A CAMCORDER-BASED RAPID MAPPING SYSTEM

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

by

Siddique Motala, B.Sc. (Survey), UND

Department of Geomatics

October 1997

DECLARATION

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

Siddique Motala

ABSTRACT

This thesis reports on investigations into the potential of off-the-shelf camcorder imagery for rapid mapping. Video technology was chosen as it has the advantage of continuous image capture, enabling quick acquisition times in environments where access may be limited. Video imagery is also very low cost compared to metric photography, and has the advantage of on-line control of the photography, which, for applications with low to medium mapping accuracies, make it an attractive and cheap alternative, well suited for application in developing countries and dynamic environments.

Two case studies are presented. In the first, Hi-8 camcorder imagery was acquired for informal settlement mapping under the UrbanModeler project being run by the Department of Geomatics at the University of Cape Town. The informal settlement of Marconi Beam in Milnerton, Cape Town serves as an application for the UrbanModeler project, whose objective is the development of image exploitation tools for automating the geo-spatial modeling of urban settlements, primarily informal (or "squatter") settlements. Rapid production of georeferenced mosaics from the video imagery is reported on, and these mosaics were used in conjunction with a GIS database of Marconi Beam. This GIS database, containing an inventory of shacks and occupants of the settlement, would ultimately help in relocating the occupants of Marconi Beam to Joe Slovo Park, a low cost formal housing settlement nearby. An analysis of the mapping potential of this data source is made with respect to image interpretability, and three-dimensional point positioning accuracy.

The second case study lies in the terrestrial photogrammetry domain, in a joint project between the Department of Geomatics and the Department of Archaeology at the University of Cape Town. In June 1996, the fieldwork for the photogrammetric mapping of a newly discovered shipwreck was carried out. A Sony Hi-8 camcorder was used for video image capture, and control for the imagery was obtained by a combination of GPS

and theodolite survey. Retroreflective targets were placed along the surface of the wreck to tie the imagery together. This thesis reports on the mapping of one of the sections of the shipwreck, and discusses the use of video imagery for the production of surface maps.

The conclusions of this project state that camcorder imagery is an inexpensive, reliable image acquisition source, that is capable of producing adequate accuracies for a host of applications.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Background and Aims

The advent of the digital image has advanced photogrammetric methods and practice greatly, and has led to the development of a whole new avenue within the discipline of photogrammetry. The ease and convenience of using digital imagery has opened up many new possibilities. It has led to an unparalleled information revolution, in which analogue imaging systems are gradually yielding to CCD imagers.

The development of digital photogrammetry has gone hand-in-hand with the development of remote sensing (Karara, 1989), and its capabilities derive from a combination of methods and technologies out of the disciplines of video technology, image processing, machine vision, and traditional photogrammetry. Digital photogrammetry is essentially a sequential process in which either film-based photographs are digitised or imagery is acquired directly with digital cameras; then the digital data is processed by computers to arrive at spatial information.

This thesis reports on the results of an investigation into the potential of off-the-shelf camcorder imagery for rapid mapping. For the purpose of this thesis, "camcorder" refers to a portable hand-held video camera. Two case studies are presented. In the first, airborne Hi-8 video remote sensing (VRS) imagery was acquired for informal settlement mapping under the UrbanModeler project (Mason et al, 1997). The second case study lies in the terrestrial photogrammetry domain in a joint project with the Department of Archaeology at the University of Cape Town, in which the mapping of a newly discovered shipwreck was carried out.

The main aims of the project are:

- the development of a low-cost, easily assembled rapid mapping tool;
- the provision of a system that would be easily operable by people who are relatively unskilled in photogrammetric mapping, thus enabling an archaeologist, for example, to perform the fieldwork for artifact mapping.

1.2 Literature Review

The use of video imagery for photogrammetric purposes has prompted the development of new terms describing these types of applications. The terms videography (Panja, 1995), videometry (Lenz and Fritsch, 1990), videogrammetry, stereovideogrammetry (Gifford,1997) and Video Remote Sensing or VRS (Thomas, 1996) all refer to this branch of photogrammetry.

Airborne VRS technology has been applied widely as a tool for natural resource management, e.g. as a tool for the detection of forest pest infestations (Linden, 1995, Thomas, 1996, Everitt et al, 1995). Conventional colour aerial video imagery was evaluated for its effectiveness in mapping vegetation in a riverine wetland in Minnesota (Sersland et al, 1995). This study also incorporated the use of airborne GPS, which was needed for georeferencing of the imagery. The result was a successful, comprehensive classification of the wetland plant communities.

Video technology is widely used in non-metric applications. In underwater archaeological work, video cameras have been used on shipwreck sites. In some cases, underwater video cameras have been placed on board remotely operated vehicles (ROVs) and used in reconnaissance and deep-water site inspection. Two examples of the use of this type of system are the discovery of the Titanic in the mid-Atlantic and the discovery of the HMS Breadalbane in northern Canada (Green, 1990). A video camera was also used in

conjunction with a CAMEL 70mm camera to carry out the measurement of underwater structures in the offshore oil and gas industry in the North Sea. The video camera served as a viewfinder, and was attached with the other camera to the front of a submersible (Newton, 1989).

The US Navy undertook some of the earliest research applying photogrammetric mapping of stereo imagery to large scale mapping of the ocean floor (Shipek, 1967). Rebikoff (1961), an inventor and pioneer in the use of photography and photogrammetry to record underwater archaeological sites demonstrated the use of optically precise underwater cameras. Rebikoff (1984) later published a paper outlining the potential application of video cameras for mapping underwater sites. Digital photogrammetric techniques were used by Rule (1995) to map the surface of the *Mary Rose*, a shipwreck that is located at an average depth of 12m. Gifford (1997) has used off-the-shelf video camera imagery in the development of a technique to create microtopographic maps of underwater sites. It is based on video records of underwater archaeological sites referenced to a user-defined coordinate system, followed by digitisation of imagery from video tape. The imagery is then manipulated in the software package called TNTmips. This package has the capability of grabbing video imagery, converting it to raster files, registering the raster files to an arbitrary, user-defined coordinate system, rectifying the imagery, and generating Digital Elevation Models (DEMs).

Video imagery was used by Forte (1995) in the process of simulating a virtual environment. Scientific visualisation techniques made it possible to simulate the exploration of an archaeological site in virtual space by using video animation for intra-site navigation. A great advantage of using a video camera is to monitor the progress of an archaeological excavation through analysis of video footage.

A vital aspect in the utilisation of camcorder imagery is the calibration of video sensors and geometric rectification of digitised imagery. This has been addressed in numerous

publications, including Mitchell (1995), Buechel et al (1995), Thomas (1996), and Shortis(1995), who investigated the effects of interlacing on camera calibration.

The viability and accuracy of using camcorders for various applications have been tested, e.g. in architectural work (Streilein and Gaschen, 1994) and in industrial applications (Shortis, 1995). Hoeflinger (1994) also evaluated the radiometric and geometric performance of a camcorder, and it was found that analogue image storage tends to degrade the geometric and radiometric performance significantly. It was also confirmed that best results for analogue image storage on video tape can be attained when the tape is played back on the same camcorder it was recorded with.

The suitability of using video technology in the abovementioned projects lies in the following capabilities and qualities:

- convenience and ease in assembling the components of a video surveying system;
- continuous image capture and on-line control of the photography, enabling quick data acquisition times in environments where access may be limited;
- ease in obtaining redundant observations during photogrammetric processing;
- lower cost compared to metric photography;
- the video operator can record real-time audio input as the video data is being collected;
- the video tape serves as a continuous, permanent photographic record of the site.

1.3 Case Studies

The first case study carried out formed part of the research for the UrbanModeler project being undertaken in the Department of Geomatics. The informal (or "squatter") settlement of Marconi Beam in Milnerton, Cape Town serves as the focus for this project. Due to the dynamic nature of informal settlements, frequent updating of maps is necessary, and it will be shown that the rapid production of rubber-sheeted mosaics can be

accomplished by means of utilising frame-grabbed imagery. The three-dimensional point positioning accuracy of camcorder imagery is also reported on.

The second case study deals with close-range imagery, in which the photogrammetric mapping of a shipwreck discovered in June 1996 was carried out. For archaeological purposes, a map of the surface of the wreck was required, thus a Digital Elevation Model (DEM) was produced from the acquired camcorder imagery.

1.4 Scope and Limitations

The method of using camcorder imagery for the extraction of spatial information has been investigated in both the terrestrial and aerial photogrammetry domains. A major limiting factor in the course of this investigation was the poor resolution of the frame-grabbed imagery. Due to financial constraints, a more suitable hardware configuration that would have improved the resolution of the imagery could not be obtained, thus limiting the overall accuracy and usability of the system. This inferior image resolution in the two case studies hampered the production of ortho-imagery of adequate quality, an end-product that would have been very useful in both projects.

A limitation in the shipwreck case study was the very short period of time that was available for the fieldwork and photography to be carried out. For the purpose of preservation of the shipwreck, it had to be re-buried three days after the site was excavated, so no additional photography could have been possible thereafter. Due to the rushed nature of this exercise, no prior testing of the method was possible. Important conclusions on how to improve results were drawn during the course of processing the data sets. Data handling and manipulation was difficult, particularly in the case of the shipwreck, due to the fact that the computer used was a 486 DX with an approximate hard disk drive size of 350Mb. Computation time would have been significantly decreased with a more powerful computing system.

1.5 Thesis outline

An explanation of the photogrammetric principles that are used in this project is given in chapter 2, which serves to familiarise the reader with the necessary background knowledge.

The hardware and software components used in the implementation of a system that utilises video technology are discussed in chapter 3.

Both the case studies differ mainly in the photography and photogrammetric processing of the data sets. Different end products have been produced for both examples. The case studies are discussed, assessed and presented in chapters 4, 5 and 6.

Conclusions are drawn about the research carried out and the viability of using a camcorder-based rapid mapping system is discussed in chapter 7.

CHAPTER TWO

Photogrammetric Principles

2.1 Introduction

Photogrammetry is defined as being the art, science, and technology of obtaining reliable quantitative information about physical objects and the environment through the process of recording, measuring and interpreting photographic images and patterns of radiant imagery derived from sensor systems (Karara, 1989). The primary objective is to solve the general photogrammetric problem, being the determination of the camera interior and exterior orientation parameters and the coordinates of object space points of interest (McGlone, 1989). Depending on the application, different combinations of known and unknown parameters exist, but the fundamental mathematics remains the same.

This chapter serves to familiarise the reader with the basic photogrammetric principles that were applied in this project.

2.2 Perspective Projection

The starting point for building a functional model for photogrammetry is the *perspective projection* (sometimes referred to as the *central perspective projection*). The perspective projection is illustrated in Figure 2.1. The point A in the three dimensional object space (defined by the XYZ coordinate system) is projected onto the *projection plane* by the straight line Aa from A passing through the *perspective centre* O. The perspective axis Pop intersects the projection plane at p, called the *principal point.* The distance Op from the perspective centre to the plane of projection is the *principal distance*, usually denoted by c.

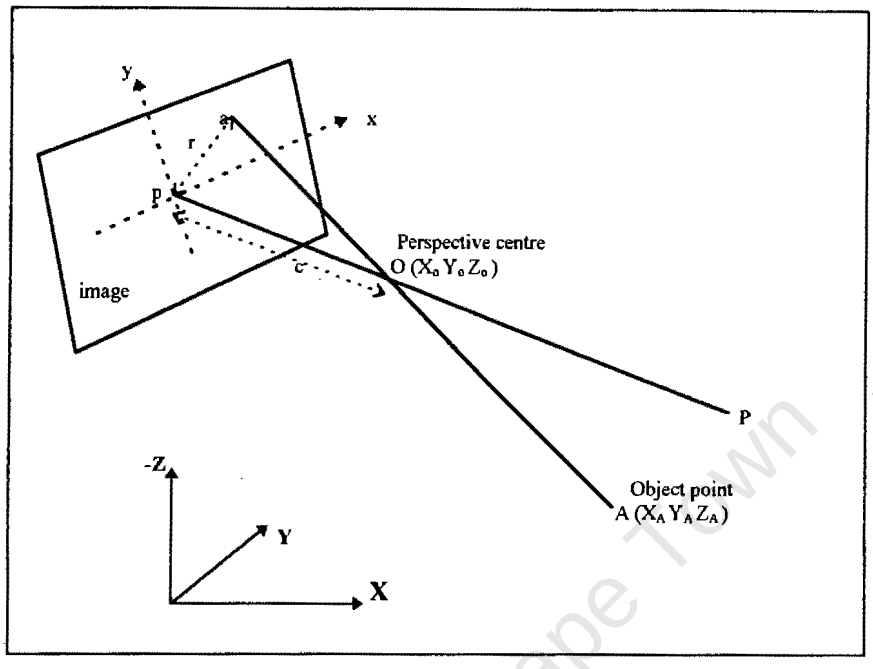

Figure 2.1 Perspective Projection (after Cooper and Robson, 1996)

2.3 Interior Orientation and Camera Calibration

Interior orientation is the term used to describe the parameters that model the passage of light rays through the lens and onto the image plane. There are two sets of parameters defining interior orientation. The first set comprises the geometric parameters of the camera itself, i.e., the principal distance c and the principal point coordinates (coordinates of point p in figure 2.1). The second set comprises the parameters that describe the various systematic errors, or distortions, which are the variations between the ideal mathematical model and the physical reality of the system. Typically, these corrections that have to be modeled are parameters for *radial lens distortion, decentering distortion, focal plane unflatness* and *focal plane distortion,* although focal plane unflatness is not considered to be an issue for small CCD chips such as those found in camcorders.

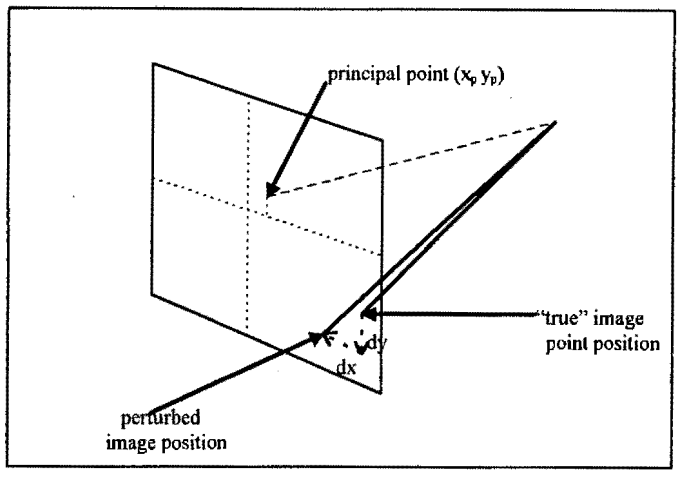

Figure 2.2 Effect of perturbations on an image point

If a point on an image is displaced from its "true" position, to a new "perturbed" position (as shown in Figure 2.2) then the net image displacement at that point will amount to the cumulative influence of each of the abovementioned perturbations. Thus,

$$dx = dx_r + dx_d + dx_u + dx_f$$
$$dy = dy_r + dy_d + dy_u + dy_f \qquad \textbf{(2.1)}$$

where the subscript r is for radial distortion, d for decentering distortion, u for focal plane unflatness and f for focal plane distortion.

Calibration may be achieved by the multiple capture of images of a calibration control frame, from different perspectives. The control frame that was used for this project was developed and constructed in the Department of Geomatics. This frame has a large number (90) of well-distributed, accurately determined control points, represented by circular retro-reflective targets on a black background (see Figure 2.3).

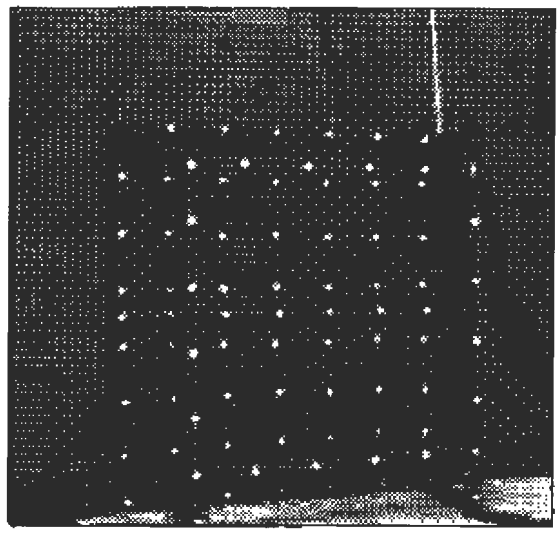

Figure 2.3 Control Frame for Camera Calibration

For calibration of the camcorder, the control frame is recorded from different positions, and images displaying these different perspectives are frame-grabbed. It should be noted that the focus of the camera must be preserved, so that the internal geometry of the camera remains constant throughout the procedure.

Due to projective coupling of interior orientation parameters with the exterior orientation parameters, it is essential that the field of control points be well distributed in three dimensions. This will allow for the recovery of the interior orientation elements with less influence from the exterior orientation parameters. The inclusion of camera stations with varying image scales (without changing the focal setting) and multiple roll angles greatly enhances the determinability of the interior orientation parameters by decreasing their correlation with the exterior orientation parameters (Fraser, 1992).

Semi-automated analysis of the calibration images by target centering and target identification algorithms, together with a constrained bundle adjustment software package (all developed in the Department of Geomatics) provide the calibration parameters. The names and short descriptions of these programs are given in Table 3.1 on page 38.

The first step is the detection of targets in the imagery. Automatic target detection is done by a "search" program, which uses a weighted centre of gravity algorithm to detect the target centres.

Target recognition is carried out by using the "idcon" program. This program requires eight control points on the calibration frame to be manually identified, which are then related to their known XYZ coordinates, defined in a file containing the coordinates of the targets. These points are used to compute the 11 DLT parameters (see section 2.7) which relate object space to image space. These values are then used to calculate the positions in which the other targets are expected to be located. The automatic identification of these targets then follows. The output of "idcon" are the DLT parameters, and the provisional interior and exterior orientation parameters.

The final step in the calibration is the refinement of the abovementioned parameters. The values obtained from "idcon" are used as provisional values in a bundle adjustment. The bundle adjustment program is called "photonet" (van der Vlugt, 1995), and it allows the user to choose which parameters are to be solved for. For calibration, the following parameters are usually solved for : interior and exterior orientation parameters, including radial and decentering distortion, and a scale factor.

2.4 Coordinate Systems and Geodetic Projection

2.4.1 Pixel and Image coordinate systems

It is necessary to make the distinction between pixel coordinates and image coordinates and to establish a relationship between the two. Image measurements are often expressed in terms of pixels or fractions thereof, however in the equations relating to the geometry of image and object space, coordinates are required in metric units related to the perspective centre of the image plane. The x-axis of the pixel coordinate system runs from

left to right and the y-axis from top to bottom. All of the algorithms used in this project use the origin of the pixel coordinate system as (0,0) for the centre of the top-left pixel. The origin of the image coordinate system is usually placed close to the principal point (see Figure 2.4).

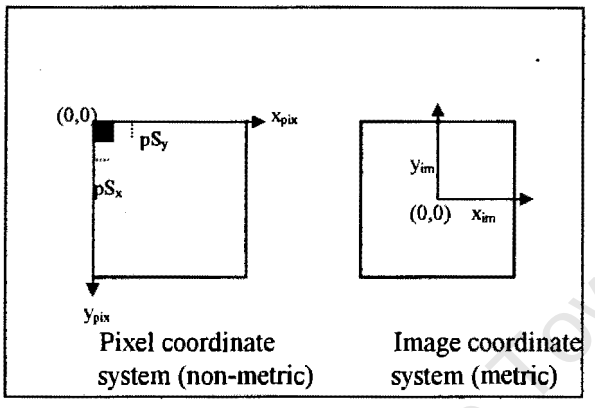

Figure 2.4 Pixel and image coordinate systems

Due to metric units being required for the geometric calculations relating image space to object space, we need to be able to transform between pixel coordinates and metric image coordinates. The conversion from pixel to image coordinates is given in equations (2.2) below:

$$x_{im} = [(x_{pix} - 1) - \frac{(num_x - 1)}{2}]pS_x$$
$$y_{im} = [\frac{(num_y - 1)}{2} - (y_{pix} - 1)]pS_y \qquad \textbf{(2.2)}$$

where

x_{im}, y_{im} are the image coordinates of the point,

x_{pix}, y_{pix} are the pixel coordinates of the point,

num_x the number of horizontal pixels in the image,

num_y the number of vertical pixels in the image,

pS_x, pS_y the pixel cell sizes in the x and y directions.

2.4.2 Object coordinate system and Geodetic Projection

The object coordinate system that was used for both the case studies in this project is the ground coordinate system defined by transforming the coordinates obtained from GPS surveys. The coordinates of the GPS points are defined within the WGS84 coordinate system, and these had to be transformed into the object space coordinate system, which is called the Gauss Conform coordinate system. The Gauss Conform system is a left-handed coordinate system, based on the Gauss Conform geodetic projection, a conformal map projection that is the national coordinate system used in South Africa.

The height system that was used has the geoid as a reference surface. These heights are called orthometric heights. In the GPS-derived WGS84 coordinate system, the reference surface for heights is the ellipsoid, hence the term ellipsoidal heights. A transformation from ellipsoidal to orthometric heights had to be carried out for all GPS coordinates obtained.

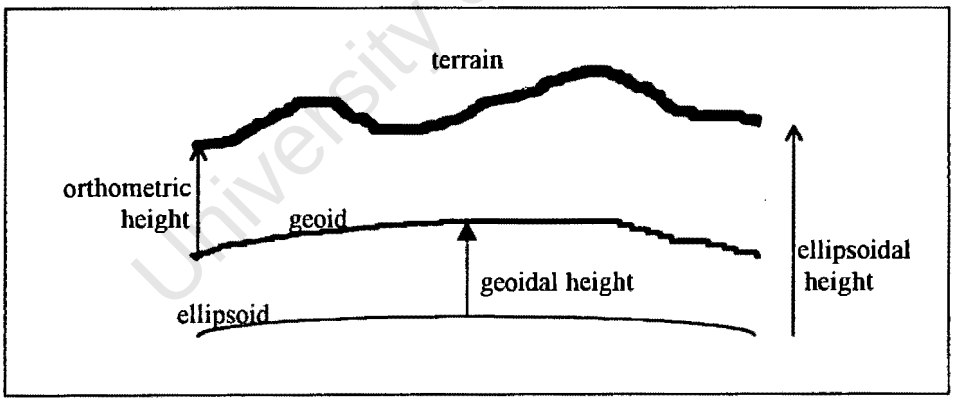

Figure 2.5 Ellipsoidal and Orthometric heights

2.5 Exterior Orientation

Exterior orientation refers to the position of the camera when the image was taken, and to the angular relationship between the image and the object space coordinate system. This relationship for the terrestrial example (chapter 5) is shown in Figure 2.6:

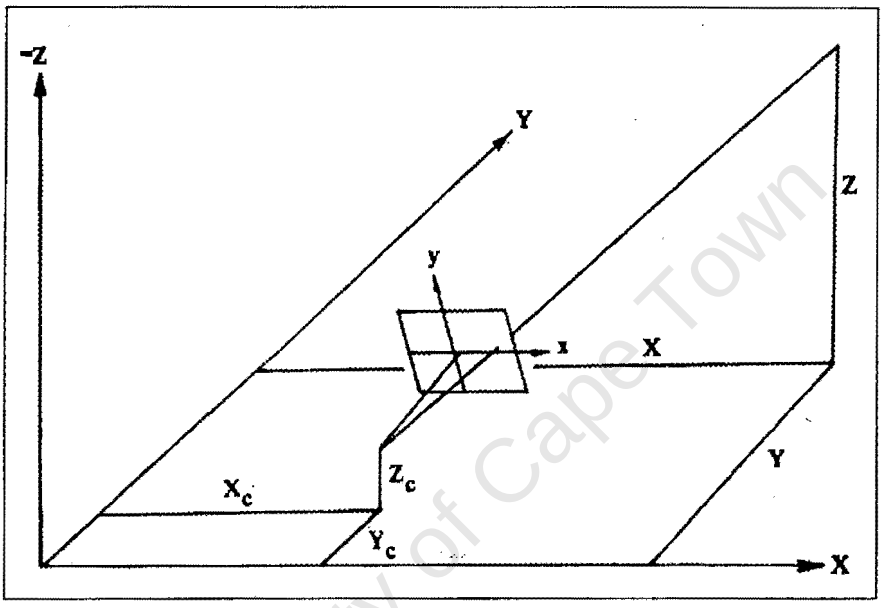

Figure 2.6 Object space (XYZ) and image (xy) coordinate systems (after McGlone, 1989)

The angular relationship between the image and object space coordinate systems is described by a 3x3 orthogonal rotation matrix. The rotation matrix R is defined by 3 sequential rotations about angles ω,φ and κ around the X, Y and Z axes respectively.

$$R = R_{XYZ} = R_{\kappa}R_{\varphi}R_{\omega} \tag{2.3}$$

or

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{2.4}$$

with

$$
\begin{aligned}
r_{11} &= \cos\varphi\cos\kappa \\
r_{12} &= \sin\omega\sin\varphi\cos\kappa + \cos\omega\sin\kappa \\
r_{13} &= -\cos\omega\sin\varphi\cos\kappa + \sin\omega\sin\kappa \\
r_{21} &= -\cos\varphi\sin\kappa \\
r_{22} &= -\sin\omega\sin\varphi\sin\kappa + \cos\omega\cos\kappa \\
r_{23} &= \cos\omega\sin\varphi\sin\kappa + \sin\omega\cos\kappa \\
r_{31} &= \sin\varphi \\
r_{32} &= -\sin\omega\cos\varphi \\
r_{33} &= \cos\omega\cos\varphi
\end{aligned}
\qquad \textbf{(2.5)}
$$

The six parameters that define the elements of exterior orientation are the coordinates ($X_cY_cZ_c$) of the perspective centre in the object space coordinate system and the angular rotations ω, φ, and κ. This angular relationship is applied in the collinearity equations (see section 2.6) when formulating the relationship between image space and object space, through the imaging process.

2.6 The Collinearity Equations

The collinearity equations, derived from the perspective transformation (see Figure 2.1) are based on the fundamental assumption that the perspective centre (O), the object point (A), and its corresponding image point (a) all lie on a straight line, i.e. they are *collinear*. The image vector, r, expressed in the image coordinate system, is

$$
r = \begin{bmatrix} x_a - x_p \\ y_a - y_p \\ c \end{bmatrix} \qquad \textbf{(2.6)}
$$

where

x_a, y_a are the image coordinates of the interest point,

x_p, y_p are the principal point coordinates, and

c is the principal distance.

The vector A, from the perspective centre to the interest point (see Figure 2.1), expressed in the object space coordinate system, is

$$A = \begin{bmatrix} X_A - X_O \\ Y_A - Y_O \\ Z_A - Z_O \end{bmatrix} \tag{2.7}$$

where

X_A, Y_A, Z_A are the object space coordinates of the interest point, and

$X_O\ Y_O\ Z_O$ are the object space coordinates of the perspective centre.

The imaging process assumes that the image and object rays are collinear, thus we multiply the object space vector by the rotation matrix R (see Equation 2.4), to bring it into the same coordinate system. A scale factor k, is then included, giving:

$$\begin{aligned} a &= kRA \\ \begin{bmatrix} x_a - x_p \\ y_a - y_p \\ c \end{bmatrix} &= kR \begin{bmatrix} X_A - X_O \\ Y_A - Y_O \\ Z_A - Z_O \end{bmatrix} \end{aligned} \tag{2.8}$$

If the first and second rows of the equation system are divided by the third row then the scale factor is eliminated. If the distortion parameters dx and dy (see Equation 2.1) are included in the equations, they are then rearranged to yield the most commonly used form of the collinearity equations:

$$\begin{aligned} x_a &= x_p - dx + c\frac{r_{11}(X_A - X_O) + r_{12}(Y_A - Y_O) + r_{13}(Z_A - Z_O)}{r_{31}(X_A - X_O) + r_{32}(Y_A - Y_O) + r_{33}(Z_A - Z_O)} \\ y_a &= y_p - dy + c\frac{r_{21}(X_A - X_O) + r_{22}(Y_A - Y_O) + r_{23}(Z_A - Z_O)}{r_{31}(X_A - X_O) + r_{32}(Y_A - Y_O) + r_{33}(Z_A - Z_O)} \end{aligned} \tag{2.9}$$

The nine 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, a parametric least squares adjustment can be carried out to solve for the unknowns.

2.7 The Direct Linear Transformation (DLT)

An alternate formulation of the analytical orientation problem is the *Direct Linear Transformation*, or *DLT*, as originally proposed by Abdel-Aziz and Karara (1971). The main advantages of the method are that it does not require a calibrated camera, or one with fiducial marks, and can be solved without supplying initial approximations for the parameters (McGlone, 1989). The DLT can be used as a direct method of calculating transformation parameters between image space and object space coordinates for a single image. After algebraic manipulation of the collinearity equations (2.9), the following formulation can be obtained:

$$\begin{aligned} x_a - dx &= \frac{L_1X + L_2Y + L_3Z + L_4}{L_9X + L_{10}Y + L_{11}Z + 1} \\ y_a - dy &= \frac{L_5X + L_6Y + L_7Z + L_8}{L_9X + L_{10}Y + L_{11}Z + 1} \end{aligned} \qquad \textbf{(2.10)}$$

where

x_a, y_a are the image coordinates of the point of interest,

dx, dy are the distortion parameters,

$L_1 \ldots L_{11}$ are the 11 DLT parameters:

$$
\begin{aligned}
L_1 &= \frac{(x_p r_{31} - c r_{11})}{L} \\
L_2 &= \frac{(x_p r_{32} - c r_{12})}{L} \\
L_3 &= \frac{(x_p r_{33} - c r_{13})}{L} \\
L_4 &= \frac{x_p + c(r_{11} X_c + r_{12} Y_c + r_{13} Z_c)}{L} \\
L_5 &= \frac{(y_p r_{31} - c r_{21})}{L} \\
L_6 &= \frac{(y_p r_{32} - c r_{22})}{L} \\
L_7 &= \frac{(y_p r_{33} - c r_{23})}{L} \\
L_8 &= \frac{y_p + c(r_{21} X_o + r_{22} Y_o + r_{23} Z_o)}{L} \\
L_9 &= \frac{r_{31}}{L} \\
L_{10} &= \frac{r_{32}}{L} \\
L_{11} &= \frac{r_{33}}{L} \\
L &= -(r_{31} X_o + r_{32} Y_o + r_{33} Z_o)
\end{aligned}
\qquad \textbf{(2.11)}
$$

where

x_p, y_p	are the principal point coordinates,
c	is the principal distance,
X_o, Y_o, Z_o	are the coordinates of the perspective centre, and
$r_{11}...r_{33}$	are the rotation matrix elements.

(2.10) is a linearly dependent set of equations containing nine linearly independent parameters and two linearly dependent parameters. Ignoring the linear dependence of the parameters allows for an approximation.

The distortion parameters (dx and dy) in equation (2.10) are formulated as (Beyer, 1990):

$$dx = \Delta x_p - \frac{\bar{x}}{c}\Delta c - \bar{x}s_x + \bar{y}a + \bar{x}r^2K_1 + \bar{x}r^4K_2 + \bar{x}r^6K_3 + (r^2 + 2\bar{x}^2)P_1 + 2\bar{x}\bar{y}P_2$$

$$dy = \Delta y_p - \frac{\bar{y}}{c}\Delta c + \bar{x}a + \bar{y}r^2K_1 + \bar{y}r^4K_2 + \bar{y}r^6K_3 + (r^2 + 2\bar{y}^2)P_2 + 2\bar{x}\bar{y}P_1$$

(2.12)

where

$\bar{x} = x_a - x_p$

$\bar{y} = y_a - y_p$

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

$\Delta x_p, \Delta y_p, \Delta c$	are changes in the interior orientation elements;
s_x	is the scale factor in x;
a	is the shear;
K_1, K_2, K_3	are the first three parameters of radial symmetric distortion;
P_1, P_2	are the first two parameters of decentering distortion.

An advantage of the DLT lies in its fast convergence even with poor initial coordinates. The DLT is often used for determining provisional values that are then refined using the bundle adjustment (section 2.8).

For the calibration of a camera, a minimum of six 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 observations are available. If two or more images are used a space intersection can be performed to measure points in object space. With more observations than unknowns, a parametric least squares adjustment can be used to calculate the 3-D coordinates of the points. The minimum solution of the intersection problem requires only three ordinates measured in image space. This, however, must include a combination of both x and y ordinates.

The 11 DLT parameters can be converted into the nine linearly dependent physical parameters of interior and exterior orientation. The camera parameters are defined according to the following algebraic solution (McGlone, 1989):

$$
\begin{aligned}
L &= \frac{-1}{\sqrt{(L_9^2 + L_{10}^2 + L_{11}^2)}} \\
x_p &= (L_1 L_9 + L_2 L_{10} + L_3 L_{11}) L^2 \\
y_p &= (L_5 L_9 + L_6 L_{10} + L_7 L_{11}) L^2 \\
c_x &= \sqrt{[(L_1^2 + L_2^2 + L_3^2) L^2 - x_p^2]} \\
c_y &= \sqrt{[(L_5^2 + L_6^2 + L_7^2) L^2 - y_p^2]} \\
c &= (c_x + c_y)^2 \\
\phi &= \sin^{-1}(L_9 L) \\
\omega &= \tan^{-1}(-L_{10} / L_{11}) \\
r_{11} &= L(x_p L_9 - L_1) / c_x \\
\kappa &= \cos^{-1}(r_{11}) / \cos\phi \\
\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} &= \begin{bmatrix} L_1 & L_2 & L_3 \\ L_5 & L_6 & L_7 \\ L_9 & L_{10} & L_{11} \end{bmatrix}^{-1} \begin{bmatrix} L_4 \\ L_8 \\ 1 \end{bmatrix}
\end{aligned}
\qquad \textbf{(2.13)}
$$

where

c_x , c_y are interpreted as the principal distances in the x and y directions.

2.8 Bundle Adjustment

The collinearity equations can be used to solve for the camera orientation parameters provided good approximations to the unknowns are available and sufficient control points are used. Systematic errors (as defined by equation 2.1) can cause departures from collinearity and are included in the equations. 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* (Brown, 1958). The bundle adjustment uses the collinearity equations for each ground point, and from multiple camera stations combined into a simultaneous solution for the unknown parameters using a least squares adjustment. The procedure is the same as that used for standard block adjustment in aerial triangulation. The physical situation for three cameras stations and four control points is illustrated in Figure 2.7:

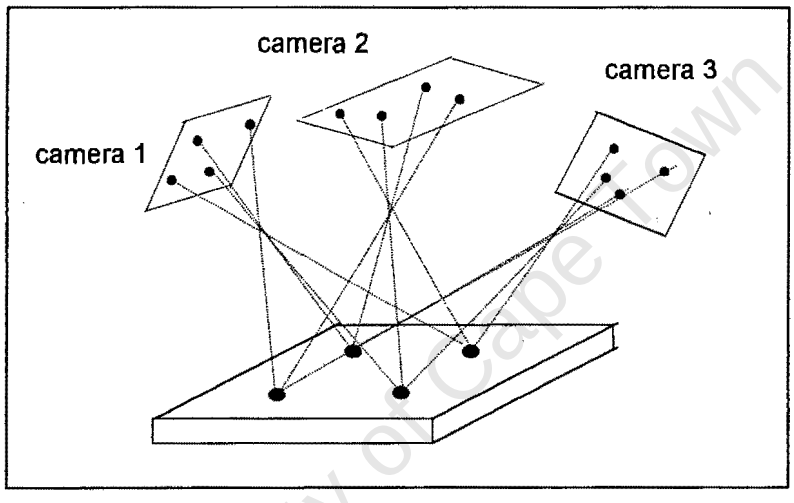

Figure 2.7 Multi-station bundle configuration

The observation equations for a least squares bundle adjustment may be obtained by linearising the collinearity equations through a Taylor series expansion. The observation equations then form part of a standard parametric least squares formulation.

In order for a simultaneous solution of the bundle adjustment to be carried out, the following conditions must be satisfied (van der Vlugt, 1995):

- every point must be visible in at least two images,
- a minimum of five well distributed points in image space must be visible on each image,
- 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,

- 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. This can be done by increasing the number of images used or by increasing the number of control points in the network. The solution for the adjustment is iterated with the provisional values being continually updated by the corrections to the unknowns, until the correction reaches a prespecified tolerance. Karara(1989) notes that in using the bundle adjustment, various parameters may be treated as known or unknown by manipulating their weights in the adjustment.

2.9 Image Matching

The basic task that has to be solved by photogrammetry is to derive a complete, precise and reliable description of the shape of the surface of the object by means of a *Digital Elevation Model* (DEM). The DEM is a dense matrix of points, usually regularly spaced in the X and Y directions, for each of which the height of the surface being modeled above some datum is given.

Image matching is the process of locating corresponding points of interest in overlapping images for the purpose of three-dimensional point position location in object space. Two or more digital images are given, covering approximately the same region in object space, along with the exterior orientation parameters for the camera.

The type of image matching carried out by the software used in this project is called *area-based matching.* The observations in area-based matching are grey levels. Here, the idea is to compare the grey level distribution of a small sub-image, called an *image patch*, with its counterpart in another image. The *reference patch* is the image patch that remains in a fixed position in one of the images, usually called the *template* image. The *search patch*

refers to the search space within which the image patches are compared with the reference patch.

2.9.1 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 falls along a unique line in the second image. Such a line is referred to as an *epipolar line*. It is the line of intersection between the image plane and the plane passing through the object point and the perspective centres of the two images, as illustrated in Figure 2.8.

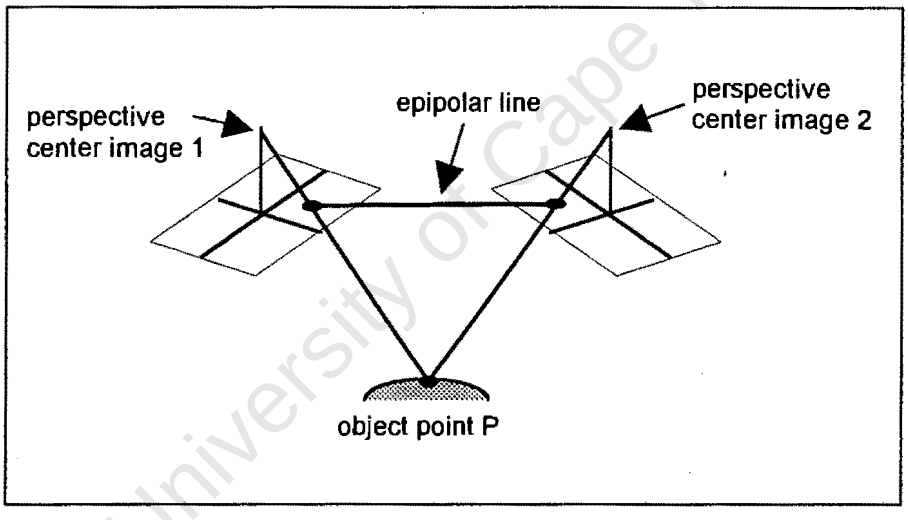

Figure 2.8 Epipolar geometry

2.9.2 Epipolar Lines using the Collinearity Equations

From perspective geometry the relationship between the target image and the object can be described in terms of the collinearity equations. If the image coordinates of a point in the first image is known, as well as the camera orientation parameters of both images, 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 and Z coordinates, the image space

coordinates in 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 point is known, the length of the epipolar line can be restricted. A certain tolerance can be added to the width of the epipolar line to form a narrow two-dimensional window in image space within which a search for the conjugate point can take place, allowing for a smaller search space in the epipolar matching, thus decreasing the computational time. This, together with area-based least squares matching form the basis for a *multi-photo geometrically constrained* (MPGC) least squares matching routine which was used for all the image matching done in this project.

The MPGC approach to matching exploits known camera orientation information considerably improving the matching performance (Baltsavias, 1991). Image matching has been approached in two distinct but interlinked processes. Epipolar geometry supplies estimates to the initial image coordinates of the corresponding points in the conjugate images, which is then followed up by least squares, grey-scale area-based matching with imposed geometric constraints to determine the final matching image positions as well as XYZ object space coordinates. The initial estimate of corresponding pixels in conjugate images, as provided by the epipolar line condition, is required to within one or two pixels (the "pull-in range") in order to achieve a solution in the subsequent matching process. In order to obtain an optimum accuracy for the final point position, the area-based least squares algorithm is supported by affine patch shaping and geometric constraints in the form of the collinearity equations.

The result of the image matching algorithm is not just the corresponding image coordinates of interest points in conjugate images, but also a dense point cloud of three-dimensional object space coordinates of successfully matched interest points. The algorithm produces both digital elevation model (DEM) data in the form of X, Y and Z coordinates and error estimates for the object space coordinates as a consequence of the least squares estimation model.

CHAPTER THREE

Components of a Camcorder-based Mapping System

Figure 3.1 gives a schematic representation of the various components of the video mapping system, along with some of the related tasks involved.

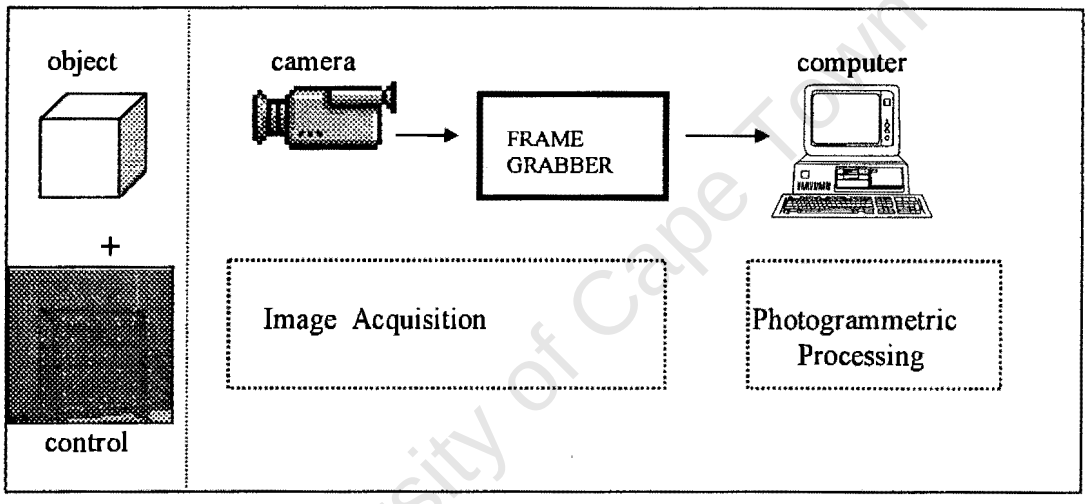

Figure 3.1 Components of camcorder-based mapping system

3.1 Cameras

Two different Hi-8 hand-held video cameras were chosen for the project. For the UrbanModeler project, a Grundig Hi-8 lc 180 HE camcorder was used, and for the mapping of the shipwreck, a Sony TR880e Hi-8 camcorder. These high-band 8mm (Hi-8) camcorders were chosen over standard VHS because they have a higher and wider carrier frequency range. In the Hi-8 system, the FM carrier frequency range of the Y (luminance) signal is shifted from the 4.2-5.4 Mhz range of standard 8mm video up to 5.7-7.7 Mhz, which increases the picture information capacity of the signal (Gifford, 1993).

The video signal is PAL colour. All the information that is obtained by means of the solid state sensor is stored on 8mm magnetic tape.

3.1.1 Magnetic Tape

The magnetic tape on which the information is archived consists of a magnetic pigment within a polymer binder that is supported on a backing film (see Figure 3.2).

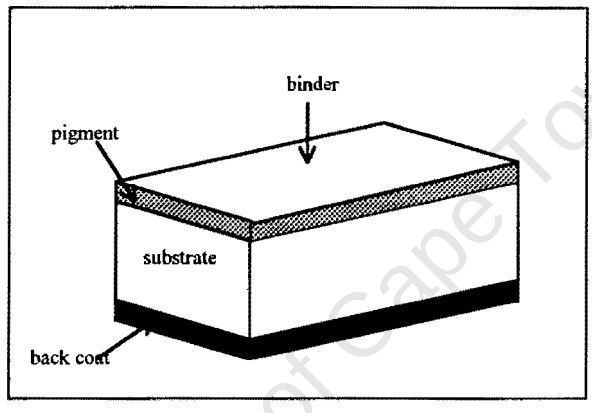

Figure 3.2 Cross-section of a magnetic tape (after Kolbl, 1996)

The magnetic pigment is responsible for storing recorded information magnetically. If there is any change in the magnetic properties of the pigment, data can be irretrievably lost. Changes can be brought about by an increase in temperature, humidity, or exposing the tape to a strong magnetic field. The coercivity of the pigment characterises the pigment's ability to resist demagnetisation. Demagnetisation can result from an externally applied field, or from self-demagnetisation via thermally induced magnetic reversals.
The chips contained in the cameras used are CCD solid-state image sensors.

3.1.2 Charge Coupled Devices

A charge coupled device (CCD) is the most commonly used device for recording the amount of light falling onto a surface. The CCD chip in a camera contains an array of

light sensitive electro-optical elements, called pixels. CCDs are arranged in linear arrays (used to scan a scene), or in two-dimensional arrays as in a camera. The two dimensional arrays provide a record of the light falling onto a two dimensional surface at a particular instant of time.

The charge in each of the individual CCD elements is a function of the incoming light and due to interconnections between the elements (coupling) the charge can be transferred pixel to pixel at high speed from the chip to a frame-grabber. Individual pixel charges are then translated from an analogue to a digital form (by the frame-grabber) resulting in a string of digital values (typically 8 bit bytes) which represents a digital image. An image is thus represented by an array of pixel values ranging from 0 (black) to 255 (white) for an 8 bit resolution.

Depending on the relative arrangement of the elements on the CCD chip, the chip is either of the frame transfer or interline transfer type (Shortis et al, 1996).

3.1.2.1 Frame Transfer CCD

The frame transfer chip has a solid block of photosensitive elements which occupies one portion of the chip while the transfer elements are grouped together on the remaining chip area (see Figure 3.3).

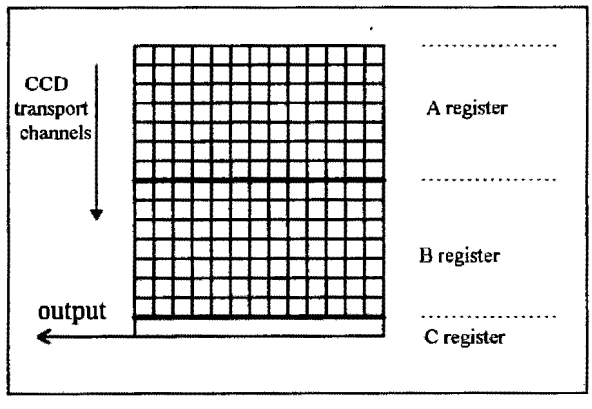

Figure 3.3 Frame Transfer CCD chip

The chip consists of three registers. The A register is light sensitive. The contents of this resister are transferred into the B register, which is covered by an opaque metal layer. The charge packets constituting the image are protected from light in the B register. The image is transferred line after line to the C register, a serial read-out register. The charge transfer process for the frame transfer chip requires that steps must be taken to prevent significant smearing of the image. Smear is caused by the fact that the sensor elements are exposed to light during the read out process, and the same sensor elements are used to expose the image and transfer the charge. A technique for eliminating smear is a mechanical shutter to cover the sensor during charge read out.

3.1.2.2 Interline transfer CCD

The interline transfer chip overcomes the problem of smear by using different elements to accumulate and transfer the charge (see Figure 3.4)

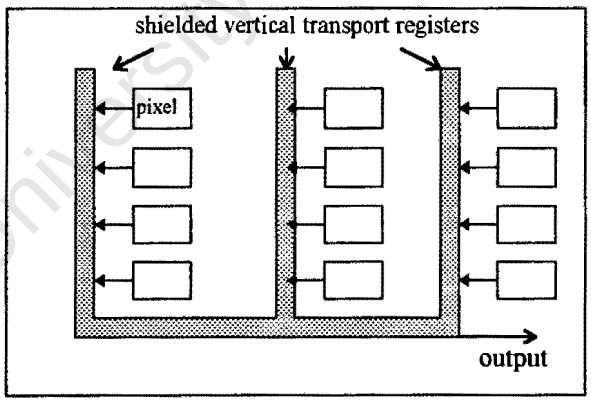

Figure 3.4. Interline transfer CCD chip (after Kolbl, 1996)

Each pixel consists of an individual photodiode, which is connected with a transfer gate to a linear CCD column. An opaque metal layer that is insensitive to light covers these columns. The transfer of photocharges occurs shielded from light, and so the image is not

corrupted during the transfer process. Today, interline transfer CCDs are the most common image sensors, found in most video cameras.

The chip in the Hi-8 camera is an interline transfer chip with a resolution of 752 x 582 pixels. This represents a much higher resolution than a standard VHS chip that has a CCD array of 300 x 210 pixels.

The basic structure of the CCD chip found in both the cameras used in this project is represented in Figure 3.5.

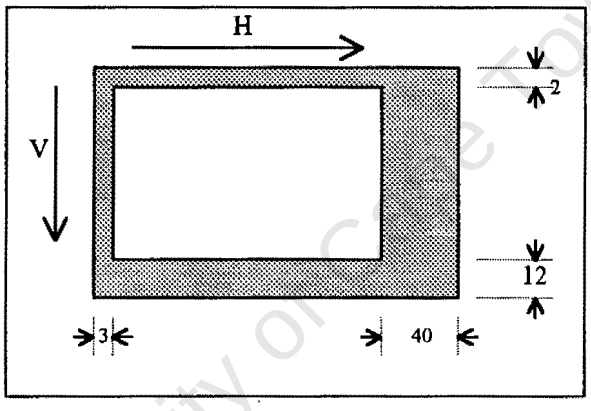

Figure 3.5 CCD chip in video camera

The chip size is 7.95 x 6.45 mm (H x V). The size of each pixel is 8.6 x 8.3 μm (H x V). The number of effective pixels are 752(H) x 582(V), which amounts to approximately 440000 pixels. The total number of pixels is 795(H) x 596(V). It has an optical black of 3 pixels in the front and 40 pixels in the back in the horizontal direction, and 12 front and 2 back in the vertical direction. The material that makes up the chip is primarily Silicon.

The video signal of the camera is transmitted via the CCIR standard. This standard was originally defined by the Comite Consultatif International des Radiocommunications (CCIR) for black-and-white television in the 1950s (Beyer, 1992). The CCIR video standard uses the principle of interlacing, depicted in Figure 3.6.

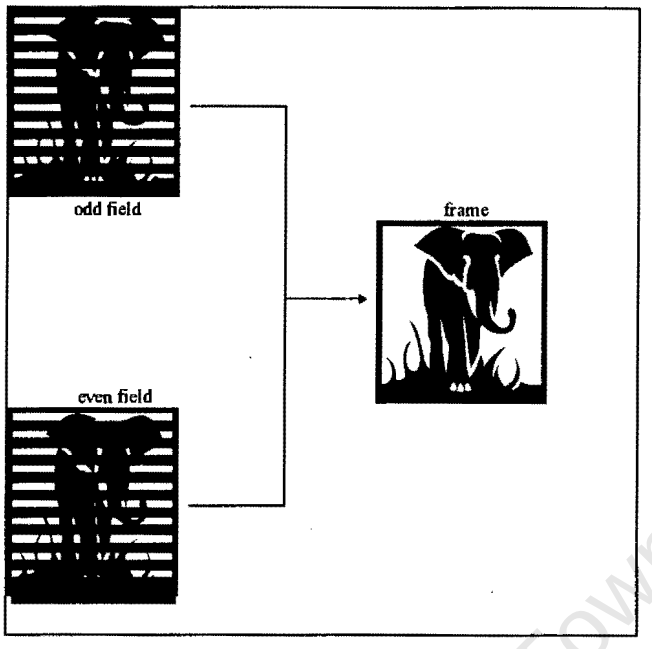

Figure 3.6 Frame and fields in interlaced video

An interlaced image, or a frame, consists of two fields, the even and the odd field. The even field contains all lines with even line numbers and the odd field those with odd line numbers.

3.2 Frame-Grabbing

A frame-grabber is a circuit board that resides in a host computer, which is designed to instantaneously sample the output from a solid state sensor transmitting standard analogue video. The sample collected by the frame-grabber is loaded into onboard solid state memory, which can then be accessed by the host computer.

The basic functional elements of all frame-grabbers are shown in Figure 3.7.

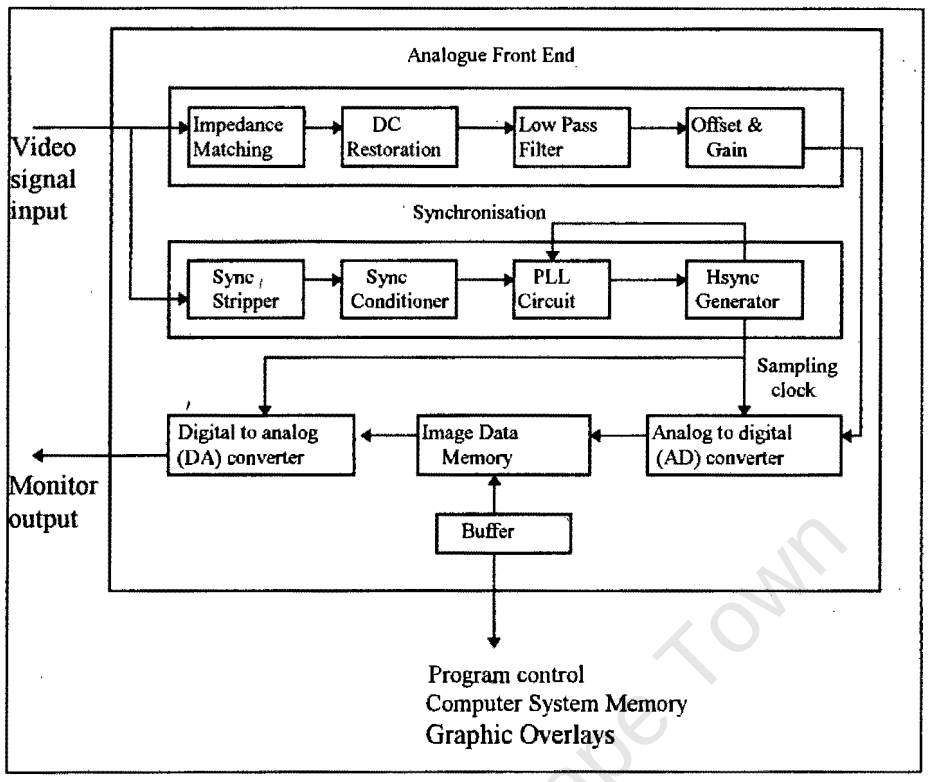

Figure 3.7 Basic frame-grabber elements (after Shortis et al, 1996)

The first functional element of the frame-grabber is called the analogue front end. The analogue front end performs a number of important analogue pre-processing tasks. It receives the signals from the camera, splits the signal into one path going through an analogue preprocessing stage to the analogue-to-digital converter and the other to the sync detection and synchronisation elements. DC restoration (or black level clamping) serves to adjust the video signal to a defined voltage level. A low pass filter is then applied to the signal to remove high frequency disturbances. The analogue offset and gain serve to adapt the analogue video signal to the input range of the analogue-to-digital converter. These adjustments also are necessary to adjust the average scene brightness prior to further processing in the frame-grabber.

The synchronisation signals used to specify the start of a field is known as the vertical sync or Vsync, and the start of a line is known as the horizontal sync or Hsync. If the

synchronisation signals are combined with the video signal, then the transmission is known as composite video. The synchronisation module first removes the video image information using a synchronisation separator, also known as a sync stripper. The signal from the sync stripper is directed to circuitry which detects the Vsync pulses using a combination of a low pass filter and edge detector. The Hsync pulses are detected using a phase locked loop (PLL) circuit. This generates the (internal) sampling clock and is controlled through the feedback of an internal Hsync. Imprecision in the PLL will result in the phenomenon known as line jitter (see section 3.2.1.1). Line jitter causes displacement of horizontal lines within the image.

Both the sampling clock and the preprocessed video signal are passed to the analogue to digital (AD) converter for digitisation. The AD converter translates the incoming, analogue video signal into discrete values which can be loaded into a solid state memory array as digital data. The digital image is constructed from individual pixels, sampled at a rate based on the internal or pixel clock, and other synchronisation signals. The digitised data and synchronisation signals are transmitted to the data buffer. The final function of the frame-grabber is digital to analogue conversion for output of live or grabbed images to a video monitor. The majority of frame-grabbers have additional capabilities such as graphic overlays or colourisation to identify features and a cursor to select areas of interest(Shortis et al, 1996).

3.2.1 Problems associated with frame-grabbing

3.2.1.1 Line Jitter

The clock frequency generated by the PLL (described in section 3.2) will fluctuate for some time after image capture until a stable lock on the Hsync pulses is obtained. Any imprecision in the phase comparison or PLL will result in the phenomenon known as *line jitter*. Line jitter causes displacements of horizontal lines within the image. Typically, a

precision of less than 10ns for the Hsync detection results in line to line jitter of less than 0.1 pixels (Shortis et al, 1996).

3.2.1.2 Line displacement

Line displacement is closely related to line jitter, but it results in a horizontal displacement by whole pixels of lines in the image. Typically, this effect causes the displacement of a complete line by one pixel to the right or left of where it should be. Line displacement can occur when the Hsync is derived from the composite video. When pixel synchronous sampling is used to eliminate line jitter, sub-pixel jitter is eliminated because the Hsync pulse can be isolated to a single pixel clock pulse. This, however, may cause line displacement. With an appropriate selection of the pixel clock phase to be used, line displacement can be eliminated.

3.2.1.3 Interline shifting

Interline shifting is a particularly common problem for airborne videography, although it can be observed in any application where the camera is in motion during image capture. Odd and even image lines should interlace together perfectly if recorded at the same moment, but an error is introduced because of the forward motion of the sensor. Realistically, the odd and even recordings do not occur simultaneously, and therefore any events that take place during this fractional amount of time will have some effect on the interlacing of the lines. In the case of airborne videography, these events primarily consist of the forward velocity of the aircraft and any changes in its angular orientation (Mitchell et al, 1995). The translation of this motion from the aircraft to the image takes the form of pixel shifting, which is particularly seen at straight-line feature boundaries.

3.2.1.4 Colour to monochrome conversion

Experiments were carried out (Gold, 1991) to test the effect of capturing monochrome images from a colour composite video signal. It was found that if a colour video camera is used, the video signal needs to be decoded in order to remove the colour modulation before being used as input into a black-and-white (B/W) frame-grabber. To demonstrate this, images were captured by the B/W frame-grabber, and the grey level profile for 1 row of pixels of each image were analysed.

The composite video signal was firstly inputted directly into the B/W frame-grabber, and an image was captured. It was noted that regular, incorrect striping patterns occurred in certain parts of the image, similar to the patterns shown in Figure 3.11. Figure 3.8 shows the grey level profile for 1 pixel row of this image:

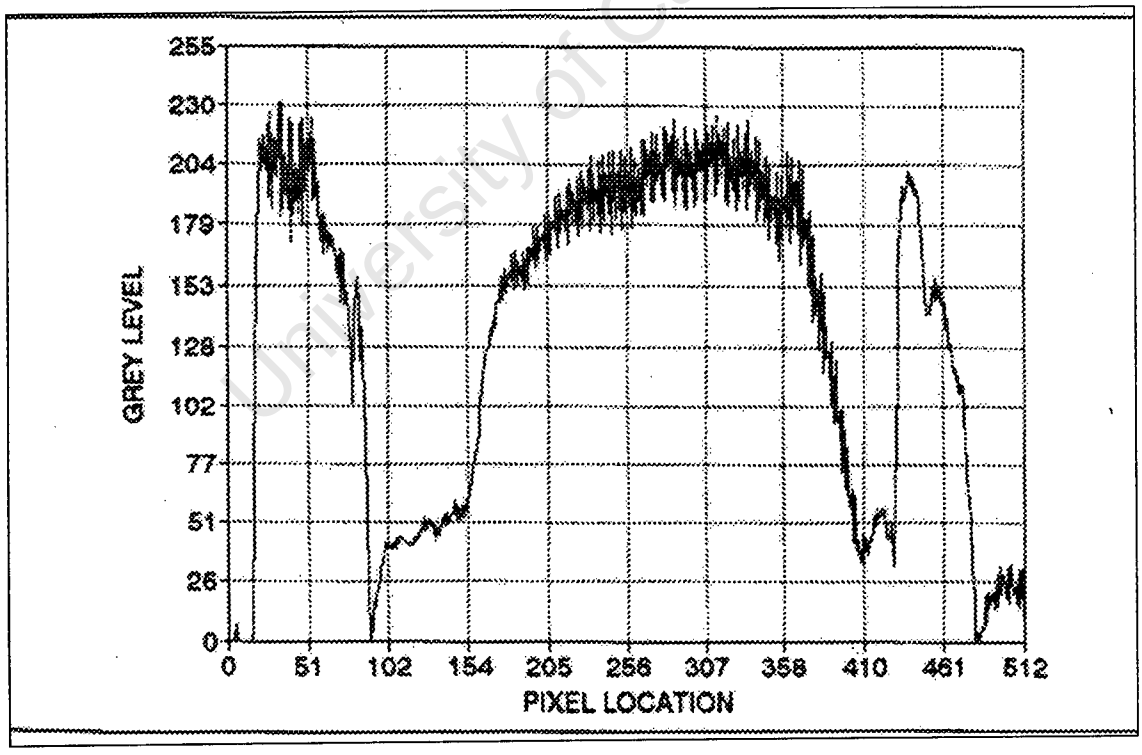

Figure 3.8 Grey level profile for 1 pixel row of a frame obtained by connecting camera directly into B/W frame-grabber (after Gold, 1991)

The incorrect striping occurred in areas where the grey level profile showed high frequency variations in the grey levels, i.e., pixels located between 25 - 95 and 165 - 358. These high frequency variations were the modulated colour information in the signal.

In order to have an accurate representation of the grey level profile of the object being imaged, the colour modulation needed to be removed from the composite video signal. This was acheived by a *Red Green Blue (RGB) decoder*. The RGB decoder connects the colour video camera to the B/W frame-grabber. The colour composite video signal is input into the RGB decoder, which decodes the signal and removes the colour modulation before outputting it to the frame-grabber. The decoded grey level profile for the same row of pixels shown in Figure 3.8 is shown in Figure 3.9:

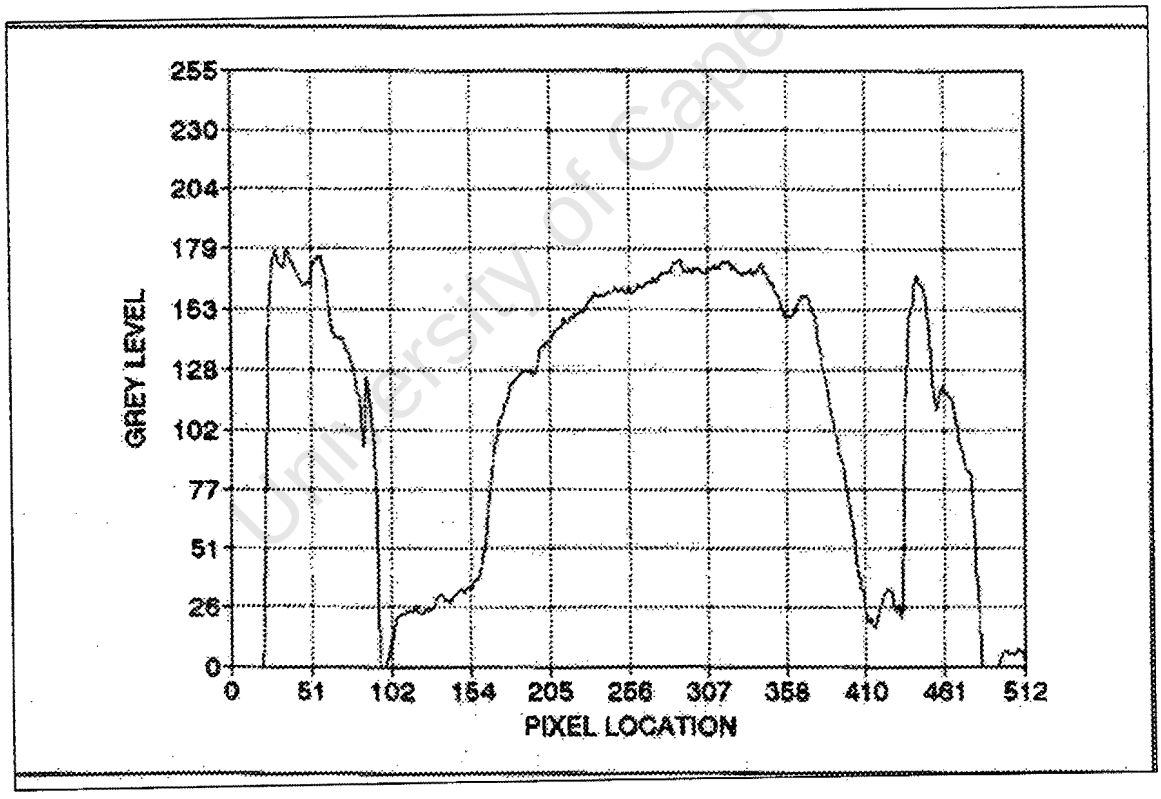

Figure 3.9 Grey level profile for 1 pixel row of a frame obtained by connecting camera to RGB decoder to frame-grabber (after Gold, 1991)

All the high frequency colour modulation was removed from the signal, leaving only the correct brightness value of the object being imaged. The conclusions drawn from these tests are that monochrome images should not be derived directly by a B/W frame-grabber from an RGB signal, and that the colour video signal has to be decoded first before being used as an input for such a frame-grabber.

3.2.2 Frame-grabber used in this project

A Silicon Video frame-grabbing card, manufactured by Epix, Inc. was initially intended to be used for the digitizing of images. Unfortunately, the software that was available with the board allowed only monochrome images to be captured. This presented problems when images were captured from the video cameras, as all the information is stored in colour on the video tapes, and capturing monochrome images directly from these tapes caused problematic effects in the resulting imagery. Firstly, line-displacement (described in section 3.2.1.2) was enhanced (see Figure 3.10).

Figure 3.10 Example of line displacement

Figure 3.10 shows a magnified section of an image of a calibration control frame. The image shows how a round retroreflective target is distorted by line displacement.

Another effect that was observed was regular, incorrect striping patterns in the image (see Figure 3.11).

Figure 3.11 Striping effect

Figure 3.8 shows a section of an image that is supposed to be relatively homogeneous in terms of texture and colour, and how this section is distorted by the striping effect. This resulted from the fact that a colour image was being digitised directly into monochrome, by the frame-grabber (see 3.2.1.4). A method of minimising this is to capture the image in colour, and then convert it into an 8-bit greyscale image. Bearing this in mind, the frame-grabber that was chosen and better suited to this application than the EPIX card for the project was a Video Blaster SE, manufactured by Creative Labs. Although it allows far less flexibility in specifying frame-grabbing options, its main feature is that it allows images to be captured in colour, which is essential in overcoming the problems encountered with the EPIX card. The size of the captured images were automatically set by the card at 628 x 372 pixels, which is less than the ideal size of 752 x 582 (the size of the CCD chip in the camera). This caused a loss of image information in both the horizontal and vertical directions. Horizontally, the 752 pixels of information that is produced by the chip is resampled into 628 pixels. This, however, is not the case in the vertical direction, as the first 372 lines of information in the image are preserved, and the remaining 256 lines are effectively "chopped" off the bottom of the image.

3.3 Data Processing

Most of the photogrammetric software that was used was developed in the Department of Geomatics. These algorithms allowed the calibration of the cameras to be carried out, as well as the final production of 3-D coordinates of points on the surface of the shipwreck. The names and a short description of these programs are listed in Table 3.1:

Program Name	Description
viewimg	Viewer for image files
search	Target location with sequential numbering of targets, uses weighted centre of gravity to locate target centres
idcon	Auto-identification of of control points according to a control file and manual identification of other targets
idsearch	Manual location and identification of target points
sobel	Sub-pixel edge detection using the sobel kernel and moment preserving
gradedge	Edge location by local maximum gradient in a 3x3 neighbourhood
exclude	Exclude edges from the edge file to be matched, within a selected polygon
match3_u.out	UNIX version for image matching using epipolar correlation and least squares multi-image geometrically constrained greyscale matching
photonet	Bundle adjustment program to calculate final orientation parameters
mtcimg2	Displays an image including the overlaid matched points on the template image
dltint	DLT intersection
dltprov	DLT resection

Table 3.1 Photogrammetry and Image Processing Software

For the rubber-sheeting, resampling and mosaicing of the Marconi Beam images, "ERDAS Imagine" software was used. "ArcView GIS" was used for overlaying the outlines of the shacks from the video mosaic onto the orthoimage produced of the settlement. "Surfer" software was used for the final representation of the Digital Elevation Model (DEM), as well as the production of a contour map, of the shipwreck. "Corel Draw!" was used for the manual mosaicing of a sequence of shipwreck images. Processing of the GPS data was carried out by Ashtech's "Prism" software.

3.4 Implementation

In order to generate spatial information of an object, in object-space coordinates, the steps to be taken are depicted in Figure 3.12:

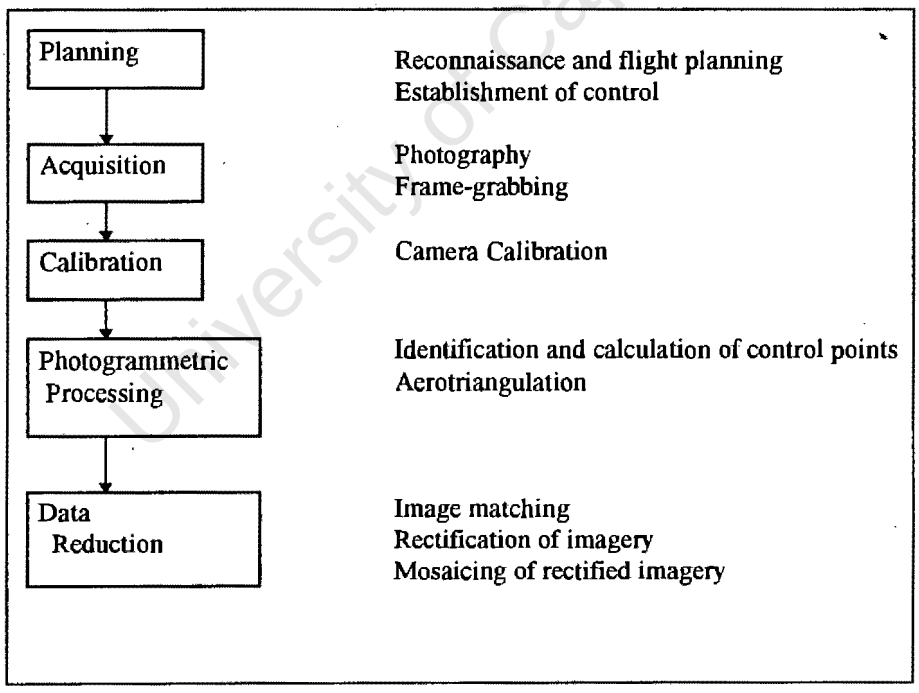

Figure 3.12 Steps in the implementation of system

Both projects will be discussed separately in chapters 4 and 5; however, there are aspects in the implementation of both projects that are similar.

CHAPTER 4

UrbanModeler - Mapping of Marconi Beam

4.1 Introduction and Background

Figure 4.1 View of Cape Town with Marconi Beam in foreground

The informal settlement at Marconi Beam in Milnerton serves as an application for the UrbanModeler project, whose objective is the development of image exploitation tools and methodologies for automating the geo-spatial modelling of urban settlements, primarily informal settlements. *Informal settlements* (or "squatter" settlements) may be defined as human settlements, which have developed in a predominantly unstructured and

unplanned manner. They are typically the result of an urgent need for shelter by the urban poor and can be characterised by a dense proliferation of small shelters composed of diverse materials (e.g. plastic and metal sheeting) and the absence of basic services. At present, the photogrammetric mapping of settlements, as done by the Cape Town City Council, is carried out manually by using a stereoplotter. This method has limitations in terms of cost and turn-around time, making it often unsuitable for the frequent updating of mapping needed for informal settlements. Digital photogrammetry, with imagery obtained by VRS was evaluated and has proved to be a quick, cheap alternative.

The Marconi Beam settlement is situated in an industrial park in the Milnerton area of Cape Town. Its dimensions are approximately 465m x 220m covering an area of roughly 7.8 ha. Saff (1996) gives an overview of the historical development of Marconi Beam. Currently, the settlement is composed of some 1200 shacks with some 6000 inhabitants. In 1994 a mapping of the settlement was carried out to determine the location of each shack. An accompanying social survey identified the owners of these shacks. One of the activities of the UrbanModeler project is support of the Marconi Beam Development Trust (MBDT) and the non-government organisation (NGO) the Development Action Group (DAG) in the process of relocating squatters in Marconi Beam to new low-cost residential housing called Joe Slovo Park, situated nearby. By agreement with the state body donating adjacent land for low-cost formal housing, these shack owners were vested with the right to relocate to Joe Slovo Park provided that the community, through the MBDT, ensured that:

- no newcomers would be allowed into the settlement;
- that each shack was demolished before individual relocation took place (Mason et al, 1997).

Fulfillment of these agreements has been subject to a number of pressures including housing shortages, a series of fires, changes in shack ownership and the implementation of the relocation process by a lottery system. The effect of these pressures is to require records of the spatial development of the settlement over time in order to be able to minimise the inevitable conflicts arising over "rights to move" as relocation proceeds. To

this end, a system for inventorising the shack stand and regularly recording the spatial changes in the settlement was required. The Department of Geomatics' involvement in Marconi Beam began in May 1996 in the form of acquisition of digital aerial imagery (see section 4.2) and updating of shack database records to establish a GIS of the settlement in ArcView. This GIS has been linked to a Microsoft Access database containing records of settlement residents and demographic and social attributes. Shack footprints as of December 1994 were also digitised based on aerial photography.

Because of the dynamic nature of the settlement, a low cost solution was deemed desirable in order to demonstrate feasibility for frequent updating, and this led to the choice of the sensors used.

4.2 Flight Planning and Photography

For the purposes of this project, two cameras, a Kodak DCS460c still video digital camera, and a Grundig lc 180 HE camcorder, were used. Mason et al (1997) reports on an investigation into the use of the DCS460 for low-cost, local area mapping. Flight planning was carried out, bearing in mind the sensor characteristics of both the cameras. The primary factors that influenced the flight plan for the settlement were:

- the dimensions of the objects of interest,
- the density of the settlement,
- the mapping accuracy requirements.

The objects of interest are primarily the shacks in the settlement, but it was also a goal to be able to interpret and measure settlement features such as public toilet facilities, these being useful indicators for quality of life studies using the resultant GIS. Thus, the imagery required for interpretation of the abovementioned settlement detail had a pixel size of approximately 20cm. The entire settlement could be covered in one strip of photography at flying heights of between approximately 300 - 1000m, therefore side overlap of photographs was of no concern. Because of the continuous image capture by

the camcorder, forward overlap of images could be controlled during the frame-grabbing process.

For the DCS460, as shacks are only 2-3m apart in the denser sections of the settlement, high forward overlap, i.e. 80% was deemed necessary for comfortable stereo viewing and measurement of the terrain and ground structures. The CCD chip in the DCS460 has a size of 27.6 x 18.4mm, made up of 3060 x 2036 pixels. The focal length is 27mm. Typical calculated flight parameters for one strip of imagery acquired by the DCS460 are listed in Table 4.1:

Parameter	**Value**
Image scale	1 : 19000
Flying height above ground (m)	530
Average terrain elevation (m)	15
Forward overlap	80%
Base length (m)	107
Image footprint on ground (m)	525 x 350
Number of strips	1
Number of photographs	6
Flying velocity (km/h)	160
Time between photographs (sec)	2.4

Table 4.1 Flight parameters for the Kodak DCS460 camera

The aerial photography was carried out on 26 May 1996. Five runs of photography, all at different flying heights (between 300 and 1000m) were carried out using both the cameras. The cameras were mounted in the aircraft, a Piper Arrow 200, using a custom-built mount fitted to a porthole below the passenger seat. The cameras were positioned such that the longer side of the chip was perpendicular to the flying direction of the plane. GPS

navigation provided coarse positioning for the strip, however, actual image acquisition had to rely on the operator's judgement due to low flying height and imprecision (±100m) of the navigation device. During the photography, it was important that the focal length remained constant, hence, on the camcorder, the "auto-focus" option had to be disabled, and the focussing ring on the DCS460 was firmly taped down at infinity focus. The camcorder's focal length was set for the widest angle of photography.

The rate of image capture of the camcorder is 30 frames/second, so a flying speed of 160 km/h will result in approximately 1 frame captured every 1.5 metres. Interline shifting (see section 3.2.1.3) was introduced because of the forward motion of the plane. It could not however be observed in the frame-grabbed imagery because the resolution of the imagery was too coarse; thus this effect was ignored.

4.3 Camera Calibration

The aim of camera calibration is to calculate the interior orientation parameters of the camera, i.e. principal distance, principal point and additional lens distortion parameters (see Section 2.3). For applications such as rubber-sheeting, calibration is not needed. The DCS460 imagery would later be used for the production of an orthoimage, so this camera had to be calibrated. Calibration of the camcorder was carried out in the way described in Section 2.3.

4.4 Establishment of Control

It was originally intended to use town survey marks located around Marconi Beam for control, however, the resolution of the camcorder (and Kodak DCS) made it impossible to identify these marks in the imagery. After the photography was completed, the images were analysed for features that could be used as control points, and these points were coordinated by means of a GPS survey. Typical features include road markings, corners

of shacks, and corners of public toilets. Two kinematic GPS surveys were carried out for this purpose, and a third was done in order to map the boundary of a fire that had swept through Marconi Beam on 9 October 1996. The control points were later be used in the rubber-sheeting of the imagery (see section 4.5).

4.5 Rubber-sheeting

The images of the squatter settlement were distorted primarily by the orientation of the camera and the undulations in the terrain. Therefore, these images needed to be geometrically corrected so that the area being photographed could be represented on a planar surface and conform to the other images of the settlement. *Rubber-sheeting* is the process of transforming the data from one coordinate system into another coordinate system by using an *n*th order polynomial. It is one of the methods of rectifying distorted imagery,and involves establishing mathematical relationships between the addresses of pixels in an image and the corresponding coordinates of those points on the ground.

Rubber-sheeting involves the conversion of image coordinates to some other coordinate system, called the reference system, in this case, the ground coordinate system, defined in section 2.4.2. This process of assigning map coordinates to image data is called *georeferencing*. There are several reasons for rubber-sheeting imagery, e.g. the creation of an accurately scaled photomap, extracting accurate distance and area measurements, mosaicing images, overlaying an image with vector data, or performing any other analyses that require precise geographic locations. Rubber-sheeting is only necessary if there is some distortion in the image. For example, an image produced by scanning a paper map need not be rectified unless there is some skew or rotation of the image. This sort of image only needs to be georeferenced to be able to extract spatial information from it.

Rubber-sheeting an image involves the following three steps:

1. Location of ground control points (GCPs) in the imagery.

2. Computation of a transformation matrix.
3. Creation of an output image that results from the original image being resampled to conform to the new coordinate system.

4.5.1 Ground Control Points (GCPs)

All the rubber-sheeting and mosaicing done on the Marconi Beam images were accomplished by using the "ERDAS IMAGINE" software package. GCPs are specific pixels in the image for which the ground coordinates are known, in this case, they are the control points whose coordinates were obtained by GPS surveys. When working with GCPs in ERDAS, two sets of coordinates need to be differentiated between, the *source* coordinates and the *reference* coordinates (ERDAS, 1994). Source coordinates are the X and Y image coordinates of the GCPs in the image being rectified. Reference coordinates are the ground coordinates of the GCPs to which the source image is being rubber-sheeted. Using the "GCP Editor" function in ERDAS, the GCPs are identified in the imagery, and their corresponding ground coordinates are entered.

4.5.2 Transformation Matrix

The transformation matrix is derived from polynomial equations that are used to convert source coordinates into rectified coordinates. Depending upon the distortion of the imagery and the number of GCPs used, complex polynomial equations may be required to express the needed transformation. The degree of complexity of the polynomial is expressed as the *order* of the polynomial; it is simply the highest exponent used in the polynomial. The *transformation matrix* that is calculated consists of coefficients that are used in the polynomial equations to convert the coordinates. The goal in calculating the coefficients of the transformation matrix is to derive the polynomial equations for which there is the least possible error when they are used to in the transformation of the source

coordinates into the reference coordinates. The minimum number of GCPs required to perform a transformation of order *n* is:

$$(n+1)(n+2)/2 \qquad \textbf{(4.1)}$$

Hence, the minimum number of GCPs required for a first order transformation is three, six for a second order transformation, and ten for a third order transformation. It is nearly never possible to derive coefficients that produce no error in the transformation if redundant observations are present.

The coefficients of the polynomial represented by the curve will obviously be influenced by every GCP. *Residuals* are the distances between the source and retransformed coordinates in X and Y directions, denoted by d_x and d_y . These residuals are used for accuracy assessment by calculating a root-mean-square (RMS) error. The RMS error in the x direction is calculated as :

$$RMS_x = [\tfrac{1}{n}\sum_{i=1}^{n}(d_{x_i})^2]^{1/2} \qquad \textbf{(4.2)}$$

where

d_{xi} is the residual of the I^{th} GCP and

n is the number of GCPs.

The RMS error in the y direction is calculated similarly and the x and y RMS errors can be combined to yield one planimetric RMS error :

$$RMS_{xy} = [RMS_x^2 + RMS_y^2]^{1/2} \qquad \textbf{(4.3)}$$

RMS error is expressed as a distance in the source coordinate system, (the image coordinate system) i.e., in number of pixels. For example, an RMS error of 2 means that the reference pixel is 2 pixels away from the retransformed pixel.

The order of the transformation matrix depends on the imagery being rectified. A first order transformation is a linear transformation, while nonlinear transformations correct nonlinear distortions. Examples of the use of second order transformations are converting latitude/longitude data to a planar projection and correcting distorted imagery due to camera lens distortions. Third order transformations are used with distorted aerial photographs, on scans of warped maps and with radar imagery. Fourth order transformations can be used on very distorted aerial photographs. For rectification of the DCS and video imagery, second order transformations were used for the rubber-sheeting. A second order transformation, according to Equation 4.1, requires a minimum of 6 GCPs, while a third order transformation requires 10 GCPs per image. Since the number of GCPs in each image was not always greater than 10, the second order transformation was used, and sufficed. Sometimes orders higher than third are used but care must be taken to avoid the introduction of worse errors than those to be corrected. The effects of nonlinear transformations are illustrated in Figure 4.2:

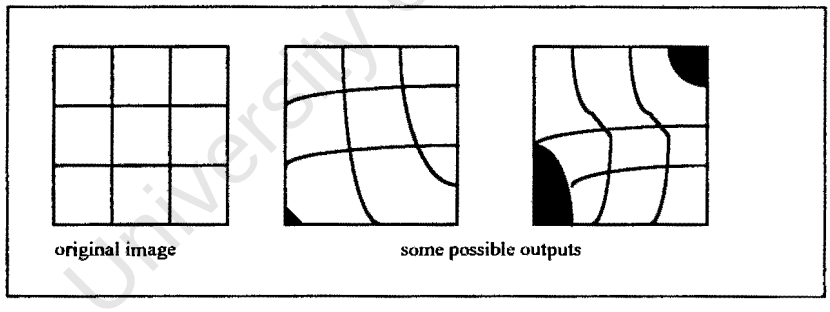

Figure 4.2 Effect of nonlinear transformations

4.5.3 Resampling

Since the pixels of the new coordinate system may not align with the pixels of the original grid, the pixels must be *resampled*. Resampling is the process of interpolating grey values for the pixels on the new grid from the values of the source pixels. Having determined the mapping polynomials by using the GCPs, the next step is to find points in the output image

corresponding to each location in the pixel grid of the original image. There are two main methods of resampling an image, i.e. the direct and indirect method. In the direct method, the grey values of the original image pixels are transferred to the grid of the output image, as shown in Figure 4.3 :

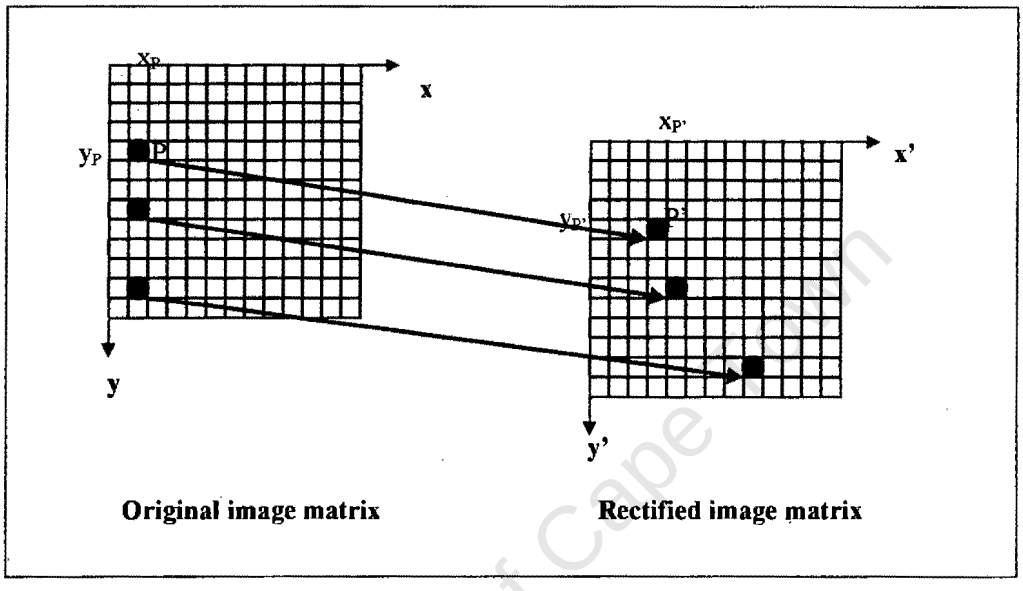

Figure 4.3 Direct method

The principle of the indirect method is that for each pixel of the rectified image, the grey values are taken from the original image, as shown in Figure 4.4 :

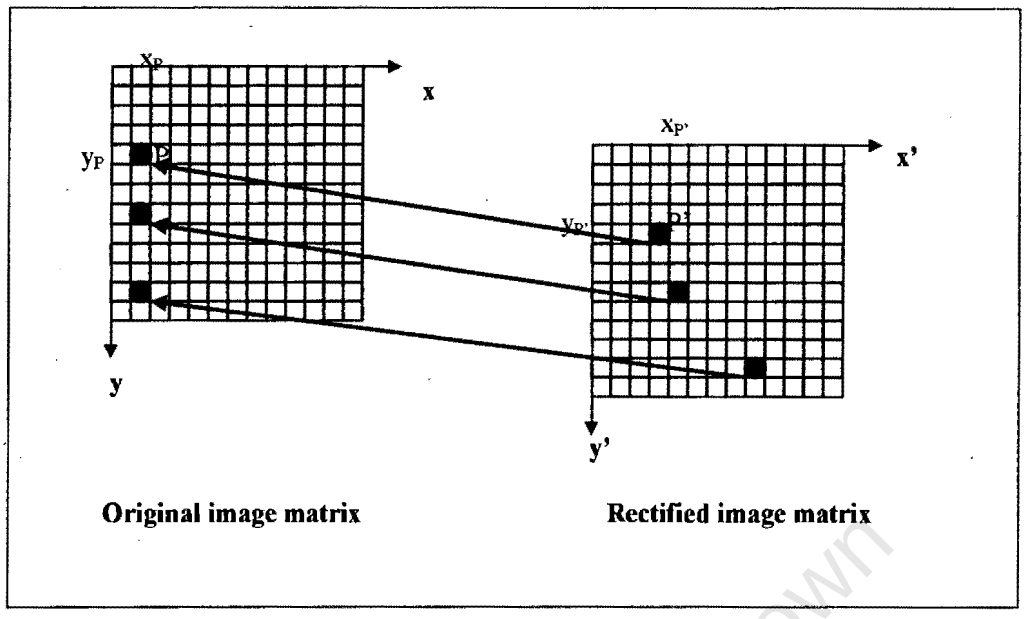

Figure 4.4 Indirect method

The most widely used methods of resampling are:

- nearest neighbour,
- bilinear interpolation, and
- cubic convolution.

4.5.3.1 Nearest neighbour

This type of resampling simply chooses the actual pixel that has its centre nearest the point located in the image. This pixel is then transferred to the corresponding display output grid location. The new image then consists simply of rearranged pixels that have their original brightnesses.

4.5.3.2 Bilinear interpolation

Bilinear interpolation uses three linear interpolations over the four pixels that surround the point found in the image corresponding to a given display grid position. The process is illustrated in Figure 4.5 :

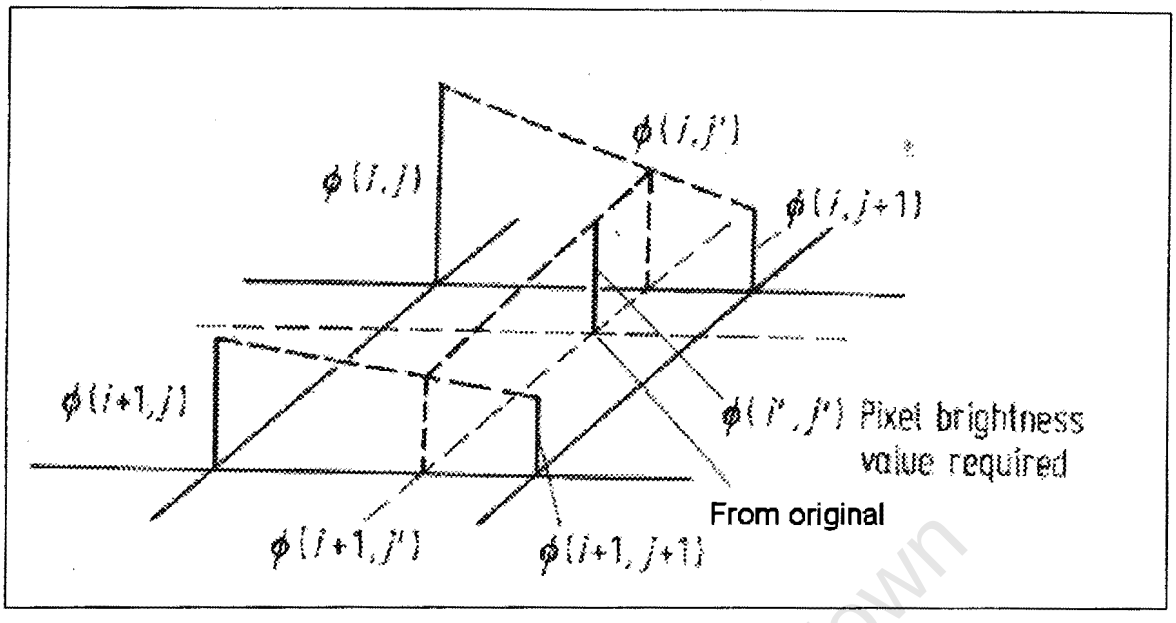

Figure 4.5 Bilinear interpolation

Two linear interpolations are performed along the scan lines to find the interpolants $\phi(i, j')$ and $\phi(i+1, j')$. These are given by

$$\phi(i, j') = j'\phi(i, j+1) + (1-j')\phi(i, j)$$
$$\phi(i+1, j') = j'\phi(i+1, j+1) + (1-j')\phi(i+1, j) \quad \textbf{(4.3)}$$

where ϕ is the grey level and (i', j') is the position at which an interpolated grey level is required. The position is measured with respect to (i, j). The final step is to interpolate linearly over $\phi(i, j')$ and $\phi(i+1, j')$ to give

$$\phi(i', j') = (1-i')\{j'\phi(i, j+1) + (1-j')\phi(i, j)\} + i'\{j'\phi(i+1, j+1) + (1-j')\phi(i+1, j)\} \quad \textbf{(4.4)}$$

4.5.3.3 Cubic convolution

Cubic convolution uses the surrounding sixteen pixels around the pixel of interest. Cubic polynomials are fitted along the four lines of four pixels surrounding the point in the image to form four interpolants. A fifth cubic polynomial is then fitted through these to synthesise a grey value for the corresponding location in the output grid, as shown in Figure 4.6 :

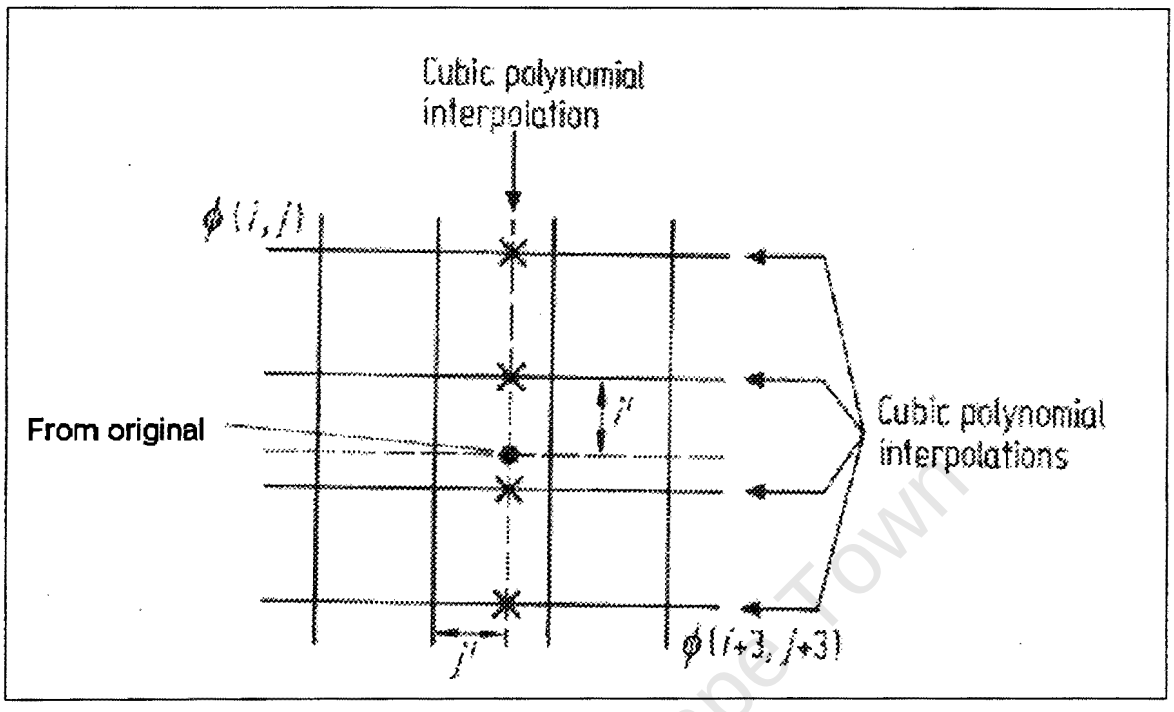

Figure 4.6 Cubic convolution

The algorithm that is used to perform cubic convolution interpolation is :

$$\phi(i', j') = j'\{j'(j'[\phi(i, j+3) - \phi(i, j+2) + \phi(i, j+1) - \phi(i, j)] + [\phi(i, j+2) - \phi(i, j+3) - 2\phi(i, j+1) + 2\phi(i, j)]) + [\phi(i, j+2) - \phi(i, j)]\} + \phi(i, j+1) \quad \textbf{(4.5)}$$

This expression is evaluated for each of the four lines of four pixels shown in Figure 4. To yield the four interpolants $\phi(i, j')$, $\phi(i+1, j')$, $\phi(i+2, j')$ and $\phi(i+3, j')$. These are then interpolated vertically according to

$$\phi(i', j') = i'\{i'(i'[\phi(i+3, j') - \phi(i+2, j') + \phi(i+1, j') - \phi(i, j')] + [\phi(i+2, j') - \phi(i+3, j') - 2\phi(i+1, j') + 2\phi(i, j')]) + [\phi(i+2, j') - \phi(i, j')]\} + \phi(i+1, j') \quad \textbf{(4.6)}$$

(after Richards, 1986)

All three methods of resampling were tested on the Marconi Beam imagery by inspection, and the resulting imagery obtained by bilinear interpolation proved to be most appropriate, as most settlement detail was preserved.

4.6 Mosaicing of Rubber-sheeted Imagery

After the rubber-sheeting of all the adjacent images of a run of photography was done, they were mosaiced together to form one large georeferenced mosaic of the settlement. ERDAS allows the user a number of options in controlling the mosaicing of the imagery. One can carry out the function of *contrast matching*, which serves to match the contrast of the adjacent images. Contrast matching can either be performed in the overlapping area between adjacent imagery, or over the entire area being mosaiced. This prevents any obvious differences in contrast showing up in the final mosaic. The most effective contrast matching option that was used was the "whole image" option.

Because of the much higher resolution of the DCS imagery, the GCPs in this imagery could be more accurately identified, thus the RMS errors derived from the transformations were lower than those of the video imagery. For the video imagery derived from the greatest flying height (resolution of approximately 1 pixel = 0.75m on ground), two images were sufficient to cover the entire settlement, and this is shown in Figure 4.7:

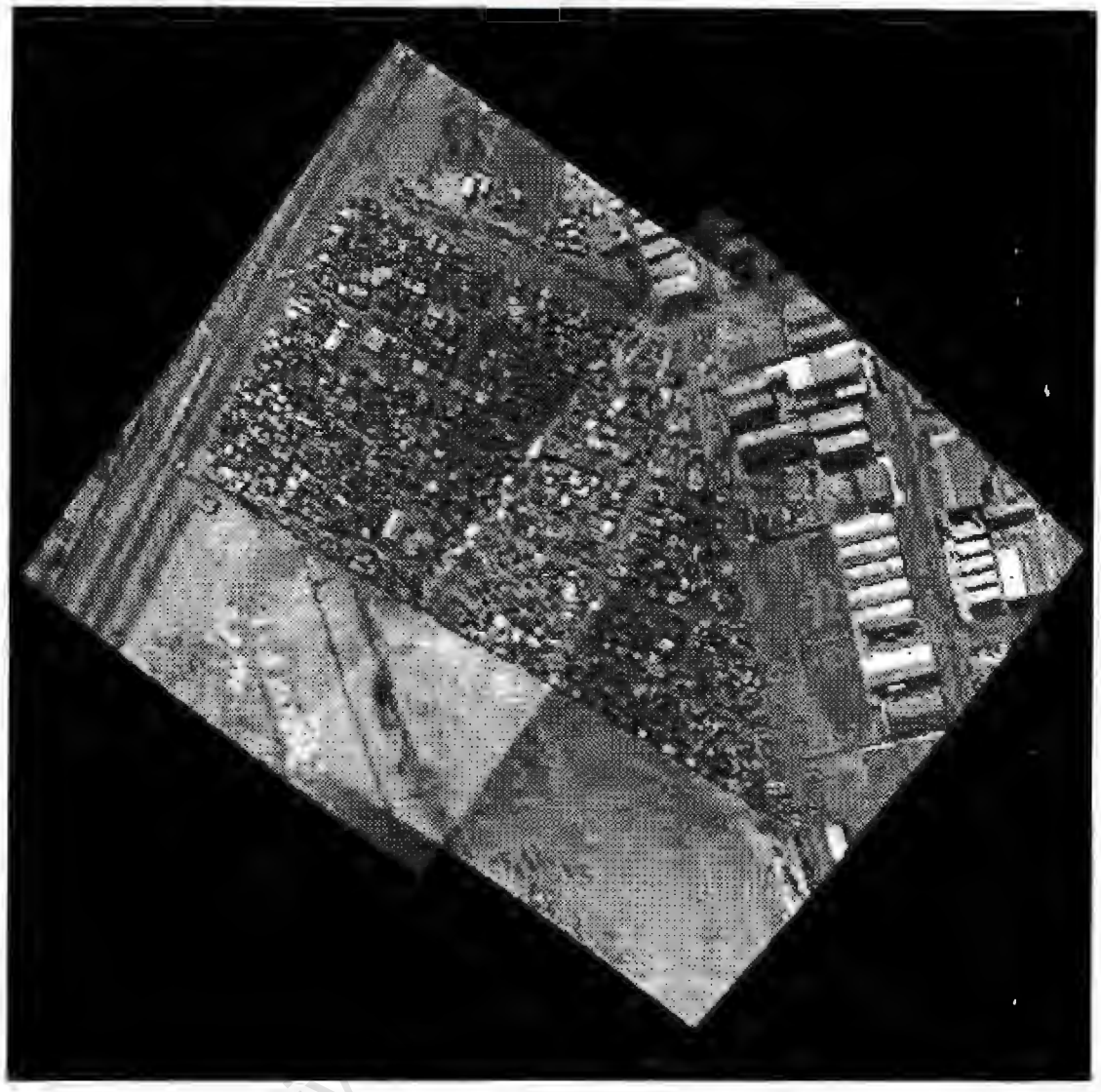

Figure 4.7 Mosaic of two rubber-sheeted images

The RMS errors for the two images used in this mosaic are shown in Table 4.2:

Image No.	X RMS	Y RMS	Total RMS
1	1.21	0.45	1.3
2	0.37	0.24	0.44

Table 4.2 RMS errors for 2 image mosaic (in pixels)

The RMS errors are expressed in pixels, so the highest RMS of 1.3 represents an error of approximately 0.9m on the ground.

Another mosaic was produced for imagery from a lower flying height. This is shown in Figure 4.8. Three images had to be used for the production of this mosaic.

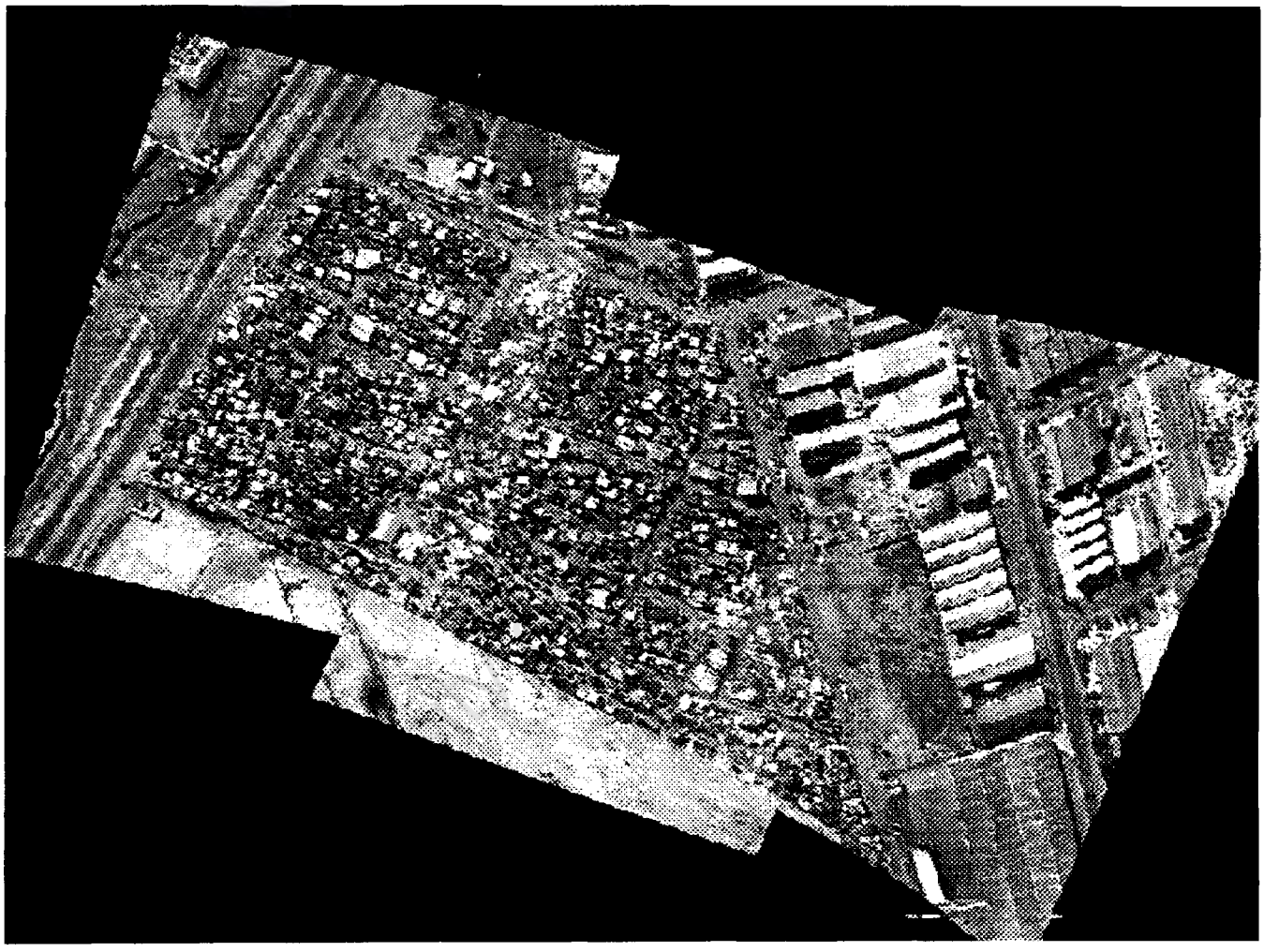

Figure 4.8 Mosaic of three rubber-sheeted images

The RMS errors produced from the transformation of these three images are shown in Table 4.3:

Image no.	X RMS	Y RMS	Total RMS
1	0.66	0.64	0.92
2	0.56	0.33	0.65
3	0.48	0.26	0.55

Table 4.3 RMS errors for 3 image mosaic (in pixels)

Table 4.3 shows that the attainable positioning accuracy of this mosaic is better than the two-image mosaic, as the imagery has a lower total RMS error, and the level of detail of this mosaic is higher due to the lower flying height. It should be noted that the RMS errors shown in Table 4.2 and 4.3 indicate the accuracy of the polynomial fit to the coordinates of the GCPs, and not the accuracy of point positioning of the mosaic. One can, however, make a rough judgement of the overall positioning accuracy that can be expected from the mosaic. In this respect, one can expect the attainable positioning accuracy of the mosaic to be within a metre. This may not, however, be the case in areas where there is a large terrain distortion which might not have been adequately compensated for in the modeling process.

4.7 Accuracy Assessment

The accuracy of using camcorder imagery for the determination of three-dimensional object space coordinates was also tested. Firstly, initial exterior orientation parameters of three overlapping images were calculated by using a DLT resection algorithm (see "dltprov" in Table 3.1), using 11 GCPs in image 1, 7 GCPs in image 2, and 7 GCPs in image 3. These orientation parameters were then refined by means of the bundle adjustment program called "photonet". Two other GPS points that were observed in all three images were used as check points (i.e. their coordinates were calculated by intersection, and then checked against their known coordinates). The residuals of the calculated object space coordinates are shown in Table 4.4:

Point No.	X Residual	Y Residual	Z Residual
1	1.13	0.12	1.28
2	0.83	0.39	1.32

Table 4.4 Residuals of check point coordinates (in m)

The resulting RMS errors, as calculated by equation 4.2, are shown in Table 4.5:

X RMS	Y RMS	Z RMS
0.99	0.29	1.30

Table 4.5 RMS errors of check point coordinates (in m)

This confirms that the attainable planimetric positioning precision is within a metre. The worst RMS is the Z coordinate RMS, and this was expected due to the fact that the imaging geometry is worst for calculating depth.

4.8 Overlays

In order to assess the usability of the mosaic, it was imported into "Arcview GIS", and was overlaid with different types of data. Firstly, the outline of the fire determined by kinematic GPS survey (see section 4.4) was overlaid over the mosaic. This is shown in Figure 4.9. This end product is a good tool in damage assessment, as the number of shacks and people affected by the fire can be easily determined. This is accomplished by means of manipulating the coverages in the database.

A coverage (or *theme*) is essentially a layer of information that may be defined by a set of points, lines, or polygons. The shack footprints were defined as a set of polygons, with each polygon linked to its specific attributes, e.g. the address and names of its occupants. Once the outline of the fire was clipped to the coverage of the shack footprints, the database was queried and a list of the people who lived within the fire boundary was ascertained.

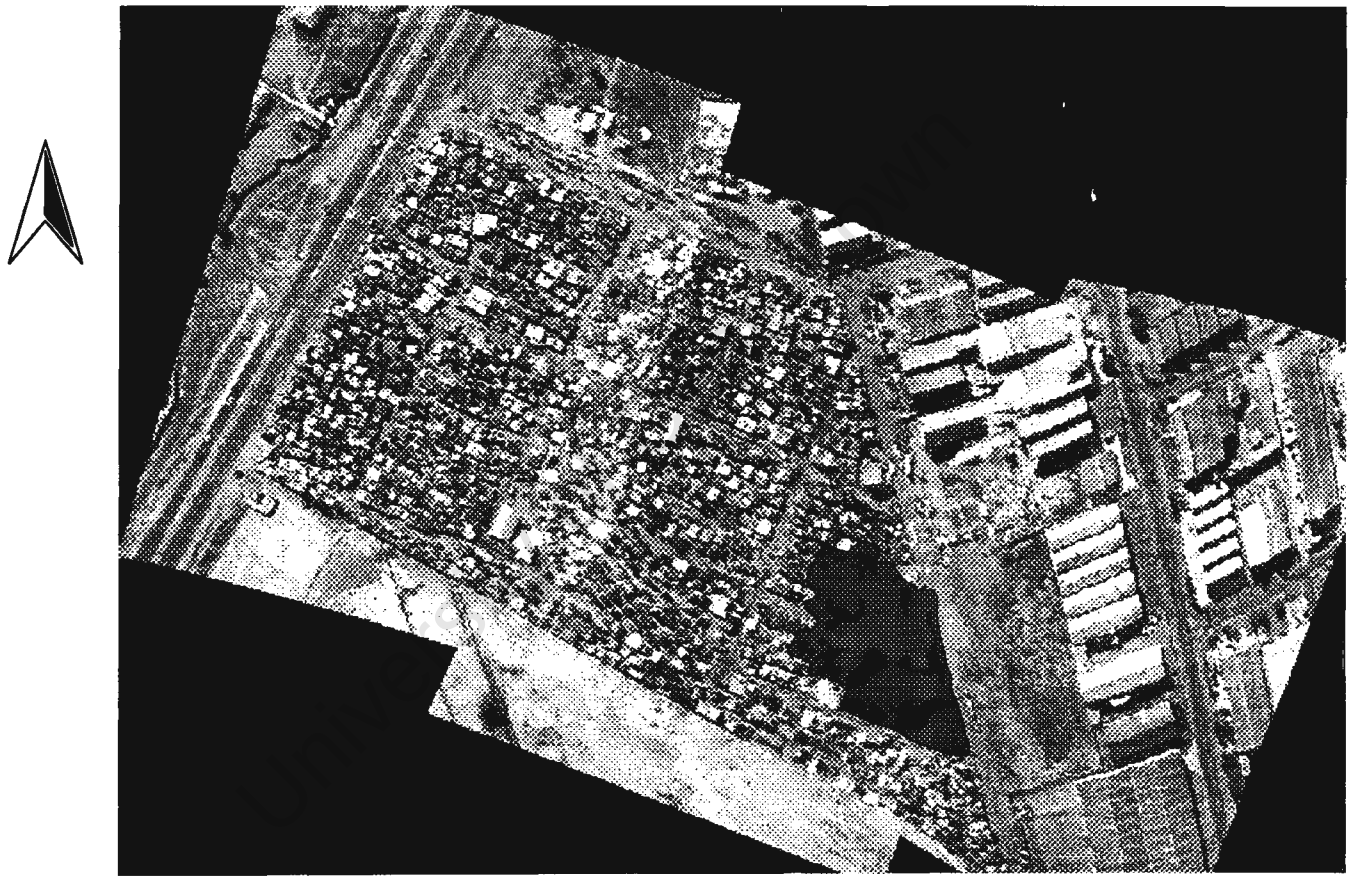

Figure 4.9 Fire boundary overlay

The plan of a road that is to be built through Marconi Beam was also overlaid over the mosaic, and this is shown in Figure 4.10. The people occupying the shacks that were located in the path of the road were to be relocated first, and this number of people was

easily obtained. This was carried out by querying the database for a list of people who lived within the boundary of the proposed road.

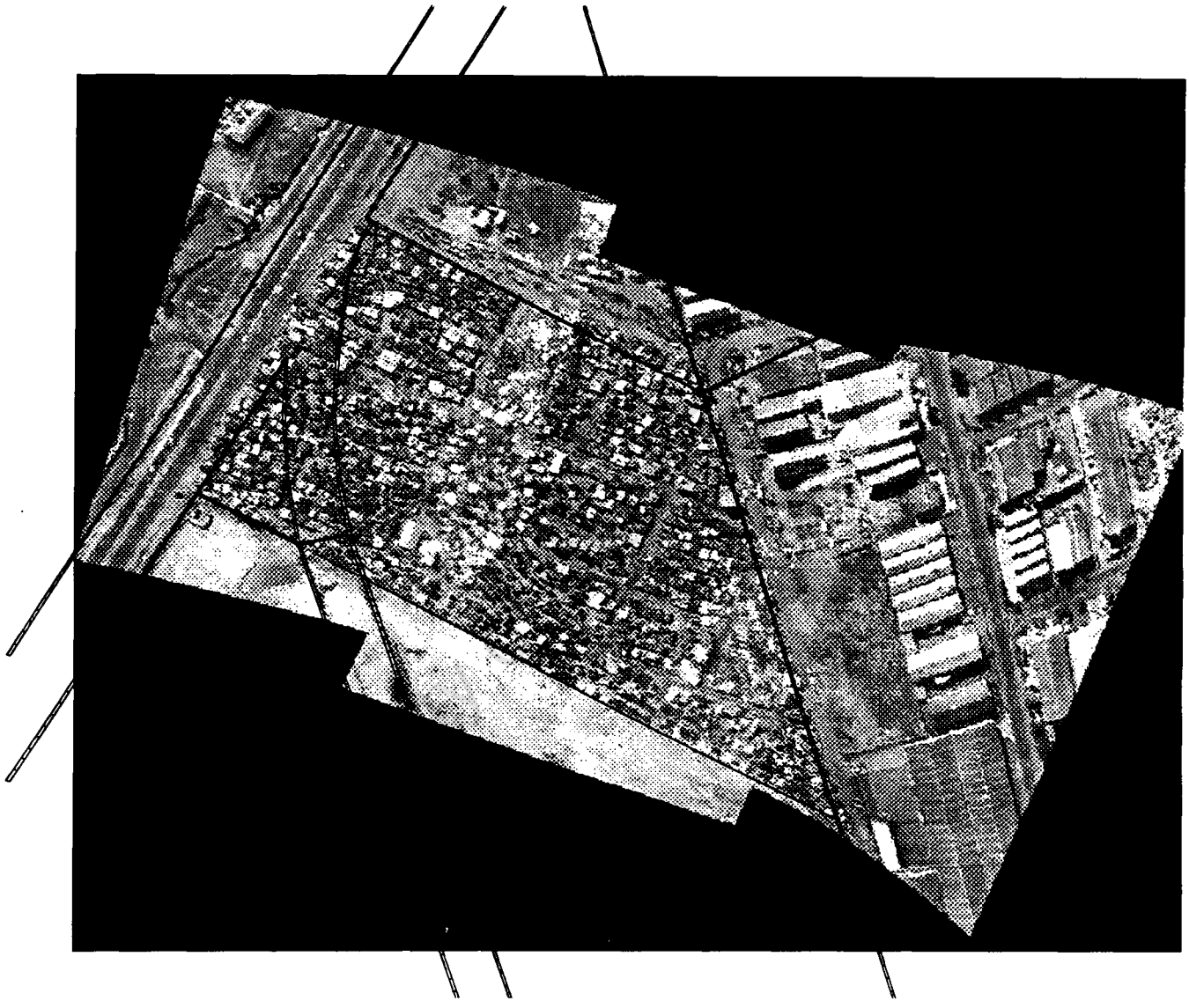

Figure 4.10 Proposed road overlay

To test the accuracy of the mosaic, the outlines of a number of shacks in the south-west section (where the road passes through the settlement) were digitised manually by tracing their outlines on the screen with a mouse. This was also done for an orthophoto of the

settlement produced by orthorectification of the DCS460 imagery. These two sets of outlines were overlaid, and this is shown in Figure 4.11:

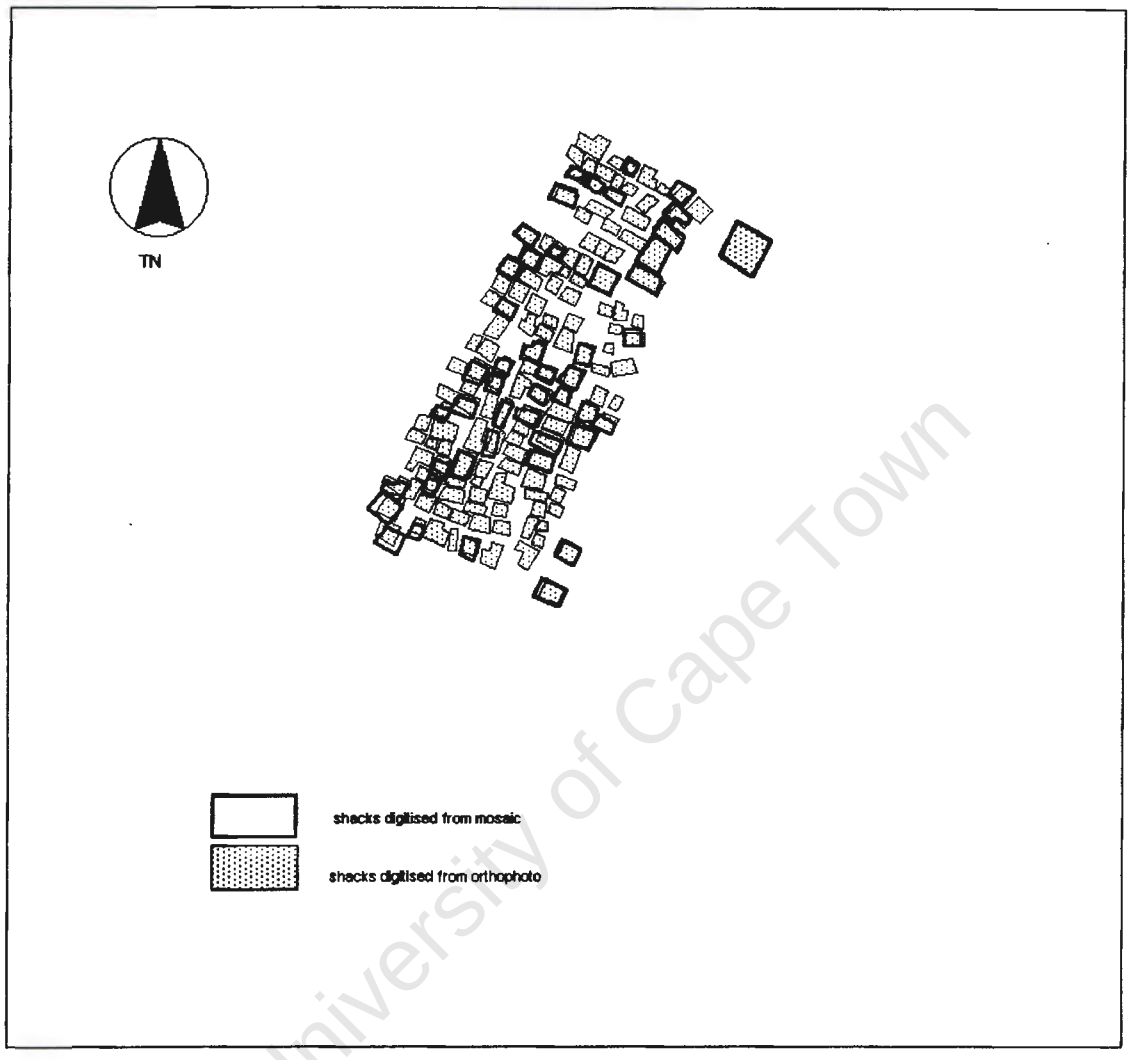

Figure 4.11 Outlines of shacks digitised from mosaic and orthophoto

It can be seen that the mosaic produced from rubber-sheeted imagery can be used for the identification of individual shacks, although not all the shacks in this part of the settlement could be identified. Of the 143 shacks that were identified from the DCS imagery, only 47 were identified in the mosaic. This poor (33%) success rate of shack identification can be expected to increase with an increase in image resolution. The resolution of the camcorder imagery (628 x 372) was lower than the resolution that could have been attained if a frame-grabber that allows one to specify the size of the output image was

used. The optimal size of the image would be the size of the CCD chip in the camcorder (752 x 582 pixels).

The effect of resolution can be seen in Figure 4.12:

Figure 4.12 Comparison of camcorder and DCS460 imagery

Figure 4.12 shows a close-up of the same area of the squatter settlement obtained from the camcorder imagery and the Kodak DCS460 imagery. The resolution of the Kodak DCS460 still video camera (3060 x 2036) is much higher than the resolution of the camcorder (752 x 582), thus making the task of shack identification much easier.

CHAPTER 5

The Mapping of the Bloubergstrand Shipwreck

5.1 Introduction

In June 1996, the author was approached by the Archaeology department at UCT to carry out the photogrammetric mapping of a newly discovered shipwreck. The shipwreck was uncovered on the shore of Dolphin Beach at Bloubergstrand on the Cape West coast, after the sand level along a stretch of the beach was scoured by storms. Milnerton Municipality officials were contacted by members of the public concerning large iron nails protruding from the surface of the sand. The site was excavated and the keelson and some of the side planking of the ship was found. Two large broken pieces of timber were found on either side of the middle section of the wreck, and was identified as the keelson of the ship. Most of the loose timbers that were found were identified as ribs of the ship and were originally attached to it but had been broken off during the excavation of the site by the Municipality. Parts of the hull planking were discovered on both sides of the keelson but not extending further than 1.5m from it. Thickness of the planking ranged from 4.5 to 7.5cm.

Video image capture was carried out with a Sony TR880e Hi-8 video camera that was purchased by the Archaeology department for the purpose of recording shipwrecks. This approach in image capture for archaeological documentation enables the acquisition of appropriate imagery after image capture. Provided the video image capture was designed accordingly, any desired stereo overlap with a continuous image coverage of the shipwreck could be attained.

5.2 Reconnaissance

The fieldwork, including the preliminary reconnaissance, lasted three successive days. The first day was taken up by a reconnaissance of the site and preparations for the next two days. The wreck was on the beach, approximately 10m from the high water line, and had been excavated by the local municipality and the archaeologists. It should be noted that the main constraint for the fieldwork was that the shipwreck could be left exposed only for a very short period of time. A cold front was approaching and since the wreck was so close to the water, stormy weather would cause the tide to flood the site, thus the wreck would be re-buried by the municipality 3 days later.

5.3 Network Design

In order to optimise the photogrammetric measurement of an object in terms of accuracy and reliability, particular attention must be given to the design of the network. Photogrammetric networks are comprised essentially of three forms of spatial knowledge (see Figure 5.1).

1. a description of the workspace and any obstructions in it,
2. the position and orientation of the features to be measured, and
3. the configuration of the camera stations in the network (Mason, 1994).

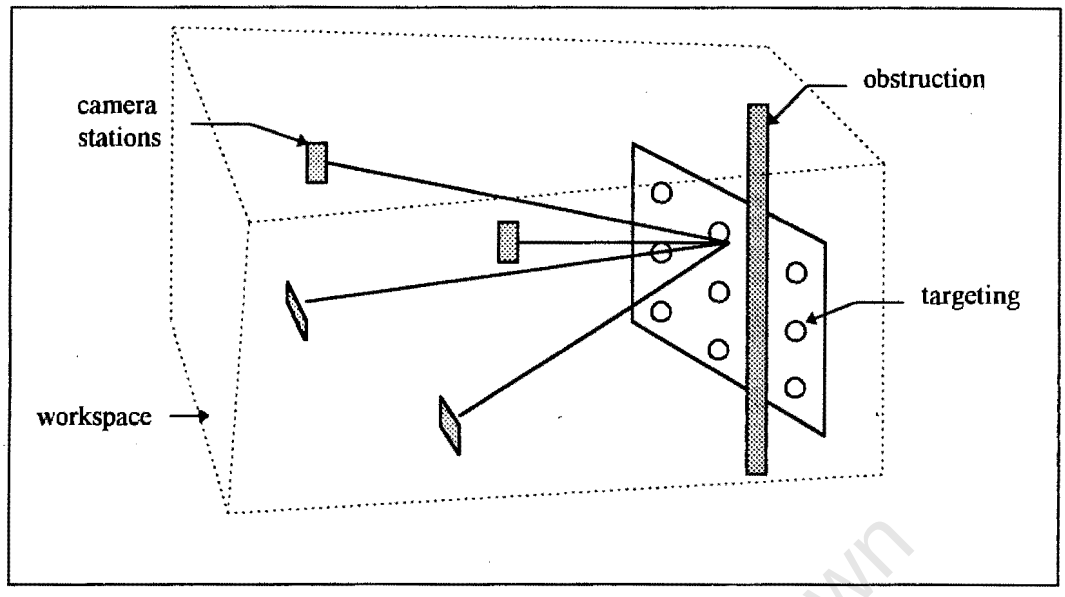

Figure 5.1 Forms of spatial knowledge in network design

Firstly, selecting an observation scheme involves consideration of any workspace restrictions. Such restrictions affect selection of the optimal camera-lens combination for the design. Secondly, targeting an object requires marking the features that characterise the object. Thirdly, placement of camera stations involves reasoning within the constraints.

5.3.1 Targets

All the targets used for the project were circular, retroreflective targets which showed up well in the imagery. Retroreflective targets possess a number of desirable properties and are widely used. They produce uniformly exposed, high contrast images with little influence from any ambient lighting. The use of targeted points continues to be recommended wherever possible. The time spent attaching and removing the targets is more than compensated for by the increase in accuracy and the removal of misidentification errors at all stages of the photogrammetric process. When dealing with archaeological artifacts, the greatest possible care should be taken so as not to disturb or damage the object being measured.

The total number of targets used and their distribution is of importance for the measurement process. The minimum number of targets is influenced by the number required for the bundle adjustment. The maximum number of targets is set by the maximum practical density of targeting, and/or considerations of design and implementation economy. A sufficient number of well-distributed target points in the imagery is needed for an adequate solution to the bundle adjustment. In object space, the targets should ideally be distributed in three dimensions. In image space, it is not absolutely necessary to achieve a dense, uniform distribution of target images in each image, but over the sequence of images forming the network.

5.3.2 Camera Placement

When dealing with video imagery, the exterior orientation of the frame-grabbed image has to be calculated, representing the position of the camcorder in object space at the time of the original capture of the frame-grabbed image. Unlike still photography, the camcorder allows for continuous image capture, thus making the camera placement task a less deliberate action because placement of the video camera in one position essentially allows for many images of the object to be captured. A great advantage of this is that if the geometry of the imagery is not adequate for a proper bundle adjustment solution, one can always play back the video tape through the frame-grabber and capture images that allows for a better network geometry, assuming that a good range of perspectives was captured in the photography.

Configuring a strong imaging geometry is fundamental in designing a photogrammetric network. By imaging geometry is meant the geometry of the intersection of rays at the object points. This geometry has a major influence on the precision of triangulation. To visualise the effects of precision caused by the placement of camera stations, see Figure 5.2. The impact of changes in network design for the measurement of an antenna dish is

shown. The point determination precision is reflected by variations in the shape and dimensions of the point error ellipsoids. The result of the inclusion of extra image stations and convergence of image axes leads to an improvement in accuracy (Mason, 1994).

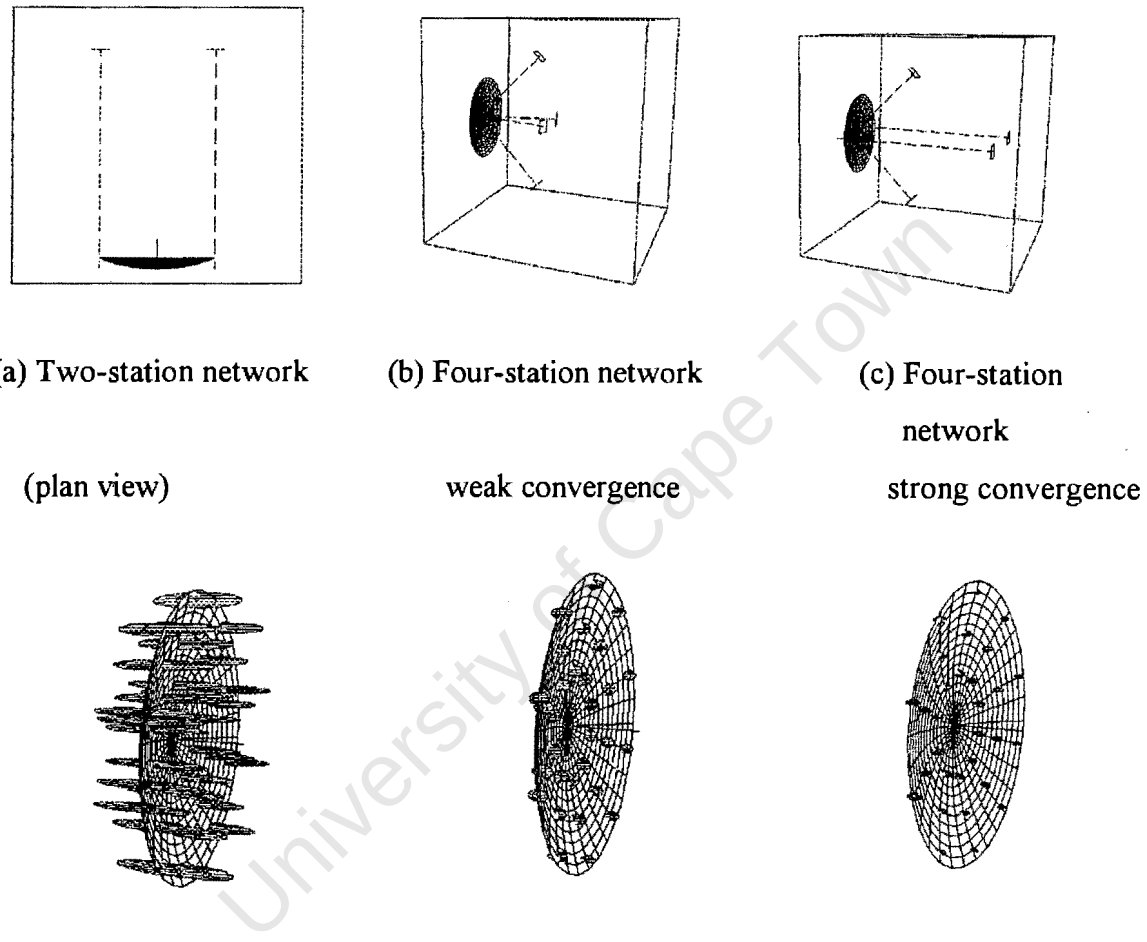

Figure 5.2 Effect of camera placement on accuracy (after Mason, 1994)

A *generic* network describes an ideal configuration of camera stations that can be employed to provide a strong imaging geometry for a class of network design problems. Generic network designs have been suggested for a number of primitive surfaces, e.g. planar surfaces like long and short walls, and the external and internal surfaces of a hemisphere. For mapping of the shipwreck, the primitive surface deemed most suitable

was the long wall. For mapping of a long wall, the network shown in Figure 5.3 is most suitable.

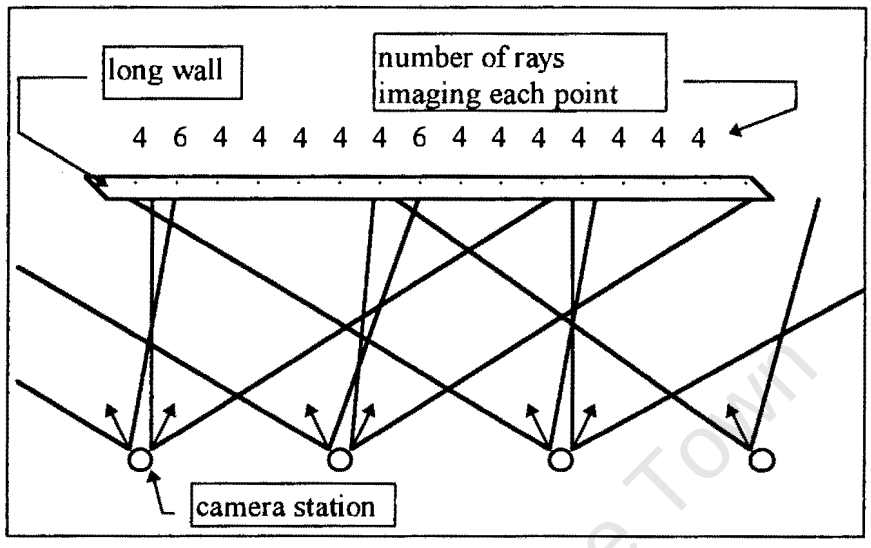

Figure 5.3 Measurement of a long wall (after Schlögelhofer, 1989)

Each point on the wall (or shipwreck, depending on the application) is imaged by at least 4 rays. The use of convergent imagery, as opposed to a strip-like network similar to that employed in aerial photogrammetry, serves two purposes. First, strongly convergent intersections at each target are ensured. Second, the greater longitudinal coverage of each of the convergent images coupled with the strong geometry, reduces the danger of warping in the long network, a problem encountered in the shipwreck project.

5.4 Fieldwork

It was decided that a network of control points would be set up around the wreck, in order to be able to reference all the mapping to the Gauss Conform (L_o 19) coordinate system. Figure 5.4 shows the distribution of control points around the wreck.

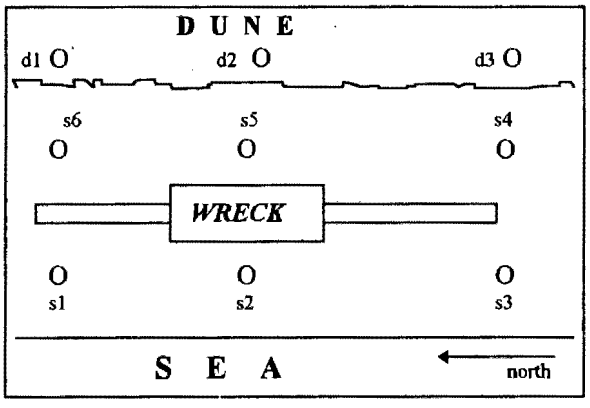

Figure 5.4 Distribution of control points around shipwreck

A kinematic GPS survey was carried out to fix three points, d1, d2 and d3, which were placed on a sand dune overlooking the site. These three points were, in turn, used to fix 6 other points, s1 - s6 by means of a theodolite survey.

The shipwreck itself is approximately 20m long, and is made up of 3 distinct sections, two 5m logs forming the outer sections, and a bigger (10m) middle section comprising lots of broken pieces of wood. In order to transfer control onto the surface of the wreck, 2 control frames of the type described in section 4.4 were placed at each end of the individual sections, and a few points on them were surveyed in from points s1 - s6. The rest of the points on the calibration frame were transformed into the ground coordinate system. In order to tie the imagery together, circular retro-reflective targets were glued to drawing pins, and these were placed on the wreck. The results of mapping one of the sides of section 1(the outer section on the northern side) is presented.

For section 1, a total of 62 uniformly distributed targets were placed on the surface of the wreck; these serve the same purpose as "pass-" or "tie-points" in an aerotriangulation exercise.

5.5 Photography

As mentioned earlier, the focal length of the video camera must remain constant, and was fixed at an appropriate distance. Image capture started with the filming of one of the coordinate reference frames, then the section of the wreck, and finished off on the second coordinate reference frame. The photography of the control frames at the beginning and end of the section would be used for four purposes:

1. for calibrating the camera;
2. defining the coordinate system;
3. transferring control to the tie-points on the wreck;
4. providing a set of control points on which to begin and end the photogrammetric triangulation.

The photography was done in an analogous way to the aerial photography in section 4, in which the camera axis was kept approximately perpendicular to the wreck while image acquisition took place.

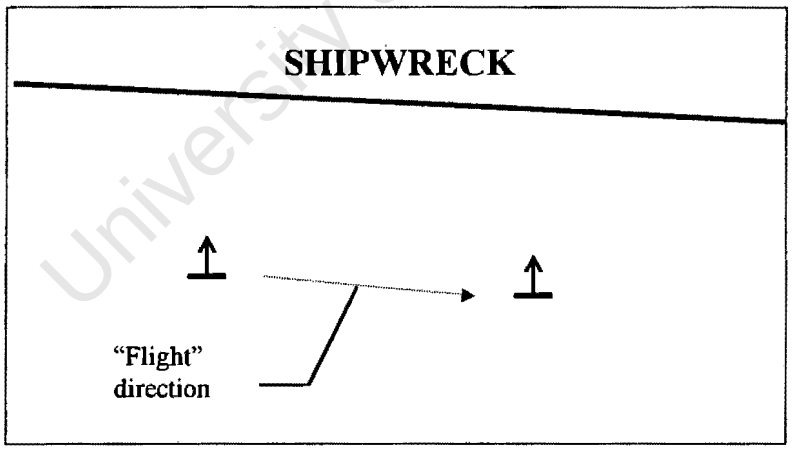

Figure 5.5 Perpendicular photography

It was later realised that doing the photography in a "panning" motion (as shown in Figure 5.6) over the object yields more favorable results.

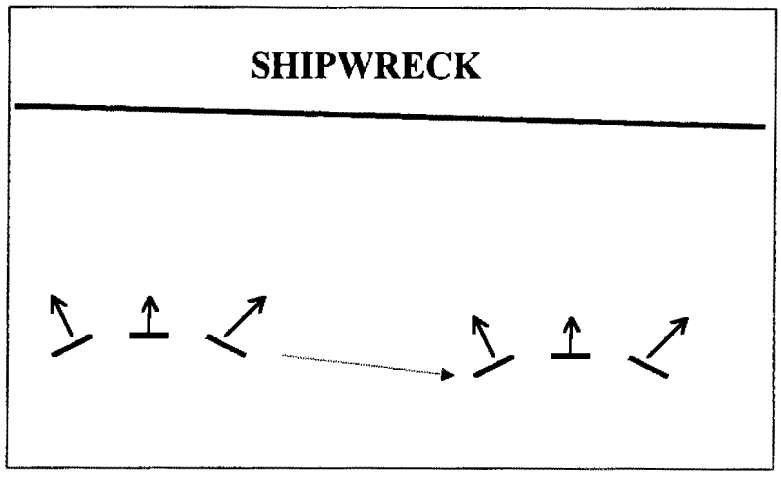

Figure 5.6 Panning motion

This improvement in the results is due to the better resulting geometry of the intersecting rays, as shown in figures 5.7:

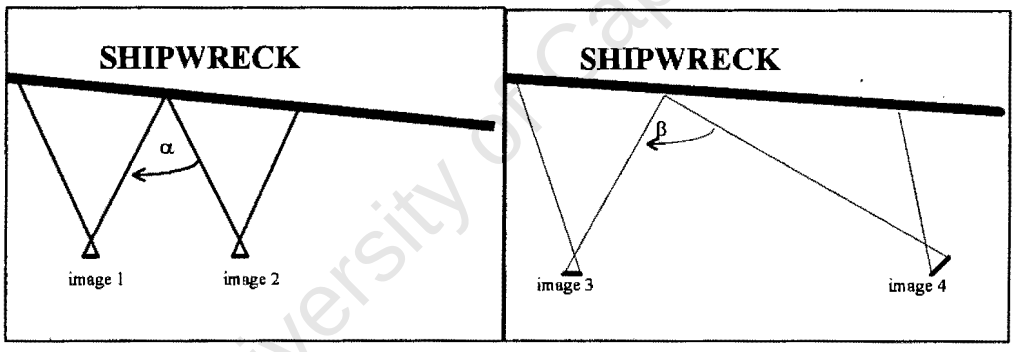

Figure 5.7 Geometry of intersecting rays

If the photography is done in a panning motion, then the resulting intersection of corresponding rays yields an angle of intersection that is closer to 90°, and this represents a better geometry for the calculation of space intersections and therefore a more stable solution in the final bundle adjustment. Another important point is that photography should be done from as many different perspectives as possible (whilst still being economically acceptable), however, if a mosaic is going to be produced, vertical photography of the shipwreck should also be done. Thus, a mix of convergent imagery (for proper tie point determination in the aerotriangulation), and vertical imagery (for the production of a mosaic) is needed.

5.6 Frame-grabbing and Camera Calibration

After the photography, a suitable strip of images of the section of the wreck was frame-grabbed, along with images of the control frame for the calibration. Calibration was carried out in the same way as described in Section 2.3. The strip of images must be captured such that there is adequate overlap between adjacent images, and as many different views as possible of the same portion of the object should be sought. 37 images of the section were acquired, with varying degrees of overlap between them. Nine images of the control frame were sufficient to carry out the calibration.

5.7 Calculation of tie points and Camera Position Location

The calculation of the preliminary exterior orientation parameters and tie point positions are determined by means of a sequential "folding-in, folding-out" process (Smit, 1996). In an initial step, the camera positions of the first few images of the strip are calculated by means of a DLT resection, using at least six reference frame points. In the second stage, the object space coordinates of tie points visible in the same imagery are calculated by means of a DLT intersection. These tie points serve as control points for the adjacent imagery, and the process is continued for all the images. The process is completed by ending on the second coordinate reference frame at the end of the strip.

Finally, a block bundle adjustment on all the imagery in a strip is carried out. This is done as a constrained network adjustment, with the calibrated interior orientation camera parameters, and the coordinates of the control frames held fixed. The bundle adjustment program that was used, is called "photonet"(see Table 3.1). The bundle adjustment produces final object space coordinates of the tie points, as well as the final camera positions. From the bundle adjustment, the root mean square (RMS) standard deviations of the coordinates of the 62 tie points (in mm) were:

RMS of X coordinate	RMS of Y coordinate	RMS of Z coordinate
2.7	4.5	2.0

Table 5.1 RMS errors of tie points (in mm)

Since the Gauss Conform coordinate system was used, the XZ plane roughly corresponded to the plane of the imagery. The worst RMS (4.5 mm) corresponds to the depth dimension (Y coordinate) due to the fact that the imaging geometry is worst for calculating depth. This result was expected as the accuracy of the X and Z coordinate calculations is much higher because of the good (X and Z coordinate) distribution of tie points in the imagery.

5.8 Feature Extraction / Edge Detection

To generate a DEM of the surface of the wreck, a dense cloud of interest points in the images needs to be extracted. Such points can be found on the basis of changes in grey level gradients in the image. They can be automatically extracted from an image by means of an interest operator. Edges can be seen as distinguishing features of an object and consequently can be used to define the shape of textured surfaces. Abrupt grey-scale value changes could indicate the location of an edge in an image. Three interest operators were tested on imagery, namely the Canny Filter, the Sobel Filter and the Maximum Gradient Filter.

In the selection of points of interest, which represent the surface being measured, point density must be balanced against the demand on computational time during the subsequent image matching process (section 5.9). The number of selected points must be sufficient to accurately represent the surface while avoiding unnecessary point density leading to

unacceptable computational times. Tests showed that unwanted point accumulations can occur in local areas as a result of unsuitable interest operators (Smit, 1997). The three abovementioned filters were tested by inspection, and the *Maximum Gradient Filter* and the *Sobel Filter* were found to be most effective in extracting interest points from the imagery.

5.8.1 Maximum Gradient Interest Operator

The Maximum Gradient Filter makes use of the neighbouring pixels surrounding the pixel of interest to the locate the magnitude and direction of the maximum gradient at the pixel of interest. The maximum gradient in a 3x3 neighbourhood around the pixel of interest is located, as follows:

$$gradient = \frac{grey_j - grey_i}{2} \qquad \textbf{(5.1)}$$

where

$i = 1, 2, 3, 4$, and

$j = 9, 8, 7, 6$

for the gradients calculated in four image directions as seen in Figure 5.8:

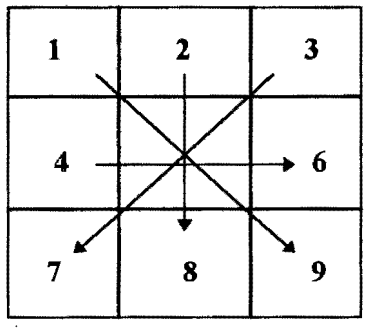

Figure 5.8 Gradient directions used

The direction in which the maximum gradient occurs is used as being perpendicular to the edge direction, at the pixel of interest, for subsequent sub-pixel edge location.

5.8.2 Sobel Interest Operator

The Sobel Filter uses a 3x3 pixel area to compute the gradient at the pixel of interest in two edge directions. It is a low level edge detector that convolves two 3x3 pixel masks (also called *kernels*) for the x and y gradients. It weights the pixels closest to the centre of the kernel to provide smoothing, decreasing its sensitivity to noise in the image. The masks for the Sobel edge detector are shown in Figure 5.9:

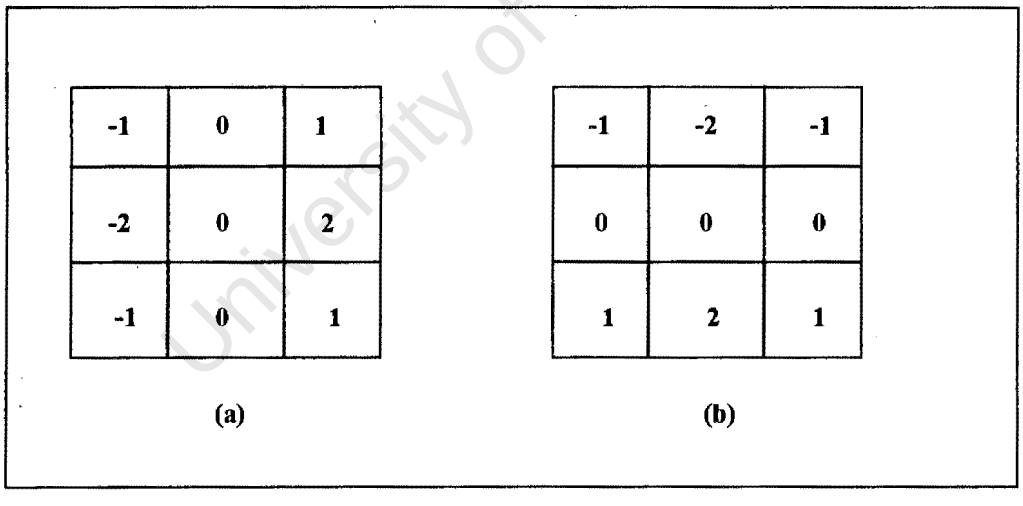

Figure 5.9 Masks used by the Sobel operator

The mask to be used for the x gradient is shown in Figure 5.9(a) and the mask for the y gradient is shown in Figure 5.9(b). Mathematically the convolution can be expressed in the form:

$$G(x_0, y_0) = \sum_{j=-1}^{1} \sum_{i=-1}^{1} M(i,j) g(x_0 + i, y_o + j) \quad \textbf{(5.2)}$$

where

$G(x_0, y_o)$ is the gradient in either the x or y direction,

M(i,j) is the Sobel value at point (i,j), and

g(x,y) is the grey-scale value at point (i,j).

The Sobel operator is convolved with the image to produce a gradient image in both horizontal and vertical image directions. These gradient images' intensity profiles are searched for maxima and minima, which correspond to the edge locations.

5.8.3 Sub-Pixel Edge Location

The Maximum Gradient Filter and the Sobel Filter are able to locate edges to single pixel accuracy as well as providing edge directions. Sub-pixel edge determination was adopted on the assumption that increased subsequent image matching accuracies would be achieved by resampling the search patch to the sub-pixel position of an edge point. The *method of preservation of moments* is subsequently used to locate edge positions to sub-pixel accuracy (Tabatabai and Mikhail, 1984).

A sample scan line across a step edge in the absence of noise is characterised by a set of grey values x_i's, $i = 1, 2, ..., n$, that are either monotonically non-decreasing or non-increasing. On the other hand, an ideal edge is a sequence of one brightness value h_1, followed by a sequence of another brightness value h_2 (see figure 5.10):

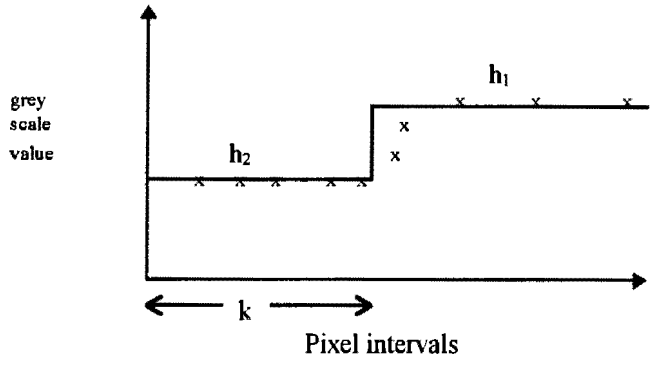

Figure 5.10 An ideal step edge matched to sample data

Here an operator is defined that, when applied to obtained edge data, locates an ideal step edge, such that the first three sample moments of the input data sequence, defined as

$$\overline{m}_i = \frac{1}{n}\sum_{j=1}^{n} x^i{}_j \qquad \textbf{(5.3)}$$

for $i = 1, 2, 3$ are preserved.

Letting k denote the number of h_1 values from the beginning of the sample to the ideal edge, there are three equations to solve for the unknowns h_1, h_2 and k (Davis, 1988):

$$\overline{m}_1 = \frac{kh_1 + (n-k)h_2}{n} \qquad \textbf{(5.4)}$$

$$\overline{m}_2 = \frac{kh_1^2 + (n-k)h_2^2}{n} \qquad \textbf{(5.5)}$$

$$\overline{m}_3 = \frac{kh_1^3 + (n-k)h_2^3}{n} \qquad \textbf{(5.6)}$$

The edge of the first pixel is located at $j = 1/2$ and subsequent pixels have a spacing of one. We are then able to obtain a sub-pixel measurement k of edge location. The solution is

$$k = \frac{1}{2}n(1 - \frac{c}{\sqrt{4+c^2}}) \tag{5.7}$$

where

$$c = \frac{3\overline{m_1 m_2} - 2\overline{m_1}^3 - \overline{m_3}}{\sigma^3} \tag{5.8}$$

and

$$\sigma = \sqrt{\overline{m_2} - \overline{m_1}^2} \tag{5.9}$$

$$h_1 = \overline{m_1} + \beta\sigma \tag{5.10}$$

$$h_2 = \overline{m_1} - \frac{\sigma}{\beta} \tag{5.11}$$

where

$$\beta = \sqrt{\frac{k}{n-k}} \tag{5.12}$$

5.9 Image Matching

Once the points of interest have been detected in the "template" image, they need to be matched with corresponding points in the "search" images to calculate their object space coordinates (Smit, 1997). The matching algorithm uses epipolar geometry, which supplies estimates to the initial image coordinates of the corresponding points in the search images. This is followed up by least squares greyscale matching to determine the final positions. This is the basis of a Multi-Photo Geometrically Constrained Image Matching (MPGC) algorithm that was used, as described in section 2.9.2.

The result of MPGC is a DEM file with the object-space X, Y and Z coordinates of the final matched points. The matching was done in sections, and the DEMs produced for all these sections were combined to form one final DEM of the entire section of the wreck. This, along with a contour plan of the wreck is shown in Figure 5.11:

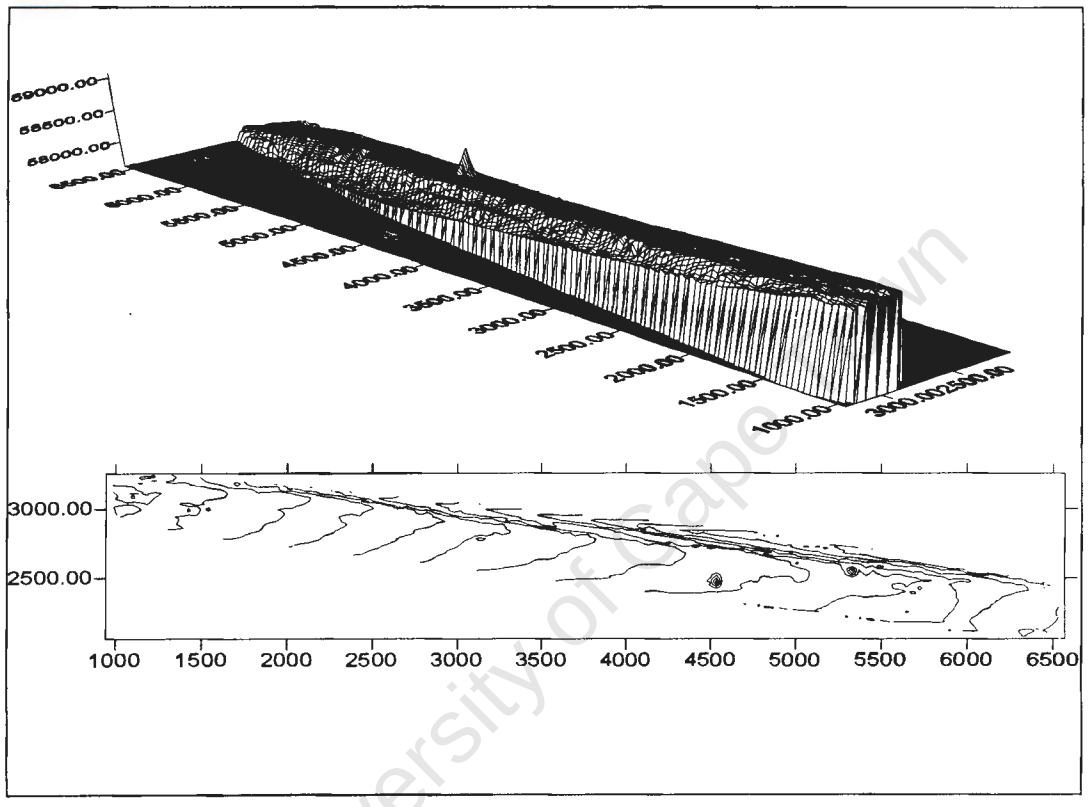

Figure 5.11 DEM and contour plan of section of shipwreck

The "spike" in the DEM represents a nail that was protruding out of the surface of the shipwreck. Corel Draw! was used to manually mosaic all the template images together, and this is shown in Figure 5.12:

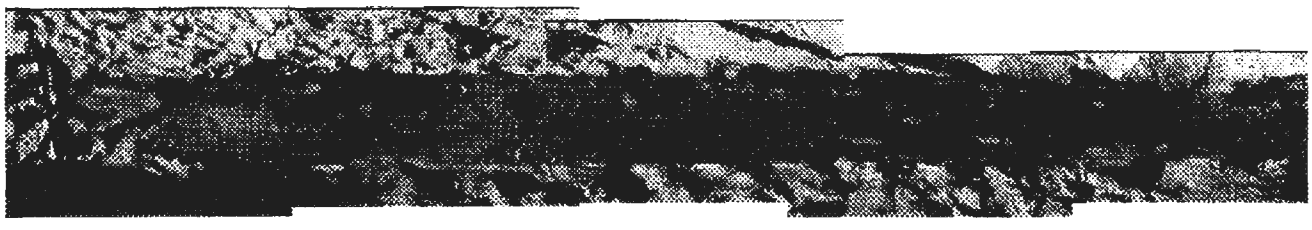

Figure 5.12 Mosaic of mapped section of shipwreck

This mosaic enables one to obtain a good overall impression of the structure of the shipwreck, and along with the DEM and contour plan, could be used by archaeologists to obtain spatial information about features on the wreck. A description of how a GIS can be used for archaeological analysis is given in section 6.3.

CHAPTER 6

Analysis and Discussion

6.1 End Products

The technique of using video imagery for mapping proves to be a cheap, quick alternative for both aerial and close-range photogrammetry, and a great advantage is the wide variety of end products that can be produced. Hard-copy photographs or images, even without any processing, are the simplest and most direct photogrammetric products. We can characterize the products as either graphical or digital. Graphical products can be obtained directly on-line from plotters, or they can be printed from digital data stored on disk. Digital data can be used to make a variety of different graphic plots, depending on their purpose.

Graphical products are maps and drawings of various types such as contours, profiles, cross-sections and different engineering projections such as plans, elevations and side views. Digital products could include all of these graphical products, but in digital form, in addition to lists of object point coordinates and DEMs. Given a DEM, it is possible to derive a number of other digital products, including sections, projections, and various views such as a perspective. Other useful data can also be digitally extracted, including areas, volumes, and orientations of planes.

6.1.1 Rectified Imagery

Rectified (or rubber-sheeted) imagery provides a very useful source of information, as was shown with the rubber-sheeted imagery produced of Marconi Beam (chapter 4).

Some care should be taken when carrying out the photography if the desired end-product is a rectified image, or a mosaic of rectified imagery. A strip of perpendicular photography of the object should be done, i.e. the camera's focal plane should be aligned parallel to the face of the object being measured. In this regard, a perpendicular image will always be a more productive source of information than a significantly tilted one. This product is of great value where a high level of surface detail is present, thus making a traditional survey too time-consuming.

6.1.2 Overlays

Overlays have proven to be valuable end products, as in chapter 4 and 5. Overlaying a mosaic produced by rapid mapping with various other types of data proves its usefulness in terms of quick assessment of a situation. Damage assessment, relocation logistics and updating of databases can be accomplished in minimal time using the technique outlined in chapter 4.

Overlays of different types of archaeological data could assist an archaeologist in drawing important conclusions about an artifact, such as the dating of a shipwreck (see section 6.3).

6.2 Advantages of Digital Techniques

The use and processing of digital imagery has many advantages over traditional, analogue techniques. Digital imagery provides for greater data exploitation than manual interpretation, e.g. although the human eye is very good at distinguishing colours (up to about 5 million), computers are still better, being able to distinguish 16 million colours.

When considering film photography, information is gained from both objective and subjective analysis. Objective data comes from such things as granularity, density, density

range, exposure scale, wavelength and resolution. Subjective data is based mainly on human visual perception, and includes terms such as contrast, definition, sharpness, colour, graininess, and just noticeable differences of tone or colour. We feel more comfortable with objective values mainly because they can be quantified and compared, but since all photographic images are viewed with the human eye, the subjective element is unavoidable. The main strength and purpose of digital imagery is its capacity for computer processing, where subjectivity is precluded from the analysis. Whether originally digital or digitised, an image in this form provides greater flexibility. A host of digital image processing operations can be employed either to improve the quality of the image in some way or to automatically extract data from it.

Digital image processing is consistent, while manual interpretation results, being subjective, may vary from day to day based on interpreter attitude, stress level, or health. Different interpreters often produce significantly different results. In contrast, the computer's results are usually exactly repeatable. The ability to store and manipulate large amounts of tedious information often allows the computer to do a superior job in this regard.

One of the most important advantages of digital photogrammetric techniques is automation of processes that are otherwise time consuming, tedious and repetitive. Softcopy photogrammetry allows for automation in the point measurement process, as described in section 5.7. This greater ease in point measurement means that it is economical to measure more points and therefore have a more robust, accurate and reliable solution. Digital imagery also allows for greater accuracy in the point transfer process. It has been reported that with points in areas of good image texture, manual transfer accuracy of 1/3 to 1/5 of a pixel, and automated transfers of 1/10 of a pixel are typical (Kaiser et al, 1996). Automation is also accomplished with other tasks in the digital photogrammetric process, including:

- orientation of imagery (interior and relative)
- DTM generation by automated image matching,

6.5 Overall Assessment of System

6.5.1 Problems encountered and Possible Solutions

6.5.1.1 UrbanModeler

In the UrbanModeler project, the initial aim was to produce a Digital Elevation Model from the camcorder imagery, but this was not able to be carried out because the resolution of the frame-grabbed imagery was too low. Individual shack identification from the mosaic was difficult because of the low resolution. A solution to this could be to acquire the imagery at a lower flying height, but that would mean that a greater amount of GCPs would be needed to rubber sheet the imagery, as each image would cover a smaller area on the ground, and a second order polynomial transformation, for example, requires a minimum of 6 GCPs (with x and y coordinates) in each image. The obvious solution to this shortcoming of the video imagery would be the use of a higher resolution video camera or more flexible frame-grabber, but this in turn would mean a higher cost of equipment.

6.5.1.2 Shipwreck

The principal problem encountered in the shipwreck project was the lack of sufficiently convergent imagery. This was due to the fact that the photography was not executed in a panning motion, as described in section 5.5. The most important part of an exercise of this nature is the design of the image acquisition geometry, and every effort should be made to ensure that the geometry of the imagery to be processed is strongly convergent, and there are enough tie points in the imagery.

In some sections of the shipwreck, the sequential "folding-in, folding-out" process, used in conjunction with the bundle adjustment which served to refine the calculated camera positions, failed to deliver a solution. This is either because of the weak geometry of intersecting rays, or a lack of an adequate number of tie points in the imagery. If this occurred, then additional imagery had to be obtained by frame-grabbing from the video tape, and these new images had to be included. This then led to a solution, due to the greater redundancy and the better geometry introduced.

In an aerotriangulation exercise, although the required number of tie points in one image is 6, it was found in this project that approximately 10 tie points per image were required to deliver good results. One should thus remember that adding extra tie points onto the surface of the measured object (in this case, the retroreflective targets placed on the wreck) can only increase the accuracy of the final solution. The little extra time spent in placing targets onto the object is adequately compensated for in the time saved during data processing.

Another problem encountered was in the transformation of the coordinates of the points on the control frames into the Gauss Conform coordinate system. The transformation of the frame targets into the ground coordinate system was unnecessary in that it would have been a better idea to work in a local coordinate system. The referencing of the wreck to the national coordinate system could have been accomplished by means of coordinating three points on the wreck (one on either end, and one in the middle) by means of GPS or theodolite survey. Thus the whole process of surveying in the control frames need not have been done, cutting down fieldwork time considerably.

6.5.2 Ease of Use

6.5.2.1 UrbanModeler

In the UrbanModeler project, the main software package that was used for the production of the mosaics was ERDAS. ERDAS was found to be a user-friendly package, and the georeferencing and mosaicing of imagery posed no difficulties. The only requirements were the coordinates of the GCPs and the equivalent original imagery as obtained by frame-grabbing. All imagery used in ERDAS is first converted to ERDAS format (with a ".img" extension). The ERDAS image converter can handle a wide variety of image formats. The mosaics of Marconi Beam took a few hours to produce, so ERDAS has proved, in this project, to be an ideal aid for rapid mapping.

6.5.2.2 Shipwreck

Fieldwork and photography for a project such as the shipwreck mapping can be accomplished relatively easily if the principles relating to obtaining convergent imagery and adequate tie points are adhered to.

Photogrammetric processing by means of the in-house software developed in the Department of Geomatics is not as easy to carry out as it is using the ERDAS package, as the former was not designed as a commercial package with a user-friendly interface. Arriving at object-space coordinates of points on the surface of the wreck required the use of a number of DOS-based programs, all listed in Table 3.1. Surface mapping can be accomplished rapidly if the user is familiar with the use of this software.

CHAPTER 7

Conclusions

The technique of using video imagery for mapping proved to be a cheap, quick, low accuracy alternative in the two test cases. For both the case studies, the cost of mapping was far less than it would have been using traditional, film-based mapping techniques. With a suitable camera - framegrabber system, one can attain adequate mapping accuracy for a host of applications in both the aerial and close-range photogrammetry domains. The test cases exhibited sub-metre planimetric point positioning accuracies for the informal settlement mapping, and sub-centimetre accuracies for the production of surface maps of the shipwreck.

From the final mosaic produced of the squatter settlement, individual shacks can be identified. This is an invaluable tool for developers in managing the relocation of thousands of people in a squatter camp. Overlaying the boundary of the fire over the mosaic is a good tool in damage assessment, and an estimate of the number of people affected can easily be inferred.

Attaining sub-centimetre accuracy for such a large amount of points as was calculated on the surface of the shipwreck, as well as the speed with which it was accomplished, indicates that the video mapping system is ideally suited for artifact mapping, as well as other applications that require minimal contact with the object in the shortest possible time. End-products like the mosaic in Figure 5.12, together with the contour plan or DEM (Figure 5.11) can be used by archaeologists to measure surfaces and dimensions of artifacts, which are helpful aids in identification and archaeological analysis. The same method can be extended to underwater work, although there are a few additional aspects that have to be considered (see section 6.4).

Digital imagery can be easily captured, stored, manipulated and analysed. Moreover, CCDs are geometrically stable sensors that allow reliable and accurate measurement and are relatively immune to environmental effects. A further benefit of video rates of image acquisition allows the possibility of real time vision systems for the recording of dynamic events, such as tracking an object in free flight or monitoring a production line system (Shortis, 1995).

Some care should be taken when choosing the hardware components of such a system, in particular, the camera-framegrabber combination. A framegrabber should be able to capture images at the exact resolution of the CCD chip in the camera. This will overcome the problem of the image information being lost in the frame-grabbing process. Also, it has been proven (Hoeflinger, 1994) that best results for analogue image storage on video tape can be attained when the tape is played back on the same camera it was recorded with. In conclusion, the results obtained are very good for a tool that was not designed for image metrology.

BIBLIOGRAPHY

Abdel-Aziz, Y.I., Karara, H.M. : "Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry" (1971) In: Proceedings of ASP/UI Symposium on Close Range Photogrammetry, Urbana, Illinois, USA, 1-8

Baltsavias, E. : "Multiphoto geometrically constrained matching" (1991) Phd Thesis, Institut für Geodäsie und Photogrammetrie ETH, Zürich

Beyer, H.A. : "Calibration of CCD Cameras for Machine Vision Robotics" (1990) In: Society of Photo-Optical Instrumentation Engineers, Vol 1197. Automated Inspection and High Speed Architectures III, 88-98

Beyer, H.A : "Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System" (1992) Phd Thesis, Institut für Geodäsie und Photogrammetrie ETH, Zürich

Brown, D.C. : "A Solution to the General Problem of Multiple Station Analytical Stereotriangulation" (1958) RCA Data Reduction Technical Report No. 43, Patrick Air Force Base, Florida, USA

Buechel, S., Nichols, J., Kruse, B., Binkley, K. : "Noise Characterization and Effects on the Image Processing of Video Grabs" (1995) In: 15th Biennial Workshop on Videography and Color Photography in Resource Assessment, American Society for Photogrammetry and Remote Sensing, Terre Haute, Indiana, 346-355

Clark, D. : "Plane and Geodetic Surveying for Engineers" (1961) Whitefriars Press Ltd, London

Cooper, M.A.R., Robson, S. : "Theory of Close Range Photogrammetry", (1996) In: Close Range Photogrammetry and Machine Vision, Whittles Publishing, Scotland

Davis, M. : "Target Centreing and Image Matching in a Near Real-Time Photogrammetric System" (1988) unpublished Honours Thesis, University of Cape Town

Dahler, J. : "Problems in Digital Image Acquisition with CCD cameras" (1987), Proceedings of the Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, 48-59

Dowman, I.J. : "Fundamentals of Digital Photogrammetry" (1996) In: Close Range Photogrammetry and Machine Vision, Whittles Publishing, Scotland

ERDAS, Inc : "ERDAS Field Guide", (1994) Third Edition

Everitt, J.H., Escobar, D.E., Davis, M.R., Summy, K.R., Judd, F.W. : " Using Spatial Information Technologies for Detecting and Mapping Insect Infestations and Plant Species" (1995) In: 15th Biennial Workshop on Videography and Color Photography in Resource Assessment, American Society for Photogrammetry and Remote Sensing, Terre Haute, Indiana, 63-72

Forte, M. : "Scientific visualization and archaeological landscape : the case studey of a terramara, Italy" (1995) In: Archaeology and Geographical Information Systems, Taylor and Francis, London

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

Gifford, J.A : "Videography and Geographical Information Systems for recording the excavation of a prehistoric underwater site" (1993) The International Journal of Nautical Archaeology and Underwater Exploration, 22.2 : 167-172

Gifford, J.A. : "Mapping Shipwreck Sites by Digital Stereovideogrammetry" (1997) Proceedings, CUA Meeting, Corpus Christi

Gold, D. : "The Investigation and Design of a Machine Vision System for the Detection and Control of the Separation in a Spiral Ore Concentrator" (1991) unpublished MSc thesis, University of Cape Town

Green, J. : "Maritime Archaeology - a technical handbook" (1990) Academic Press, London

Hoeflinger, W., Beyer, H. : " Radiometric and Geometric Performance Evaluation of a S-VHS Camcorder" (1994), Proceedings of ISPRS Commission V Symposium, March 1994, Melbourne, Vol 30, Part 5, 191-197

Kaiser, D., Miller, S.B. : "Comparison of Softcopy vs. Hardcopy Photogrammetry for DTM Collection, Editing and Quality Control" (1996) In: Digital Photogrammetry, An Addendum to the Manual of Photogrammetry, American Society for Photogrammetry and Remote Sensing, Maryland, 142-144

Karara, H.M. : "An Introduction to Non-Topographic Photogrammetry" (1989) In: Non-Topographic Photogrammetry, American Society for Photogrammetry and Remote Sensing, 1-5

Kolbl, O. : "Scanning and State-of-the-Art Scanners" (1996) In: Digital Photogrammetry, An Addendum to the Manual of Photogrammetry, American Society for Photogrammetry and Remote Sensing, Maryland, 3-39

Lenz, R., Fritsch, D. : "Accuracy of videometry with CCD sensors" (1990) ISPRS Journal of Photogrammetry and Remote Sensing, Vol 45, 90-110

Linden, D., Hoffer, R., Pywell, H., Bobbe, T. : "Automated Digital Mosaicing of Airborne Videography" (1995) draft report for Department of Forest Sciences

Marble, D.F. : "Geographical Information Systems: an overview" (1987) In: Ripple, W.J. (ed.) Geographical Information Systems for Resource Management: A Compendium. American Society for Photogrammetry and Remote Sensing, 2-8

Mason, S. : "Expert System-Based Design of Photogrammetric Networks" (1994) Institut für Geodäsie und Photogrammetrie ETH, Zürich

Mason, S., Rüther, H., Smit, J. : "Investigation of the Kodak DCS460 for low-cost, local area mapping" (1997) submitted to ISPRS Journal of Photogrammetry and Remote Sensing, Jan. 1997

McGlone, J.C. : "Analytic Data-Reduction Schemes in Non-Topographic Photogrammetry"(1989) In: Non-Topographic Photogrammetry, American Society for Photogrammetry and Remote Sensing, 37-55

Mitchell, T.A., Neale, C.M.U., Schowengerdt, R.A. : "Geometric Rectification of Multi-Temporal Multi-Band Videographic Imagery" (1995) In: 15th Biennial Workshop on Videography and Color Photography in Resource Assessment, American Society for Photogrammetry and Remote Sensing, Terre Haute, Indiana, 100-105

Motala, S., Mason, S., Rüther, H., Smit, J.L. : "Rapid Mapping using Camcorder Imagery" (1997) CONSAS '97, Durban

Muessig, H : "Mapping Archaeological Sites using Ground-Based Photogrammetry" Close-Range Photogrammetry and Surveying : State-of-the-Art, American Congress on Surveying and Mapping; 1984 Fall Convention : 394-404

Newton, I. : "Underwater Photogrammetry" (1989) In: Non-Topographic Photogrammetry, American Society for Photogrammetry and Remote Sensing, 147-163

Panja, K.V., Hardy, T.B., Neale, C.M.U. : "Comparison of Multispectral Videography Based Classification of Mesoscale Habitats and Ground Based Mapping under Turbid Riverine Conditions" (1995) In: 15th Biennial Workshop on Videography and Color Photography in Resource Assessment, American Society for Photogrammetry and Remote Sensing, Terre Haute, Indiana, 158-166

Raynor, J.M., Seitz, P. : "The Technology and Practical Problems of Pixel-Synchronous CCD Data Acquisition for Optical Metrology Applications" (1990), SPIE Vol. 1395, Close-Range Photogrammetry Meets Machine Vision, 96-103

Rebikoff, D. : "Underwater Archaeological Exploration" (1961), Paper presented at the Third International Congress on Underwater Archaeology, Barcelona, Spain. Reprinted 1971 by the International Institute of Ligurian Studies, Bordighera, Italy

Rebikoff, D. : "Two Eyes: Stereovideo and Videogrammetry" (1984), Keynote paper presented at Marine Technology Society ROV Conference, San Diego, CA

Richards, J.A. : "Remote Sensing Digital Image Analysis, An Introduction" (1986), Springer-Verlag, Berlin

Rule, N : "The Direct Survey Method (DSM) of underwater survey, and its application underwater" The International Journal of Nautical Archaeology and Underwater Exploration (1989), 18.2 : 157-162

Rule, N : "Some techniques for cost-effective three-dimensional mapping of underwater sites" (1995) In: Computer Applications and Quantitative Mehods in Archaeology, Bar International Series, Staffordshire University

Saff, G : "Claiming a Space in a Changing South Africa : The "Squatters" of Marconi Beam, Cape Town" (1996), Annals of the Assoc. Americ. Geographers, 86(2) : 235-255

Sersland, C.A., Johnston, C.A., Bonde, J. : "Assessing Wetland Vegetation with GPS-Linked Color Video Image Mosaics" (1995) In: 15th Biennial Workshop on Videography and Color Photography in Resource Assessment, American Society for Photogrammetry and Remote Sensing, Terre Haute, Indiana, 53-62

Sharfman, J. : "The Application of a Geographical Information System on a Maritime Archaeological Project : The Example of the Excavation of the VOC Ship 'Oosterland'(1697)" (1994) unpublished Honours Thesis, Department of Archaeology, University of Cape Town

Shipek, C.J. : "Deep-Sea Photography in Support of Underwater Acoustic Research" (1967) In: Deep-Sea Photography, J.B. Hershey, editor. John Hopkins Press, Baltimore, MD

Shortis, M.R., Snow, W.L. : "Calibration of CCD cameras for field and frame capture modes" (1995) Conference on Digital Photogrammetry and Remote Sensing '95, St. Petersburg-Great Lakes, Russia

Shortis, M.R., Beyer, H.A. : "Sensor technology for digital photogrammetry and machine vision" (1996) In: Close Range Photogrammetry and Machine Vision, Whittles Publishing, Scotland, 106-155

Smit, J.L; Rüther, H; Siebrits : "The 3D Mapping of a Textured Surface using digital Photogrammetric Techniques" (1996), ISPRS Vol. 31(B4), 728-733

Smit, J.L. : "The Three Dimensional Measurement of Textured Surfaces Using Digital Photogrammetric Techniques" (1997), unpublished Phd thesis, University of Cape Town

Streilein, A., Gaschen, S. : "Comparison of a S-VHS Camcorder and a High-Resolution CCD-Camera for use in Architectural Photogrammetry" (1994) Proceedings of ISPRS Commission V Symposium, March 1994, Melbourne, Vol 30, Part 5, 382-389

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

Thomas, M.: "Georeferencing Airborne Video Imagery Independently from maps using differentially-corrected Global Positioning Coordinates" (1996) South African Journal of Geo-Information, 17(2), 27-34

Thomas, M. : "The Development and Assessment of a GPS integrative technique for georeferencing video remote sensing imagery through trials" FOR-I 650, unpublished CSIR report

van der Vlugt, G : "Algorithms and Design Aspects of an Automated Vision Based 3-D Surface Measurement System" (1995) unpublished Phd thesis, University of Cape Town

Warner, W.S., Graham, R.W., Read, R.E. : "Small Format Aerial Photography" (1996), Whittles Publishing, Caithness, Scotland

Werz, B.E.J.S; Martin, C.G : "Hydrographic Surveys in Aid of Maritime Archaeological Research : The Example of the Table Bay Project" The Hydrographic Journal (1994) 72 : 3-15

Werz, B.E.J.S : "The Excavation of the Oosterland in Table Bay : The first systematic exercise in maritime archaeology in Southern Africa" South African Journal of Science, (Feb 1992) 88 : 85-89

Werz, B.E.J.S : "A Maritime Archaeological Project in Table Bay" South African Archaeological Bulletin (1990) 45 : 121

Werz, B.E.J.S : "Maritime Archaeological Project Table Bay : Aspects of the First Field Season" South African Archaeological Society Goodwin Series (1993) 7 : 33-39