14%
Z	39.9	14.6 - 61.2	6%

Table 5.8: Summary of the perspective centre errors resulting from the Smith method using nine images with four non-coplanar points on the control frame, where the coordinates have been randomly varied by up to 2,5% of the control frame size.

	RMS Error (degrees)	Error Range (degrees)	Angle Range (degrees)
ω	4.04	0.11 - 10.16	17.82
κ	5.69	0.52 - 9.08	43.25
ϕ	1.59	0.00 - 3.30	3.50

Table 5.9: Summary of the rotation angle errors resulting from the Smith method using nine images with four non-coplanar points on the control frame, where the coordinates have been randomly varied by up to 2,5% of the control frame size.

Smith's method of exterior orientation provides an accurate method of space resection, provided that four non-collinear control points, well distributed in image space and with high accuracy object space coordinates, are used. The configuration of the control points for this method does not require a good distribution of points in three dimensions, thus producing significant advantages over certain methods, such as the DLT. However, the requirement of good control point coordinates limits the usefulness of this technique in applications where only poor control point coordinates is available.

5.4 THE COLLINEARITY EQUATIONS

The previously discussed methods of image orientation all proved unsatisfactory for use in this project. A method of orienting single images based on the collinearity equations was found to be a suitable alternative. This method was developed specifically for this project and was based on a similar technique described by Professor Clive Fraser of the

University of Melbourne in personal correspondence. Although the collinearity equation method of single image orientation, described in this section, is not an elegant technique, it outperforms the other methods that were investigated in terms of the features of the orientation method that was sought. These criteria were listed towards the beginning of this chapter.

The collinearity equation method described here, uses previously determined interior orientation parameters, as well as very approximate perspective centre coordinates and object space coordinates of the control points, to calculate the rotation angles of exterior orientation that are used as initial approximations in a subsequent bundle adjustment. Only the rotation angles are calculated with the collinearity equation method. All of the other parameters must be obtained using other techniques. The interior orientation parameters can be determined by a process of camera calibration, as described in detail in section 7.5. It will be seen later in this section that even with the coordinates of the control points and perspective centres accurate to 10% of the size of the surface being measured, good results are produced. The coordinates of the perspective centres can be determined very rapidly using a tape measure. It was found that it was sufficient to estimate the coordinates of the control points by inspecting the image taken from the central camera station (i.e. making use of the image with the least perspective distortion). The only additional information required to aid in the coordination of the control points is the approximate height of the surface to be measured, as well as the approximate distance from the camera stations to the targets. In this way the time required for the field component of the photogrammetric process is kept to a minimum.

The collinearity equations are non-linear and thus need to linearised by means of a Taylor Series expansion and then formulated in a least squares adjustment (where redundant control is used). Approximations to the unknown rotation parameters need to be determined. If the surface of an object to be measured is approximately planar and only one dominant surface needs to be mapped, such as is typical with a rock face (as shown in Figure 5.1), the rotation angles for all the convergent images necessary to map the surface, can be estimated very approximately. Using the object space and image space coordinate systems shown in Figure 5.3, all three exterior orientation rotation angles, for all of the images, can be initially estimated to be zero. (For the measurement of more complex shapes, where more than one dominant side of an object needs to be measured, the rotation angles would be too variable to be estimated, even very roughly, and this method of image orientation would be unsuitable.)

The interior orientation parameters, as well as the coordinates of the perspective centre of the camera and the coordinates of the control points, are held fixed. The only free parameters are the exterior orientation rotation angles. The use of the collinearity equations by this method can be viewed as an adaptation of the bundle adjustment, solving for a single image at a time. Bundle adjustment software, written by Ph.D. candidate Graeme van der Vlugt of the Department of Surveying and Geodetic Engineering at UCT, was used for this purpose. If additional parameters are available from the camera calibration procedure, these can be included in the adjustment, although

this is not necessary considering that only initial approximations to the unknown parameters are required.

This method was tested using the same nine real images of the control frame that were used for the tests described earlier in this chapter. A bundle adjustment solution using all nine images simultaneously was used for purposes of comparison. To simulate inaccurate control point coordinates, the coordinates on the control frame and perspective centre were randomly altered by up to 10% of the control frame size. Using a coordinate system as shown in Figure 5.3, the approximate rotation angles were initialised to zero for all three angles. Three non-coplanar control points were used in a configuration similar to that shown in Figure 5.1, using the two outermost points in the top row, and the centre point in the bottom row. Each image was adjusted separately, solving for only the rotation angles. A summary of the results is given in Table 5.10.

	RMS Error (degrees)	Error Range (degrees)	Angle Range (degrees)
ω	2.60	0.54 - 3.57	17.82
κ	3.23	0.21 - 5.36	43.25
ϕ	3.26	1.97 - 4.11	3.50

Table 5.10: Summary of the rotation angle errors resulting from the collinearity equation method of image orientation using nine images with three control points on the control frame, where the coordinates have been randomly varied by up to 10% of the control frame size.

The results shown in Table 5.10 indicate that the maximum error was less than 5.5°, with a RMS error for all the angles of less than 3.5°. Initial approximations of rotation angles for the subsequent bundle adjustment need only be accurate to within approximately 10°. Thus this method was shown to be more than adequate for orienting single images within the physical constraints of an open cast mining environment. The method satisfies the ideal criteria for an image orientation method that was listed at the beginning of this chapter. Namely, it produces a good solution using only a few control points, it solves in a poor configuration that would be expected in an open cast mining environment, and it allows for the use of poor control point coordinates.

Although a minimum of two control points are required to solve for the three unknown parameters, this produces results that deteriorate significantly compared to the three point adjustment. When more than three control points were used, the results did not appear to improve significantly. A substantial improvement in the accuracy of the rotation angles will occur only if the accuracy of the control points, or other parameters held fixed, improve. For the tests of excavation surfaces that are described in chapter 8, eight control points were used for the following two reasons. Firstly, the redundancy of the network is increased as the number of control points increases, thus making the solution more reliable. Secondly, it is often the case that some targets are occluded or not detected due

to, for example, unfavourable illumination. Redundant control points are useful to make sure that sufficient targets are always available.

5.5 FURTHER METHODS OF CAMERA ORIENTATION

A few further methods of camera orientation are briefly mentioned in this section. A camera calibration technique that is sometimes employed is that of using only distances in a free network calibration adjustment to obtain relative orientation at the correct scale. This method has the advantage of not requiring a control point field. Fraser (1992) describes moving a bar, with targets placed at both ends, into various positions within the area to be measured. Multiple cameras are used to take images for each position of the bar. The result is a number of distances in object space that allows for the determination of the relative orientation parameters of the cameras. Heikkilä (1990) reports on the successful practical implementation of such a method. This approach is more suitable, however, to small scale industrial measurement tasks than to large scale engineering applications. Furthermore, multiple cameras are required, since images must be acquired from all the camera stations for each position of the bar. For this thesis it was practical to use only a single camera, in terms of the cost of the cameras, as well as in terms of synchronising the cameras to take the images simultaneously. Moving a large bar around an excavation site is also impractical.

Another method of relative orientation is that of the Wunderlich approach (Shih, 1994). This method requires four conjugate points per image pair, but is restricted by the need for a high degree of coplanarity of the points in object space.

Strunz (1993) describes the use of points, lines and surfaces as control information for camera orientation. Further research on these methods is still required before robust methods are fully developed (Petsa and Patias, 1994).

6. THE SELECTION OF FEATURES OF INTEREST

The extraction of pixels of interest on an image is an important component of automatic and semi-automatic digital photogrammetric processes. Pixels of interest can be considered to be those selected pixels on an image that are used to adequately describe the object to be measured. Such pixels are typically associated with the edges of the object or the texture of the surface of the object. The extraction of these pixels of interest has a further function of reducing the number of pixels to be processed. One method of selecting pixels to be mapped is that of using a suitable feature extractor. A good feature extractor should identify a representative selection of points on the surface of the object to be mapped. In order for this to be achieved the points should have a good distribution and an adequate density. The number of pixels chosen and their distribution is dependant on the nature of the object to be measured. A simple object would, for example, require considerably fewer points to adequately represent the object, compared to a more complex object. The number of points selected is also influenced by time constraints since image matching is very computationally intensive. For example, in chapter 8 it is stated that image matching of approximately 2800 pixels, using five images, took approximately seven hours to process.

A second method of selecting pixels to be matched is that of matching pixels at regular intervals across the image. This method is not based on the selection of interest pixels and consequently has a number of disadvantages. Firstly, the selection of points to be matched will not necessarily allow for a good reconstruction of the surface. Secondly, since the selection of pixels is not a function of the grey scale values on the image, many of the windows surrounding the selected pixels may not contain much grey scale variation, thus resulting in a greater number of unmatched points than would be expected if an interest operator were used to select the points to be matched.

For the purpose of mapping a complex, textured surface, such as an irregular rock surface, a feature extractor was required that identifies a large number of well distributed points. A further attribute of the feature extractor that was sought was that of allowing the user flexibility to alter the number of selected pixels while still maintaining a good distribution of points. This would allow the user some control over the length of time necessary for the image matching process. Two general classes of feature extractors were investigated, namely edge operators and interest operators. The attributes of these different types of feature extractors were investigated with respect to the selection of a representative number of interest points on a variety of test images of rock surfaces. Although the rock surfaces differed in terms of type of rock, colour and fracture patterns, no significant differences were found in the comparison of the performance of the various feature extractors. For this reason the comparison of the feature extractors in this section will be limited to only a single test image, shown in Figure 6.1(a). Although the sub-pixel accuracy moment preserving edge operator was investigated during the course of

this thesis, it's results are not included in this comparison since the image matching software that was used could only accept pixels to be matched to single pixel accuracy (see section 7.4.7).

6.1 CANNY EDGE OPERATOR

An efficient feature extractor is the Canny edge detector, previously described in section 4.2.2, which uses the first derivative of the Gaussian as a kernel to detect abrupt grey scale changes in an image. Gaussian smoothing allows for low pass filtering characteristics that reduces the inevitable influence of noise. Two variables effect the performance of this operator:

(i) the space constant σ, which determines the level of smoothing and the size of the convolution mask;
(ii) a threshold value, that is used to identify edgels as those pixels whose convolution magnitude, determined as a function of the x and y convolution values, is above the threshold value.

Flexible, user friendly, software was written to implement this algorithm. The value of σ chosen is image dependent and must be selected by the user. Once the convolution magnitude has been calculated for each pixel, the edgels are selected based on the threshold value and are initially displayed as a binary edge image. A second edge image is now created where the edgels are displayed proportionally to their convolution magnitudes (i.e. the stronger edges have proportionally higher grey scale values). The user now has the option of interactively thresholding out the weaker edges. Once the user is satisfied with the remaining thresholded edges, a thinning algorithm is used to select edgels with local convolution maxima values. The thinning process implemented is the same as that used by the Förstner interest operator for the suppression of local non-maxima pixels, as described in section 4.3.1. The size of the window used in the thinning process effects the number of interest pixels selected. The thinned pixels can be viewed overlaid on the edge image or can be viewed as plotted points against a uniform background. If the user is not satisfied with the results, the whole process can be repeated by again interactively thresholding out of the weaker edges, or by selecting a different size window used in the thinning process.

The combination of a low threshold value, allowing many edges to be detected, and a suitably large thinning window size, ensures a good distribution of thinned interest points even in areas on the image with low illumination, such as shadow regions. The Canny operator was tested on the test image of the rock surface containing significant amounts of shadow regions. In order to identify interest pixels in these regions with high local convolution values, but low values when considered over the entire image, it was necessary to use a threshold value as low as 1/20. This means that on a scale based on the difference from the lowest to the highest convolution value, all pixels with convolution

values of less than 1/20 th of this scale are rejected as edgels, while pixels with convolution values of greater than or equal to 1/20 th of the convolution range are regarded as edgels. In most images the majority of pixels are in regions of low grey scale variation, usually associated with non-edgels. The results of the Canny edge detection process on the 512 x 512 pixel test image are shown in Figure 6.1(a)-(d), using a threshold value of 1/20 and a σ value of 1. This σ value relates to a convolution mask size in the x and y directions of nine pixels.

A good distribution of approximately 3100 thinned points was produced using a window size for the thinning process of 7 x 7 pixels. From Figure 6.1(d) it can be seen that the extracted points were evenly distributed and that even in the darker, shadow regions of the image interest points were detected. Depending on the nature of the surface to be mapped, the time constraints to be considered for the image matching process, and the sampling interval required to create a representative DEM, the size of the thinning window can be varied.

6.2 FÖRSTNER INTEREST OPERATOR

The Förstner operator, previously described in section 4.3.1, is designed to detect points, corners and centres of circular features. The mathematical models to detect each of these classes of features are slightly different. For the purpose of this thesis computer code was written to detect corners (the intersection of straight lines). Five variables are required to implement this operator:

(i) a σ value, used in the formation of a Gaussian convolution mask;
(ii) a critical value c, determined empirically;
(iii)a roundness parameter, q_{lim};
(iv)the size of the window used to evaluate the pixel at the centre of the window;
(v) the size of the window used for the local non-maxima suppression of interest pixels.

Experiments with the test image found that the combination of $\sigma = 1$, $c = 2$, $q_{lim} = 0.5$ and with a 5 x 5 pixel window to evaluate the centre pixel, good results were produced. Using a 7 x 7 pixel window for the suppression of local non-maxima pixels, approximately 2250 interest points were extracted. This was roughly 10% of the original number of selected pixels that were located prior to the local non-maxima suppression process. Although the distribution of extracted pixels was fairly good, all the selected points were in well illuminated areas on the image, with very few selected pixels in the darker regions of the image (relating to shadows on the rock surface). The number of extracted points can easily be controlled by changing the size of the window used for the suppression of local non-maxima interest pixels. The distribution of the detected Förstner interest points is shown in Figure 6.2(b) alongside the original image shown in Figure 6.2(a).

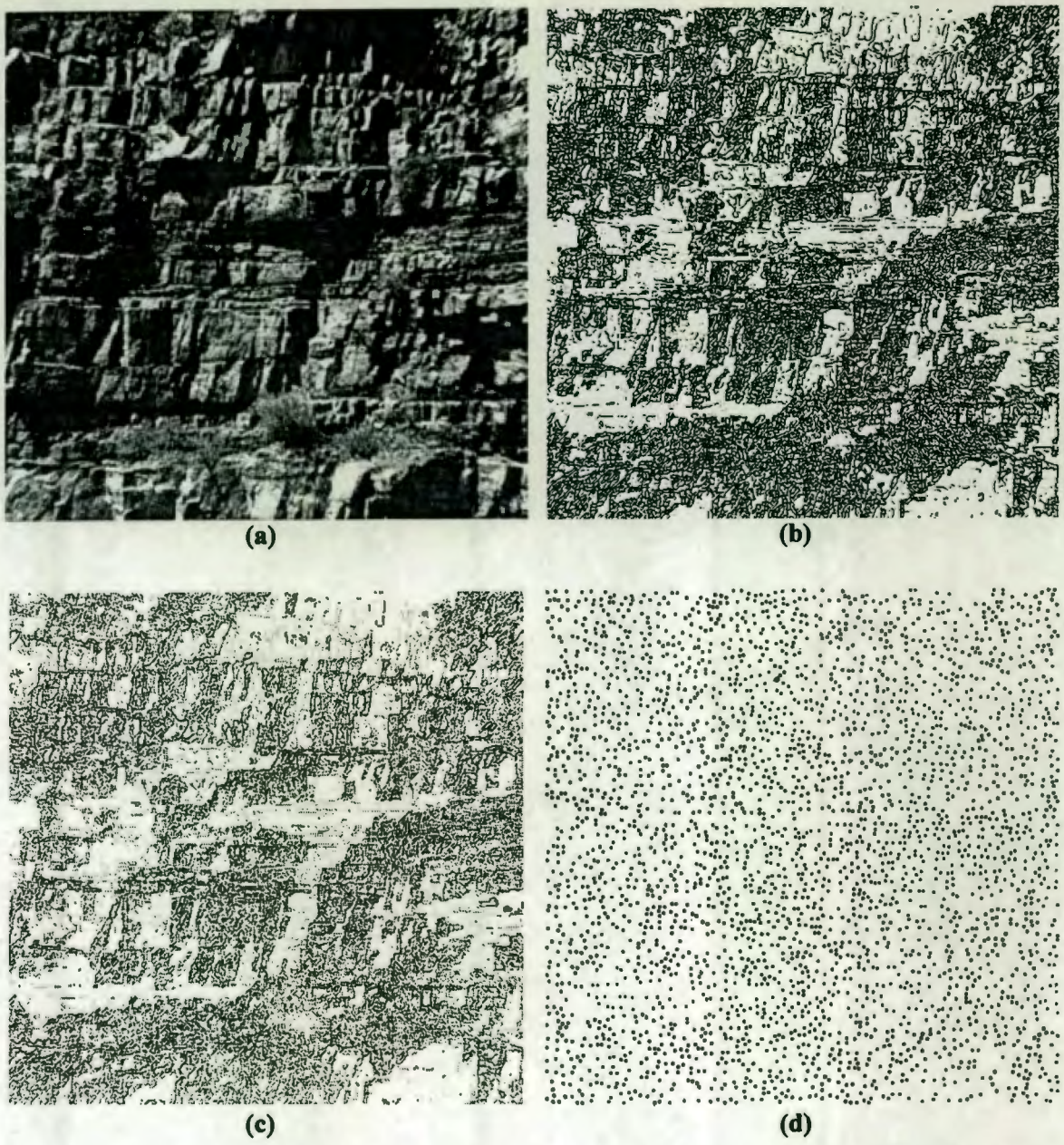
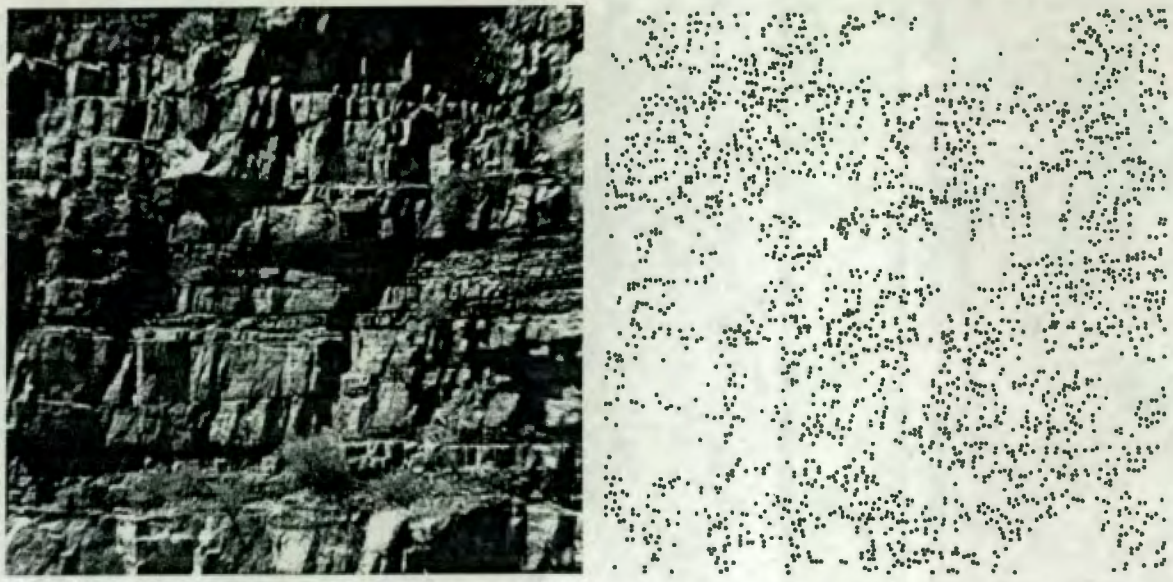


Figure 6.1: The sequence of the Canny edge detection process. (a) The original test image. (b) The detected binary edges. (c) The detected edges shown as a grey scale, proportional to their strength. (d) The thinned points using a 7 x 7 pixel window.

The effect of varying the input parameters on the number of pixels extracted is shown in a series of graphs in Figure 6.3(a)-(d). For the purpose of comparison, except for the variable under consideration, the remaining variables were held constant. When the parameters were held constant the values were set to $\sigma = 1$, $c = 2$, $q_{lim} = 0.5$, with a 5 x 5 pixel window to evaluate the centre pixel. A 7 x 7 pixel window for the suppression of local non-maxima pixels was used for all the graphs.



(a) (b)

Figure 6.2: (a) The original test image. (b) The detected Förstner interest points.

Referring to Figure 6.3(a) and (b) it can be seen that a decrease in the q_{lim} and c parameters produced a greater quantity of selected points. This result is to be expected, since as these values are decreased the conditions for selecting optimal windows that contain the pixels of interest become more flexible. These conditions, that the error ellipse should be both small and near-isotropic, relate to parameters c and q_{lim} respectively. The variations in c had a significant effect on the number of pixels selected whereas the variations in q_{lim} had a marginal effect. This indicates that the condition relating to the size of the error ellipse is more significant for the selection of points than the condition relating to the form of the error ellipse. Although an increased number of points can be selected by decreasing the values of c and q_{lim}, this results in the additional selected points having lower interest values. Thus the additional points are weaker interest points than those previously selected. An better strategy for increasing the number of points selected would be to use a smaller window for the suppression of non-maxima interest pixels.

From Figure 6.3(c) it can be seen that decreasing the size of the window used in evaluating the centre pixel has an effect of increasing the number of points selected. However, a window size as small as three pixels, that results in a large number of selected points, may not produce satisfactory results due to the small size of this convolution window. Varying the σ value has a marginal effect on the number of interest points selected, as seen in Figure 6.3(d).

A good feature extractor should produce a good distribution of points in addition to an adequate density of points. The effect of varying the input parameters on the spatial distribution of selected points was considered. Because the development of a mathematical function to compare the spatial distribution of points is complex, a visual

comparison of the distribution of points was made. Within areas of high illumination on the image the distribution of points increased as the values of c and the window size used to evaluate the centre pixel decreased. However, in regions of low illumination on the image, the distribution of points remained poor. Figure 6.4(a)-(b) indicates the different distribution of points for c=1 and c=2. For this comparison the values of the remaining four parameters were the same as those for the comparison of how the varying of the input parameters effected the number of pixels extracted. A decrease in the window size used to evaluate the centre pixel produced similar results. Variations in σ and q_{lim} had little effect on the distribution of points.

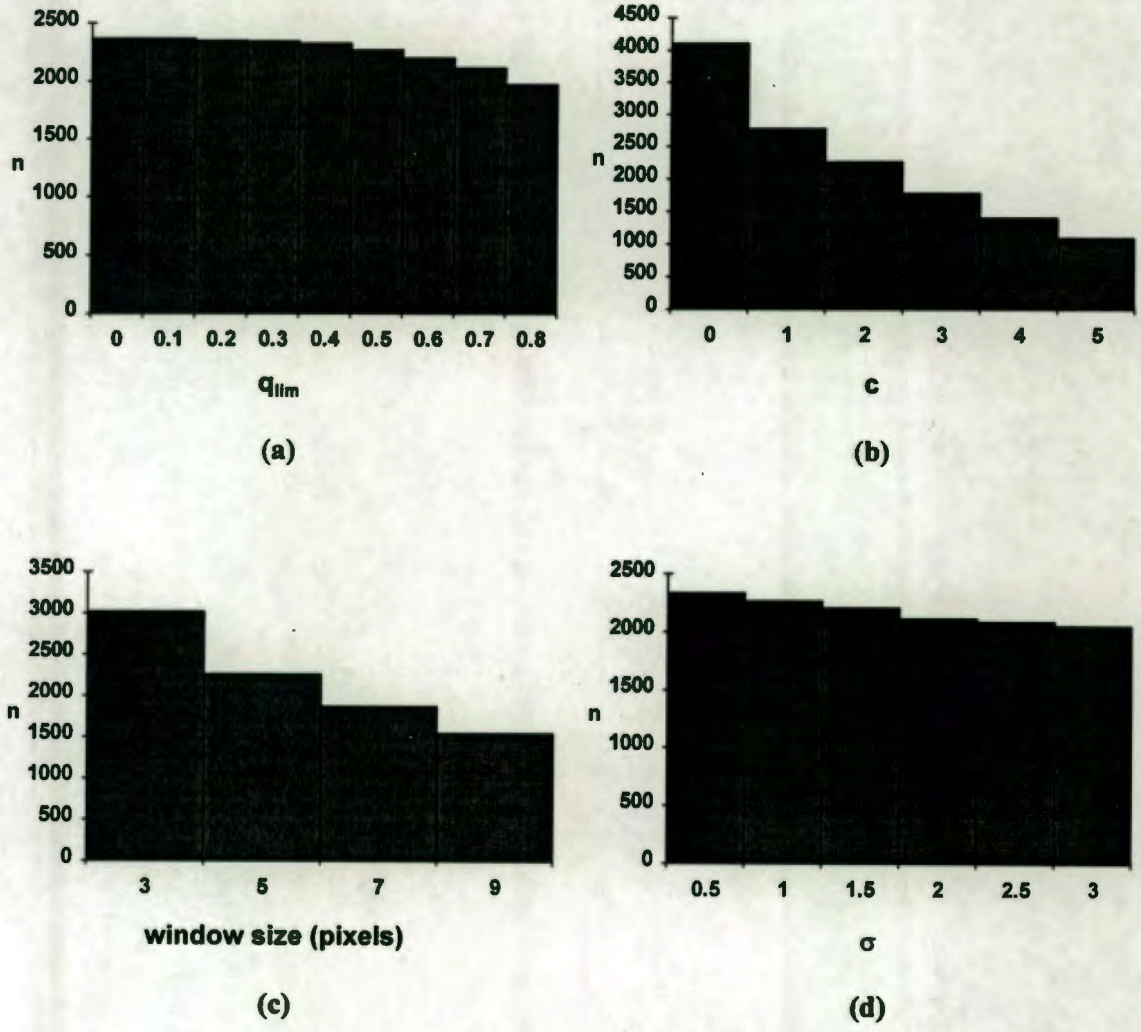


Figure 6.3: Graphs indicating the effect on the number of pixels selected, n, by varying (a) qlim, (b) c, (c) the window size used to evaluate the centre pixel, (d) σ.

A method of increasing the density of points to be matched in the darker regions on the image that contain a sparse distribution of selected interest points, is to select pixels at

regular intervals within these darker regions. However, this method has certain disadvantages that were discussed towards the beginning of this chapter.

Figure 6.4: The distribution of the detected Förstner interest points for (a) c=1 and (b) c=2, applied to the test image in Figure 6.2(a).

It should be noted that in many of the shadow regions there is indeed useful grey scale information that is important for adequately describing the rock surface (although this detail is not easily seen on the printed images such as in Figure 6.2(a)). It is important that feature extractors have the ability to extract points of interest in these regions. When comparing the spatial distribution of extracted points using the Canny edge operator, in Figure 6.1(d), to the spatial distribution using the Förstner interest operator, in Figure 6.2(b), it can easily be seen that the former operator produces a far superior distribution of points than the latter operator.

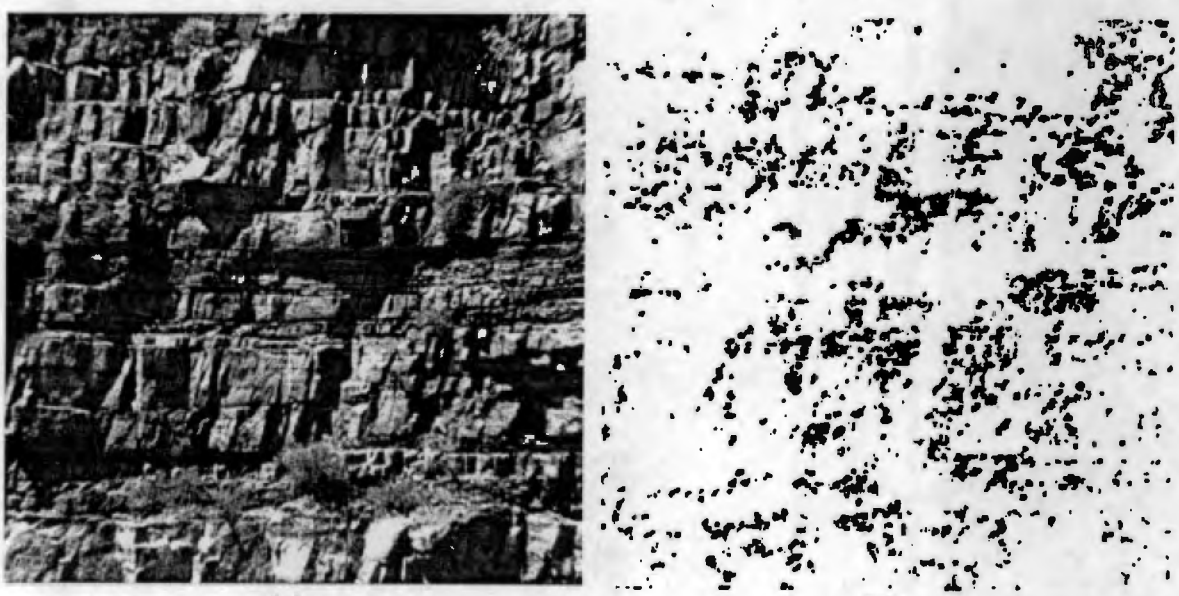
6.3 MORAVEC INTEREST OPERATOR

Software to test the Moravec interest operator was written by Ph.D. candidate Michelangelo Calitz of the Department of Surveying and Geodetic Engineering at UCT. This operator, described previously in section 4.3.2, identifies regions on an image with large grey scale variations. The selection of interest points is influenced by two variables:

(i) the size of the window of the square region on the image within which the operator evaluates the interest value of the pixel at the centre of the window;
(ii) a σ value relating to a threshold function that identifies the pixels with high interest values. The number of pixels chosen as interest points is a function of the area under the normal distribution function between the σ value and positive infinity.

Many feature extractors tend to be highly directional, vertically and horizontally, due to their use of convolution processes in the x and y directions. The Moravec operator, however, makes use of the four principal directions within a given window to determine the interest value at the centre pixel. When determining points of interest on textured surfaces that tend to have multi-directional aspects to it, this should be an advantage over the more two directional operators.

The result of passing this operator over the test image with a 5 x 5 pixel window and a σ value of 1 is shown in Figure 6.5(b) alongside the test image in Figure 6.5(a). A σ value of 1 results in a threshold being created whereby 15.9% of the pixels with the highest interest values are selected as interest points. The results shown in Figure 6.5(b) show the distribution of interest clumps (patches of interest points). A thinning routine for the suppression of local non-maxima pixels could be included if single interest points are required. (The software that was used did not contain a thinning function.)



(a) (b)

Figure 6.5: (a) The original test image. (b) Interest clumps detected using the Moravec interest operator.

This operator gave a sparse distribution of interest points that appeared to be highly regional within the image, with no interest points identified in the darker regions of the image. The method that was used to threshold out non-interest pixels limits the selection of interest points to those pixels above the given threshold value (calculated using the σ value). The same threshold value is used for the entire image. Thus pixels with high local interest values, but which are lower than the threshold level used over the whole image, are excluded. This results in a poor distribution of interest points, especially in areas on the image with poor illumination, due to among other factors, shadows or darker variations of rock type. The result is patches of interest clumps on the image with the areas between the patches devoid of interest pixels. One possible method to address this problem is that of tessellation. The image could be divided up into regularly shaped sub-regions with a separate threshold value being calculated for each region. A more effective variation of this would be to distinguish between the darker and lighter regions on the image and apply two separate threshold values to the respective regions.

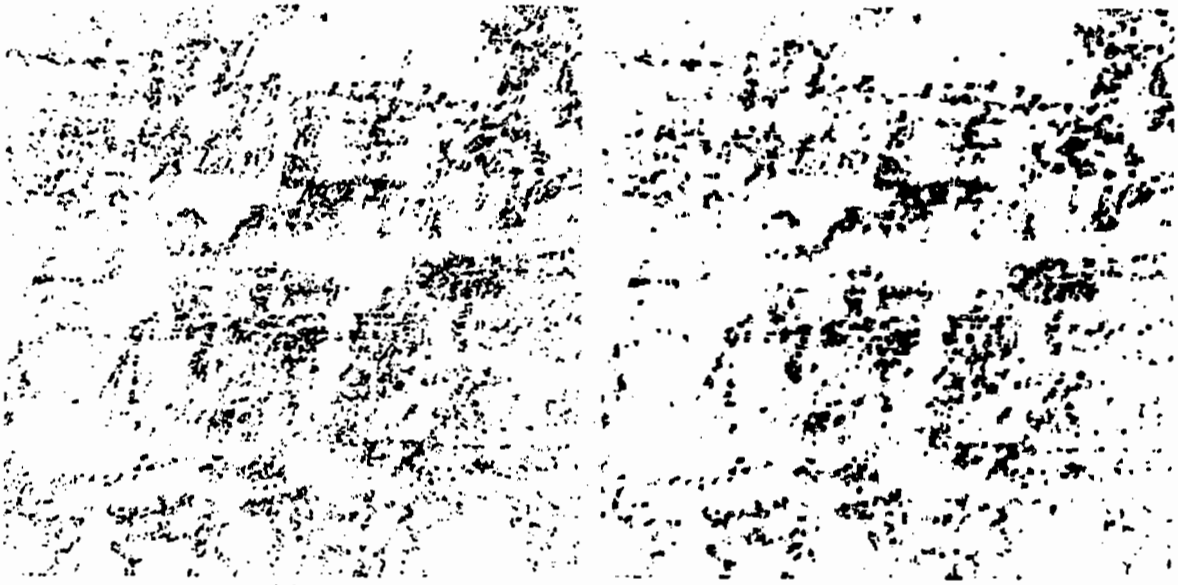


Figure 6.6: The distribution of the detected Moravec interest clumps with the size of the window used to evaluate the interest value at the centre pixel, equal to (a) 3 x 3 pixels and (b) 5 x 5 pixels, applied to the test image.

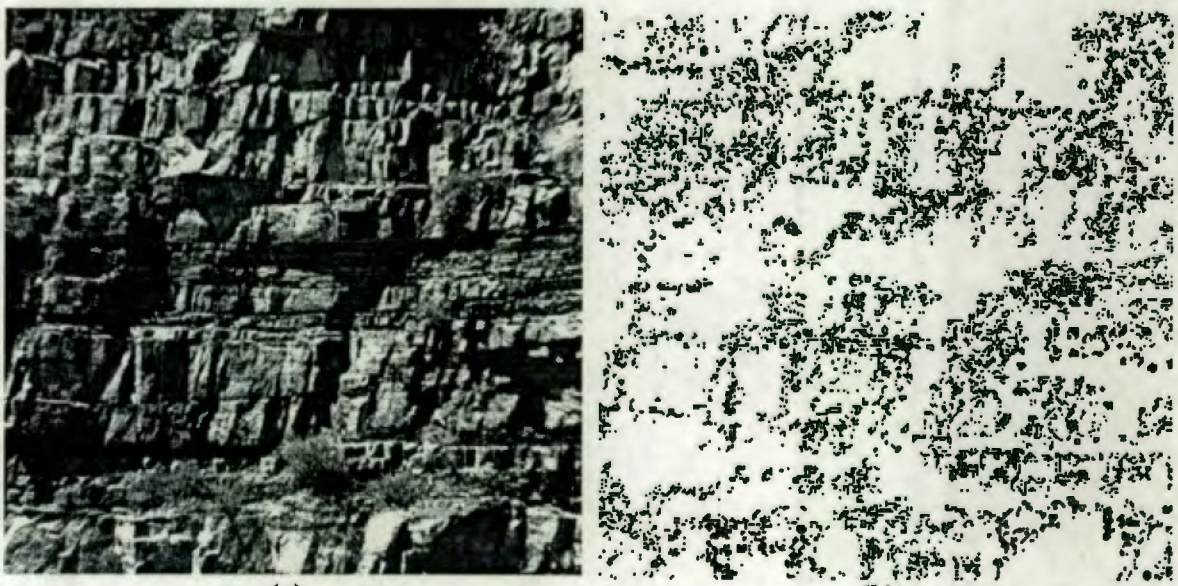
Decreasing the threshold level results in more interest pixels being selected in regions with already high distributions of points, while regions with poor distributions of points do not benefit. Varying the window size has no effect on the number of pixels selected since this is entirely dependent on the σ value. The window size does, however, have an effect on the distribution of selected pixels. As can be seen in Figure 6.6(a)-(b), the interest clumps tend to be smaller and better distributed when a smaller window size, of 3 x 3 pixels as opposed to 5 x 5 pixels, is used. An increase in the window size results in

higher interest values and produces a 'spill over' effect at the edges of regions with large grey scale variations. This effect results in high interest values occurring further from the true location of the regions of large grey scale variation than would be the case for smaller windows. The smaller the window size the more accurate the operator becomes by reducing this 'spill over' effect.

Altering the variables of this operator had little effect on increasing the distribution of the selected interest pixels. On the basis of the poor distribution of the interest points when applied to the test image, the Moravec interest operator was rejected for the application of detecting interest points on images of rock surfaces.

6.4 ANGLE BASED INTEREST OPERATOR

The angle based interest operator, described in section 4.3.3, was also tested using software designed and written by Michelangelo Calitz. This operator detects interest pixels based on significant changes in the direction of grey scale gradients. The edges on the image are found using the Sobel edge detector, described in section 4.2.1. (The threshold value used to identify the pixels as edgels is determined using the Otsu thresholding technique (Otsu, 1979). Although any thresholding technique can be used, this method has the advantage of being fully automatic.) The gradients in the x and y directions at every edge pixel are determined from the Sobel masks in the x and y directions.



(a) (b)

Figure 6.7: (a) The original test image. (b) Interest clumps detected using the angle based interest operator.

Referring to Figure 6.7(b) it can be seen that when the operator was passed over the test image of the rock surface a good distribution of points was produced in regions on the image with high grey scale values (i.e. lighter regions). However, in the darker, shadow regions of the image where no edgels had been located, no interest points were identified. This problem relates again to the thresholding problem discussed with respect to the Moravec operator in section 6.3. Edges that are locally strong, but weak when considered over the entire image, are discarded. A more efficient thresholding algorithm that retains locally strong edges could possibly produce a better distribution of points. For example, thresholding with the aid of tessellation, described briefly in the previous section, is more efficient than using one threshold value for the whole image.

A further limitation of this operator is that it does not produce interest values. Thus the suppression of local non-maxima pixels in the interest clumps is not possible and single, isolated interest pixels cannot be identified. This operator was rejected for its application to images of rock surfaces since it is produces poor distributions of points in areas of shadow, and since it does not produce single, isolated interest pixels.

6.5 CONCLUSIONS OF CHAPTER

All of the feature extractors discussed here are image dependant and sensitive to radiometric variations as well as noise. Image processing techniques to reduce the effects of noise and radiometric variations can be implemented. Using certain operators, the influence of noise can be reduced by increasing the convolution mask size, as in the case of the Canny and Förstner operators, or by using larger windows to determine the interest values. This effectively increases the low pass filtering effect, but has the disadvantage of causing signal reduction. All of the feature extractors investigated here, and probably most other feature extractors not included in this investigation, will, to some degree, incorrectly some select pixels that have little interest value, and will ignore some pixels that should be selected as interest points. This can be seen in Figure 6.8, where the Förstner interest operator was passed over an aerial image of urban environment with obvious interesting points.

Many image processing techniques exist incorporating further algorithms for noise removal, image sharpening, local enhancement, contrast manipulation and other optimisation techniques. The computer programming and processing time involved for implementation of these techniques would be significant. With the determination of interest pixels to single pixel accuracy, these additional algorithms were not considered necessary.

Figure 6.8: An aerial image of an urban environment with Förstner interest points overlaid, to illustrate the effect of the detection of non interest points and the non detection of obvious interest points.

Figure 6.9: A comparison of the distribution of selected pixels using the four feature extractors investigated, based on a portion of the test image. (a) Canny edge operator. (b) Förstner interest operator. (c) Moravec interest operator. (d) Angle based interest operator. These examples were extracted from Figures 6.1(d), 6.2(b), 6.6(a) and 6.7(b) respectively. In (a) and (b) each point represents a single interest pixel. In (c) and (d) interest clumps are shown.

A comparison of the distributions of interest points determined using the four feature extractors investigated in this thesis is shown in Figure 6.9(a)-(d). The bottom left hand corner of the test image was used for this comparison. The Canny edge operator proved the most successful in identifying pixels of interest of sufficient density and distribution. The Förstner, Moravec and angle based operators produced very similar distributions of points, with good distributions of points in well illuminated regions of the image but sparse distributions of points in the darker regions where the extraction of locally strong interest points was not achieved. The angle based interest operator had the additional

disadvantage of being able to only identify interest clumps and not interest points, since interest values were not produced. Both the Canny and Förstner operators were significantly more flexible and robust in their performance than the Moravec and angle based operators.

7. SYSTEM DESCRIPTION

Digital photogrammetric measurement requires the careful selection of both equipment and computer software. In this chapter, the system developed for the measurement of rock surfaces in open cast mining environments is discussed with respect to these two considerations. The configuration of the network to create a strong image geometry, will also be discussed. The measurement procedure, which is defined largely by the software structure, will be outlined. The methodology of the camera calibration procedure, used to determine the interior orientation parameters, is described in detail. In this thesis both a still video camera and a CCD video camera were used. The still video camera was used to acquire images for two test measurements of rock surfaces (see sections 8.1 and 8.2). A third test measurement made use of the CCD video camera to capture the images (see section 8.3). In this chapter, where relevant, the differences in the procedures relating to the use of the images acquired using the two camera systems will be noted.

7.1 CHARACTERISTICS OF THE MEASURING ENVIRONMENT

Open cast mines generally take the form of a series of rock steps, known as benches, that are excavated in a systematic manner. Depending on the size of the mine, it is divided into a number of sections and within each section only one bench is mined at a time. As the benches are mined, they retreat thereby enlarging the level area at the bottom of the bench. Once this level area has been made sufficiently large, the bench below the current level area can be excavated. In this way the benches are mined from the top to the bottom bench. Once the bottom bench has been reached the excavating can again begin from the top bench. The benches have a number of functions that are relevant to the measurement process. Firstly, they limit the height of the rock surface that is to be excavated and measured at any one time. For the mine this is important in terms of safety. Although the height of each bench is not regulated by law, the Department of Mineral Affairs encourages mines to not exceed a bench height of approximately 15 metres. Secondly, the horizontal components of the steps act as roads for trucks and other vehicles to gain access to all parts of the mine. This allows a sufficiently large area within which to position camera stations.

7.2 EQUIPMENT

The equipment used can be divided into a number of categories. The first category is the computer hardware that is required in order to download the images, if a still video camera is used, or to capture the images, if a CCD camera is used. The computer hardware is also needed to carry out the image and photogrammetric processing and for displaying the results. The second category of equipment is the type of camera system

employed. Both the still video camera and the CCD camera used, will be described. The last category is the additional equipment that is required.

7.2.1 Computer Hardware

An IBM compatible 486 DX personal computer with 8 MB RAM and a super VGA display driver, equipped with 1 MB of display memory, was used. This display memory provides a display resolution of 768 x 1024 pixels in 256 colours. Since the still video camera that was used produces images of 1012 x 1524 pixels, resulting in approximately 1.5 MB files for each image, a relatively large amount of hard drive space is required for purposes of image storage. Additional hard drive space must be available for the photogrammetric software as well as for the files that this software creates for the processing of the images.

7.2.2 Kodak Still Video CCD Camera

A single Kodak DCS 420 monocolour still video camera was used for the image acquisition of the first two test surfaces. Still video imagery has recently attracted much attention from the digital photogrammetric community and high accuracies using these cameras have been reported for a variety of industrial applications (Fraser and Shortis, 1995). Still video cameras can be distinguished from other CCD cameras by their on-board storage facilities. The DCS 420 camera has a 9.3 x 14 mm CCD sensor chip allowing for 1012 x 1524, 9 x 9 μm pixels. A removable PCMCIA 105 MB hard drive allows 65 images, of approximately 1.5 MB each, to be stored on a single drive. A PCMCIA adapter for a PC interface provides the ability to download the images to a PC. The camera body is that of a conventional camera and thus the typical wide selection of lenses can be accommodated. For the purpose of this project a wide angle Sigma lens with a 14 mm focal length was used, translating into a 37° x 53° field of view.

7.2.3 ITC CCD Video Camera

Low cost, low resolution, 'off the shelf' CCD video cameras have been widely used as image capture devices in close-range digital photogrammetry. Many practical examples exist of the ability to achieve high accuracy results, often to less than a millimetre, using these cameras. (For examples the reader is referred to Welsch (1995)). An 'off the shelf' CCD video camera manufactured by ITC was used for image capture for the third test surface measurement as described in section 8.3. The CCD sensor is of an interline transfer type and has a size of 7.95 x 6.45 mm with 795 x 596 pixels. A disadvantage of CCD video cameras, as opposed to still video CCD cameras, is that in order for it to be able to capture images the camera needs to be on-line to a computer with a framegrabber card. For this project a PIP-512B framegrabber/image processing card was used. It was

capable of digitising and capturing images with a maximum size of only 512 x 512 pixels. A video monitor was also necessary for displaying the output from the framegrabber in order to view the video images prior to capturing them. A wide angle Computar lens with a 6 mm focal length was used. This produced a 67° x 56° field of view.

7.2.4 Additional Equipment

The additional equipment that was used is discussed below.

- A tripod was used as a stable platform on which to mount the cameras.

- For the CCD video camera system an external power source was required to power the camera, computer and video monitor. A special transformer, designed to run off a cigarette lighter connection in a motor vehicle, was used to convert power from a 12 volt car battery into the required 230 volts. Still video cameras contain an internal power source.

- Circular targets, placed on the rock surface, were used as control points for camera orientation. The size and nature of these targets are discussed in section 7.3.4, relating to network design.

- A 100 metre tape measure was used to determine the very approximate positions of the camera stations and the targets.

- In order to place the middle row of targets on the excavation surface (refer to Figure 5.1), targets were attached to levelling poles which were placed against the rock surface. The use of levelling poles for this purpose is shown in Figure 8.2.

- For camera calibration for the determination of the interior orientation parameters of the camera, a control (calibration) frame, as shown in Figure 5.2, was used. This frame, of dimensions 700 x 700 x 250 mm, contained 90 circular retroreflective targets, the centres of which, had been previously coordinated to high accuracy. The targets were 8 mm in diameter. The frame was coated with black, non-reflective, matt paint (blackboard paint).

7.3 NETWORK DESIGN

Network design can be defined as 'the process by which the goal of precise, reliable and economic object measurement is achieved through the configuration of a suitable optical triangulation network' (Mason, 1995). Some of the factors effecting the design of the

network include the configuration of camera stations and control points, the number of images taken per camera station, and the size and nature of the targets, used as control points. These considerations will be discussed briefly with respect to the acquisition of images of the test surfaces. Factors relating to the network design for camera calibration, although similar to those described here, will be discussed more specifically in section 7.5.

7.3.1 Camera Configuration

Good network design requires the configuration of cameras to form a strong multi-station imaging geometry. It is generally accepted that multi-station configurations require a minimum of three non-parallel rays to intersect each target (Mason, 1995). This ensures both redundancy, for system reliability, and a good distribution of convergent rays. Mason defines a strong configuration as 'one in which all targets are intersected by at least three or four rays giving convergence angles in both the XZ and YZ coordinate planes of between 60° and 90°, or more'. This ensures that the error ellipsoids are near-homogeneous and near-isotropic. The large horizontal area at the base of the benches allowed ample space for the positioning of the camera stations in a good configuration in the XZ coordinate plane. In a mining environment it is difficult, costly and time consuming to position the cameras at different heights so as to ensure good convergence of rays in the YZ coordinate plane. The camera configuration that was used for the measurement of the test surfaces, described in chapter 8, consisted of five camera stations all at the same height. Figure 8.1 shows a typical camera configuration that was used. Although the positions of the camera stations were not optimal, they were satisfactory in providing a good multi-station camera configuration.

7.3.2 Number of Images per Camera Station

Multiple exposures at each camera station can be used to influence the precision and reliability of a network. The RMS error of object space coordinates can be approximated by (Fraser, 1995):

$$\bar{\sigma}_c = \frac{qS}{\sqrt{k}}\sigma \qquad 7.1$$

where k is the number of exposures at each camera station, 1:S is the approximate image scale, σ is the average image measurement coordinate precision, and q is an empirical factor relating to the geometric strength of the network. q can range from 0.4 to 0.7 for a reasonable convergent multi-station network geometry. From equation 7.1 it can be seen that although an increase in the number of images per camera station results in a decrease in the RMS error of the object space coordinates, the rate of increase of precision decreases as k increases. It should be noted that the increase in time required to acquire

additional images at existing camera stations is marginal. However, image matching is very computationally intensive and consideration needs to be given to the benefit of the additional accuracy compared to the significant increase in time required to process additional images. For the tests described in chapter 8, sufficient accuracy was produced using only a single image per camera station. Approximately seven hours of processing time was required for image matching using just five images.

7.3.3 Control Point Configuration

An ideal control point configuration is well distributed in three dimensions so as to minimise the correlations between the unknown parameters. Using the collinearity equation method of image orientation, as described in section 5.4, it is not necessary to spend any time surveying the positions of the targets. The number of control points used tends to increase the amount of time necessary to set up the targets, but does not significantly increase the amount of time required for their coordination, since the latter is done mainly by inspection of the image taken from the central camera station. A generally accepted network design criteria is that of imaging as many control points as possible. The number of targets per image effects both the precision and reliability of the network (Shortis and Hall, 1989). Redundant control also allows for flexibility if, as is sometimes the case, some of the targets are not detected due to factors such as occlusions or poor illumination. Thus, although the collinearity equation method of image orientation requires a minimum of three control points, eight targets were used for the test surface measurements.

The distribution of the eight control points used for the test measurements in this thesis is shown in Figure 5.1. A significant consideration in terms of the configuration is that of the physical placing of targets on the rock surface. Targets in the bottom row can be placed at ground level and at virtually any depth range. The top row of targets can typically be placed on the horizontal plane above the bench. Where the centre section of the surface, in terms of height, is not easily accessible due to the size of the surface, the centre row targets can be placed on extended levelling poles. As a result of the physical limitations, a poor distribution of control points in three dimensions was obtained. In the XY plane, however, a sufficiently good distribution was achieved.

7.3.4 Size and Nature of Targets

The size of the targets is dependent on the camera station configuration. The greater the distance between the targets and camera stations, the smaller the targets will appear on the image. The size of circular targets, in pixels on the image, has an effect on the repeatability of the measurements of these targets. Using least squares matching on targets with diameters ranging from 2 pixels to 20 pixels, Beyer (1992b) reports a significant improvement in repeatability up until a 6 pixel diameter target. Target diameters exceeding 11 pixels showed no improvement in repeatability accuracy. Beyer

concluded that a diameter of 6 pixels was the optimum target size. Similar results for target measurement using centre of mass algorithms can be expected. A rule of thumb in digital photogrammetry states that targets should have diameters in the order of 5 to 10 pixels. Fraser (1995) reports high accuracy results for a photogrammetric survey of a large scale engineering structure using circular targets to define the surface with diameters as low as 3 pixels.

The optimum diameter of a target in object space can be determined using the following equation:

$$t = d\frac{t_{pix}s}{cn} \qquad 7.2$$

where d is the object space distance from the camera station to the target, c is the principal distance of the camera lens, t_{pix} is the diameter of the target in pixels on the image, s is the physical length of the CCD sensor chip, and n is the number of pixels in each row (or column, depending on which dimension of the chip s relates to) of the sensor chip. The ratio s/n represents the length of each pixel on the sensor chip.

Retroreflective targets are preferred for most photogrammetric measurement tasks since they allow for uniformly exposed, high contrast targets. Due to the high cost of retroreflective material an alternative target material was required for this project. White paper disks placed on square sheets of cardboard that had been coated with black non-reflective, matt blackboard paint was found to provide adequate contrast. These were then mounted on hardboard for rigidity.

The size of the black board around the target is critical for the detection of the target on the image. It was found that a three to one ratio for the length of the side of the black board to the diameter of the target was sufficient for good contrast and target identification. Smaller black backgrounds often result in the targets not being detected.

7.4 SOFTWARE

The main component of digital photogrammetric systems is software. Once the images are captured and the control points coordinated, it is the software that drives the measurement process. It is for this reason that, excluding the field component of the system, the software structure is defined by the measurement procedure (see section 7.4.2). Some software was written specifically for this project and is discussed below. Other software required for this project was written by fellow postgraduate students in the Department of Surveying and Geodetic Engineering at UCT. This software is discussed more briefly. A third category of software that was used comprised commercial software.

All of the software, excluding the commercial software, that required the image files (i.e. the full grey scale images), had the limitation of using only 512 x 512 pixel image formats. This severely restricted the use of this software and limited the potential accuracy of the measurement system. The reason for using this format is a result of initially only having access to CCD video cameras that required a framegrabber card only capable of handling images with a maximum size of 512 x 512 pixels. It was only at a much later stage that a larger image format camera, in the form of the Kodak still video camera, became available. The code written specifically for this thesis was not adapted to allow for any image size, since the image matching software used, and which would have required considerable effort to change, could also only accommodate image formats of 512 x 512 pixels. All the final software used in the project was specifically written to have a user-friendly interface.

7.4.1 Programming Language

The software developed for this project, both to test the execution of the algorithms and to produce working software, was written in C. The main compiler used was GNU C, which incorporates a DOS extender to allow easy access to the full memory of the computer. This had advantages when working with and displaying images. A Borland C compiler was used for some programs that did not require images or large data files. The Borland C software had extensive help and debugging functions that the GNU C software lacked.

7.4.2 Software Structure and Measurement Procedure

The four main programs used in the surface measurement system were either written specifically for this project or were written by fellow postgraduate students in the Department of Surveying and Geodetic Engineering at UCT. These programs are listed below along with a short description of each of them.

(i) **Target measurement and DLT software.** This software was specifically written for this project. It consists of two components integrated into one package. Firstly, it locates circular targets on a single image, followed by the measurement of their centres to sub-pixel accuracy. If only the image coordinates of the targets are required, the user can exit after this stage. The second component calculates approximate interior and exterior orientation parameters of the image by means of the DLT adjustment. The object space coordinates of the targets (control points) must be known and listed in a data file along with the target name. An initial DLT adjustment is carried out using at least six detected control points that the user manually identifies. Using the initial DLT parameters the remaining targets are automatically named (using a method described in section 7.4.5) and subsequently a second DLT adjustment based on all the identified targets is calculated. For the process of camera

calibration (see section 7.5), both components of this software are used. Similar software is available within the department, but this is divided into two separate programs. The new software is more versatile and has capabilities for both flexible user interactions as well as for a high degree of automation. It is described in detail in section 7.4.5.

(ii) **Bundle adjustment software.** This highly flexible bundle adjustment package, written by Ph.D. candidate Graeme van der Vlugt of the Department of Surveying and Geodetic Engineering at UCT, was used for a number of specific functions in the measurement system. Firstly, it is used for the purposes of camera calibration in order to obtain the interior orientation parameters of the camera (see section 7.5). The software is also used for the collinearity equation method of determining approximate rotation parameters for single images of the surface to be measured (see section 5.4). The third function of this software is to carry out a final bundle adjustment, using the previously determined interior orientation parameters, to obtain final exterior orientation parameters for the images of the surface. This software is discussed further in section 7.4.6.

(iii) **Edge detection software.** Specifically written for this project, this software is used to extract features of interest based on the Canny operator. It allows for easy interactive manipulation of the edges by the user for the purpose of thresholding out weaker edges (if desired) and choosing a suitable window size for the thinning routine. This software was previously described in section 6.1.

(iv) **Multi-photo geometrically constrained (MPGC) matching software.** This software, written by Graeme van der Vlugt as part of his Ph.D. thesis, determines the positions of conjugate points in multiple images to a high degree of accuracy and produces a DEM of the surface. A description of this software can be found in section 7.4.7.

The secondary software that was used is listed below. These programs, although necessary, fall into the categories of being either commercial software or trivial to write.

(i) **IPhoto Deluxe image processing package.** This commercial software was used to download images from the PCMCIA card, when the still video camera was used. IPhoto Deluxe is also a general, low level image processing package that has a wide variety of functions, such as, tone adjustment, contrast and brightness enhancement, edge detection, cropping, rotating, zooming, resampling, etc. The downloading of images is discussed in section 7.4.3.

(ii) **Image reduction software.** This was written specifically for this project. It reduces 1012 x 1524 pixel images acquired using the Kodak still video camera, to a 512 x 512 pixel image size. The method of image reduction and its consequences are discussed in section 7.4.4.

(iii) **Surfer surface display package.** This commercial software package was used to generate plots of the test surfaces using the DEM produced from the

MPGC matching software. It contains a wide range of options for displaying surface data.

Because the photogrammetric measurement in this system is carried out digitally on a personal computer, the measurement procedure and the software structure are integrally linked. The measurement procedure, showing the relationship between the various software components, is illustrated in Figure 7.1.

7.4.3 Downloading of Images

After capturing the images using the Kodak still video camera, the images were downloaded using a PCMCIA adapter with the aid of the IPhoto Deluxe image processing package. The package produced image files in tagged image file format (TIFF). However, all the non-commercial software that was used, and that made use of image files, required these files be in a flat image format. The conversion from TIFF to flat image formats was carried out using appropriate additional software.

The drivers that were supplied by Kodak for downloading of the images were found to be incorrect. This resulted in sections of the image with grey scale values of 255 being distorted. For example, white circular targets become teardrop shaped. Ph.D. candidate Julian Smit of the Department of Surveying and Geodetic Engineering at UCT investigated this effect and found no distortion for grey scale values less than 255. The effect of this distortion on the targets appears to be limited as camera calibration, using images reduced in size to 512 x 512 pixels, commonly produced a posteriori standard deviations of unit weight as low as 0.0003 (see section 7.5). This driver error, however, is still a factor to be considered with respect to the final accuracy of the mapping of the textured surface. Unfortunately, once the images are removed from the PCMCIA hard drive, the images cannot be corrected.

7.4.4 Image Reduction

The Kodak DCS 420 still video camera produced images with a size of 1012 x 1524 pixels. In order to cater for the software that required 512 x 512 pixel image formats, the images had to be reduced. This was done by a combination of cropping the image and taking the average grey scale value of blocks of four pixels to form a single pixel. The effect of this averaging procedure is that useful image information is lost and the quality and sharpness of the image is reduced. This is equivalent to passing a low pass filter over the image. A further consequence of image reduction is the effect on the interior orientation parameters of the camera. The numerical values of these parameters are reduced by a factor of two for the averaging of blocks of four pixels.

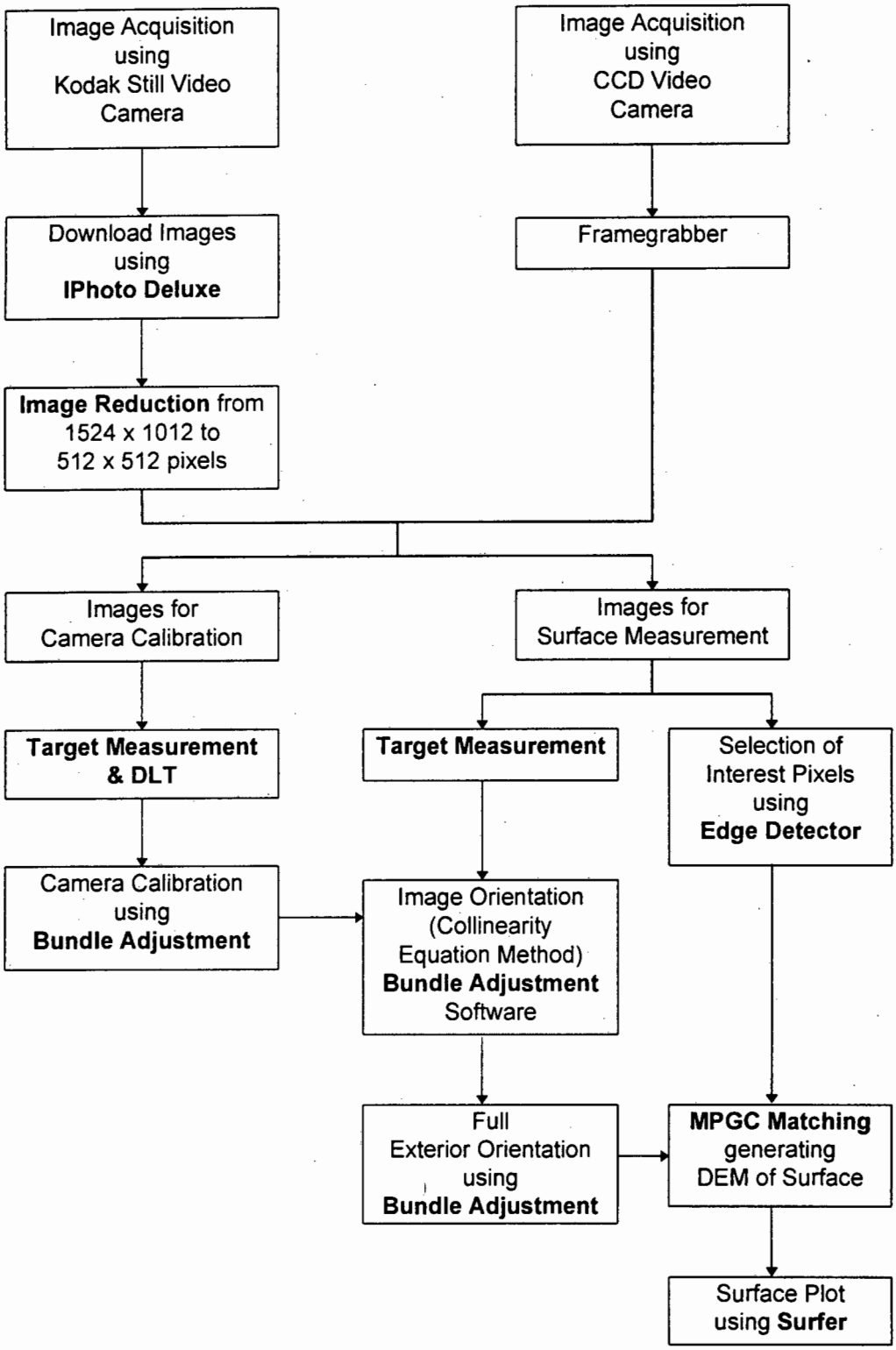


Figure 7.1: A flowchart of the measurement procedure indicating the relationship between the various software components of the measurement system. The first two vertical branches of the flowchart relate to the process of using either the still video camera or the CCD video camera. The bold print indicates the specific software used in the system.

Another image reduction technique that could have been implemented is that of resampling the image. However, regardless of the technique used, the loss of image data will always have a significant negative effect on the accuracy of the surface measurement system.

7.4.5 Target Measurement and DLT Software

This software was designed to identify white circular targets on a black background. The required 512 x 512 pixel flat format image is displayed on the computer screen. The user interactively thresholds the image until an optimal thresholded image is produced in which all the targets are displayed. All the areas on the image where the pixels contain grey scale values above the threshold level are automatically identified as potential targets. Almost always, but especially when detecting targets on images of the surface to be measured, many areas that are not the correct circular targets are identified as potential targets. Working systematically through the thresholded image, an algorithm searches for potential targets. As each potential target is detected, its extent, in terms of pixels, is determined by means of chain coding. Those that do not satisfy certain predetermined target parameters are rejected as targets. These target parameters must be included in a file required to run the software. Potential targets are rejected based on the following criteria:

(i) minimum and maximum target extents in the x and y directions;
(ii) the ratio of x to y target extents;
(iii)the ratio of the number of pixels within the detected area to the number of pixels within a rectangle, defined by the x and y target extents. This criteria rejects targets whose shape does not conform to that of a typical target, but which still satisfies the first and second criterias.

As each target is accepted or rejected, based on the above criteria, the pixels making up the target area are tagged so that they will not be searched again. In this way, even in an extreme case, as shown in Figure 7.2, where a first target is completely surrounded by a second target, both targets will be located and independently subjected to the target geometry criteria. A less efficient, but faster method of ensuring that the targets are not re-detected, is to tag all the pixels within a block as defined by the minimum and maximum extents. In many cases the latter method results in the actual targets not being correctly identified.

Once all the potential targets have been located and subjected to the criteria described above, the extents of all the accepted targets are shown by overlaying a rectangle over the original image. The user now has the option of adding target areas that were not detected or were incorrectly rejected. This is done by defining the target extents with the aid of a mouse controlled cursor. Care must be taken when defining the target extents to ensure that no pixels with grey scale values above the threshold value are included within the defined area. If any such grey scale values are included within this area, the centre of

mass of the target, as determined using the weighted centre of mass algorithm, will be displaced. It should be noted that often non-targets will pass the target acceptance criteria. The user has the option of manually, with the aid of a mouse, selecting these targets to be removed.

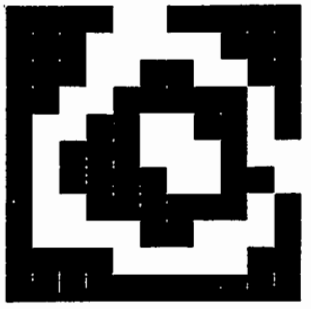


Figure 7.2: Two potential targets, a ring target and a solid target, located by means of thresholding, are subject to the target parameter criteria independently.

The sub-pixel centres of each identified target is now calculated by means of the weighted centre of mass algorithm with threshold subtraction. The threshold value used in this step can be the same value determined interactively, or can be a value determined using either of the automatic threshold methods as given in equations 4.4 and 4.5. The user now has the option of naming some or all of the remaining targets, by selecting detected targets with the aid of a mouse. At this point the program may be terminated since the target measurement and naming process has been completed. A file of the image coordinates and names, if applicable, of all the detected targets is created.

The second component of this software, that of determining the approximate interior and exterior parameters of the image, requires a file with the names and related object space coordinates of all the control points. A minimum of six targets must be selected and named by the user, as described in the previous paragraph. Based on these targets, initial DLT parameters of the image are calculated in a least squares adjustment. These DLT parameters are used to determine image coordinates of the remaining control points based on the object space coordinates in the control file. By comparing the calculated image coordinates to the measured image coordinates, within a predetermined tolerance, all the control points can be named automatically. A full DLT adjustment can now be carried out based on all the identified control points. The final DLT parameters are then converted into the nine parameters of interior and exterior orientation. In the DLT model, additional parameters were not included since the parameters determined from the DLT are merely used as approximations to the unknown orientation parameters in a subsequent bundle adjustment.

If good knowledge of the target attributes is available, the user can allow the system to run with minimal interaction. The initial target criteria parameters must be set in a file

prior to the software being run. This file also contains the tolerance value required for the automatic naming of all the control points. The minimum user interaction that is required is the interactive thresholding of the image and the naming of at least six control points.

The interactive thresholding step could be eliminated by either using an automatic thresholding algorithm, such as a method described by Otsu (1979), or by using an empirically determined threshold value. Complete automation of this software is possible if the naming of control points is not required. This could be done by including a circular bar code around the target for the automatic identification of control points (Van den Heuvel, Kroon and Le Poole, 1992). Good image quality would be required to reduce the possibility of mis-identifications.

7.4.6 Bundle Adjustment Software

The bundle adjustment software that was used is very flexible in terms of being able to adjust various types of networks under different measurement circumstances. The adjustment can take the form of either a free network or a constraint network. Intersections or resections or both can be processed. The minimum number of images that can be used is dependent only on the minimum solution of the various measurement tasks. For example, a minimum of two images is required for intersections, and for resections the minimum solution requires only a single image. The solution, using multiple images, can be adjusted using either different interior orientation parameters for each camera, or the same interior orientation parameters for all the cameras. Distance observations in object space can be included.

The software allows for the inclusion of seven additional parameters, namely the first three terms of radial lens distortion, the two terms of decentring lens distortion, a scale correction factor in the x direction and an image shear parameter. The model used is described in section 3.8. A full error analysis is also provided. This includes all of the standard deviations of the unknown parameters, the degrees of freedom, the a posteriori standard deviation of unit weight, and the residuals of all the observations. The correlations between the camera parameters, and between the camera parameters and the station parameters, are also determined. A list of the observations that can be considered as outliers, where the residual of the observation is greater than three times the a posteriori standard deviation of unit, is also produced.

7.4.7 MPGC Matching Software

The MPGC matching software that was used was based on the procedure developed by Van der Vlugt and Rüther (1994), described previously in section 4.5. The advantage of this method of matching over other methods of MPGC matching, such as that developed by Baltsavias (1991), is that the former is specifically designed to overcome the problem of determining good provisional positions of the conjugate points when the surface to be

measured is complex and discontinuous. Rock surfaces found at excavation sites typically fall into this category, since they tend to be sharp and jagged in nature, with large variations in depth within small areas. The latter method is better suited to smooth, continuous surfaces.

The software, Digital Image Measurement System (DIMS), was designed by Graeme van der Vlugt as one of the major components of a digital photogrammetric surface measurement system. Besides image matching, the software has a number of other image processing functions. These include target detection, target centring and feature extraction. Only the image matching component of DIMS was used in the measurement system developed for this thesis. DIMS was limited in two significant ways. Firstly, it could only use 512 x 512 pixel images, thus reducing the potential accuracy of the measurement system. Secondly, the software only allowed for the image coordinates of the interest points to be matched, to be defined to single pixel accuracy. In many applications it is desirable to match features whose image coordinates have been determined to sub-pixel accuracy. Another image matching program, written by Ph.D. candidate Julian Smit of the Department of Surveying and Geodetic Engineering at UCT, was also available. This software overcame the limitations of DIMS, but was limited to using a maximum of four images. DIMS was written using a GNU C compiler and thus is able to make use of the full memory of the computer. The number of images that can be processed is limited only by computer hardware and time constraints.

The image matching component of DIMS can either be manually driven, where the user uses a mouse to indicate which point on a reference image is to be matched, or it can be automatically driven using a file of the image coordinates of interest points on the reference image. DIMS has the ability to create a file of the image coordinates of the interest points using one of its two in-built feature extractors, namely the Förstner interest operator and a simple gradient edge operator. An externally created file of the interest points can also be used to drive the automated matching process. For this thesis a file of interest points determined using the Canny edge operator was used (see section 6.1). Although any image of the surface can be used as the reference image, in this thesis the most 'normal' image (i.e. the central image) was used for this purpose.

The extent of the search along the epipolar line (determined using the collinearity equation method - see section 3.7.2), for MPGC matching, and along the reference ray, for MIC, is limited by the minimum and maximum Z coordinates of the control points. Since the top row of control points (one of which contains the maximum Z coordinate) is placed directly on the surface to be measured, the search along the epipolar line and the reference ray, based on this maximum Z coordinate, is not always sufficient to match the entire surface. To avoid this problem a fictitious control point was created, specifically for this matching software, with a Z coordinate slightly greater than the maximum Z coordinate of the real control points. Good correlations of the conjugate points in every image was required in order for the match to be successful.

The model used to account for additional parameters, was the same as that used in the bundle adjustment (see section 7.4.6). In addition to creating a file of the surface DEM coordinates, DIMS produces a comprehensive analysis of every successfully matched point. This includes the standard deviations of the object space coordinates of each point, image coordinates and their standard deviations of all the conjugate points, the correlation values with respect to the reference patch, a σ_0 value as well as the number of iterations taken for matching each point.

7.5 CAMERA CALIBRATION PROCEDURE

The method of camera calibration used in this thesis to determine the interior orientation parameters of the camera will be described. A procedure generally referred to as pre-calibration or test-range calibration, as opposed to self-calibration, was employed. In spite of the implications of the name, the process of pre-calibration does not necessarily have to precede the acquisition of images of the surface to be measured, and for this thesis camera calibration was always performed after the images of the surface had been acquired. The reason for this is that it is often necessary to correctly focus the camera on the surface to be measured to ensure optimum image quality. The camera is then calibrated for this focal setting. It should be noted, however, that if the distance from the camera stations to the surface always exceeds the minimum distance that is required for the camera to be focused to infinity, as in the case for this thesis, or if the approximate principal distance is known from previous image acquisitions, calibration can take place in the strictly pre-calibration sense. The 14 mm Sigma lens used with the Kodak still video camera allowed for the camera to be focused to infinity for distances greater than approximately 2 metres. For the 6 mm Computar lens used with the ITC CCD camera, this distance was approximately 0.5 metres.

Once the camera has been focused, the principal distance must remain constant for the duration of the image acquisition of the surface to be measured as well as for the image acquisition of the control frame necessary for camera calibration. This was done by securely wrapping masking tape around the focus adjustment dial of the camera so as to prevent any movement of the focal setting.

During the image acquisition of the calibration frame the illumination of the frame is of critical importance to ensure good contrast between the targets and the background and to reduce the presence of non-target regions on the image that contain high grey scale values. Good contrast between the retroreflective targets and the background is possible if the frame is surrounded by non-reflective black cloth. By positioning the frame in a dark room and by using a camera flash, for the Kodak still video camera, or a strong light positioned behind the camera, for the ITC CCD camera, near optimum image quality can be attained.

The repeatability of measurements in image space, also known as internal precision, based on images taken with a CCD video camera, is affected by the length of time between switching on the camera and capturing the images (Beyer, 1992b). The camera only reaches a stable condition after it has warmed up sufficiently. The ITC CCD video camera was warmed up for two hours prior to capturing the images of the calibration frame. Still video cameras are not impeded by this effect.

For camera calibration a set of nine images of the frame is taken in a convergent, symmetrical configuration, using three rows of images at different heights. The camera positions are chosen so as to maximise the control frame in the field of view. The scale of the calibration frame results in very close range images of the frame with the result that the images may be slightly out of focus, particularly when using the 14 mm Sigma lens with the Kodak still video camera. For the circular targets on the calibration frame the blurring is symmetrical about the centre of the target. The accuracy of the weighted centre of mass target centring algorithm with threshold subtraction is not affected by such symmetrical effects, so long as an appropriate threshold level is chosen (see section 4.1.3).

Using the target measurement and DLT software, described in section 7.4.5, initial interior and exterior orientation parameters are determined separately for each image. These values are used as approximations in a subsequent free network bundle adjustment, adjusting all the images simultaneously. The adjustment solves for only a single set of interior orientation parameters, relating to the fact that a single camera was used for the capture of all the images. One of the distances between two of the control points was held fixed in order to provide a scale. The first parameter of radial lens distortion, k_1, the two decentring parameters, P_1 and P_2, the scale correction factor in the x direction, s_x, and the image shear parameter, a, were solved for. (Solving for all seven additional parameters and studying the correlations between the interior orientation parameters and the additional parameters provides a useful method of determining which additional parameters are necessary to be included in the adjustment. The second and third parameters of radial lens distortion are very highly correlated to k_1, and thus it is not necessary to include k_2 and k_3 in the adjustment. Although P_1 and P_2 are relatively highly correlated to x_p and y_p respectively, they were included to be rigorous. k_1, s_x and a have low correlations to any other parameters and were thus included in the adjustment.)

As a result of image perspective, calibration frame design, and shadows, it frequently occurs that some targets are partly occluded. This results in the displacement of the image coordinates of the occluded targets. These targets are, however, often correctly identified in the automatic naming routine of the target measurement and DLT software. Such targets can easily be identified after the bundle adjustment software has been run, by inspecting the observation outliers which are listed in the output file. Outliers are considered to be observations where either the x or y image coordinate residuals are greater than three times the a posteriori standard deviation of unit weight. The image coordinates of the observation outliers can be eliminated and the whole bundle adjustment recalculated. This process can be iterated until no more observation outliers

exist or until the value of the a posteriori standard deviation of unit weight is sufficiently low. Typically an average of five iterations is required until this stage is reached.

A summary of the final results of three typical camera calibrations using nine images of the control frame is shown in Table 7.1. The first two adjustments were carried out using images acquired with the Kodak still video camera. In the first calibration adjustment no additional parameters were included. The second calibration included five additional parameters. The improvement factor 1 indicates the improvement of the second calibration over the first. For both of these adjustments the same set of 512 x 512 pixel images, reduced from a 1012 x 1524 image format, were used. The third calibration used images acquired with the ITC CCD camera. The five additional parameters were included in the adjustment. The improvement factor 2 indicates the improvement of the adjustment using the Kodak still video camera, and adjusting for the additional parameters, over the calibration adjustment using the ITC CCD camera.

		Kodak No AP's	Kodak 5 AP's	ITC 5 AP's	Improvement Factor 1	Improvement Factor 2
	σ_0	0.0014	0.0003	0.0005	4.7	1.7
RMS Error (mm)	σ_X	0.155	0.070	0.062	2.2	0.9
	σ_Y	0.156	0.061	0.059	2.6	1.0
	σ_Z	0.346	0.098	0.085	3.5	0.9
	v_x	0.0012	0.0002	0.0004	6.0	2.0
	v_y	0.0013	0.0004	0.0004	3.3	1.0
Std. Deviation (mm)	σ_{xp}	0.012	0.0077	0.0053	1.6	0.7
	σ_{yp}	0.016	0.0073	0.0055	2.2	0.8
	σ_c	0.011	0.0029	0.0020	3.8	0.7

Table 7.1: Summary results of typical camera calibration adjustments using images acquired with the Kodak still video camera and the ITC CCD video camera. The first improvement factor serves as a comparison between the two calibration sets using the Kodak images. The second improvement factor indicates the improvement of the second calibration set over the third.

Considering the two calibration adjustments carried out using the images acquired with the Kodak still video camera, the improvement factor 1 clearly shows a significant improvement of the results when the five additional parameters are included. No conclusive improvement was noted between the second and third adjustments.

Referring to the calibration using the images acquired with the Kodak still video camera where the additional parameters were included, the first time the bundle adjustment was run 18 observation outliers were detected. All the outliers were inspected on the respective images and 15 were found to be due to occlusions. It is possible that the remaining outliers were a result of the tear drop effect of the targets, distorting the correct

target centres (see section 7.4.3). After five iterations of eliminating the observation outliers no more outliers were identified and no further increase in accuracy was attained. The degrees of freedom for the final adjustment was 1089. The data presented in Table 7.1 indicates the relatively high accuracy to which the camera can be calibrated.

Two factors that affect the accuracy of the calibration results:

(i) The tear drop effect, that results in displaced centres of the circular targets. This effect can be eliminated by using the correct drivers when downloading the images;
(ii) The size of the circular targets on the image. As explained in section 7.3.4 the size of the targets have an effect on target centring algorithms. To achieve the maximum target centre accuracy, target diameters should exceed five to six pixels. For the calibrations used for the purpose of this thesis, using the reduced images of the 512 x 512 pixel formats, many of the target diameters were as small as four pixels. By using the full 1012 x 1524 images all of the targets would have diameters that would result in the maximum possible target centring accuracy.

In order to compare the effect of target size on the calibration process a calibration was carried out using both the 1012 x 1524 images and their associated 512 x 512 images acquired with the Kodak still video camera. Software written in the Department of Surveying and Geodetic Engineering, that could accommodate any size images, was used for both calibration adjustments. When comparing the RMS errors of the standard deviations of the control points, improvement factors of 1.2, 1.2 and 1.1 were obtained for X, Y and Z respectively. The RMS error of the residuals for the image coordinates was almost exactly different by a factor of two. This relates to the reduction in image size by a factor of two in the x and y directions. The results of this comparison showed a small, yet significant improvement in the accuracy of the calibration with targets sizes exceeding the recommended minimum size.

The accuracy of the interior orientation parameters can be further improved through the use of more images resulting in greater redundancy in the system. The determinability of the interior orientation parameters can be enhanced by using different image scales and by increasing the camera roll angles (Shortis and Hall, 1989).

8. SURFACE MEASUREMENT TEST RESULTS

8.1 CAPE TOWN HARBOUR EXCAVATION

An excavation site for a new yacht basin at the Cape Town harbour provided a good example of a quarry surface for the surface measurement system to be tested on. The type of rock at this site was Malmesbury Shale, which is very dark in colour. A large amount of rubble was situated at the base of the vertical section of the bench. The height of the bench that was measured was approximately seven metres. A large flat area below the bench provided ample area for the positioning of camera stations in a suitable configuration. A simple network geometry using five camera stations, with one image per station, was used. The camera stations were positioned in a row roughly parallel to the surface. This was to allow for fast and easy determination of very approximate camera station coordinates. Provisional values for the object space coordinates of the targets were determined by inspecting the image taken from the central camera station (see section 5.4). The distance to the cameras from the rock surface was chosen so as to fill the image frame taken from the central camera station, with the surface to be measured. The positioning of the camera stations allowed for a high degree of image convergency. Although this was not an optimal network design, it ensured a simple and rapid image acquisition procedure that could easily be implemented by a non-photogrammetric expert. The network geometry is illustrated in Figure 8.1. The coordinate system used, was that as previously illustrated in Figure 5.3. A seven by seven metre extent of the rock surface was measured.

Eight circular targets, with 10 cm diameters, were used. The configuration of these points on the rock surface is shown in Figure 8.2. The bottom row of targets was placed towards the camera stations to provide some degree of depth for the control point configuration. The difference in the depth between the top and bottom rows of targets was approximately 5 metres. Extended levelling poles were used to position the centre row of targets against the surface.

The images were captured using the Kodak DCS 420 camera with a 14 mm Sigma lens. The interior orientation parameters, including five additional parameters (k_1, P_1, P_2, s_x, a), were calculated by means of the camera calibration procedure described in section 7.5. All the observations were set to unit weight for this adjustment and the unit of measure was millimetres. An a posteriori standard deviation of unit weight of 0.0003 was attained. This relates to an average standard deviation of the image coordinates of 0.3 μm. Approximations for the exterior orientation parameters of the images of the rock surface were obtained using the collinearity equation method, as described in section 5.4. These parameters were used as provisional values in a free network bundle adjustment in which the images of the rock surface were oriented. A single distance between two of the targets was held fixed to provide a scale for the network. The results of the free network

bundle adjustment are summarised in Table 8.1. In this table σ_X, σ_Y and σ_Z refer to the standard deviations of the object point coordinates, whereas σ_{Xc}, σ_{Yc} and σ_{Zc} refer to the standard deviations of the perspective centre coordinates of the camera stations. Again, all the observations had unit weight and the unit of measure was millimetres. The σ_0 value of 0.0012 relates to an average standard deviation of the image coordinates of 1.2 μm. The results shown in Table 8.1 indicate relatively high accuracies for the orientation of the images.

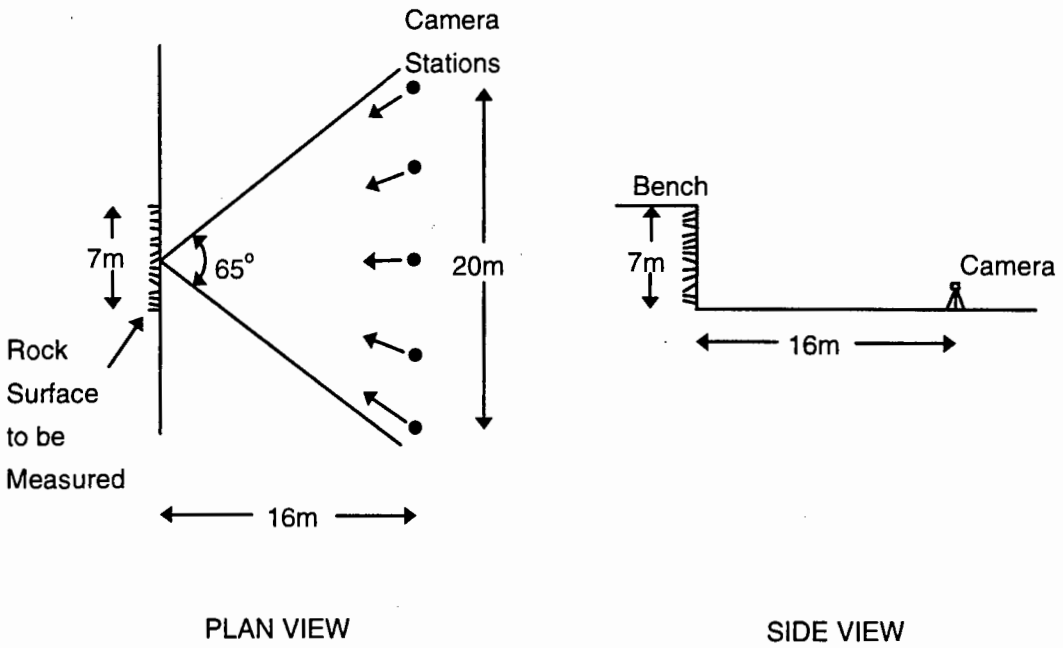


Figure 8.1: The network geometry of the Cape Town harbour test surface (dimensions are approximate).

The Canny edge operator with a subsequent thinning routine, that used a 7 x 7 pixel window, was used to select the interest pixels to be matched. 2807 well distributed interest points were selected using the image taken from the central camera station as the reference image. The MPGC matching was carried out using the software written by Ph.D. candidate Graeme van der Vlugt of the Department of Surveying and Geodetic Engineering at UCT (see section 7.4.7), with a reference patch size of 5 x 5 pixels and with the maximum number of iterations set to 30. A summary of the results of the MPGC matching of the points on the rock surface is shown in Table 8.2. The approximate accuracies in X and Y relative to the surface size and the approximate accuracy in Z relative to the average distance from the camera stations to the targets, is given in Table 8.3.

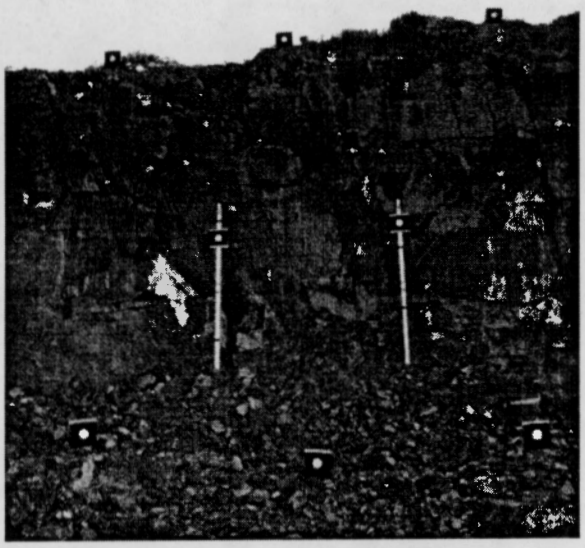


Figure 8.2: Digital image of the test surface at the harbour excavation site as seen from the central camera station.

	df	31
	σ_0	0.0012
RMS (mm)	v_x	0.0006
	v_y	0.0008
	σ_X	2.87
	σ_Y	2.37
	σ_Z	2.81
	σ_{Xc}	17.0
	σ_{Yc}	17.0
	σ_{Zc}	15.2
RMS (degrees)	σ_ω	0.069
	σ_κ	0.068
	σ_ϕ	0.027

Table 8.1: Summary results of the free network bundle adjustment used to orient the 5 images of the harbour test surface.

Average Iterations	RMS					
	σ_0	σ_x (pixels)	σ_y (pixels)	σ_X (mm)	σ_Y (mm)	σ_Z (mm)
22.4	4.0	0.06	0.01	0.53	0.43	2.22

Table 8.2: Summary results of the MPGC matching on the harbour test surface.

X	Y	Z
1: 13200	1: 16300	1: 7200

Table 8.3: The approximate relative accuracies in the X, Y and Z directions for the harbour test surface.

Of the 2807 points that were entered into the matching routine, 1641 points were successfully matched. The large number of unsuccessfully matched points can be primarily attributed to the large amount of lose rock, or scree, especially at the base of the vertical section of the bench, where steep edges of individual rocks prevented many points from being seen from all the image perspectives. On the vertical section of the rock surface, most points could be seen in all images, and a much lower incidence of unsuccessful matches occurred. The matching took approximately seven hours to run. Approximately 30 points were successfully matched per square metre of the rock surface. Although a denser distribution of points is preferable, the additional processing time must be carefully considered.

An orthographic plot of the test surface, generated from the DEM that was produced by the matching software, is shown in Figure 8.3. This was created using the surface display package Surfer.

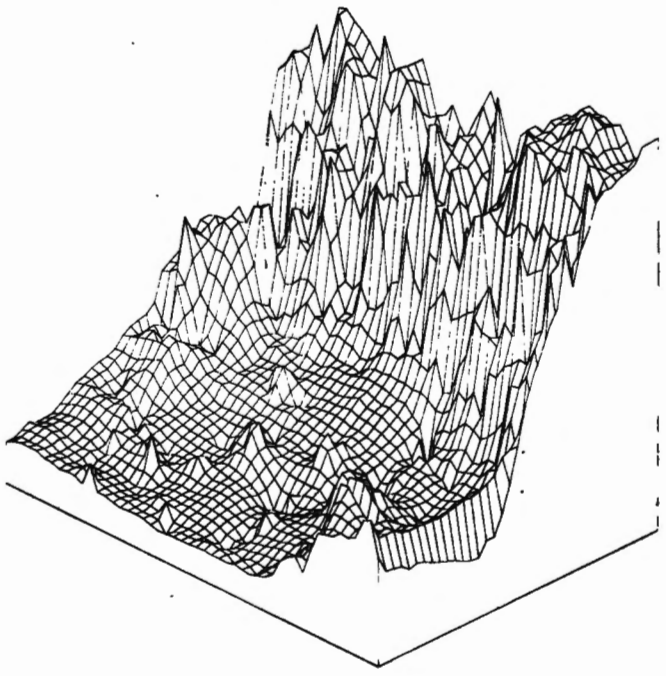


Figure 8.3: An orthographic plot of the harbour excavation test surface generated by the Surfer surface display package.

8.2 ROCK WALL

A retaining wall below Rhodes Memorial in Cape Town provided another test surface for the surface measurement system. The wall consisted of large, lightly coloured sandstone blocks, and sloped slightly towards the cameras, as indicated in the side view of Figure 8.4. Five camera stations, imaging one frame per station, were used. Due to a sharp drop-off of the ground close to the wall, the height of the perspective centres of the cameras was approximately at the same height as the bottom row of targets. Although this led to a network configuration that was far from optimal, sufficient accuracies were still obtained. The network geometry is depicted in Figure 8.4.

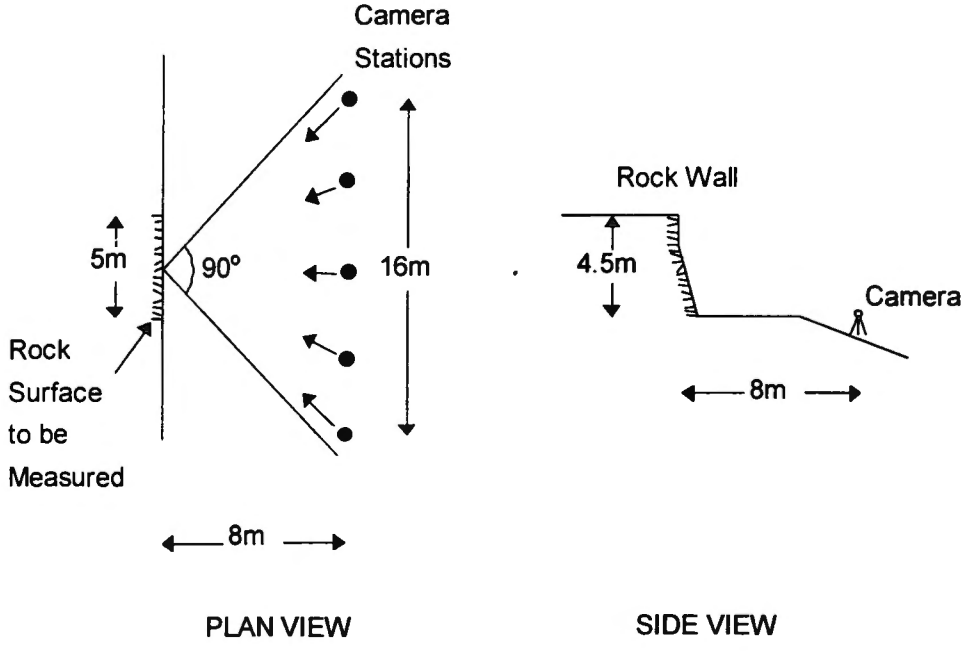


Figure 8.4: The network geometry of the rock wall test surface (dimensions are approximate).

A similar control point configuration as that used for the Cape Town harbour test, was employed. All of the targets had 10 cm diameters. Figure 8.5 shows the test surface with the target configuration. The difference in the depth between the top row and bottom row of targets was approximately 2 metres.

The images were captured using the Kodak DCS 420 camera with the 14 mm Sigma lens. Camera calibration, solving for the interior orientation parameters as well as five additional parameters, yielded an a posteriori standard deviation of unit weight of 0.0003. For the bundle adjustments for both camera calibration and the orientation of the images of the rock wall, all the observations were of unit weight and the unit of measure was millimetres. A summary of the free network bundle adjustment used for the orientation of the images of the rock surface is shown in Table 8.4. These results indicate a higher accuracy compared to the images of the harbour excavation site. This can partly be

attributed to the fact that for the harbour images some of the targets furthest away from the camera stations had diameters as low as four pixels, thereby resulting in lower accuracies of the target centres. For the images of the rock wall, all the targets had diameters of at least six pixels. The σ_0 value of 0.0006 indicates an average standard deviation of the image coordinates of 0.6 μm.

Figure 8.5: Digital image of the rock wall test surface as seen from the central camera station. (The dark area towards the top right hand side of the image is a shadow of a light pole.)

	df	33
	σ_0	0.0006
RMS (mm)	v_x	0.0004
	v_y	0.0003
	σ_X	0.84
	σ_Y	0.80
	σ_Z	0.71
	σ_{Xc}	6.3
	σ_{Yc}	6.2
	σ_{Zc}	6.1
RMS (degrees)	σ_ω	0.043
	σ_κ	0.044
	σ_ϕ	0.012

Table 8.4: Summary results of the free network bundle adjustment used to orient the 5 images of the rock wall test surface.

Average Iterations	RMS					
	σ_0	σ_x (pixels)	σ_y (pixels)	σ_X (mm)	σ_Y (mm)	σ_Z (mm)
21.3	3.4	0.06	0.01	0.25	0.36	1.26

Table 8.5: Summary results of the MPGC matching on the rock wall test surface.

X	Y	Z
1: 20000	1: 12500	1: 6300

Table 8.6: The approximate relative accuracies in the X, Y and Z directions for the rock wall test surface.

As in the previous example, the points to be matched were selected using the Canny edge operator with a 7 x 7 pixel window size for the thinning routine. The image taken from the central camera station was used as the reference image. 2651 well distributed interest points were located and selected. MPGC matching was carried out with a 5 x 5 pixel reference patch size and with the iteration maximum set to 30. The summary results of the MPGC matching of the rock wall is given in Table 8.5. The approximate accuracies in X and Y relative to the surface size and the approximate accuracy in Z relative to the average distance from the camera stations to the targets, is shown in Table 8.6.

1898 points were successfully matched. As with the surface at the harbour, many of the points that were unsuccessfully matched were distributed towards the bottom of the image, where the grass resulted in the effect that many of the interest points could not be seen from all image perspectives. The matching took approximately seven hours to run. On average 80 points were successfully matched per square metre of the test surface. An orthographic plot of the test surface, generated from the DEM, is shown in Figure 8.6.

A comparison of the results of the photogrammetric measurement system with independent measurements, made using a theodolite, was carried out. Twenty four discrete points on the rock wall that could be identified on the image taken from the central camera station were measured using the two methods. The points were chosen to fall in a profile strip running down the centre of the section of the rock wall that had been imaged. The photogrammetric measurement of the points was made using the manual option in the DIMS image matching software (described in section 7.4.7), whereby a cursor controlled mouse can be used to identify pixels on the surface to be matched. The theodolite measurements took the form of intersections to targets placed on or as close as possible to the points. The targets consisted of stickers with a cross in the centre. The theodolite measurements and the placing of the targets on the surface was done after the acquisition of the images of the surface and thus the stickers did not contaminate the images. A three dimensional transformation of the set of points measured with the

theodolite into the set of points measured photogrammetrically was carried out using software written by Ph.D. candidate Graeme van der Vlugt, of the Department of Surveying and Geodetic Engineering at UCT. This software used Rodrigues parameters to model the transformation. Due to the difficulty of placing the cursor on exactly the same point on the image that was intersected using the theodolite, six of the points were identified as having poor corresponding positions based on their large residuals after the transformation of all 24 points. The transformation was carried out again, using the remaining 18 points. The RMS errors of the residuals of the coordinates was 8 mm, 7 mm and 5 mm for X, Y and Z respectively. Considering the approximate method by which the points were identified, as well as the possible errors of the theodolite observations, the results of the transformation can be considered to be conservative. This comparison demonstrates a high degree of agreement between the results obtained by means of the two methods of measurement.

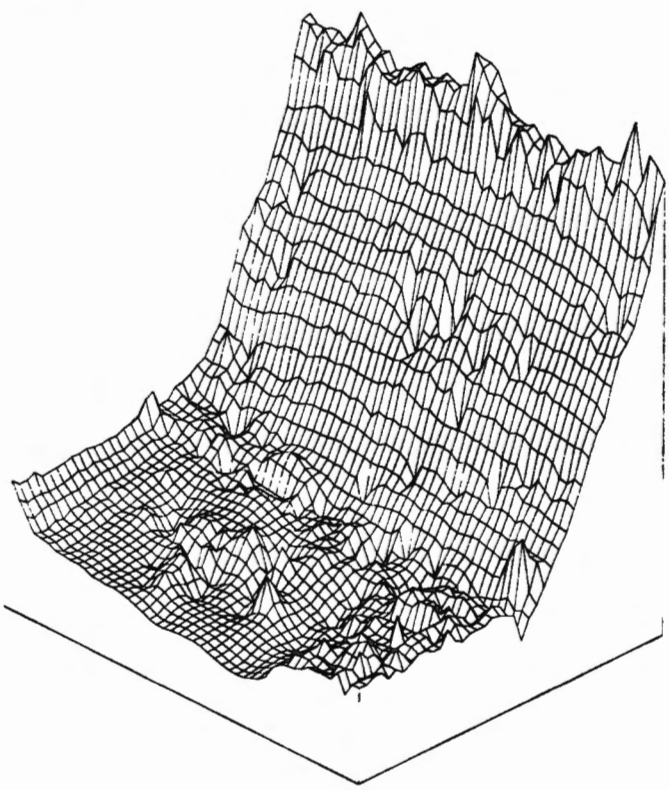


Figure 8.6: An orthographic plot of the rock wall test surface generated by the Surfer surface display package.

8.3 TYGERBERG HILLS QUARRY SITE

The third set of test measurements was taken at a large working quarry in the Tygerberg Hills just outside of Cape Town. The rock that is mined here is Malmesbury Shale and is used primarily for road aggregate. The benches at this quarry range in height from

approximately 5 metres to 12 metres. For the test measurement a section of the rock 12 metres high was chosen. A large amount of rubble was located at the base of the vertical section of the rock surface. The flat area at the base of the bench provided ample space to position the camera stations. The configuration of the network is shown in Figure 8.7. The ITC CCD video camera with a 6 mm Computar lens was used to capture the images. As reported in section 7.2.3, in order to acquire the images, the camera needs to be on-line to a computer. A video monitor is also necessary to view the images prior to capturing them. This computer hardware remained in a motor vehicle for the duration of the image acquisition process. The flat area at the base of the bench proved suitably level for manoeuvring the vehicle to any required position.

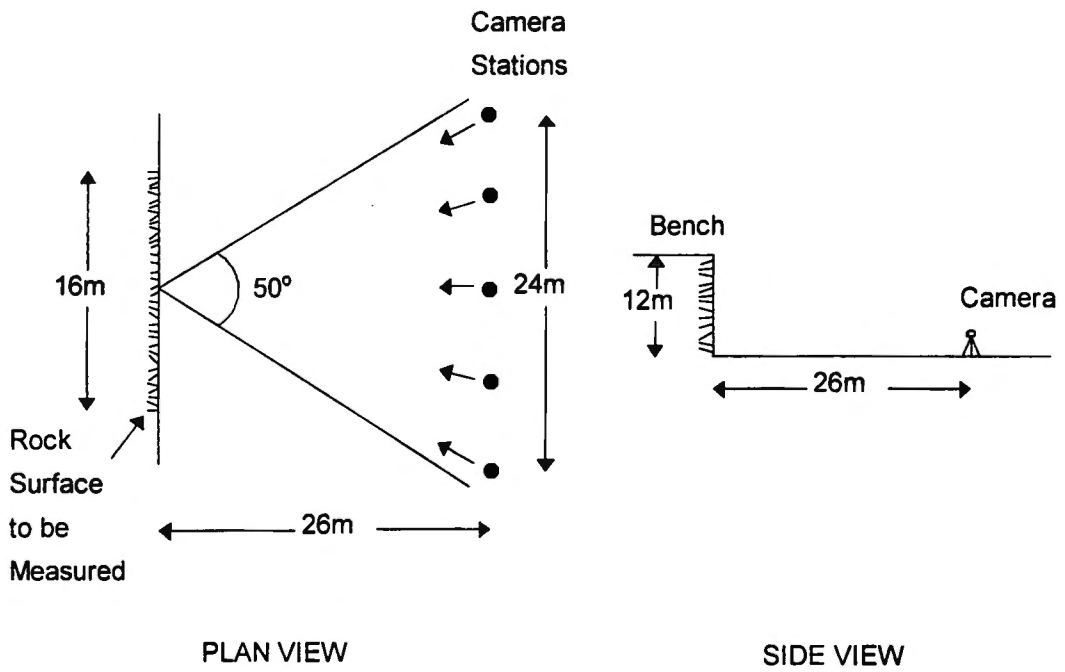


Figure 8.7: The network geometry of the Tygerberg Hills quarry test surface (dimensions are approximate).

The control point configuration was the same as that used for the previous two test measurements. All of the targets had diameters of 18 cm. Figure 8.8 shows the test surface with the target configuration. The difference in the depth between the top and bottom rows of targets was approximately 10 metres.

Prior to capturing the images of both the calibration frame and the test surface, the camera was warmed up for approximately two hours in order to allow the camera to reach a stable condition. (This camera warm-up effect was mentioned briefly in section 7.5. A thorough investigation of this effect can be found in Beyer (1992b).) Camera calibration, solving for the interior orientation parameters as well as five additional parameters, resulted in an a posteriori standard deviation of unit weight of 0.0005. As with the

previous two test measurements all the observations were set to unit weight and the unit of measure was millimetres.

Figure 8.8: Digital image of the test surface at the Tygerberg Hills quarry site.

The ITC CCD video camera tended to be very sensitive to the lighting conditions at the quarry. The weather conditions on the day on which the images were taken was that of strong sunshine interrupted by short periods of cloud cover. When the surface was in direct sunlight it was impossible to acquire suitable images of the rock surface due to the problem of too much reflectance. It was only during the periods of cloud cover that good images could be captured. The final image taken was acquired from the camera station on the extreme left (looking at the surface) and was captured just as the sun reappeared from behind the clouds. (At this point it was clear that no further cloud cover would be occurring in the short to medium time frame.) As a result this image proved unsuitable for image matching. However, the targets in this image were sufficiently well defined for their image coordinates to be included in the free network bundle adjustment that was used to determine the exterior orientation parameters of the images of the surface. A better solution for the exterior orientation parameters was produced when all five images were used, compared to when the fifth image was excluded from the adjustment. A summary of the results using the five images is given in Table 8.7. The σ_0 value of 0.0033 indicates an average standard deviation of the image coordinates of 3.3 μm.

The lower accuracy of the results shown in Table 8.7 compared to the results shown in sections 8.1 and 8.2 can be attributed to a number of factors:

(i) The images taken with the ITC CCD video camera were of a lower resolution compared to the images taken with the Kodak still video camera;
(ii) The quarry network geometry was not as strong as the geometry used in the first two test measurements;
(iii)Some of the targets on the images of the quarry test surface have diameters as low as four pixels, resulting in lower accuracies of the target centres;
(iv)The scale of the surface measured for the quarry is significantly larger than the surfaces measured in the previous two test measurements.

	df	31
	σ_0	0.0033
RMS (mm)	v_x	0.0022
	v_y	0.0019
	σ_X	22.7
	σ_Y	18.9
	σ_Z	23.0
	σ_{Xc}	95.8
	σ_{Yc}	64.0
	σ_{Zc}	82.9
RMS (degrees)	σ_ω	0.18
	σ_κ	0.27
	σ_ϕ	0.07

Table 8.7: Summary results of the free network bundle adjustment used to orient the 5 images of the quarry test surface.

Average Iterations	RMS					
	σ_0	σ_x (pixels)	σ_y (pixels)	σ_X (mm)	σ_Y (mm)	σ_Z (mm)
23.6	5.0	0.06	0.01	3.2	2.9	9.9

Table 8.8: Summary results of the MPGC matching on the quarry test surface.

X	Y	Z
1: 5100	1: 4200	1: 2600

Table 8.9: The approximate relative accuracies in the X, Y and Z directions for the quarry test surface.

The Canny edge operator as well as a subsequent thinning routine, using a 7 x 7 pixel window, selected 2238 interest pixels to be matched. The image taken from the central camera station was used as the reference image. MPGC matching was carried out using only four images and with a reference patch size of 5 x 5 pixels. The results of the matching are summarised in Table 8.8. The approximate accuracies in X and Y relative to the surface size and the approximate accuracy in Z relative to the average distance from the camera stations to the targets is given in Table 8.9.

Only 1282 points, out of the 2238 interest points that were entered into the matching routine, were successfully matched. The large number of unsuccessfully matched points can be primarily attributed to the plentiful amount of loose rock and scree at the base of the vertical section of the bench, where many interest points, selected using the Canny operator, could only be seen from one or two image perspectives. The matching took approximately six hours to run. An average of only 7 points were successfully matched per square metre of the test surface. Such a density of points is inadequate for all but the most approximate representations of the surface. A higher density of points can easily be chosen for matching by decreasing the size of the window used in the process of thinning the interest points that were selected by means of the Canny edge operator. However, the time that would be necessary to process these additional points would be considerable. An orthographic plot of the test surface, generated from the DEM, is shown in Figure 8.9.

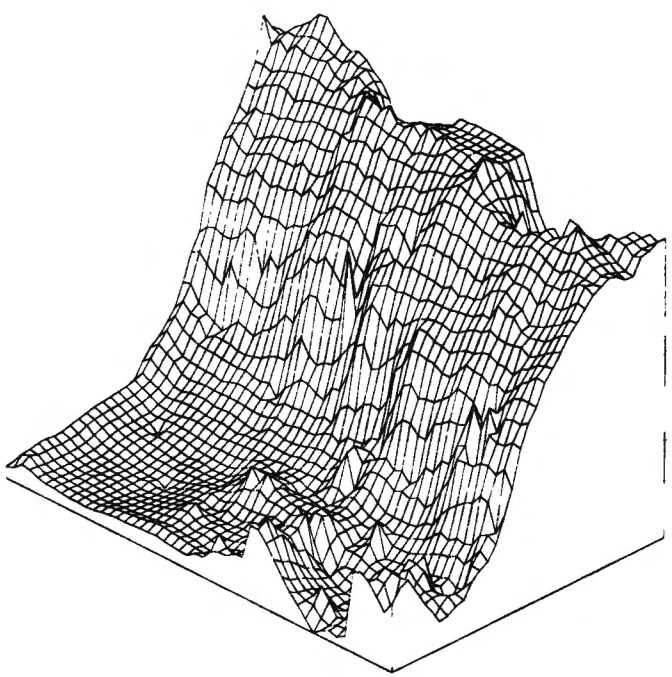


Figure 8.9: An orthographic plot of the Tygerberg Hills quarry test surface generated by the Surfer surface display package.

9. CONCLUSIONS

In this thesis a digital photogrammetric surface measurement system capable of measuring textured surfaces for industrial applications, was developed. The system was specifically designed to measure rock surfaces in open cast mining environments. The conclusions drawn from the research for this thesis are discussed below.

The mapping of all three test surfaces was successfully performed. The results indicate that the system is capable of relatively high accuracies. For the measurement of the selected points on the first two test surfaces (when the Kodak still video camera was employed for image capture) sub-millimetre precisions were achieved. The average relative precision in the XY plane was of the order of 1:15000 and for the depth, 1:7000. The measurement of the third test surface (using images captured with the ITC CCD video camera) yielded sub-centimetre results with a relative precision in the XY plane of approximately 1:4500 and 1:2600 in the depth.

The design of the system took into account three significant considerations. Firstly, the amount of time necessary for the field component of the system was minimised to reduce possible downtime at the excavation site. Secondly, the system was designed to be used by a non-photogrammetrist, and thirdly, a high degree of automation was required. As a result of the successful implementation of these considerations, the system, although relatively costly to set up in terms of the camera and computer equipment, is relatively inexpensive to run.

The Kodak still video camera had significant advantages over the ITC CCD video camera. The still video camera, with its onboard image storage ability and internal power source, allowed for very convenient image capture as opposed to the CCD video camera which was very cumbersome to use in the field. The images produced by the still video camera were of a better quality and a significantly better resolution than the CCD video camera. A major limitation of the system was the inability of the software to make use of the large image format of the still video camera. An obvious improvement of the system that would be to adapt the software to cater for any image size.

The speed of the image matching process was a limiting factor in terms of the number of images that could be processed, thus affecting the accuracy of the system, and in terms of the number of interest points that were used for matching. With five images and 2800 points, the matching process took approximately seven hours to run on a 486 DX personal computer with 8 MB RAM. Although approximately 1300 points on the Tygerberg Hills test surface were successfully matched, this resulted in an average of only seven points per square metre. A significantly greater density of points is necessary to adequately represent the surface. For the Cape Town harbour excavation and rock wall test surfaces, which covered smaller areas than the Tygerberg Hills test surface, the density of successfully matched points was greater, with 30 and 80 points per square metre respectively. One method of overcoming the

limitation of speed is to use a more powerful computer. Another method is to employ more efficient computer algorithms.

For the measurement of both the Cape Town harbour and the Tygerberg Hills test surfaces, over 40% of the interest points could not be matched. This can largely be attributed to the fact that many of the interest points were not visible from all the image perspectives due to surface discontinuities (especially at the base of the vertical section of the bench). As a consequence of this not many points occurring in depressions on the surface were successfully matched, resulting in the plotted surface appearing too smooth. For the rock wall test surface, where the surface is relatively smooth, only 28% of the interest points were unsuccessfully matched.

Although the measurement system described here was specifically designed for the mensuration of surfaces in open cast mining environments, the same methods that were employed can be applied to the measurement of other textured surfaces.

REFERENCES

Baltsavias, E. P., 1991. **Multiphoto Geometrically Constrained Matching**, Dissertation No. 9561, ETH-Zurich.

Beyer, H. A., 1992a. **Automated Dimensional Inspection with Real-Time Photogrammetry**, International Archives of Photogrammetry and Remote Sensing, 29(B5), pp 722-727.

Beyer, H. A., 1992b. **Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System**, Dissertation No. 9701, ETH-Zurich.

Calitz, M. F.,1995. **Image Understanding and Feature Extraction for Applications in Industry and Mapping**, Ph.D. Dissertation, University of Cape Town.

Canny, J., 1986. **A Computational Approach to Edge Detection, Transactions on Pattern Analysis and Machine Intelligence**, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6, pp 679-698.

Dold, J., and Maas, H., 1994. **An Application of Epipolar Line Intersection in a Hybrid Close Range Photogrammetric System**, International Archives of Photogrammetry and Remote Sensing, 30(5), pp 65-70.

Förstner, W., and Gülch, E., 1987. **A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features**, Proceedings of the Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, June 2 - 4, 1987, pp 281-305.

Fraser, C. S., 1992. **Photogrammetric Camera Component Configuration: A Review of Analytical Techniques**, Workshop on Calibration and Orientation of Cameras in Computer Vision (TU-1), August 2, 1992, Washington DC.

Fraser, C. S., 1995. **Deformation Measurement of a Large Coal Dredger by Digital Photogrammetry**, International FIG Symposium Proceedings, Commission 6, Cape Town, pp 67-75.

Fraser, C. S. and Shortis, M. R., 1995. **Metric Exploitation of Still Video Imagery**, Photogrammetric Record, 15(85), pp 107-122.

Fuller, M. E., and Ehlers, M. **An Automated System for Image Co-Registration Using Interest Clump Matching.**[1]

Gruen, A. W., 1985. **Adaptive Least Squares Correlation: A Powerful Image Matching Technique**, South African Journal of Photogrammetry, Remote Sensing and Cartography, 14(3), pp 175-187.

Gruen, A. W., 1994. **Digital Close-Range Photogrammetry: Progress through Automation**, International Archives of Photogrammetry and Remote Sensing, 30(5), pp 122-135.

Haralick, R. and Shapiro, L., 1992. **COMPUTER AND ROBOT VISION**, Vol. 1-2, Addison Wesley.

Heikkilä, J., 1990. **Update Calibration of a Photogrammetric Station**, SPIE Vol. 1395, Close-Range Photogrammetry Meets Machine Vision, pp 1234-1241.

Huertas, A., and Medioni, G., 1986. **Detection of Intensity Changes with Subpixel Accuracy Using Laplacian-Gaussian Masks**, Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 5, pp 651-664.

Maas, H., 1991. **Digital Photogrammetry for Determination of Tracer Particle Coordinates in Turbulent Flow Research**, Photogrammetric Engineering and Remote Sensing, 57(12), pp 1593-1597.

Mason, S. O., 1995. **Conceptual Model of the Convergent Multistation network Configuration Task**, Photogrammetric Record, 15(86), pp 277-299.

McGlone, J. C., 1989. **Analytic Data-Reduction Schemes in Non-Topographic Photogrammetry**, NON-TOPOGRAPHIC PHOTOGRAMMETRY, Second Edition, Editor: H. M. Karara, American Society for Photogrammetry and Remote Sensing.

Otsu, N., 1979. **A Threshold Selection Method for Grey-Level Histograms**, IEEE Transactions on Systems, Man and Cybernetics, Vol. 9, pp 62-69.

Petsa, E., and Patias, P., 1994. **Relative Orientation of Image Triples using Straight Linear Features**, International Archives of Photogrammetry and Remote Sensing, 30(3/2), pp 663-669.

Rubinstein, M., 1990. **Assessing Target Centring Algorithms for use in Near-Real-Time Photogrammetry**, M.Sc. Thesis, University of Cape Town.

[1] An attempt to locate the name and date of the publication in which this paper appeared was unsuccessful.

Rubinstein, M., and Rüther, H., 1991. **Generating Synthetic Targets**. Conference Proceedings. First Australian Photogrammetric Conference. University of New South Wales, Sydney, Australia, 7-9 November 1991. 11 pages.

Shih, T. Y., 1994. **RLT: A Closed Form Solution for Relative Orientation**, International Archives of Photogrammetry and Remote Sensing, 30(5), pp 357-363.

Shortis, M. R., and Hall, C. J., 1989. **Network Design Methods for Close-Range Photogrammetry**, Australian Journal of Geodesy, Photogrammetry and Surveying, No. 50, June 1989, pp 51-72.

Smith, A. D. N., 1965. **The Explicit Solution of the Single Picture Resection Problem. with a Least Squares Adjustment to Redundant Control**, The Photogrammetric Record, 5(26), pp 113-121.

Strunz, G., 1993. **Bildorientierung und Objektrekonstruktion mit Punkten, Linien und Flächen**, Dr.-Ing Dissertation, Technical University of Munich, p 4.

Tabatabai, A. J., and Mitchell, O. R., 1984. **Edge Location to Subpixel Values in Digital Imagery**, Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 2, pp 188-201.

Van den Heuvel, F. A., Kroon, R. J. G. A., and Le Poole, R. S., 1992. **Digital Close-Range Photogrammetry using Artificial Targets**, International Archives of Photogrammetry and Remote Sensing, 29(B5), pp 222-229.

Van der Merwe, N., and Rüther, H, 1994. **Image Matching Through a Combination of Feature and Area Based Matching**, International Archives of Photogrammetry and Remote Sensing, 30(5), pp 407-413.

Van der Vlugt, G., 1991. **The Development of a Real-Time Photogrammetric System for Patient Positioning in Proton Therapy**, M.Sc. Thesis, University of Cape Town.

Van der Vlugt, G., and Rüther, H., 1994. **The Development of an Automated Surface Measurement System**, International Archives of Photogrammetry and Remote Sensing, 30(5), pp 414-419.

Van der Vlugt, G.,1995. **Algorithms and Design Aspects of an Automated Vision Based 3-D Surface Measurement System**, Ph.D. Dissertation, University of Cape Town.

Welsch, W. M., 1995. **Engineering Surveying in our Days: State of the Art, Trends and Prospects**, International FIG Symposium Proceedings, Commission 6, Cape Town, pp 272-290.

Wong, K. W., and Ho, W., 1986. **Close-Range Mapping with a Solid State Camera**, Photogrammetric Engineering and Remote Sensing, 52(1), pp 67-74.

Wong, K. W., Ke, Y., Slaughter, M., and Gretebeck, R., 1992. **A Computer Vision System for Mapping Human Bodies**, International Archives of Photogrammetry and Remote Sensing, 29(B5), pp 850-855.

BIBLIOGRAPHY

Beyer, H. A., 1994. **IMETRICS TP200 - A System for High Accuracy 3D Image Metrology**, Paper presented at the Symposium on Optics for Productivity in Manufacturing, 20 - 24 June 1994, Frankfurt.

Chang, B., 1986. **The Formulae of the Relative Orientation for Non-Metric Cameras**, International Archives of Photogrammetry and Remote Sensing, 26(5), pp 14-22.

Haggrén, H., and Heikkilä J., 1989. **Calibration of Close-Range Photogrammetric Stations using a Free Network Bundle Adjustment**, The Photogrammetric Journal of Finland, 11(2), pp 21-31.

Heipke, C., Englisch, A., Speer, T., Stier, S., and Kutka, R., 1994. **Semi-Automatic Extraction of Roads from Aerial Images**, International Archives of Photogrammetry and Remote Sensing, 30(5), pp 353-360.

Karara, H. M.(Editor), 1989. **NON-TOPOGRAPHIC PHOTOGRAMMETRY**, Second Edition, American Society for Photogrammetry and Remote Sensing.

Maas, H., 1992. **Complexity Analysis for the Establishment of Image Correspondences of Dense Spatial Target Fields**, International Archives of Photogrammetry and Remote Sensing, 28(B5), pp 102-107.

Matthews, M., 1994. **A Machine Vision Based Non-Contact Measuring Tool for Mensuration in Industry and Manufacturing**, M.Sc. Thesis, University of Cape Town.

Peipe, J. and Schneider, C. T., 1995. **High Resolution Still Video Camera for Industrial Photogrammetry**, Photogrammetric Record, 15(85), pp 135-139.

Schenk, T., 1992. **Algorithms and Software Concepts for Digital Photogrammetric Workstations**, Collection of papers presented at the XVII Congress of ISPRS, pp 17-36.

Van der Merwe, N., 1995. **Development of an Image Matching Scheme using Feature and Area Based Matching Techniques**, Ph.D. Dissertation, University of Cape Town.